б 28
К64

Профессиональное образование

Учебник

А. С. Коничев
Г. А. Севастьянова

МОЛЕКУЛЯРНАЯ БИОЛОГИЯ

4-е издание



ПЕДАГОГИЧЕСКОЕ ОБРАЗОВАНИЕ

ACADEM'A

ИЗДАТЕЛЬСКИЙ ЦЕНТР

5 28
К64

Высшее профессиональное образование

БАКАЛАВРИАТ

А. С. КОНИЧЕВ, Г. А. СЕВАСТЬЯНОВА

МОЛЕКУЛЯРНАЯ БИОЛОГИЯ

Учебник
для студентов учреждений
высшего профессионального образования,
обучающихся по направлению подготовки
«Педагогическое образование» профиль «Биология»

4-е издание, переработанное и дополненное

Қ. БӘЙСЕМБАЕВ АТЫНДАҒЫ ҚЫЛЫМИ КІТАПХАНА
ОҚУ ЗАЛЫ
ЧИТАЛЬНЫЙ ЗАЛ
НАУЧНАЯ БИБЛИОТЕКА им. К. БЕЙСЕМБАЕВА

ACADEM'A

Москва
Издательский центр «Академия»
2012

б 21.07Оя73

УДК 577.2(075.8)
ББК 28.007я73
К64

Рецензенты:
чл.-кор. РАН, д-р биол. наук, проф. *А. П. Рысков*;
д-р пед. наук, проф. *В. В. Пасечник* (зав. кафедрой методики преподавания географии, естествознания и экологии Московского педагогического университета)

Коничев А.С.

К64 Молекулярная биология : учебник для студ. учреждений высш. пед. проф. образования / А. С. Коничев, Г. А. Севастьянова. — 4-е изд., перераб. и доп. — М. : Издательский центр «Академия», 2012. — 400 с. — (Сер. Бакалавриат).

ISBN 978-5-7695-9147-1

Учебник создан в соответствии с Федеральным государственным образовательным стандартом по направлению подготовки «Педагогическое образование» профиль «Биология» (квалификация «бакалавр»).

В учебнике кратко изложены история и методы молекулярной биологии, подробно рассмотрены основные направления изучения нуклеиновых кислот и белков: структура геномов про- и эукариот, вирусов и фагов, митохондрий и хлоропластов; подвижные генетические элементы; повреждения и репарация структуры ДНК; молекулярные основы генетической рекомбинации; структура, процессинг и функции различных видов РНК; механизмы и принципы регуляции основных молекулярно-генетических процессов (репликации, транскрипции, трансляции); структура и фолдинг белков; белково-нуклеиновые взаимодействия; молекулярные механизмы регуляции клеточного цикла, канцерогенеза и программируемой клеточной смерти. Значительное место отведено методам генетической инженерии, ее достижениям и перспективам.

Для студентов учреждений высшего педагогического профессионального образования. Может быть полезен преподавателям колледжей и средних школ с углубленным изучением биологии.

661378

УДК 577.2(075.8)
ББК 28.007я73

С. Торайғыров
атындағы ҚМУ-дің
академик С. Бейсембаев
атындағы ғылыми
КІТАПХАНАСЫ

Оригинал-макет данного издания является собственностью Издательского центра «Академия», и его воспроизведение любым способом без согласия правообладателя запрещается

© Коничев А. С., Севастьянова Г. А., 2008
© Коничев А. С., Севастьянова Г. А., 2012, с изменениями
© Образовательно-издательский центр «Академия», 2012
© Оформление. Издательский центр «Академия», 2012

ISBN 978-5-7695-9147-1

ПРЕДИСЛОВИЕ

Настоящая книга представляет собой первый опыт создания учебника для бакалавров естественно-научного образования (профиль «Биология»). Необходимость такого издания связана с включением дисциплины «Молекулярная биология» в Федеральный государственный образовательный стандарт третьего поколения.

В соответствии с принятым стандартом составлена программа по молекулярной биологии для студентов-бакалавров, обучающихся по педагогическим специальностям, в соответствии с которой и подготовлен данный учебник.

В основе учебника лежат материалы, опубликованные авторами в книге «Молекулярная биология», изданной ранее в Издательском центре «Академия» в 2003, 2005, 2008 годах. Вместе с тем предлагаемый вниманию читателей учебник существенно дополнен новейшими сведениями в области ряда разделов молекулярной биологии, включая изучение структуры геномов, строения рибосом, РНК-интерференции и др.

При подготовке книги авторы использовали свой многолетний опыт чтения курса лекций по молекулярной биологии в Московском педагогическом государственном университете, Московском государственном областном университете и Московском городском педагогическом университете.

Предлагаемый учебник не велик по объему и, естественно, не охватывает все содержание современной молекулярной биологии. Тем не менее в нем отражены наиболее фундаментальные достижения этой науки, имеющей исключительно важное значение для познания живой природы.

Освоение содержания учебника требует определенной подготовки в области органической химии и биохимии, однако авторы, сохранив насколько возможно фундаментальный научный подход, старались сделать учебник до-

ступным не только для студентов вузов, но и для учителей и учащихся старших классов специализированных школ, включив в него, в частности, максимальное количество иллюстраций, облегчающих восприятие весьма непростых молекулярно-биологических понятий.

В конце учебника приведен список литературы, рекомендуемой читателям, которые заинтересованы в углубленном изучении отдельных вопросов молекулярной биологии.

Авторы надеются, что учебник поможет студентам получить необходимый минимум знаний по важному разделу современной биологии, и будут искренне благодарны за все отзывы и пожелания.

ВВЕДЕНИЕ

Молекулярная биология возникла во второй половине XX в. Название этой науки чаще всего связывают с именем У. Эстбюри, который в 1939 г. назвал себя «молекулярным биологом». Через два года он же получил первую рентгенограмму ДНК и тем самым положил начало изучению тонкой структуры «самой главной молекулы», впервые выявленной Ф. Мишером еще в 1869 г. Первое официальное упоминание о молекулярной биологии, вероятно, принадлежит У. Уиверу, руководившему отделом естественных наук Рокфеллеровского фонда, который в 1938 г. написал: «В тех пограничных областях, где *химия* и *физика* пересекаются с *биологией*, постепенно возникает новый раздел науки — *молекулярная биология*, начинающая приоткрывать завесы над многими тайнами, окутывающими основные элементы живой *клетки*» (курсив наш. — *Авторы*).

Таким образом было постулировано возникновение нового направления современной биологии, которое интегрировало усилия биологов, химиков и физиков в области изучения объектов живой природы.

Разработка тонких физических и химических методов анализа структуры и функций молекул, свойственных всем живым системам и прежде всего клетке как элементарной и универсальной составляющей всех организмов, имела определяющее значение для рождения молекулярной биологии. Мощным стимулом для ее развития стали успехи и еще в большей мере нерешенные проблемы биохимии, цитологии и генетики. Эти сформировавшиеся к середине XX в. биологические науки создали почву для молекулярной биологии, которая была призвана заняться изучением жизни на молекулярном уровне.

Центром молекулярно-биологических исследований стали работы в области изучения материальных основ наследственности, природы генов и механизмов передачи наследственных признаков из поколения в поколение.

Именно под влиянием генетиков новой формации (Т. Моргана, Н. К. Кольцова, Н. В. Тимофеева-Ресовского и др.) физики-теоретики и экспериментаторы, эмигрировавшие в конце 30-х годов из Европы в США, организовали там так называемую «фаговую группу» во главе

с М. Дельбрюком, которая начала исследования в области молекулярного строения и мутагенеза вирусов и бактериофагов.

Позднее эти работы были существенно развиты в нашей стране Б. Ф. Поглазовым, Н. А. Киселевым и другими учеными. Еще раньше В. А. Энгельгардт совместно с М. Н. Любимовой обосновали молекулярные механизмы мышечного сокращения, а А. Н. Белозерский впервые выделил ДНК из растений, что также относится к фундаментальным вехам становления молекулярной биологии.

Впоследствии именами этих замечательных ученых были названы крупнейшие научно-исследовательские центры: Институт молекулярной биологии АН СССР им. В. А. Энгельгардта и Институт физико-химической биологии им. А. Н. Белозерского.

К началу 50-х годов XX в. в недрах биохимии были получены фундаментальные данные об элементарном строении белков и нуклеиновых кислот, включая сведения о способах организации полипептидных цепей белков, и, что особенно важно, о структуре нуклеотидов и закономерностях их количественного содержания в молекулах ДНК и РНК (Э. Чаргафф).

Именно указанные работы, а также биофизические исследования структуры ДНК, выполненные в Англии методом рентгеноструктурного анализа Розалиндой Франклин и Морисом Уилкинсом, вплотную подвели воспитанника «фаговой группы» Джеймса Уотсона и английского физика Френсиса Крика к раскрытию молекулярной природы генов и механизма их воспроизведения (репликации) в составе ДНК.

Создание модели двойной спирали ДНК и открытие принципа комплементарности стали важнейшим событием современной биологии, вскрывшим фундаментальные принципы функционирования живых систем и определившим дальнейшие направления исследований современной биологии. Современное естествознание обязано именно молекулярной биологии тем, что в период с середины 50-х до середины 70-х годов XX в. с невероятной быстротой были раскрыты природа и основные пути передачи и реализации генетической информации.

Основополагающие открытия молекулярной биологии

1869 — Ф. Мишер (F. Miesher) впервые выделил ДНК из лейкоцитов и молок лосося.

1935 — А. Н. Белозерский выделил ДНК из растений.

1939 — В. А. Энгельгардт открыл АТРазную активность миозина.

1940 — У. Эстбюри (W. Astbury) получил первую рентгенограмму ДНК.

1944 — О. Т. Эвери (O. T. Avery) установил, что ДНК (а не белок, как полагали ранее) является носителем генетической информации.

1951 — Л. Полинг и Р. Кори (L. Pauling, R. Corey) обосновали существование основных типов укладки аминокислотных остатков в полипептидных цепях белков (α-спираль и складчатый β-слой).

1953 — Дж. Уотсон и Ф. Крик (J. Watson, F. Crick) создали модель двойной спирали ДНК на основе рентгенограмм, полученных Р. Франклин и М. Уилкинсом (R. Franklin, M. Wilkins).

1953 — Ф. Сангер (F. Sanger) расшифровал первичную структуру инсулина быка.

1956 — А. Корнберг (A. Kornberg) открыл ДНК-полимеразу.

1957 — А. Н. Белозерский и А. С. Спирин предсказали существование мРНК.

1960 — Дж. Кедрью (J. Kendrew) впервые описал трехмерную структуру миоглобина кашалота, а М. Перутц (M. Perutz) — структуру гемоглобина.

1960 — одновременно в нескольких лабораториях был открыт фермент транскрипции — РНК-полимераза.

1961 — Ф. Жакоб и Дж. Моно (F. Jakob, J. Monod) разработали модель оперона.

1965—1967 — Р. Холли (R. Holley) выяснил первичную структуру аланиновой тРНК, а А. А. Баев — валиновой тРНК.

1966 — М. Ниренберг, С. Очоа и Х.-Г. Корана (M. Nirenberg, S. Ochoa, H.-G. Khorana) расшифровали генетический код.

1967 —М. Геллерт (M. Gellert) открыл ДНК-лигазу — фермент, способный соединять фрагменты ДНК.

1970 — Г. Темин и Д. Балтимор (H. Temin, D. Baltimor) открыли обратную транскриптазу (РНК-зависимую ДНК-полимеразу) в онкогенных вирусах.

1972 — П. Боэр, С. Коэн и П. Берг (P. Boyer, S. Cohen, P. Berg) разработали технологию клонирования ДНК, заложили основы генетической инженерии.

1972 — Х.-Г. Корана (H.-G. Khorana) осуществил химический синтез гена аланиновой тРНК.

1975—1977 — Ф. Сангер (F. Sanger), а также А. Максам и У. Гилберт (A. Maxam, W. Gilbert) разработали методы быстрого определения первичной структуры ДНК.

1976 — Ф. Сангер (F. Sanger) расшифровал нуклеотидную последовательность ДНК фага φX174.

1976 — У. Гилберт (W. Gilbert) открыл мозаичное строение генов эукариот.

1976 — С. Ким, А. Рич и А. Клуг (S. Kim, A. Rich, A. Klug) определили третичную структуру тРНК.

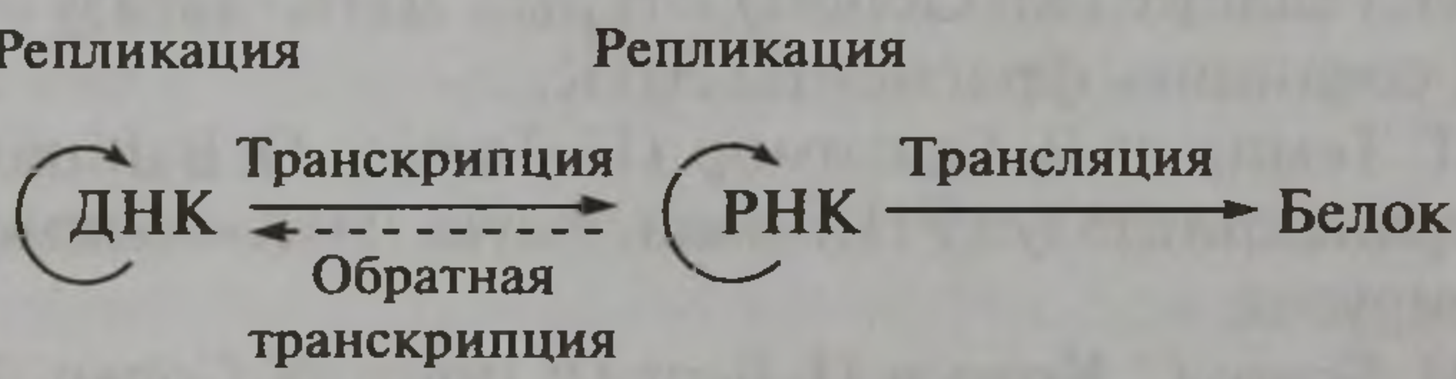
В результате выдающихся открытий Дж. Уотсона, Ф. Крика, Х.-Г. Кораны, А. Корнберга и других крупнейших молекулярных биологов, удостоенных нобелевских премий, уже в середине 60-х годов XX в. окончательно утвердился основной постулат молекулярной

генетики, формулирующий магистральный путь реализации генетической информации в клетке:

$$\text{ДНК} \rightarrow \text{РНК} \rightarrow \text{Белок}$$

Затем были детально изучены механизмы воспроизведения (репликации) ДНК, транскрипции (биосинтеза РНК) и трансляции (биосинтеза белка). Параллельно развивались работы по изучению внутриклеточной локализации этих процессов, что привело к осознанию функционального значения внутриклеточных компонентов (ядра, митохондрий, рибосом и др.) и дало основание Дж. Уотсону в 1968 г. для определения молекулярной биологии: «Молекулярная биология изучает связь структуры биологических макромолекул и основных клеточных компонентов с их функцией, а также основные принципы и механизмы саморегуляции клеток, которые опосредуют согласованность и единство всех протекающих в клетке процессов, составляющих сущность жизни».

Впоследствии центральный постулат молекулярной генетики был дополнен представлениями о существовании процесса обратной транскрипции (о биосинтезе ДНК на матрице РНК) и репликации РНК, что позволило придать ему следующий вид:



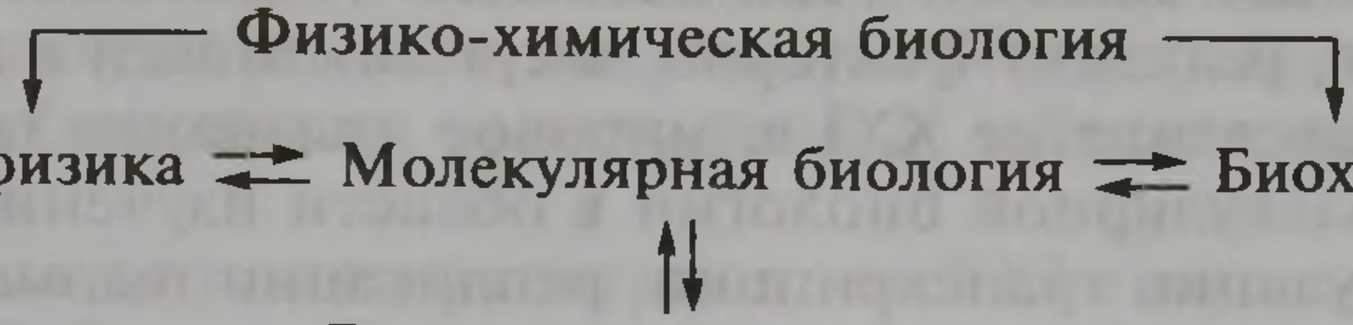
Одновременно все более детализировались представления о строении и функциях белков, необходимых для катализа (ферменты) и регуляции (регуляторные белки, пептидные гормоны) всех важнейших молекулярно-генетических процессов.

Открытие и разработка методов целенаправленного использования целого ряда ферментов (обратной транскриптазы, ДНК-рестриктаз и др.) привели к созданию технологии получения рекомбинантных ДНК, возникновению генетической инженерии, что стало поистине революционным событием в истории молекулярной биологии.

В конце 70-х и начале 80-х годов XX в. молекулярная биология вступает в период расцвета. В это время выясняются механизмы сплайсинга (В. Келлер и др.), происходит открытие РНК-ферментов (рибозимов) и аутосплайсинга (Т. Чек), активно изучаются механизмы генетической рекомбинации и подвижные генетические элементы (Д. Хогнесс, Г. П. Георгиев), на новый уровень выходят работы в области структуры ферментов и биологических мембран (Ю. А. Овчинников), начинаются работы по расшифровке структуры геномов

высших организмов (включая геном человека), создаются основы новых (генно-инженерных) биотехнологий, обнаруживаются и синтезируются каталитически активные антитела (абзимы), возникает белковая инженерия.

Постепенно молекулярная биология становится в центре наук, составляющих современную физико-химическую биологию:



Бурное развитие молекулярной биологии привело в начале 80-х годов XX в. к возникновению новой науки — *биоинформатики* (вычислительной биологии, компьютерной генетики), возникшей на стыке молекулярной генетики и информатики. Толчком к ее возникновению послужило появление быстрых методов определения нуклеотидных последовательностей ДНК, разработанных в 1975—1976 гг. Ф. Сангером и А. Коулсоном, а также А. Максамом и У. Гилбертом. В 1982 г. были организованы банки нуклеотидных последовательностей: Gen Bank в США и EMBL в Европе, в которых концентрировалась информация о расшифрованных нуклеотидных последовательностях ДНК различных организмов. Постепенно биоинформатика включилась в разработку ряда важных молекулярно-биологических проблем, включая: статистический анализ нуклеотидных последовательностей ДНК; предсказание функций по первичной структуре биополимеров (ДНК, РНК и белков); анализ (моделирование) пространственной структуры белков и нуклеиновых кислот; теорию молекулярной эволюции и систематики.

Прогресс в области определения нуклеотидных последовательностей (секвенирования) ДНК различных организмов, достигнутый в конце XX в. (в 1995 г. был секвенирован первый бактериальный геном, в 1997 г. — геном дрожжей, в 1998 г. — геном нематоды, в 2000 г. — геном дрозофилы и почти полностью — геном человека), привел к возникновению *геномики* — науки, изучающей наборы всех генов данного организма как единое целое. Одновременно возникла *протеомика* — наука, исследующая полные наборы белков, функционирующих на различных этапах развития того или иного организма.

В конце XX в. расширяются и становятся все более целенаправленными в научно-практическом отношении задачи молекулярной биологии, среди которых:

- расшифровка структуры геномов;
- создание банков генов;
- геномная дактилоскопия;

• изучение молекулярных основ эволюции, дифференцировки, биоразнообразия, развития и старения, канцерогенеза, иммунитета и др.;

• создание методов диагностики и лечения генетических болезней, вирусных заболеваний;

• создание новых биотехнологий производства пищевых продуктов и разнообразных биологически активных соединений (гормонов, антигормонов, релизинг-факторов, энергоносителей и др.).

В первое десятилетие XXI в. мировое признание получили достижения молекулярной биологии в области изучения структуры рибосом, регуляции транскрипции, репликации теломерных участков хромосом и другие исследования, удостоенные Нобелевских премий.

Нобелевские премии 2001 — 2009 гг. за достижения в области молекулярной биологии:

2001 г. — Л. Хартвелл, Т. Хант, П. Нерс «За открытие ключевых регуляторов клеточного цикла» (Нобелевская премия по физиологии и медицине).

2002 г. — С. Бреннер, Р. Хорвиц, Д. Салсон «За открытия в области генетического регулирования развития человеческих органов» (Нобелевская премия по физиологии и медицине).

2002 г. — Д. Фенн, К. Танака, К. Вютрих «За разработку методов идентификации и структурного анализа биологических макромолекул и, в частности, за разработку метода масс-спектрофометрического анализа биологических макромолекул» (Нобелевская премия по химии).

2004 г. — А. Гершко, И. Роуз «За открытие убиквитин-опосредованной деградации белков» (Нобелевская премия по химии).

2006 г. — Э. Файер, К. Мелло «За открытие РНК-интерференции — эффекта гашения активности определенных генов» (Нобелевская премия по физиологии и медицине).

2006 г. — Р. Корнберг «За исследование механизма копирования клетками генетической информации» (Нобелевская премия по химии).

2007 г. — М. Капеччи, М. Эванс, О. Смитис «За открытие принципов введения специфических генных модификаций у мышей с использованием эмбриональных стволовых клеток» (Нобелевская премия по физиологии и медицине).

2008 г. — Ф. Барре-Синусси, Л. Монтанье «За открытие ВИЧ» (Нобелевская премия по физиологии и медицине).

2008 г. — О. Симомура, М. Чалфи, Р. Тсьен «За открытие и развитие зеленого флуоресцентного белка» (Нобелевская премия по физиологии и медицине).

2009 г. — Р. Рамакришнан, Т. Стейц, А. Йонат «За исследование структуры и функций рибосомы» (Нобелевская премия по химии).

2009 г. — Э. Блекборн, К. Грейдер, Д. Шостак «За открытие механизмов защиты хромосом теломерами и фермента теломеразы» (Нобелевская премия по физиологии и медицине).

Начало нового тысячелетия ознаменовалось выдающимся событием — расшифровкой нуклеотидной последовательности генома человека, с которой связаны надежды на решение многих проблем человечества (коррекция наследственных заболеваний, продление жизни и т. д.).

Таким образом, по праву считается, что XXI в. должен стать веком молекулярной биологии и новых биотехнологий, призванных освободить человечество от тяжкого груза болезней, пороков и лишений и заложить основы его будущего процветания.

Глава 1
МЕТОДЫ МОЛЕКУЛЯРНОЙ БИОЛОГИИ

Молекулярная биология использует широкий арсенал биологических, физических и химических методов, одни из которых достались ей «в наследство» от наук-предшественниц (биохимии, цитологии, генетики и др.), а другие были созданы в процессе ее собственного развития специально для работы с молекулярными объектами. Ниже приведен перечень ряда наиболее широко используемых методов и отмечены области их применения.

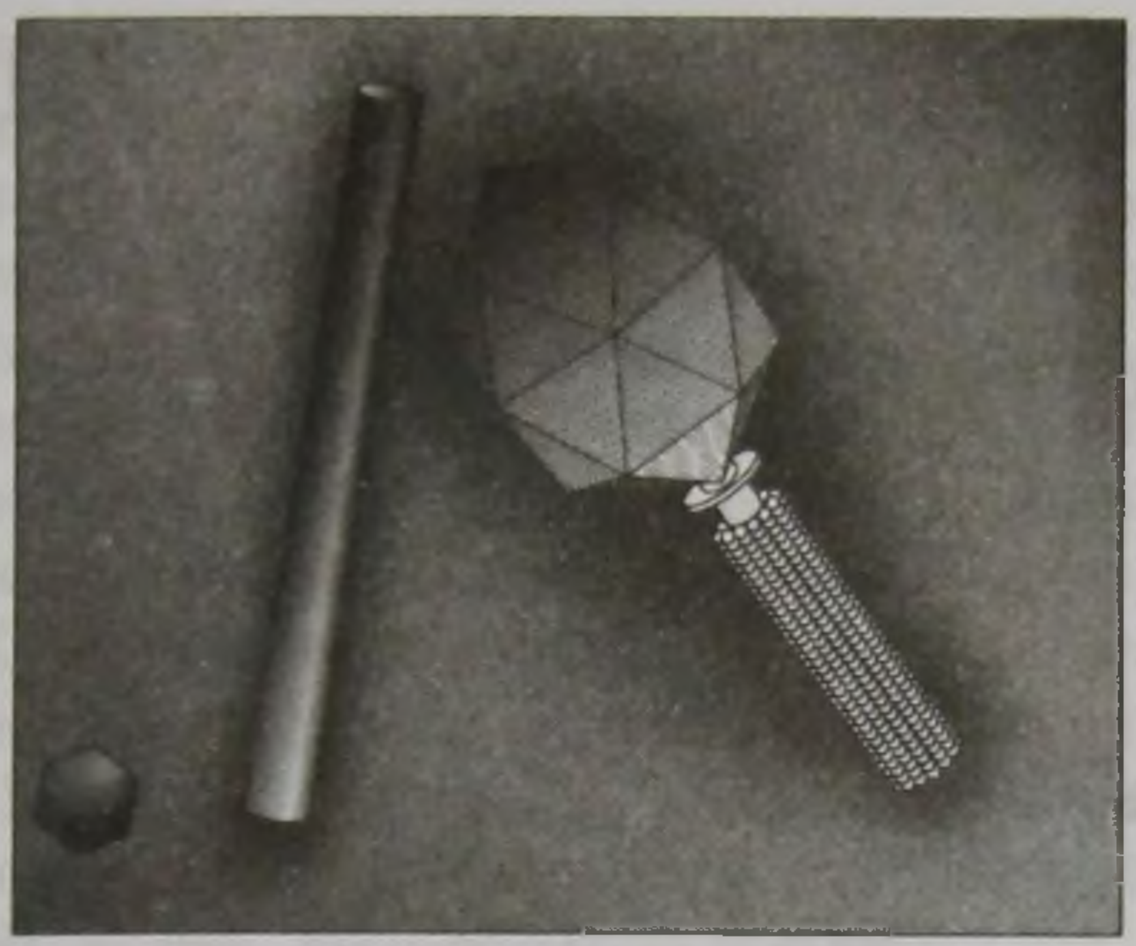
Микроскопия — имеет давнюю историю, начиная с XVII в., когда в 1611 г. Й. Кеплер предложил принцип создания светового микроскопа, а А. Левенгук впервые наблюдал с его помощью одноклеточные бактерии (1638). Именно световая микроскопия, имеющая предел разрешения 0,4—0,7 мкм, позволила М. Шлейдену и Т. Швану в 1838 г. создать клеточную теорию, согласно которой клетка, содержащая ядро, является структурной и функциональной единицей строения всех животных и растений. Затем Р. А. Колликер впервые в 1857 г. описал митохондрии, а В. Флемминг — поведение хромосом при митозе.

В развитии микроскопии существенными вехами стали изобретения интерференционного (1930), фазово-контрастного (1932) и, наконец, электронного микроскопов (1939).

Электронная микроскопия, позволяющая обычно различать объекты размером около 20 Å (2 нм) и имеющая в наиболее современных моделях электронных микроскопов предел разрешения до 0,1 нм, нашла самое широкое применение для изучения структур вирусов, внутриклеточных органелл, белково-нуклеиновых комплексов (хроматин, рибосомы, информосомы) и отдельных белковых молекул. Один из вариантов этого метода — криоэлектронная микроскопия — является в настоящее время ведущим при изучении тонкой структуры рибосом.

Наиболее впечатляющие и информативные трехмерные (объемные) изображения клеточных структур позволяют получить сканирующие электронные микроскопы (рис. 1).

Рентгеноструктурный анализ — основан на дифракции рентгеновских лучей (электромагнитного излучения с длиной волны около 10^{-10} м); позволяет выявить трехмерное расположение атомов





А *Б*

Рис. 1. Примеры изображений, получаемых методом электронной микроскопии:

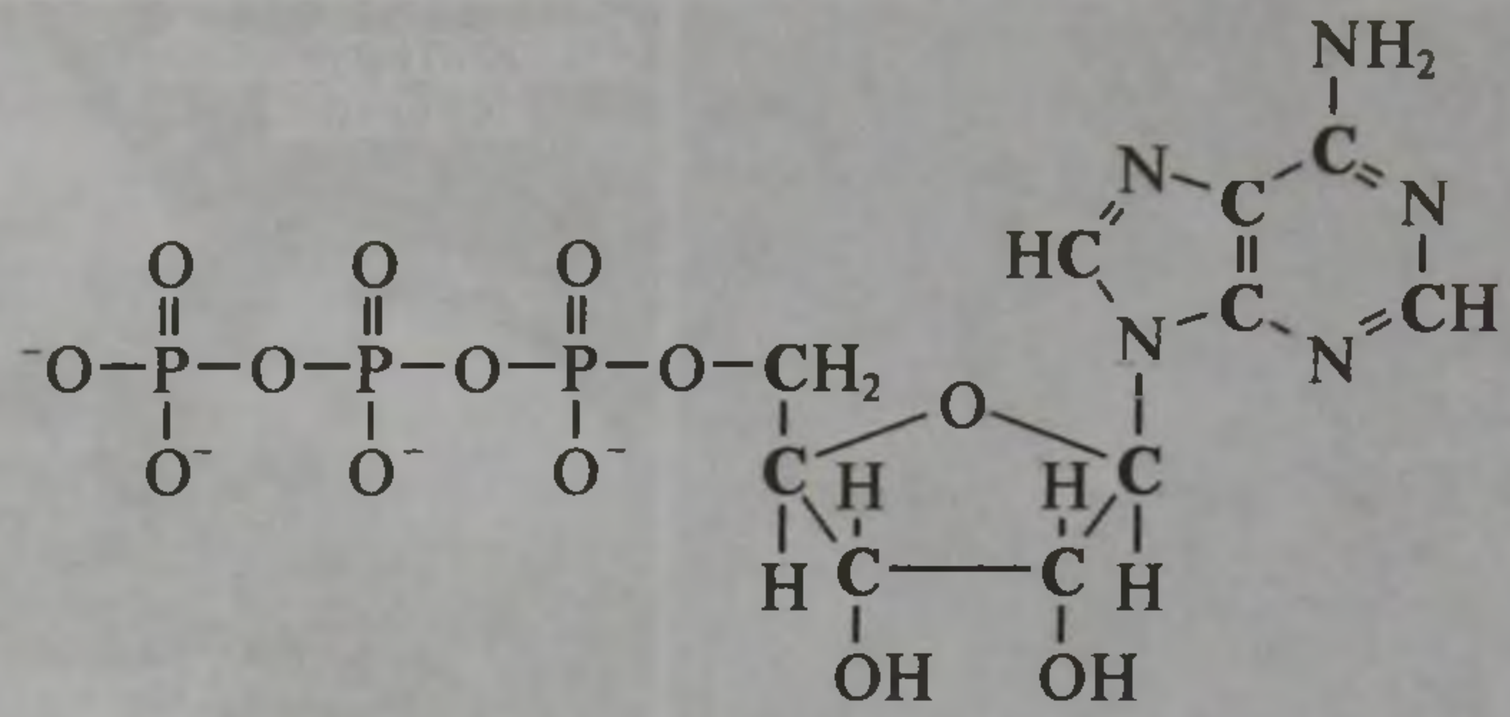
А — электронная микрофотография фаговых частиц (×200 000). Слева направо: фаг φX174, имеющий форму икосаэдра, вирус табачной мозаики (палочкообразная частица) и фаг T_4; *Б* — полученная с помощью сканирующего электронного микроскопа фотография полиморфноядерного лейкоцита, поглощающего посредством фагоцитоза дрожжевую клетку

в молекулах (разрешение составляет менее 0,1 нм). Метод был разработан в Англии Г. Брэггом и Л. Брэггом, и именно с его помощью были получены основополагающие сведения о структуре молекул белков, ДНК и РНК.

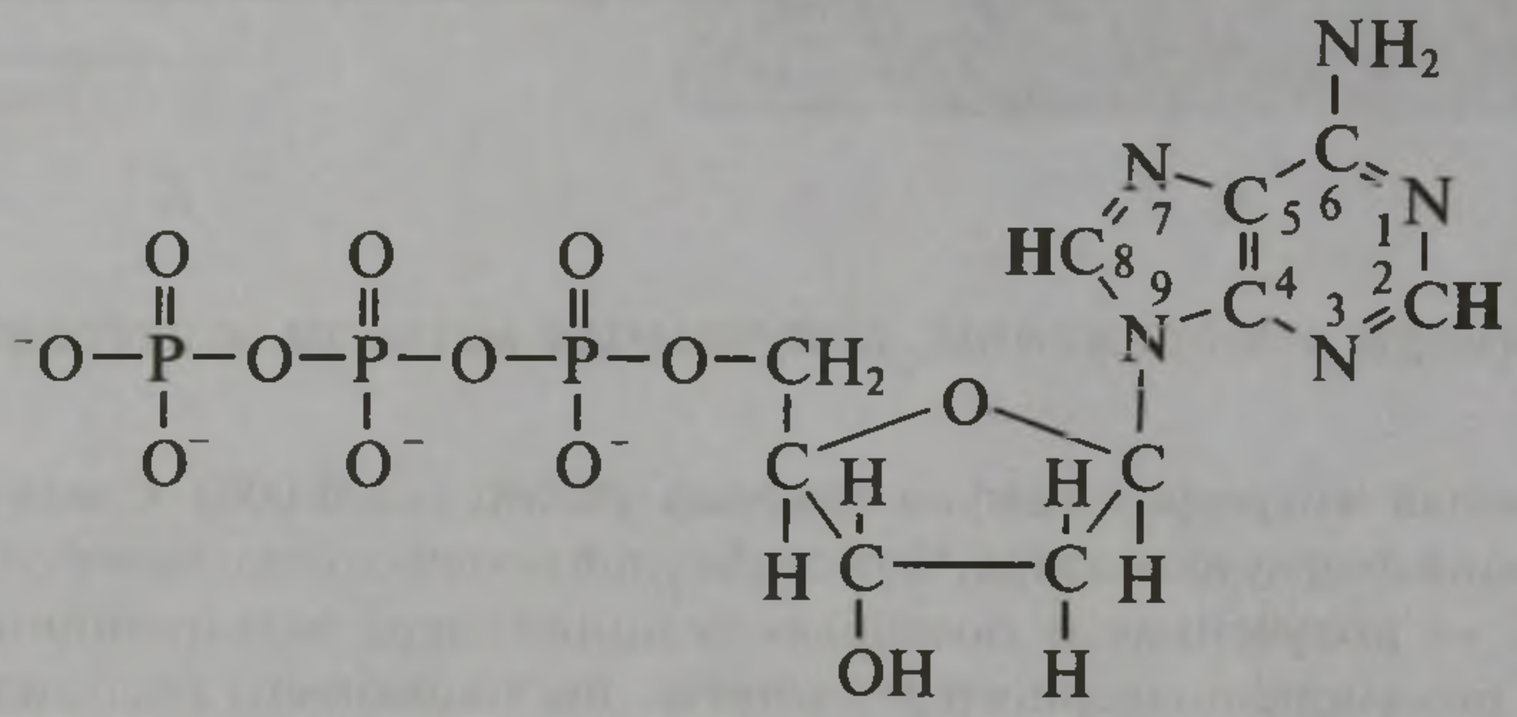
В наше время этот метод в сочетании с компьютерным анализом полученных данных является ведущим для изучения трехмерной структуры биополимеров.

Радиоактивные изотопы — широко используются для изучения нуклеиновых кислот, белков, углеводов и других молекул в живой клетке. Радиоизотопы нестабильны и подвергаются спонтанному распаду, при котором либо высвобождаются заряженные частицы — электроны, либо происходит гамма-излучение. Период полураспада варьирует от короткого, как, например, у изотопа ^{32}P (14 сут), до очень протяженного, как у ^{14}C (5 570 лет). Электроны определяют по ионизации газа в сцинцилляционном счетчике или счетчике Гейгера, либо методом радиоавтографии (по их воздействию на серебро, содержащееся в чувствительном фотоэмульсионном слое).

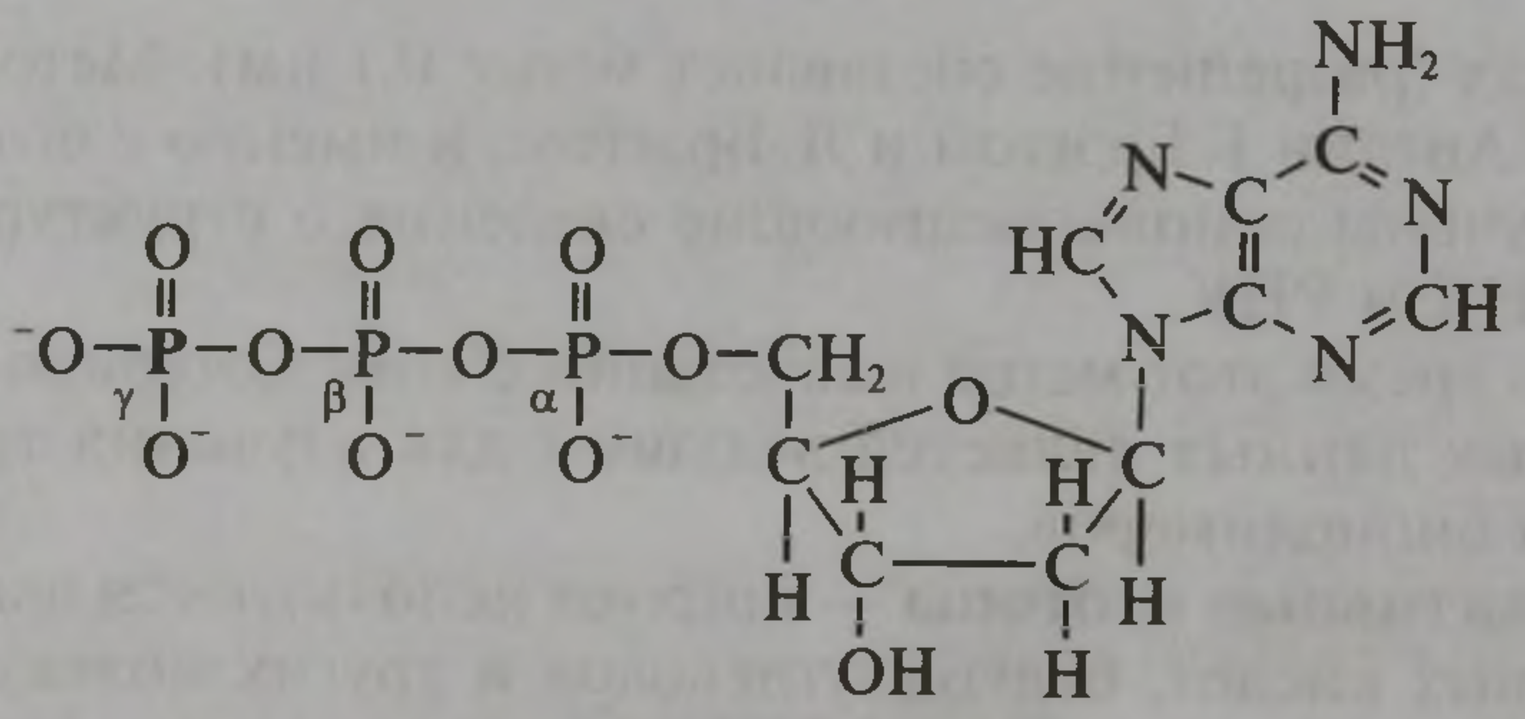
Радиоактивные молекулы используются при изучении самых разнообразных внутриклеточных процессов: синтеза молекул из их предшественников, определения внутриклеточной локализации молекул, времени их функционирования в клетке и ее отдельных компартментах, химических превращениях в отдельных участках макромолекул и т. д. Если, например, инкубировать клетки с радиоактивным предшественником РНК (^3H-уридином), то можно опреде-



ATP[^{14}C]



ATP[2,8-^{3}H]



ATP[γ-^{32}P]

Рис. 2. Три радиоактивные формы АТР.

Радиоактивные атомы выделены жирным шрифтом. Приведены обозначения, указывающие расположение и тип радиоактивных атомов

лить, что РНК синтезируется в ядре, а затем переходит в цитоплазму клеток. Радиоактивную метку можно вводить в изучаемые молекулы с помощью ферментов. Так, с использованием фосфотрансфераз и нуклеотидилтрансфераз можно вводить метку в белки, углеводы и нуклеиновые кислоты, используя для этого различные радиоактивные формы АТР (рис. 2).

Ультрацентрифугирование (седиментационный анализ) — получило широкое распространение после изобретения в 1926 г. Т. Свед-бергом аналитической ультрацентрифуги, с помощью которой он

впервые определил молекулярную массу гемоглобина (68 кДа). В этом методе скорость седиментации (осаждения) определяется размером и формой разделяемых компонентов и выражается коэффициентом седиментации S. Коэффициент седиментации измеряется в секундах и определяется по формуле: $(dx/dt)/w^2x$, где x — расстояние от центра вращения (радиус ротора центрифуги (см), определяемый от оси вращения до дна центрифужной пробирки); dx/dt — скорость седиментации (см/с); w — угловая скорость вращения ротора (рад/с). Коэффициенты седиментации обычно представляют собой очень малые значения и поэтому выражаются в единицах Сведберга (S): $1S = 1 \cdot 10^{-13}$ с. Коэффициент седиментации молекулы гемоглобина составляет 4,5S, молекулы тРНК — 4S, рибосомы кишечной палочки — 70S, лизосомы — 9 400S.

В 40 — 50-е годы XX в. А. Клод и Ж. Браше разработали метод дифференциального центрифугирования для разделения органелл клетки, с помощью которого де Дюв (de Duve) в 1953 г. впервые выделил лизосомы, а затем — пероксисомы. В 1957 г. М. Месельсон, У. Сталь и Дж. Виноград разработали метод центрифугирования в градиенте плотности хлористого цезия для разделения нуклеиновых кислот, с помощью которого было установлено, что репликация ДНК осуществляется полуконсервативным путем (см. гл. 7). Различные варианты метода ультрацентрифугирования широко используют в молекулярно-биологических исследованиях для выделения внутриклеточных компонентов и макромолекул (нуклеиновых кислот и белков), а также определения их молекулярных масс и коэффициентов седиментации.

Для более тонкого фракционирования и очистки разнообразных молекул и прежде всего белков в настоящее время применяется большая группа разнообразных физико-химических методов. Рассмотрим только наиболее широко распространенные методы фракционирования белков: хроматографию, электрофорез и изоэлектрофокусирование. Методы исследования структуры и функций нуклеиновых кислот изложены в последующих главах.

Хроматография — метод, впервые изобретенный русским ученым М. С. Цветом, который в 1906 г. фракционировал окрашенные экстракты листьев растений на колонках с порошком мела.

В настоящее время существует большое количество вариантов хроматографии, в которых используют матриксы (носители) разных типов, позволяющие разделить белки по заряду (ионообменная хроматография), размеру молекул (гель-хроматография, чаще называемая гель-фильтрацией) или способности специфически взаимодействовать с определенными химическими группами веществ, предварительно связанных с матриксом (аффинная хроматография). Из всех вариантов хроматографии наибольшей эффективностью обладает *аффинная хроматография* (хроматография по сродству), в основе которой лежит специфическое взаимодействие (узнавание)

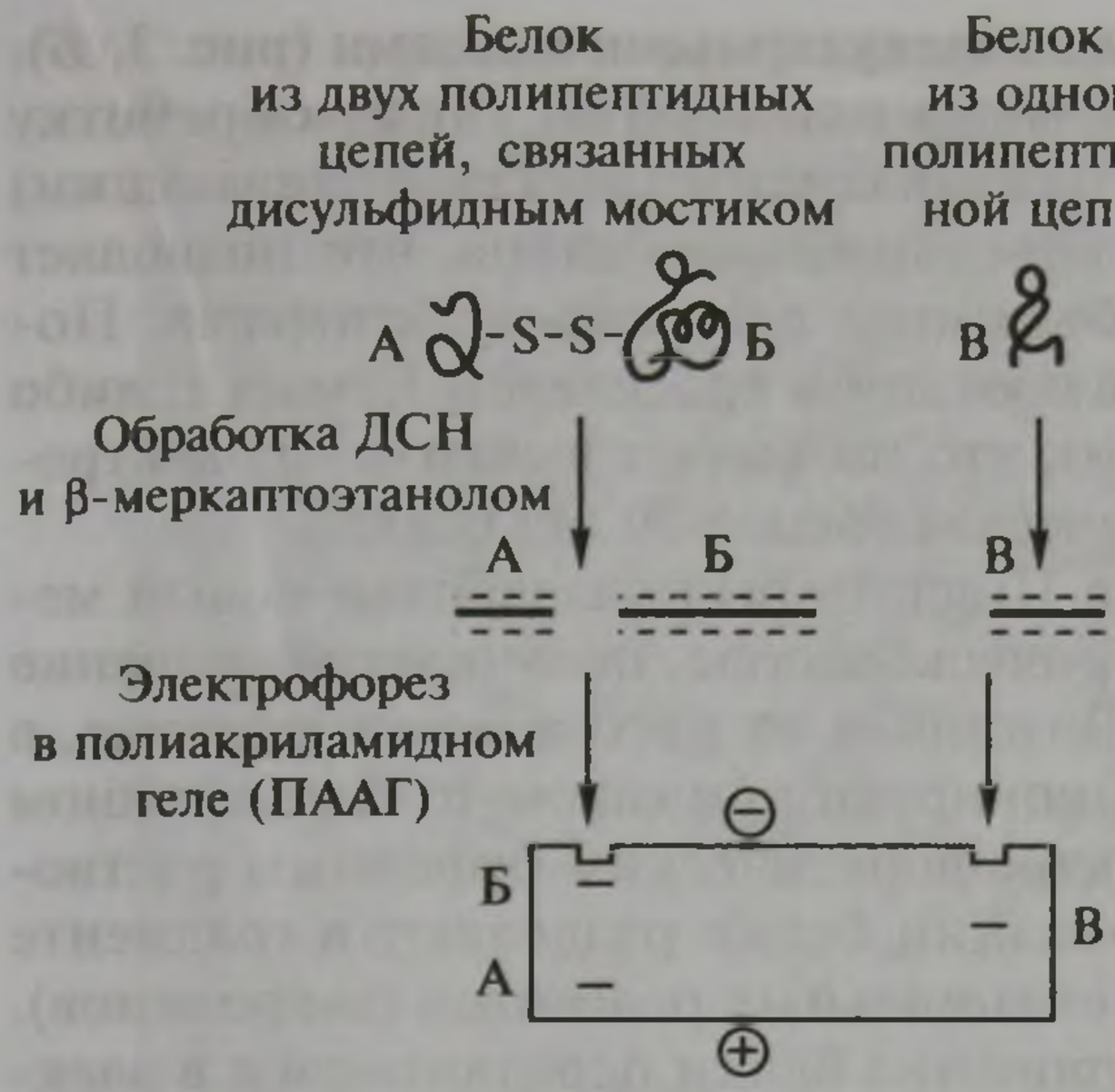
молекул взаимодействующих веществ. Например, предварительно связав с матриксом субстрат фермента, можно добиться избирательного удерживания фермента на колонке, а затем после выхода остальных («балластных») белков элюировать его в практически гомогенном состоянии. Матрикс можно оснастить и специфическими антителами и использовать такой носитель для выделения белков, распознаваемых этими антителами. Связав с матриксом фрагменты ДНК, можно выделить белки, специфически связываемые с определенными участками хромосом.

Широкое распространение получила также *высокоэффективная жидкостная хроматография* (HPLC). В этом методе используются специально разработанные кремнийорганические смолы, способные образовывать гомогенную среду в форме микросфер диаметром 3—10 мкм и позволяющие при высоком давлении осуществить быстрое равномерное протекание растворителя через колонку, при котором весь процесс хроматографии занимает считанные минуты и обеспечивает тонкое фракционирование смеси анализируемых молекул.

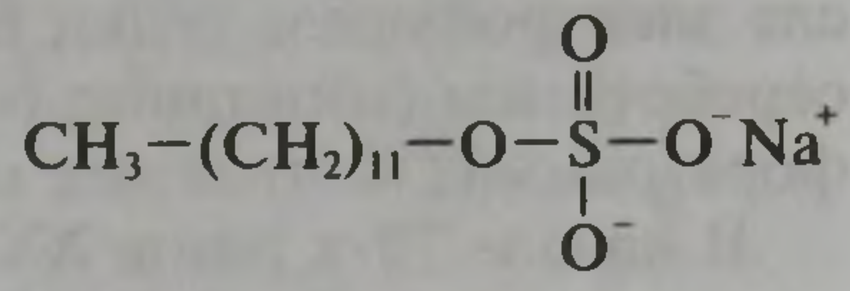
Электрофорез — в основе этого метода лежит способность белков, обладающих определенным суммарным положительным или отрицательным зарядом, перемещаться в электрическом поле в соответствии с величиной заряда, размером и формой молекул. Электрофорез можно проводить в водном (буферном) растворе, однако, как правило, его проводят в каком-либо пористом (полимерном) носителе: крахмальном, агарозном или полиакриламидном геле, на целлюлозных или нитроцеллюлозных пластинах и т. д.

Простейший метод электрофореза *на пластинах целлюлозы* был использован В. Ингрэмом в 1956 г. при фракционировании фрагментов протеолитического расщепления гемоглобина человека, полученных в результате обработки этого белка трипсином. С помощью данного метода были получены пептидные карты — «фингерпринты» (отпечатки пальцев), по которым было установлено, что заболевание серповидноклеточной анемией обусловлено заменой одной-единственной аминокислоты в β-цепи гемоглобина.

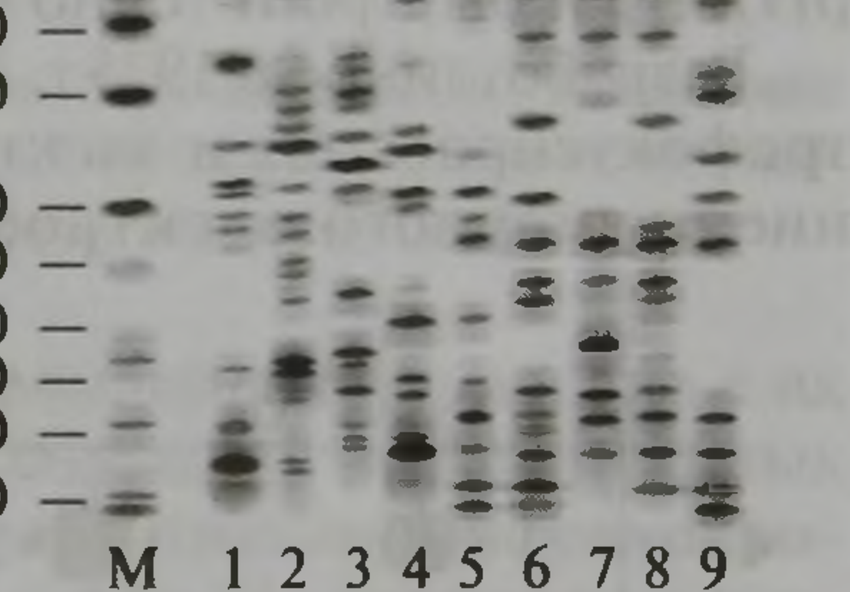
Наиболее часто в настоящее время для разделения белков используют метод электрофореза в *полиакриламидном геле* (ПААГ), который представляет собой инертный матрикс с высоким и контролируемым числом поперечных сшивок. Варьируя размеры гелевых пор, можно фракционировать белки, существенно отличающиеся по молекулярной массе (от нескольких десятков тысяч до сотен тысяч дальтон). Метод был разработан в 1959 г. С. Рэймондом, а затем усовершенствован Б. Дэвисом и Л. Орнстейном. Он позволяет осуществить тонкое фракционирование смесей белков, отличающихся по заряду и молекулярной массе. Дальнейшей модификацией этого метода стал *электрофорез в ПААГ с додецилсульфатом натрия*, предложенный в 1966 г. Дж. Майзелем. Додецилсульфат натрия (ДСН, SDS, ДС-Na) представляет собой мощный ионный детергент,



Пластина ПААГ после электрофореза и окрашивания

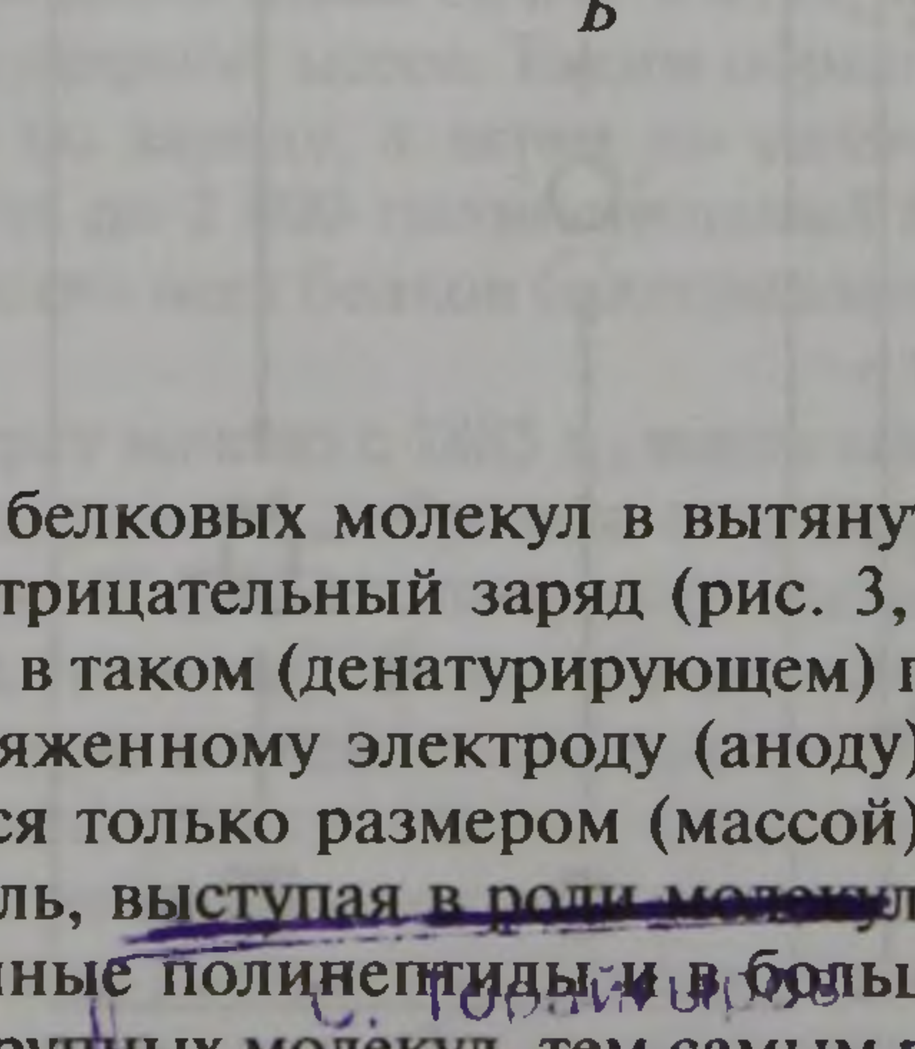


Додецилсульфат натрия (ДСН)



β-Меркаптоэтанол

А



Б

Рис. 3. Электрофорез белков с додецилсульфатом натрия:

А — обработка белков ДСН и β-меркаптоэтанолом, приводящая к разрыву ковалентных (дисульфидных) связей между отдельными полипептидными цепями, вызывающая разворачивание белковых глобул и сообщающая им избыточный отрицательный заряд (при электрофорезе такие денатурированные белки перемещаются к аноду со скоростью, обратно пропорциональной размеру (массе) их полипептидных цепей); *Б* — электрофореграмма, полученная в результате фракционирования растворимых белков листьев различных видов картофеля методом электрофореза в полиакриламидном геле с ДСН: 1—9 — различные виды картофеля; М — маркеры молекулярной массы

который вызывает разворачивание белковых молекул в вытянутые цепи и сообщает им избыточный отрицательный заряд (рис. 3, *А*). Потерявшие нативную форму белки в таком (денатурирующем) геле перемещаются к положительно заряженному электроду (аноду) со скоростью, которая детерминируется только размером (массой) их полипептидных цепей. При этом гель, выступая в роли молекулярного сита, легче пропускает некрупные полипептиды и в большей мере тормозит продвижение более крупных молекул, тем самым раз-

601378

Ғылыми кітапхана
атындағы ТІМУ-дің
академик С.Бейсембаев
атындағы ғылыми
КІТАПХАНАСЫ
17

деляя белки в соответствии с их молекулярными массами (рис. 3, *Б*). При электрофорезе с SDS обычно используют также обработку белков β-меркаптоэтанолом, разрывающим (восстанавливающим) дисульфидные связи между субъединицами белков, что позволяет определить число и массу субъединиц в белках-мультимерах. После электрофореза белки выявляют либо красителем Кумасси, либо серебрением (минорные белки), что позволяет выявить на электрофореграммах ничтожные количества (около 10 нг) белка.

В начале 70-х годов XX в. в Швеции был разработан новый метод разделения белков в электрическом поле, получивший название **изоэлектрофокусирования**. В отличие от метода электрофореза, в процессе которого белки фракционируют при каком-то определенном значении pH, задаваемом электрофоретическим буферным раствором, при изоэлектрофокусировании белки разделяют в градиенте pH, создаваемом с помощью специальных реагентов (амфолинов). В процессе изоэлектрофокусирования белки передвигаются в электрическом поле в строгом соответствии только с зарядом молекул и останавливаются (фокусируются) в тех точках градиента pH, которые соответствуют их изоэлектрическим точкам (*pI*), т. е. там, где молекулы становятся электронейтральными (рис. 4). Изоэлектрофокусирование позволяет не только исключительно тонко разделить белки по заряду (разделяются белки, отличающиеся по *pI* всего на 0,005 единиц pH), но и быстро и точно определить их изоэлектрические точки.

Разработанное в 1975 г. П. Офареллом сочетание методов изоэлектрофокусирования и электрофореза в ПААГ с SDS получило название **двумерного электрофореза**. В этой процедуре белки вначале

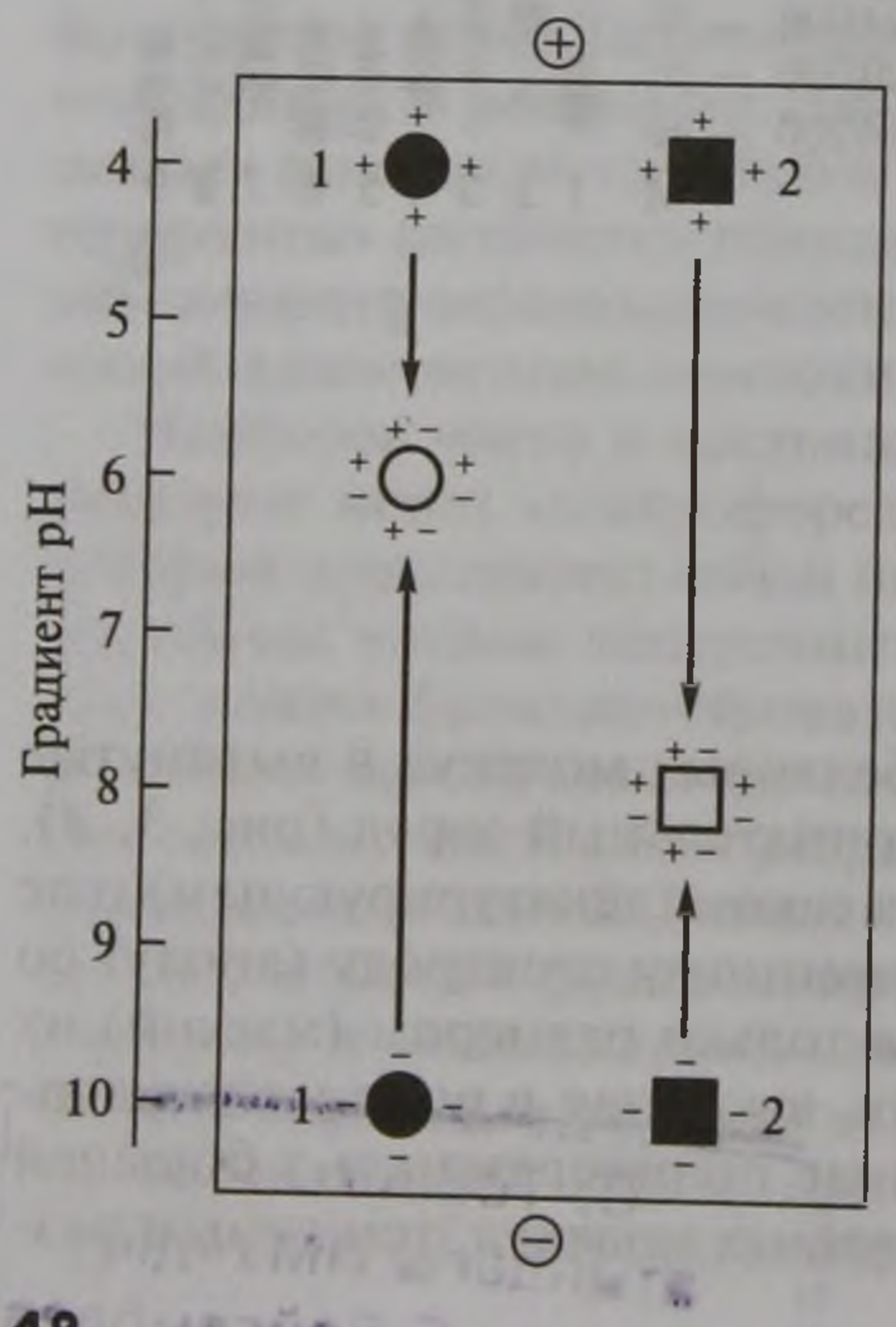


Рис. 4. Разделение белков методом изоэлектрофокусирования.

При низких значениях pH белки заряжены положительно (+), при высоких значениях pH — отрицательно (–). В электрическом поле молекулы белков мигрируют к противоположно заряженным полюсам и останавливаются в тех точках градиента pH, которые соответствуют их изоэлектрическим точкам (*pI*). Изоэлектрические точки белка 1 и белка 2 находятся при pH 6,0 и 8,0 соответственно



Рис. 5. Радиоавтограмма, иллюстрирующая результат разделения белков кишечной палочки с включенными в них ^{14}C-аминокислотами методом двумерного электрофореза.

На радиоавтограмме, полученной наложением на пластину полиакриламидного геля фотографической пленки, можно различить свыше 1 000 пятен, соответствующих отдельным полипептидным цепям белков

обрабатывают β-меркаптоэтанолом и мочевиной, что приводит к их полному растворению, денатурации и диссоциации полипептидных цепей без изменения заряда. Далее проводят изоэлектрофокусирование в ПААГ, разделяя белки по заряду, а затем (в перпендикулярном направлении) ведут их электрофорез в блоке ПААГ с SDS, при котором белки разделяются по молекулярной массе. Таким образом, сочетая тонкое разделение вначале по заряду, а затем по размеру (массе), удается за один раз разделить до 2 000 полипептидных цепей, т. е. проанализировать большинство всех белков бактериальной клетки (рис. 5).

Культура клеток — этот метод берет начало с 1885 г., когда впервые было обнаружено, что клетки куриного эмбриона сохраняют жизнеспособность в солевом растворе. С 1907 г. стали использовать культуры фрагментов тканей (эксплантантов). В настоящее время используют диссоциированные культуры клеток, из которых можно получить клетки одного типа. Иногда в культуре клеток появляются мутанты, которые могут бесконечно размножаться и образовывать клеточную линию. В отличие от раковых клеток, также способных к непрерывному делению, мутантные клетки лучше растут в контакте с

какой-то поверхностью. Клеточные линии служат источником большого количества клеток одного типа и к тому же могут храниться при -70 °С неопределенно долго, не теряя способности к пролиферации. Наиболее часто используются клеточные линии фибробластов мышей и хомячков, а также эпителиальные клетки человека. В 1952 г. была получена перевиваемая до сих пор линия клеток карциномы шейки матки человека, широко известная как линия HeLa.

Однородность клеточных линий можно повысить путем клонирования. Клон — это популяция клеток, происходящая от одной клетки-предшественницы. С помощью клонирования выделяют мутантные клеточные линии, у которых мутация затронула определенные гены. Такие клетки бывают нередко дефектны по какому-то определенному белку, что позволяет проанализировать его функцию в нормальных клетках.

Слиянием двух клеток могут быть получены *гетерокарионы* — клетки с двумя отдельными ядрами. После митотического деления гетерокарион превращается в гибридную клетку, у которой все хромосомы объединяются в одном ядре. Такие гибридные клетки дают возможность изучать функции отдельных хромосом, взаимодействие внутриклеточных компонентов (митохондрий, ядер и др.) различных клеток. Их можно клонировать и получить гибридную клеточную линию. Существенно, что гибридные клетки нестабильны. В частности, гибридные клетки человек-мышь постепенно утрачивают хромосомы человека, что позволяет судить о функциях отдельных хромосом, осуществляя таким образом их генетическое картирование.

Бесклеточные системы — незаменимы для изучения механизмов определенных молекулярных процессов, так как исключают различные побочные реакции, протекающие в клетке. В 1954 г. П. Замечник получил первую бесклеточную систему для изучения биосинтеза белка (трансляции). Возможности этих систем были далее блестяще использованы А. С. Спириным, М. Номурой и другими исследователями для изучения деталей трансляции. При этом из экстрактов клеток выделяли важнейшие компоненты белок-синтезирующей системы (мРНК, рибосомы, тРНК и др.), а затем последовательно вводили их в систему и уточняли роль каждого компонента в биосинтезе белка. С помощью бесклеточных систем был расшифрован генетический код, для чего в качестве мРНК использовали синтетические олиго- и полинуклеотиды известного состава. Бесклеточные системы используются также для изучения деталей других важнейших молекулярно-генетических процессов: репликации и транскрипции ДНК, сплайсинга РНК и др.

Моноклональные антитела — очень чувствительный инструмент выявления (идентификации) молекул в сложных смесях, и поэтому они находят очень широкое применение в молекулярной биологии. Антитела в целом представляют собой белки (иммуноглобулины), вырабатываемые позвоночными животными для защиты от чужеродных

соединений — антигенов. Клетки позвоночных животных способны синтезировать колоссальное количество (10^8) различных форм антител, отличающихся участками узнавания антигенов. Именно высокая специфичность антител определяет их уникальные возможности для выявления различных молекул в клетке. Разработанный в 1976 г. метод включает клонирование В-лимфоцитов, секретирующих только определенный вид антител, благодаря чему эти антитела можно получать в больших количествах. Но время жизни В-лимфоцитов в культуре невелико, поэтому проводят их слияние с клетками, происходящими от «бессмертной» опухоли, также полученной из В-лимфоцитов. Из образовавшейся смеси клеток отбирают гибриды, способные развиваться в культуре и синтезировать антитела определенного вида. Эти гибридные клетки называют *гибридомами*. Их клонируют по отдельности и получают клоны, каждый из которых является источником моноклональных антител. Моноклональные антитела происходят от одной-единственной клетки и обладают абсолютной специфичностью к белкам: они способны узнавать определенную конфигурацию группы из 5—6 аминокислотных остатков. Каждый тип антител можно далее использовать в качестве *зонда* для определения локализации белка в клетке и для очистки белков в аффинной хроматографии. В недалеком будущем моноклональные антитела могут стать источником для создания новой группы каталитически активных молекул — абзимов (см. гл. 2).

Каталитически активные белки (ферменты) — важнейшие инструменты биохимических методов исследования. Они нашли широкое применение в молекулярной биологии и конкретные примеры их использования равно как и других катализаторов биологической природы (рибозимов, абзимов, гибридозимов) будут приведены в последующих главах.

Глава 2
БЕЛКИ

Белки представляют собой *биополимеры* (*полипептиды*), составленные из *аминокислотных остатков*, соединенных *пептидными связями*.

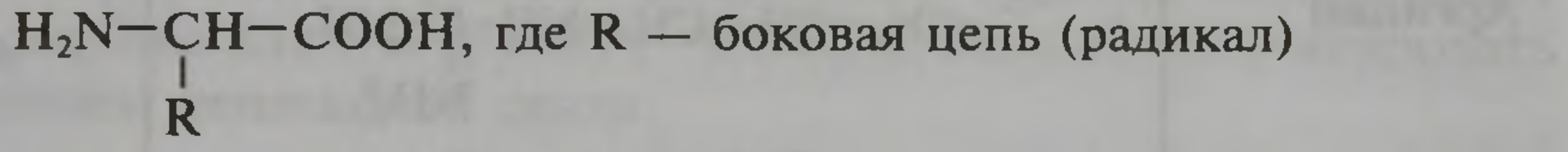
Белки — самый распространенный в живой природе класс биополимеров, составляющий большую часть сухого вещества животных (50—80 %) и, что самое главное, обладающий наибольшим по сравнению со всеми другими веществами органической природы разнообразием структуры и функций. Именно белки определяют в конечном счете фенотип любого организма, формируют облик всех живых систем. И хотя их строение запрограммировано нуклеотидной последовательностью соответствующих генов, в структурном и функциональном отношении белки несравненно более вариабельны, чем кодирующие их полинуклеотиды (ДНК и РНК). Эта особенность белков проистекает из внутреннего разнообразия составляющих их мономерных звеньев — аминокислот. Число постоянно встречающихся в живых организмах аминокислот (20 так называемых канонических, или белковых, аминокислот) в четыре раза больше, чем число канонических нуклеотидов (A, G, T, C, U), образующих молекулы нуклеиновых кислот, и поэтому «язык» белков оказывается намного богаче «языка» нуклеиновых кислот.

Белки участвуют во всех без исключения молекулярно-биологических процессах. Обладая прежде всего ярко выраженной каталитической способностью, они в качестве *ферментов* детерминируют интенсивность всех метаболических процессов в клетке и организме в целом. Разнообразные *регуляторные белки* контролируют своевременность и точность основных молекулярно-генетических процессов (белки-регуляторы репликации, транскрипции и трансляции). Белки являются непременным компонентом всех биологических мембран, составляют основу цитоскелета, входят в состав соединительных тканей, волосяного покрова, т. е. обеспечивают «строительную» функцию (*структурные белки*). Они же осуществляют иммунологическую защиту от чужеродных соединений и патогенных микроорганизмов (*защитные белки-иммуноглобулины*), обеспечивают снабжение организма кислородом и разнообразными питательными веществами (*транспортные белки*). Особая группа полипептидов контролирует

физиологическую активность тканей и органов (*белки-гормоны*). Белки формируют ионные каналы и осуществляют восприятие, трансформацию и передачу разнообразных внешних сигналов (*белки-рецепторы*). Все виды движения, начиная с движений жгутиков бактерий и кончая движениями пальцев пианиста, обеспечиваются работой «белковых моторов» (*сократительные белки*). Таков далеко не полный перечень функций белков. Уникальные свойства белков заложены в колоссальном разнообразии пространственного строения их молекул. Это разнообразие определяется огромным числом возможных сочетаний аминокислотных остатков в длинных, состоящих, как правило, из нескольких сотен остатков, полипептидных цепях белков. Далее, по мере рассмотрения уровней структурной организации белков, будут приведены конкретные примеры взаимосвязи их структуры и функций.

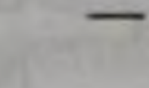
Аминокислотный состав белков

Аминокислоты, участвующие в образовании белков, имеют общую формулу:

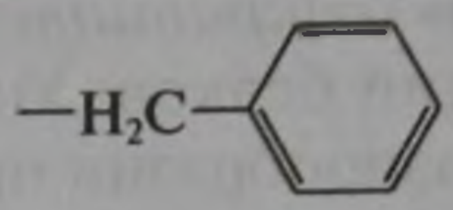
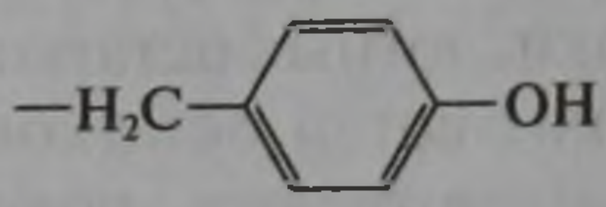
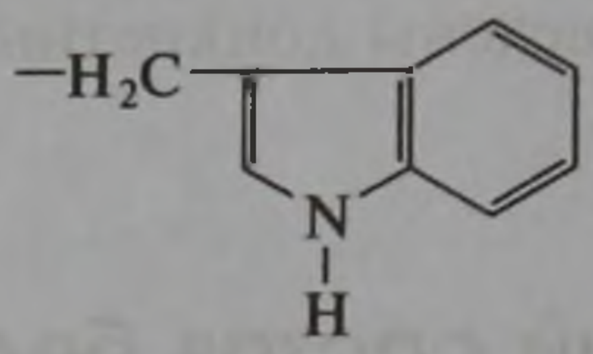
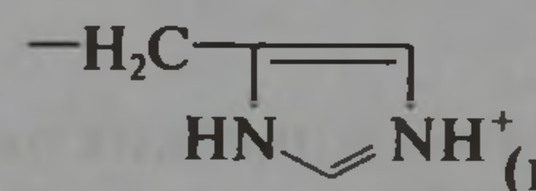


Формулы радикалов 20 канонических аминокислот (18 аминокислот и двух амидов — аспарагина и глутамина) представлены в табл. 1, где приведены также их обычно используемые трехбуквенные и одно-

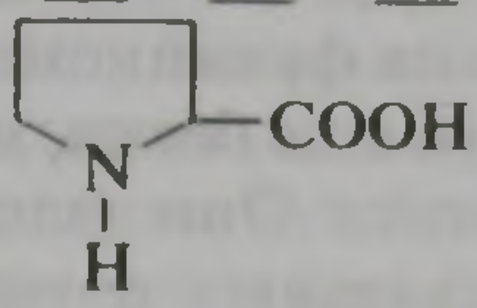
Таблица 1. **Протеиногенные аминокислоты и амиды**

Аминокислота	Структуры боковой цепи (радикал)	Сокращенное обозначение
	Алифатические	
Глицин		Gly, G
Аланин		Ala, A
Валин		Val, V
Лейцин		Leu, I
Изолейцин		Ile, I

Продолжение табл. 1

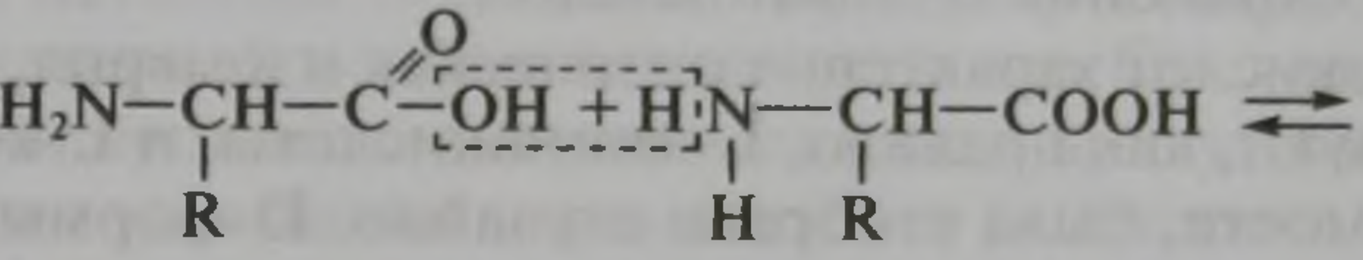
Аминокислота	Структуры боковой цепи (радикал)	Сокращенное обозначение
	Ароматические	
Фенилаланин		Phe, A
Тирозин		Tyr, Y
Триптофан		Trp, W
	Основные (несущие положительный заряд)	
Лизин	$-CH_2-CH_2-CH_2-CH_2-NH_3^+$	Lys, K
Аргинин	$-CH_2-CH_2-CH_2-NH-C(=NH_2^+)-NH_2$	Arg, R
Гистидин	 (при pH 6,0)	His, H
	Дикарбоновые (несущие отрицательный заряд)	
Аспарагиновая кислота	$-CH_2-COO^-$	Asp, D
Глутаминовая кислота	$-CH_2-CH_2-COO^-$	Glu, E
	Амиды дикарбоновых аминокислот	
Аспарагин	$-CH_2-C(=O)-NH_2$	Asn, N
Глутамин	$-CH_2-CH_2-C(=O)-NH_2$	Gln, Q
	Серосодержащие	
Цистеин	$-CH_2-S$	Cys, C
Метионин	$-CH_2-CH_2-S-CH_3$	Met, M

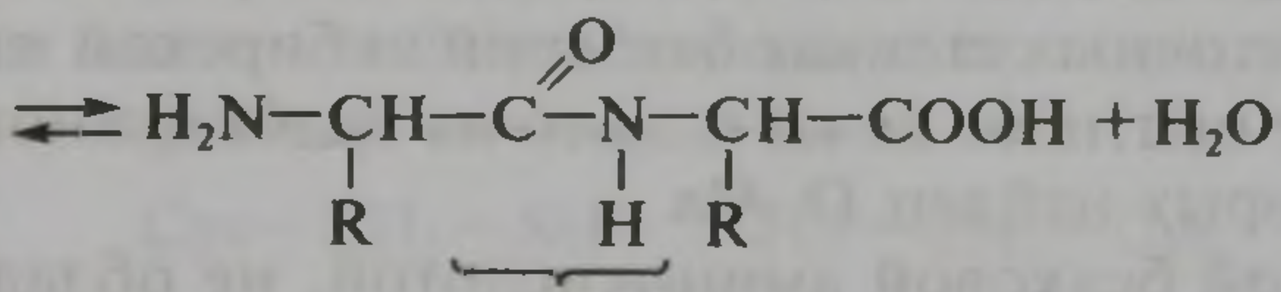
Окончание табл. 1

Аминокислота	Структуры боковой цепи (радикал)	Сокращенное обозначение
	Содержащие гидроксильную группу **(аминокислоты-спирты)**	
Серии	$-CH_2-OH$	Ser, S
Треонин	$-CH(OH)-CH_3$	Thr, T
Пролин	 (формула приведена полностью)	Pro, P

буквенные сокращенные обозначения (в дальнейшем будут использоваться более легко запоминающиеся трехбуквенные символы).

Все постоянно встречающиеся в белках (протеиногенные) аминокислоты содержат две разные химические группы: аминогруппу ($-NH_2$) и карбоксильную группу ($-COOH$), способные реагировать с образованием *пептидной связи*:





Пептидная связь в дипептиде

В водной среде равновесие такого типа реакций сдвинуто в сторону образования свободных аминокислот, и поэтому биосинтез полипептидов в клетках не сводится к простому отщеплению воды, а протекает по более сложному энергоемкому механизму с участием специфических внутриклеточных частиц и макроэргических соединений.

Единственной протеиногенной аминокислотой, не имеющей свободной аминогруппы, является пролин. Не имеет ее и содержащееся в некоторых белках производное пролина — оксипролин. Пролин и оксипролин иногда называют *иминокислотами*, жесткая конформация которых вызывает изгибы в полипептидных цепях белков.

Связанные в результате биосинтеза в длинные полипептидные цепи аминокислоты присутствуют в белках в виде аминокислотных

остатков, молекулярная масса которых варьирует от 57 до 186 Да (среднее значение принято 110), и, таким образом, белок с молекулярной массой 44 000 Да содержит примерно 400 аминокислотных остатков. Кроме амино- и карбоксильных групп в аминокислотах содержатся также боковые цепи, не участвующие в образовании пептидных связей и называемые *радикалами*. Химическая природа радикалов весьма разнообразна, что позволяет им вступать в различные (в большинстве случаев нековалентные) взаимодействия по мере синтеза и формирования пространственной структуры белковой молекулы. Именно последовательность расположения радикалов определяет особенности пространственной (трехмерной) структуры белков и соответственно их функциональные свойства.

Аминокислоты, образующие белки, можно рассматривать как производные карбоновых кислот. Они являются α-аминокислотами, так как аминогруппы в них связаны с α-углеродным атомом. Сходные с α-аминокислотами во многих отношениях и также встречающиеся в природе β-аминокислоты, тем не менее, обычно не входят в состав природных полипептидов, очевидно, в силу того, что полипептидная цепь, содержащая β-аминокислотные остатки, имела бы ковалентные связи типа C_α—C_β, которые допускают дополнительную подвижность (вращение), затрудняющую самопроизвольное свертывание цепи в регулярные структуры, свойственные природным белкам (см. ниже). Вероятно, поэтому в ходе эволюции отбор был направлен на преимущественное сохранение α-аминокислот.

Для аминокислот характерна оптическая изомерия, однако в белках присутствуют, как правило, L-аминокислоты, и L-конфигурация, по всей видимости, была отобрана случайно. D-формы аминокислот обнаруживаются в составе пептидов достаточно редко. Они выявлены, например, в клеточных стенках бактерий сибирской язвы (D-Glu), а также в составе пептидов кожи одного из видов южноамериканских лягушек, в которых найден D-Ala.

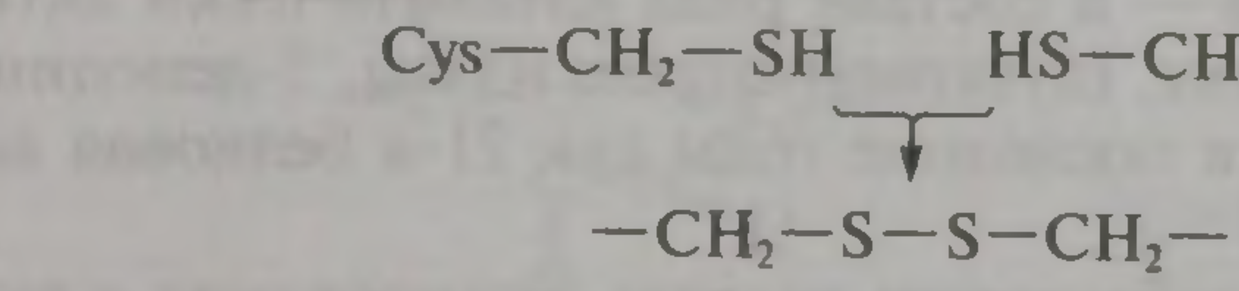
Единственной белковой аминокислотой, не обладающей оптической активностью, является глицин, не имеющий хирального углеродного атома. У глицина в качестве радикала выступает атом водорода, и само понятие радикал в данном случае имеет условный характер. Отсутствие радикала приводит к тому, что Gly увеличивает подвижность полипептидной цепи в белках и одновременно (из-за малого размера) способствует ее плотной упаковке. Остальные аминокислоты содержат в качестве радикалов разнообразные по химическому строению и свойствам группировки, по которым их обычно и классифицируют.

К аминокислотам с *алкильными* радикалами относятся аланин, валин, лейцин и изолейцин (последний содержит 2 хиральных центра и соответственно имеет 4 оптических изомера). Аланин — самая распространенная в природе аминокислота, содержание которой в белках обычно максимально, вероятно, в силу простоты ее строения

и доступности для синтеза клеточными ферментами. Разветвленные объемные и «жесткие» радикалы Val, Leu и Ile, в свою очередь, существенны для гидрофобной стабилизации белковых тел.

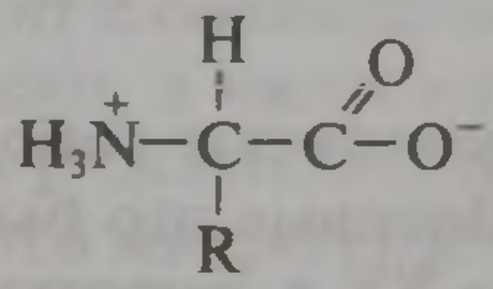
К группу *ароматических аминокислот* входят фенилаланин, тирозин и триптофан. Крупные неполярные остатки Phe и Trp наряду с Val, Ile и Pro располагаются преимущественно внутри белковых молекул. Тирозин содержит функционально активную группу OH, протон которой легко диссоциирует, что способствует образованию водородных связей. Поэтому тирозин, относящийся к полярным незаряженным аминокислотам (к ним относятся также серин, треонин, глутамин, аспарагин и цистеин), участвует в образовании многочисленных водородных связей между боковыми цепями аминокислотных остатков при формировании уникальных пространственных структур белков. По той же причине остатки Tyr, Ser и Cys часто выступают в виде функционально активных групп при ферментативном катализе. Гидроксильные группы, свойственные серину и треонину, способны образовывать эфиры с фосфорной и рядом органических кислот. Поэтому, в частности, функционально важное фосфорилирование ряда белков осуществляется чаще всего по остаткам данных аминокислот. Серин, как и тирозин, присутствует в активных центрах ряда ферментов. По гидроксильным группам Ser и Tyr происходит присоединение к белкам углеводных компонентов (гликозилирование), что имеет место при синтезе сложных белков — гликопротеинов.

Цистеин и метионин составляют группу *серосодержащих аминокислот*. Характерной особенностью остатков Cys является их способность к самопроизвольному окислению в присутствии кислорода с образованием «двойной» аминокислоты — цистина. В белках это ведет к возникновению поперечных ковалентных связей между полипептидными цепями либо внутри одной полипептидной цепи. Так возникают *дисульфидные мостики*:



Содержание Cys особенно высоко в особой группе белков, способных связывать атомы тяжелых металлов — металлотионеинах (см. ниже).

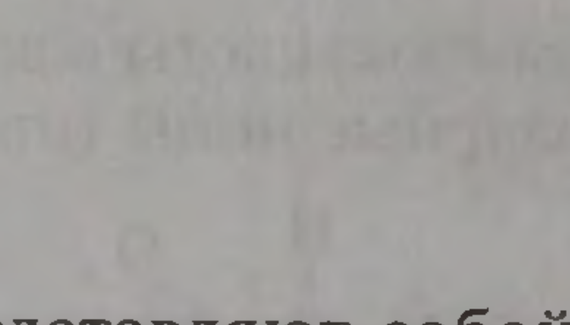
Следует помнить, что аминокислоты в водной среде живых клеток представляют собой биполярные ионы (цвиттерионы):



Поэтому большинство аминокислот, будучи моноаминомонокарбоновыми кислотами, не обладают сколько-нибудь выраженным

зарядом вблизи нейтральных значений pH. К *заряженным аминокислотам* относят те из них, которые содержат дополнительные заряженные группы в боковых цепях (радикалах). Так, к аминокислотам, обладающим кислотными свойствами, относятся аспарагиновая и глутаминовая, имеющие при нейтральном значении pH отрицательный заряд благодаря наличию в их радикалах карбоксильной группы. К заряженным положительно (основным) аминокислотам относятся лизин, аргинин и гистидин. Заряженные остатки Lys и Arg могут эффективно связывать фосфорные остатки, что лежит в основе взаимодействия положительно заряженных белков хроматина (гистонов) с молекулами ДНК (см. гл. 9). Гистидин содержит гетероциклическую (имидазольную) группу, которая в физиологическом интервале значений pH может выполнять роль акцептора протона. Поэтому His присутствует в активных центрах многих ферментов, где выполняет роль «протонной помпы».

В настоящее время пока невозможно однозначно объяснить, почему именно рассмотренные аминокислоты были отобраны в ходе эволюции для построения большинства белков. Простые по строению Gly, Ala, Glu, Ser, Asp считаются более древними в эволюции в отличие от сложных аминокислот, таких, как Asn, His. Кроме вышеохарактеризованных 20 канонических аминокислот в белках встречаются и другие аминокислоты, одна из которых, *селеноцистеин*, содержит атом селена:



Селеноцистеин обнаружен у самых разных организмов — от архибактерий до человека — в составе ряда каталитически активных белков (глицин-редуктазы, глутатион-пероксидазы, 5-деиодиназы и др.) и рассматривается в последние годы как 21-я белковая аминокислота.

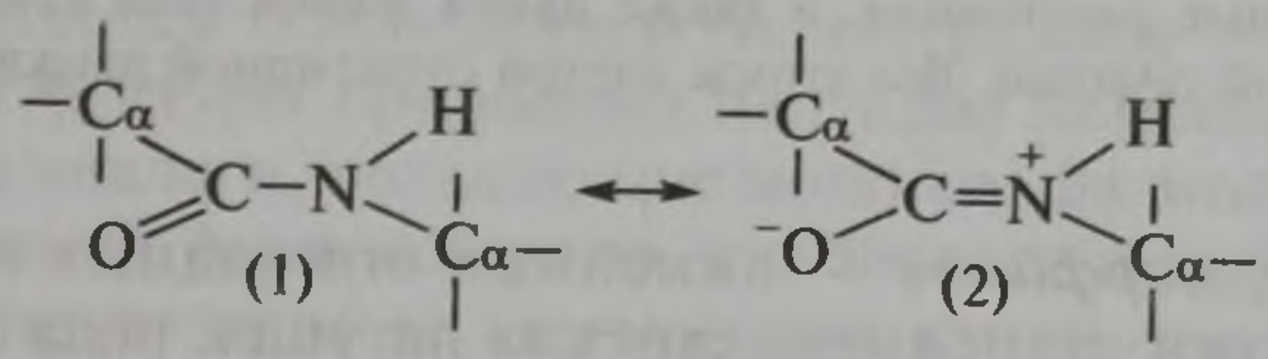
Недавно открытые особенности синтеза, кодирования и включения селеноцистеина в полипептидные цепи белков будут рассмотрены в гл. 11.

Пептиды

α-Аминокислоты представляют собой мономерные блоки, из которых строятся белки. Впервые это было доказано Э. Фишером, синтезировавшим в начале XX в. короткие пептиды, состоящие из нескольких аминокислотных остатков. В 50-х годах XX в. Р. Меррифильд разработал технологию твердофазного синтеза полипептидов

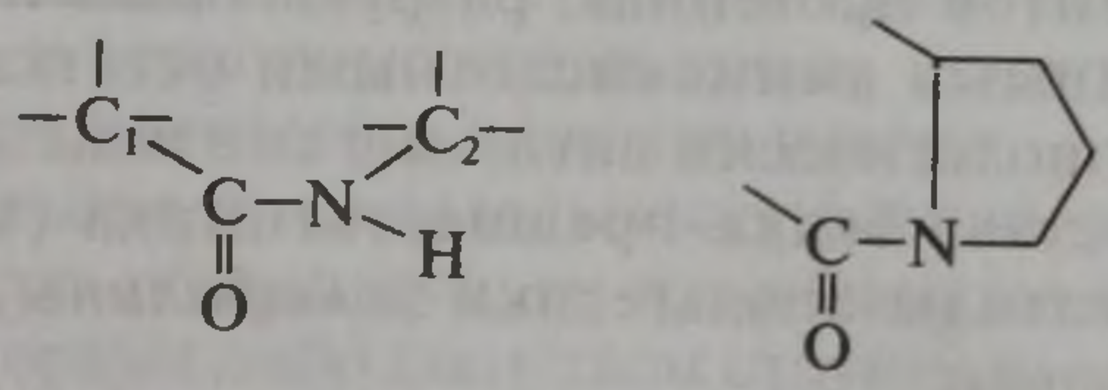
для получения полноразмерных белковых молекул, имеющих в своем составе более сотни аминокислотных остатков. Так впервые был синтезирован каталитически активный белок — рибонуклеаза, состоящая из 124 аминокислотных остатков, а затем и другие функционально активные молекулы белков, что стало абсолютным доказательством полипептидной теории их строения.

Таким образом, полипептидные цепи белков состоят из аминокислотных остатков, связанных *пептидными связями*. Структура пептидной связи приведена на рис. 6. Ее параметры (межатомные расстояния и валентные углы) были определены Л. Полингом и Р. Кори на кристаллических структурах молекул, содержащих одну или несколько пептидных связей. Оказалось, что расстояние между атомами C и N (0,132 нм) меньше обычного, тогда как связь C—O длиннее, чем в обычных органических соединениях (альдегидах и кетонах). Полинг и его сотрудники объясняли это наличием резонанса двух предельных структур пептидной связи:



Пептидная связь имеет, таким образом, частично двойной характер и состоит на 60 % из формы (1) и на 40 % из формы (2).

В белках наиболее распространена *транс*-пептидная связь, представленная на рис. 6, вращение вокруг которой заторможено, и шесть атомов (C, O, N, H, а также C_1 и C_2) находятся в одной плоскости, из чего следует, что полипептид можно рассматривать как цепочку из связанных между собой плоских звеньев. *Цис*-пептидная связь менее распространена и чаще всего образуется при атоме азота пролина:



Цис-связи обнаружены в ряде природных циклических пептидов. Хотя полипептидная цепь имеет обычно одну свободную аминогруппу на одном конце (N-конец) и одну свободную карбоксильную группу на другом конце (C-конец), иногда эти группы связываются одна с другой, что и приводит к образованию циклического пептида. Циклическое строение имеют, в частности, природные пептидные антибиотики грамицидин S и циклоспорин. Эти и некоторые другие циклопептиды синтезируются микроорганизмами по особому нематричному мультиэнзимному механизму и содержат в своем составе неканонические аминокислоты (N-метилированные аминокислотные остатки, D-изомеры аминокислот и др.). D-аланин недавно выявлен

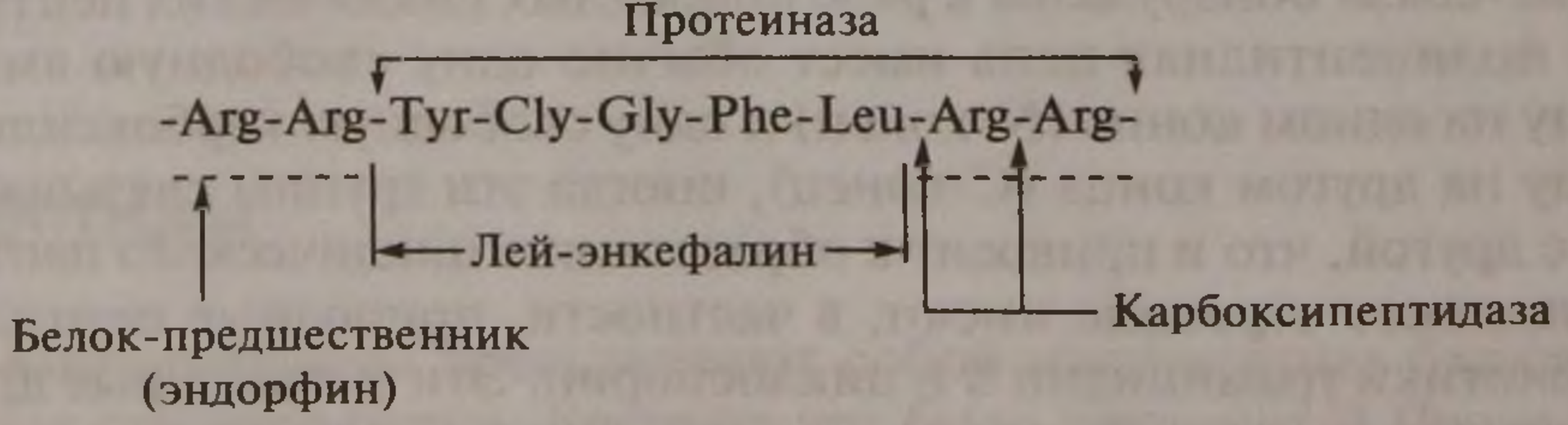


Рис. 6. Строение пептидной группы.

Указаны межатомные расстояния, а также длина водородной связи, образуемой с соседней пептидной группой. Все атомы внутри пунктирной рамки лежат примерно в одной плоскости

и в составе *дерморфинов* — линейных опиоидных пептидов, выделенных из кожи южноамериканских лягушек рода *Phyllomedusa*, у которых он находится в последовательности: Tyr-D-Ala-Phe-Gly--Tyr-Pro-Ser-амид. В составе дерморфинов D-Ala образуется из L-Ala уже после синтеза пептида в результате действия особого фермента «пептидизомеразы», который выявлен также в яде воронковых пауков. Этот фермент способен катализировать превращение L → D-Ala в составе пептидов.

Многие природные пептиды возникают из более крупных белков-предшественников в результате *ограниченного протеолиза* под действием ферментов протеиназ, разрушающих пептидные связи между определенными аминокислотными остатками, тем самым высвобождая физиологически активные пептиды. Таким образом, например, из состава белка-предшественника (эндорфина) высвобождаются пептиды-анальгетики *энкефалины*, относящиеся к опиоидным пептидам:



Из белка-предшественника могут одновременно вырезаться либо несколько молекул одного и того же пептида, либо несколько различных по структуре и обладающих различными физиологическими эф-

фектами пептидов. Например, в составе β-эндорфина обнаружен набор пептидов, получивший название пропиомеланокорпин (ПОМК), в состав которого помимо энкефалина входят вырезаемые соответствующими протеиназами три пептидных гормона: меланоцитстимулирующий гормон (МСТ), адренокортикотропный гормон (АКТГ) и липотропин. Таким образом, из одного белка-предшественника образуется целый «пакет» пептидов, модифицирующий обмен веществ в организме в стрессовых ситуациях. Как следует из вышесказанного, многие природные пептиды выполняют определенные физиологические функции. Для некоторых из них характерна полимодальность — способность влиять на многие процессы в организме. Так, тетрапептид тафцин (Tre-Lys-Pro-Arg) является одновременно и иммуностимулятором, и психостимулятором; трипептид глутатион (γ-Glu-Cys-Gly) поддерживает восстановленную форму железа (Fe^{2+}) в гемоглобине и участвует в переносе аминокислот через мембранные структуры в клетке. Ряд пептидов, вырабатываемых в мозге, способны управлять эмоциями и поведением (психотропные пептиды).

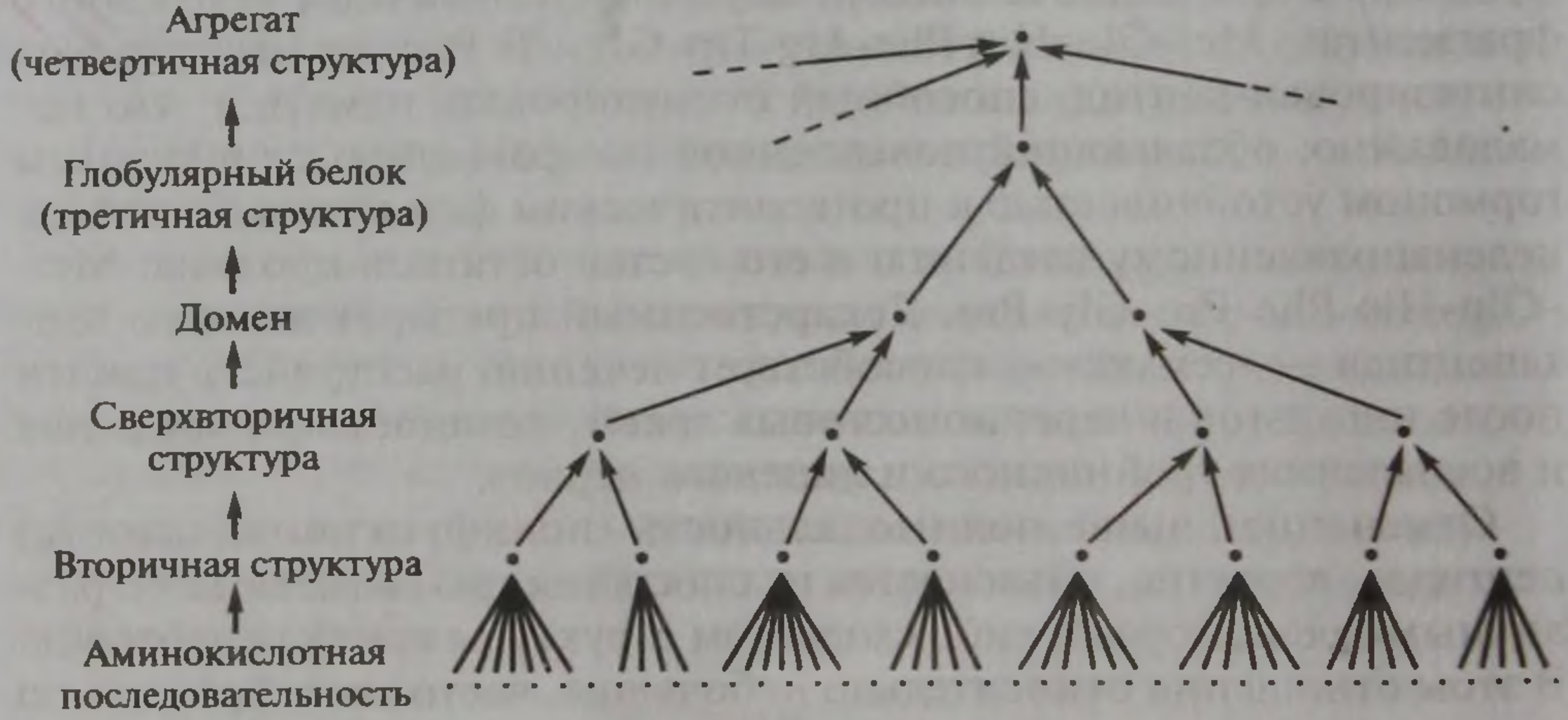
Изучение структуры природных пептидов помогает создавать их синтетические аналоги, обладающие мощным физиологическим эффектом. Например, было показано, что биологический эффект адренокортикотропного гормона (гормон пептидной природы, содержит 39 аминокислотных остатков, синтезируется в гипофизе, регулирует функцию коры надпочечников) обусловлен наличием пептидного фрагмента: -Met-Glu-His-Phe-Arg-Trp-Gly-. В России недавно был синтезирован пептид, способный стимулировать память и, что немаловажно, обладающий повышенной по сравнению с природным гормоном устойчивостью к протеолитическим ферментам благодаря целенаправленному введению в его состав остатков пролина: Met--Glu-His-Phe-Pro-Gly-Pro. Лекарственный препарат данного гептапептида — «семакс» — способствует лечению расстройств памяти после инсультов и черепномозговых травм, помогает при мигренях и воспалениях тройничного и лицевого нервов.

Отмеченная выше полимодальность (полифункциональность) пептидов, вероятно, объясняется их способностью связываться с различными рецепторами либо сходством структуры самих рецепторов. В этом отношении относительно небольшие, часто своеобразные по аминокислотному составу природные пептиды существенно отличаются от своих более крупных «собратьев» — собственно белков, или протеинов (от греч. «протос» — первичный), у которых выражено соответствие между последовательностью расположения аминокислотных остатков в полипептидной цепи и определенной пространственной структурой молекулы. Несмотря на то что некоторые пептиды способны принимать определенную конформацию, удобную для взаимодействия с рецепторами, только белки способны формировать и удерживать сложные пространственные структуры, для образования которых необходимы достаточно протяженные полипептидные цепи.

Нижним предельным размером для белков считается полипептидная цепь, состоящая из 50 аминокислотных остатков. К таким белкам относятся, в частности, фактор роста эпителия, инсулин и белки оболочки ряда бактериофагов. Чаще всего белки имеют в своем составе 100 — 400 аминокислотных остатков, образующих в пространстве определенную трехмерную структуру.

Структурная организация белков

В 1959 г. датский биохимик К. Линдерстрем-Ланг предложил различать *четыре уровня* структурной организации белков: *первичную*, *вторичную*, *третичную* и *четвертичную структуры*, которые обозначают соответственно аминокислотную последовательность, упорядоченное строение основной цепи полипептида, трехмерную структуру белка и структуры белковых агрегатов. Эта классификация господствовала до начала 80-х годов XX в., когда Г. Шульц и Р. Ширмер с учетом новых данных о структурных особенностях белков предложили дополнить ее еще двумя уровнями организации: *сверхвторичными структурами* и *доменами*. В итоге сложилось представление о шести уровнях структурной организации белков с определенной иерархией, выраженной Шульцем и Ширмером в виде схемы.



Как видно из схемы, в основе структурной организации белков лежит определенная генетически детерминированная аминокислотная последовательность полипептидной цепи (первичная структура), которая определяет все последующие более высокие уровни организации.

Первичная структура белка. Под первичной структурой белка понимают порядок чередования аминокислотных остатков в полипептидной цепи белка. Эта структура, достаточно жестко запрограммированная в определенной нуклеотидной последовательности соответствующих структурных генов ДНК, переписывается (*транс-*

крибируется) в комплементарные нуклеотидные последовательности мРНК, которые служат матрицами для биосинтеза полипептидных цепей в процессе декодирования (*трансляции*) генетической информации в рибосомной системе клетки. Тем не менее первичная структура зрелого функционально активного белка не всегда соответствует структуре мРНК, что связано с возможностями перепрограммирования в ходе трансляции (см. гл. 11), а также с процессами посттрансляционной модификации синтезированных на рибосомах полипептидных цепей.

Внушительные размеры полипептидных цепей белков, которые могут включать несколько тысяч аминокислотных остатков (цепь тиреоглобулина состоит из 2 750 аминокислотных остатков, а цепи отдельных трансмембранных белков-рецепторов содержат до 5 000 аминокислотных остатков), осложняют расшифровку их первичной структуры. Для расшифровки используются специальные методы, с которыми связан прогресс в изучении молекулярной биологии белковых тел. Первой удалось расшифровать аминокислотную последовательность инсулина (Ф. Сангер, 1953). Она состоит из 51 аминокислотного остатка и организована в виде двух полипептидных цепей (цепь А — 21 остаток и цепь В — 30 остатков), ковалентно связанных между собой двумя дисульфидными мостиками (рис. 7). Эта длительная работа, называемая иногда «ручным секвенированием» (от англ. *sequence* — последовательность, сочетание), была выполнена динитрофторбензольным методом Ф. Сангера (ДНФ-метод), с помощью которого позднее были расшифрованы первичные структуры аспартатаминотрансферазы (Ю. А. Овчинников, А. Е. Браунштейн, 1971) и еще нескольких крупных полипептидов. Анализ аминокислотных последовательностей белков резко ускорился

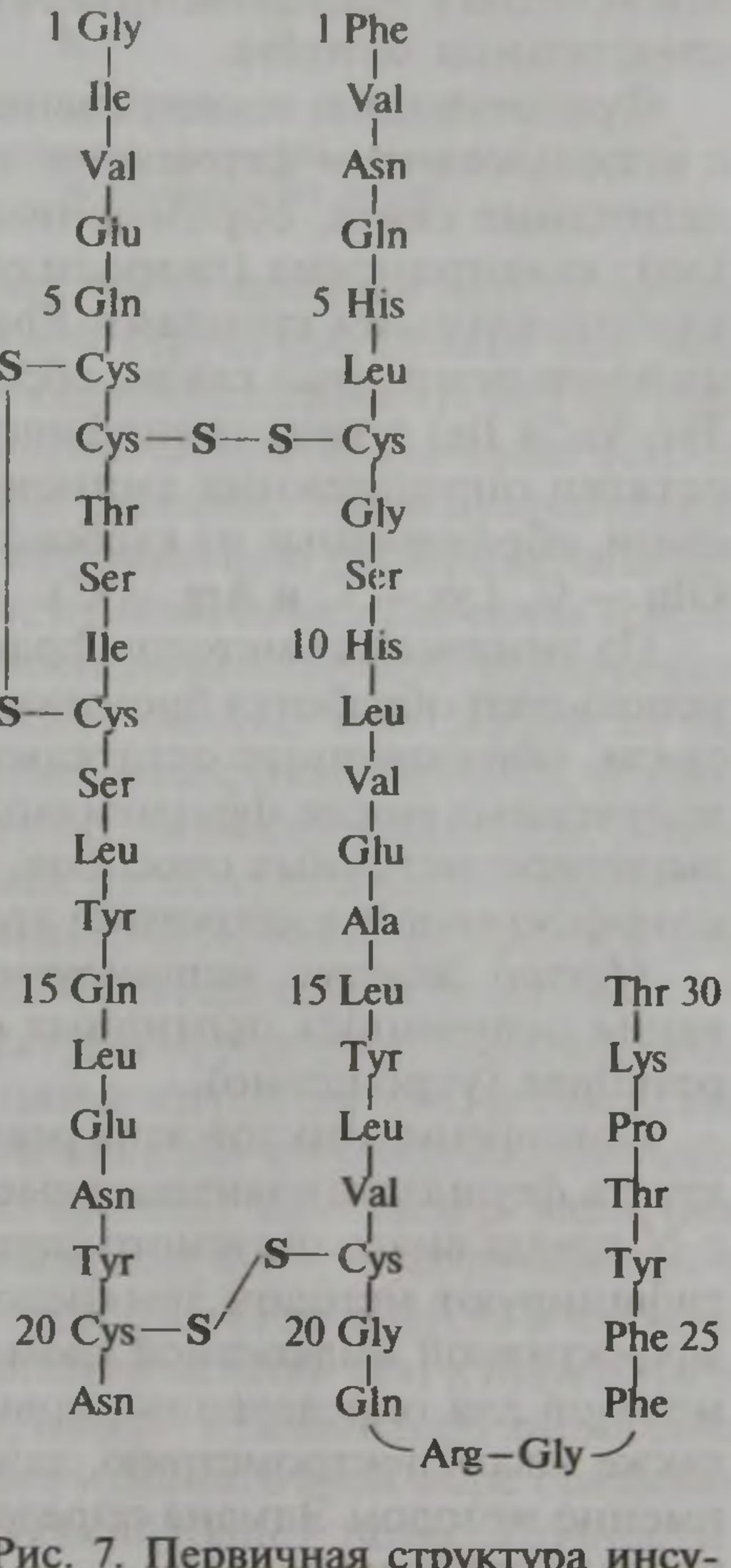


Рис. 7. Первичная структура инсулина.

Показана последовательность расположения аминокислотных остатков начиная с N-концов двух полипептидных цепей, а также расположения дисульфидных связей

после внедрения метода П. Эдмана (см. ниже) и создания автоматических приборов (секвенаторов), которые позволяют в автоматическом режиме осуществлять секвенирование фрагментов полипептидных цепей длиной в несколько десятков (до 50) аминокислотных остатков. Однако даже использование секвенаторов требует значительного времени для определения первичной структуры крупных белков, которые нередко состоят из нескольких полипептидных цепей. В этом случае вначале требуется установить число цепей, затем разделить эти цепи, расщепить их на относительно короткие фрагменты, пригодные для автоматического секвенирования по методу П. Эдмана, определить их структуру и только тогда составить представление о целостной аминокислотной последовательности белка. Все эти операции требуют специальных методов.

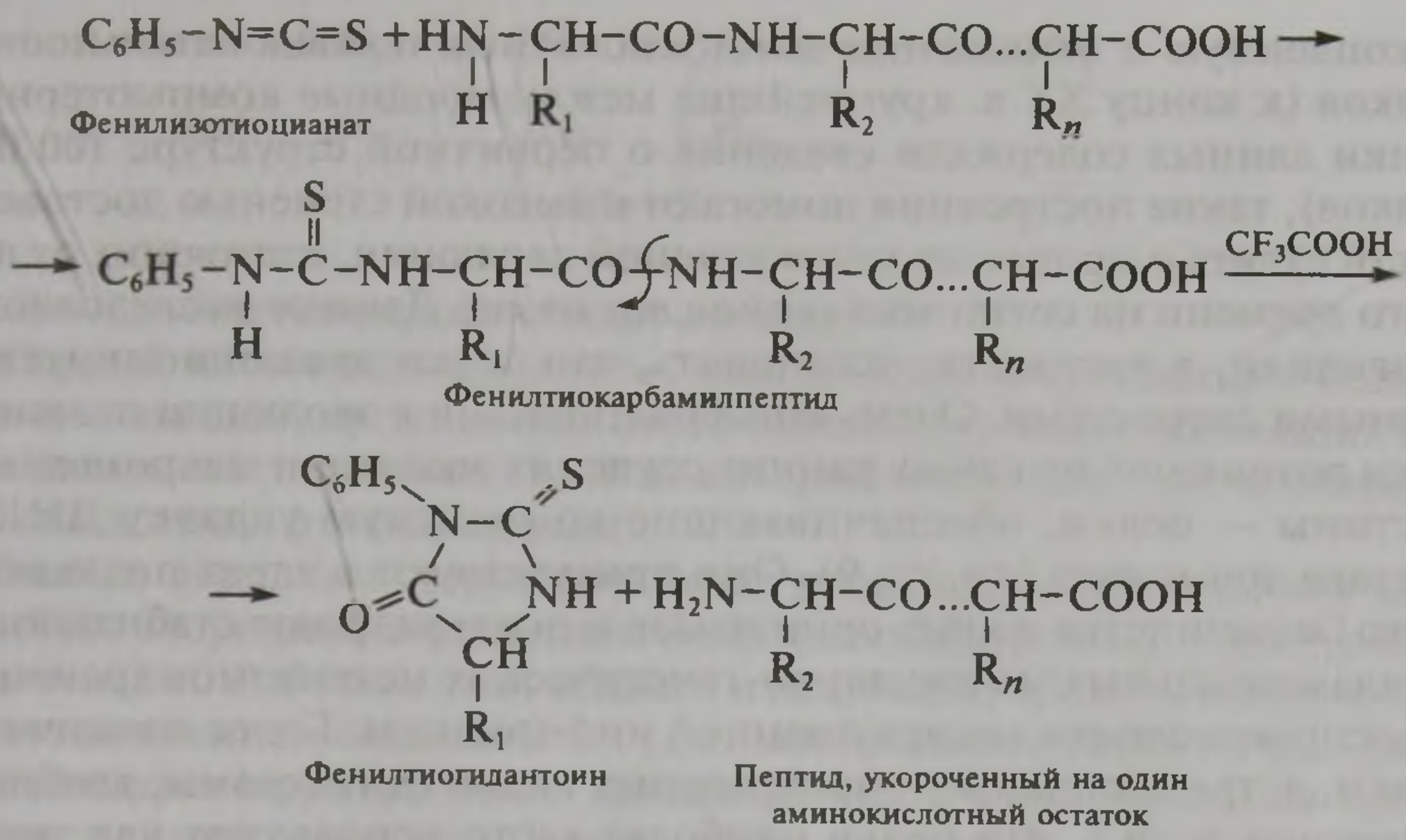
Фрагментацию изолированных полипептидных цепей проводят с использованием ферментов: трипсина (избирательно гидролизует пептидные связи, образованные карбоксильными группами Arg и Lys), химотрипсина (гидролизует пептидные связи, образованные карбоксильными группами Phe, Tyr, Trp, Leu), термолизина (расщепляет пептидные связи, образованные аминогруппами Leu, Phe, Tyr, Val и Ile) и ряда специфических эндопротеиназ, распознающих остатки определенных аминокислот и расщепляющих пептидные связи, образованные их карбоксильными группами (эндопротеиназы Glu — C, Lys — C и Arg — C).

Из химических методов фрагментации полипептидов чаще всего используют обработку бромцианом, который расщепляет пептидные связи, образованные остатками метионина. Разделение пептидов, полученных после фрагментации полипептидных цепей одним из вышеперечисленных способов, чаще всего проводят методом высокоэффективной жидкостной хроматографии (HPLC).

Метод Эдмана, используемый для автоматического секвенирования полученных пептидных фрагментов, основан на следующих реакциях (упрощенно).

Повторение циклов этих реакций позволяет последовательно получать фенилтиогидантоиновые производные аминокислот начиная с N-конца анализируемого пептида. Эти производные затем идентифицируют методом тонкослойной хроматографии либо высокоэффективной жидкостной хроматографией. Кроме вышеотмеченных методов для определения первичной структуры белков используют также масс-спектрометрию, лазерную рентгеноскопию и др. Однако именно методом Эдмана определено большинство выясненных первичных структур белков.

Даже самые прогрессивные методы белкового (прямого) секвенирования требуют получения значительных (чаще всего миллиграммовых) количеств гомогенных препаратов белков, получение которых, особенно в отношении минорных белков, содержащихся в клетках в ничтожно малом количестве, сопряжено с длительными трудоемки-



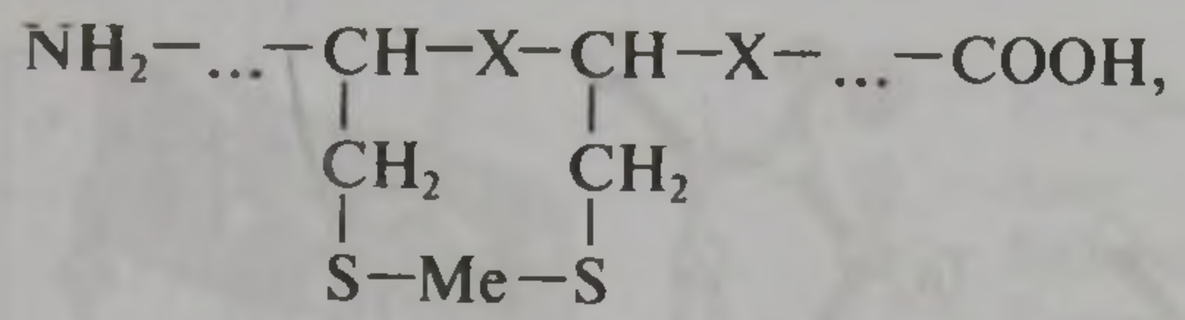
ми процедурами их выделения и очистки от примеси других белков. Поэтому для определения первичной структуры белков все чаще используют анализ нуклеотидной последовательности кДНК, применяя соответствующие методы *генетической инженерии* (см. гл. 14), что занимает значительно меньше времени. Узнав нуклеотидную последовательность кДНК или соответствующую последовательность мРНК, можно, учитывая особенности *генетического кода* (см. гл. 11), определить аминокислотную последовательность кодируемого ими белка (не имея препарата белка!). Однако этот подход не позволяет выявить посттрансляционные модификации белка, и такое (косвенное) секвенирование нельзя считать исчерпывающим. Поэтому методы прямого секвенирования на основе анализа белковых препаратов не утратили своего значения. Интенсивная работа в этом направлении имеет как научное, так и практическое значение.

Сопоставление первичных структур нормальных и аномальных белков позволяет раскрыть основы патологических процессов, включая широко распространенные генетически детерминированные заболевания. Именно этот подход выявил причину серповидно-клеточной анемии — распространенной болезни крови, которая возникает из-за точечной мутации гена, кодирующего последовательность β-цепи гемоглобина. Замена одного-единственного остатка Glu в нормальном гемоглобине типа A на остаток Val приводит к появлению аномального гемоглобина типа S, не способного в достаточной мере связывать кислород, что приводит к развитию кислородной недостаточности в тканях человека, наблюдаемой при этом заболевании.

Выявление структурного подобия (гомологии первичных структур) белков лежит в основе изучения молекулярной эволюции белков, а дендрограммы, построенные на основе первичных структур белков, широко применяются для выявления филогенетических связей различных групп организмов. Учитывая огромную информацию,

накопленную в отношении аминокислотных последовательностей белков (к концу XX в. крупнейшие международные компьютерные банки данных содержали сведения о первичной структуре 100 000 белков), такие построения помогают с высокой степенью достоверности судить о процессах молекулярной эволюции, отстоящих от нашего времени на сотни миллионов лет назад. Данные исследования позволили, в частности, установить, что белки эволюционируют с разными скоростями. Очень консервативными в эволюции оказываются возникшие на самых ранних ступенях эволюции макромолекул гистоны — белки, обеспечивающие компактную укладку ДНК в составе хроматина (см. гл. 9). Они присутствуют в ядрах подавляющего большинства живых организмов и поддерживают стабильность фундаментальных молекулярно-генетических механизмов хранения и воспроизведения наследственной информации. Более изменчивы белки, встречающиеся у определенных видов (цитохромы, глобины животных и др.). Эти белки наиболее часто используют как модели для изучения молекулярной эволюции. Сходство или различие первичных структур используются для построения молекулярно-филогенетических «деревьев» с привлечением для этого специально разработанных компьютерных программ, что позволяет установить скорость развития белков на отдельных этапах эволюции органического мира. Так, в отношении глобинов установлено, что с максимальной скоростью эти белки эволюционировали 400—500 млн лет назад, в период выхода позвоночных животных на сушу. Особенно сильно изменялась первичная структура тех участков их полипептидных цепей, которые ответственны за образование четвертичной структуры. К наиболее быстро эволюционирующим белкам принадлежат белки РНК-содержащих вирусов (вирусов гриппа, вирусов иммунодефицита человека, онкогенных вирусов). Не следует забывать, что в основе эволюции белков лежит изменчивость соответствующих генов, фиксированные нуклеотидные замены (мутации) в кодирующих областях которых и приводят к изменению первичной структуры белков. Вместе с тем естественному или индуцированному мутационному процессу в живой клетке постоянно противостоит *система репарации* (исправления повреждений нуклеотидов), которая мощно развита в отношении ДНК (см. гл. 12) и малоэффективная в отношении РНК, чем, вероятно, и объясняется необычно высокая скорость изменчивости белков, кодируемых РНК-содержащими геномами.

Анализ первичной структуры существен для выяснения и предсказания функциональных свойств белков. Так, например, специфическая группа широко распространенных в природе белков — металлотионеинов, способных связывать весьма значительное количество тяжелых металлов (Cd, Cu, Fe и др.), обладает обогащенной остатками цистеина аминокислотной последовательностью регулярной структуры:



где X — любой аминокислотный остаток.

Первый из металлотионеинов, выделенный Марголисом и Вэлли из почек лошади, представляет собой небольшой белок, состоящий из 61 аминокислотного остатка, 20 из которых — остатки Cys, служащие для связывания атомов кадмия.

Знание первичной структуры белка необходимо для создания моделей его вторичной структуры и более высоких уровней структурной организации. Работы в этой области в настоящее время выполняются с использованием специальных компьютерных программ, в которых учитываются определенные закономерности формирования пространственной структуры белков, напрямую связанные с их аминокислотными последовательностями. Фундаментальные представления о закономерностях укладки в пространстве полипептидных цепей были сформулированы в середине XX в. в основополагающих работах Л. Полинга, посвященных изучению вторичной структуры белков.

Вторичная структура белков. Упорядоченное расположение отдельных участков основной цепи полипептида, без учета расположения боковых цепей (радикалов аминокислотных остатков) — это вторичная структура белков. Исследования вторичной структуры, основанные на данных рентгеноструктурного анализа, показали наличие в белках регулярно свернутых и закономерно закрученных участков полипептидной цепи. Л. Полинг и Р. Кори в 1951 г. впервые постулировали спиральное расположение в пространстве полипептидных цепей, а затем сформулировали основные принципы построения моделей вторичных структур белков.

В соответствии с этими правилами было предложено несколько типов белковых спиралей — α, 3_{10}, π и др. (рис. 8). Дальнейший анализ структур белков подтвердил существование в природе преимущественно α-спиралей, что объясняется их высокой стабильностью. Спирали типа 3_{10} (имеют 3 аминокислотных остатка в одном витке и 10 атомов в цикле, образованном водородной связью) встречаются в природе исключительно редко и образуют только короткие участки. Вероятно, это связано с тем, что спираль 3_{10} энергетически не выгодна, так как в ней водородные связи не находятся на одной линии. Эти спирали образуют всего 1 — 2 витка и встречаются у единичных белков: в гемоглобине морской миноги, а также в белке паутины, для первичной структуры которого характерно наличие кластеров аминокислотных остатков типа Gly-Gly-X, образующих спирали этого типа. Спирали π в белках не обнаружены. Преобладающая в белках α-спираль имеет радиус 0,23 нм, расстояние между витками составляет 0,54 нм, среднее число аминокислотных остатков, приходящееся на

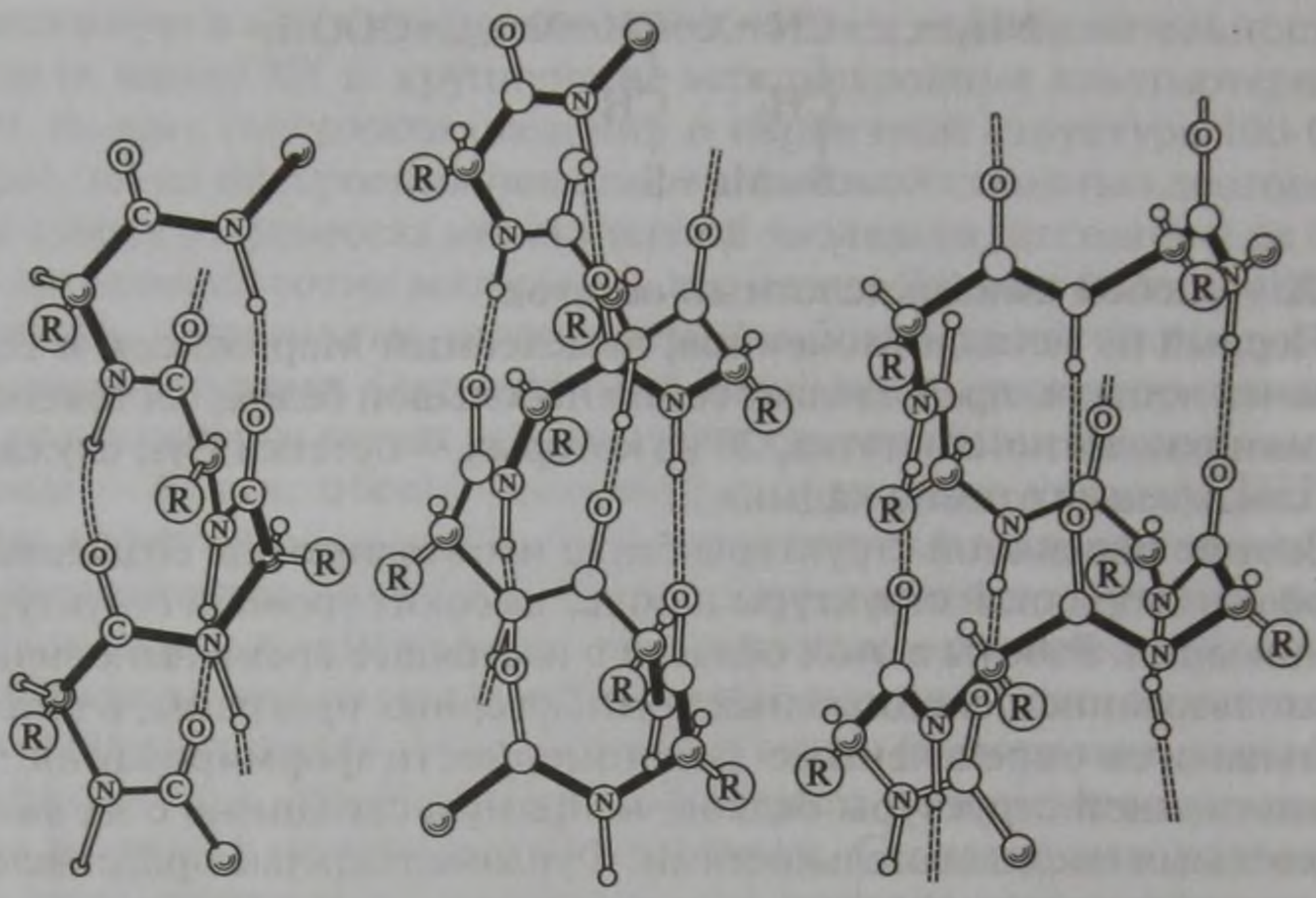


Рис. 8. Различные типы (слева направо: 3_{10}, α и π) спиралей в полипептидных цепях белков (Г. Шульц, Р. Ширмер, 1982). Пояснения см. в тексте

один оборот, — 3,6, угол подъема витка спирали — 26 °. В природных белках существуют только правозакрученные α-спирали, что связано с наличием в них аминокислот L-ряда.

Все вторичные структуры стабилизированы водородными связями между кислородом карбонильной группы и водородом имидной группы пептидных связей: C=O...H—N. Архитектоника расположения атомов азота и карбонильных групп в α-спирали такова, что создает возможность для образования максимального количества водородных связей между этими группами, поэтому спираль оказывается полностью пронизанной (насыщенной) водородными связями.

В силу этого она имеет плотно упакованную структуру, не имеющую внутреннего канала и непроницаемую даже для молекул воды (рис. 9). Протяженность α-спиральных участков в белках различна: в глобулярных белках она обычно не превышает 15 аминокислотных остатков (3—4 оборота спирали), а в фибриллярных белках гораздо длиннее, и в ней могут наблюдаться перегибы (изломы), чаще всего в местах расположения остатков пролина, не способных к образованию соответствующих водородных связей. Например, в белке паутины плотно упакованная α-спираль, образованная преимущественно остатками Ala и Gln, чередуется с менее регулярными (эластичными) участками, где располагаются кластеры аминокислот типа Gly-Pro-Gly. Длинные стержневидные полипептидные цепи, состоящие из нескольких α-спиралей, обнаружены во многих белках, в частности



Рис. 9. Схематическое изображение α-спирали (показано расположение водородных связей; боковые группы (радикалы аминокислотных остатков) не показаны)

в α-кератине, внутриклеточные волокна которого обеспечивают прочность кожи.

Особая спираль выявлена в составе коллагена, наиболее распространенного белка млекопитающих. Эта правозакрученная спираль состоит из трех левых спиралей, расположенных параллельно и образующих суперспираль. Структурной единицей цепей коллагена является триплет с общей формулой Gly-X-Y, где X чаще всего представлен остатком пролина, а Y — остатком оксипролина. Коллаген содержит 33 % глицина, а на долю пролина и оксипролина в сумме приходится 21 %.

β-***Структуры.*** Эти вторичные структуры, очень широко представленные в белках, были постулированы Л. Полингом и Р. Кори и названы *параллельными «плоскими»* и *антипараллельными* β-складчатыми листами. Экспериментальной моделью для их изучения стал β-фиброин шелка, откуда собственно и возникло название β-структур, подобно тому как название α-спираль появилось в связи с работами по изучению структуры α-кератина. Фиброин шелка состоит в основном из повторяющейся последовательности: -(Gly-Ser-Gly-Ala-Gly-Ala)$_n$-. В отличие от α-спирали в β-структурах водородные связи возникают не внутри цепи, а между параллельно или антипараллельно идущими и сближенными участками цепи (рис. 10). Эти структуры не столь насыщены водородными связями, как α-спираль, поэтому в их состав могут входить несколько участков полипептид-

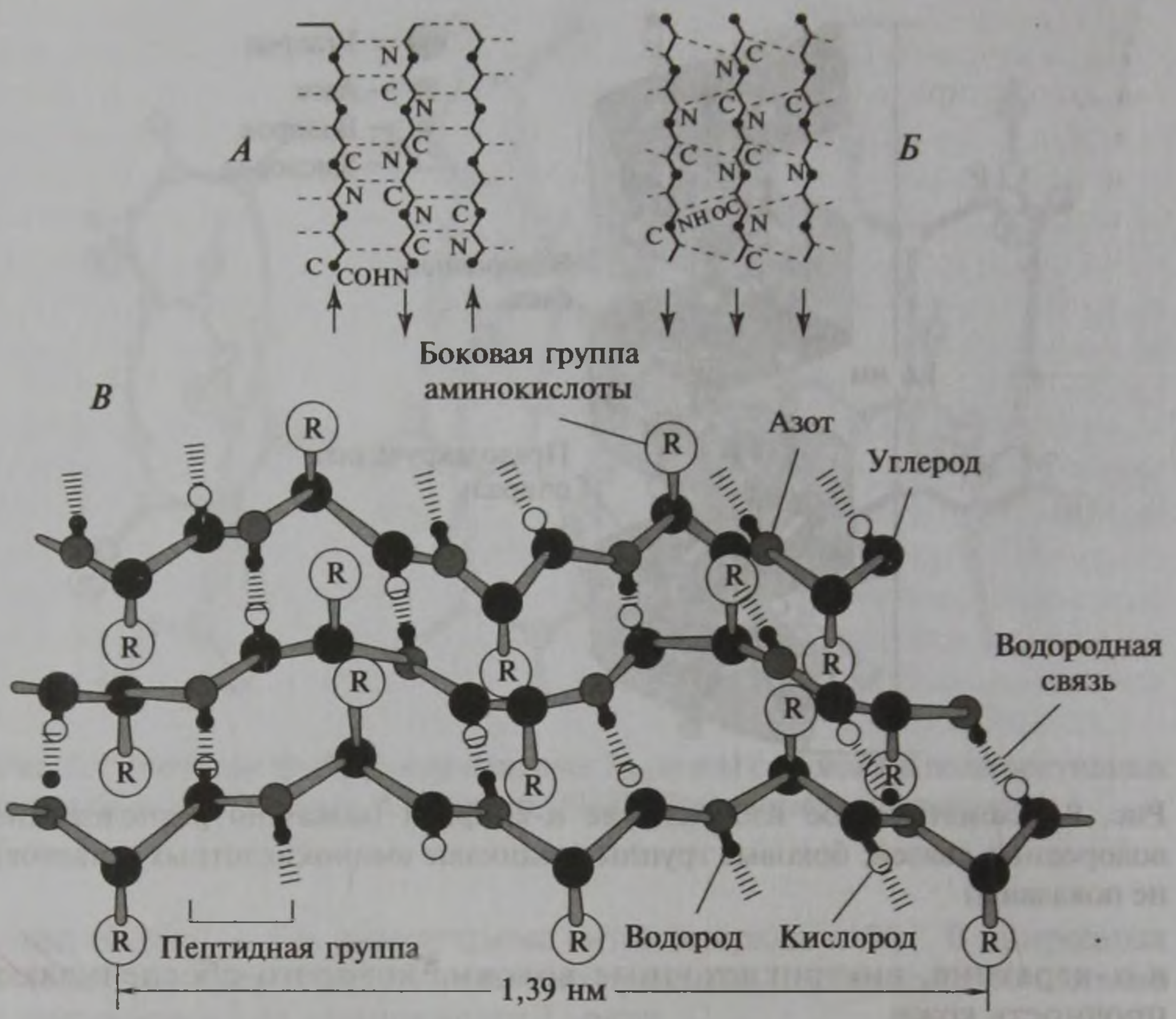


Рис. 10. β-Складчатые листы:

А — антипараллельная трехцепочечная β-структура; *Б* — параллельная трехцепочечная β-структура; *В* — более детальное изображение β-слоя. Водородные связи обозначены пунктирными линиями, направление цепей — стрелками, C_α-атомы — точками

ной цепи. Большинство складчатых листов содержат не более шести цепей, и их средняя протяженность (~2 нм) соответствует размеру структурного домена (см. ниже). У глобулярных белков β-структуры составляют ~15 %, хотя их число может быть и значительно больше (например, во вторичной структуре нейраминидазы — одного из каталитически активных белков вируса гриппа). Радикалы аминокислотных остатков расположены по разные стороны от поверхности β-слоя (см. рис. 10). Они образуют своего рода гребни, в которых нередко сосредоточиваются боковые цепи гидрофобных аминокислот, и такие β-структуры участвуют в образовании гидрофобных ядер глобулярных белков. В белках β-складчатые листы, как правило, не являются абсолютно плоскими, чаще они имеют левую закрученность. Такое закручивание происходит как в фибриллярных (фиброине шелка), так и в глобулярных (карбоангидразе человека) белках.

Наличие в белках α-спиралей и β-структур было положено в основу структурной классификации белков (M. Levitt, C. Chotia,

1976), в соответствии с которой различают белки типа α (содержат только α-спирали), типа β (обладают только β-структурами), а также белки α + β и α/β. Белки α + β содержат α-спирали и β-структуры, отчетливо разделенные в пространстве вдоль цепи. Белки α/β имеют более сложное строение с перемежающимися участками α-спиралей и β-структур (нечто напоминающее слоеный пирог или сэндвич). К наиболее известным и детально изученным белкам типа α относятся миоглобин, цитохром *c*, белок вируса табачной мозаики и инсулин; к белкам типа β — нейраминидаза вируса гриппа, супероксиддисмутаза и конканавалин А, к белкам типа α + β — рибонуклеаза, лизоцим куриного яйца и карбоангидраза, к типу α/β — карбоксипептидаза и триозофосфатизомераза. На рис. 11 представлены схематические изображения трехмерных структур белков с различным представи-

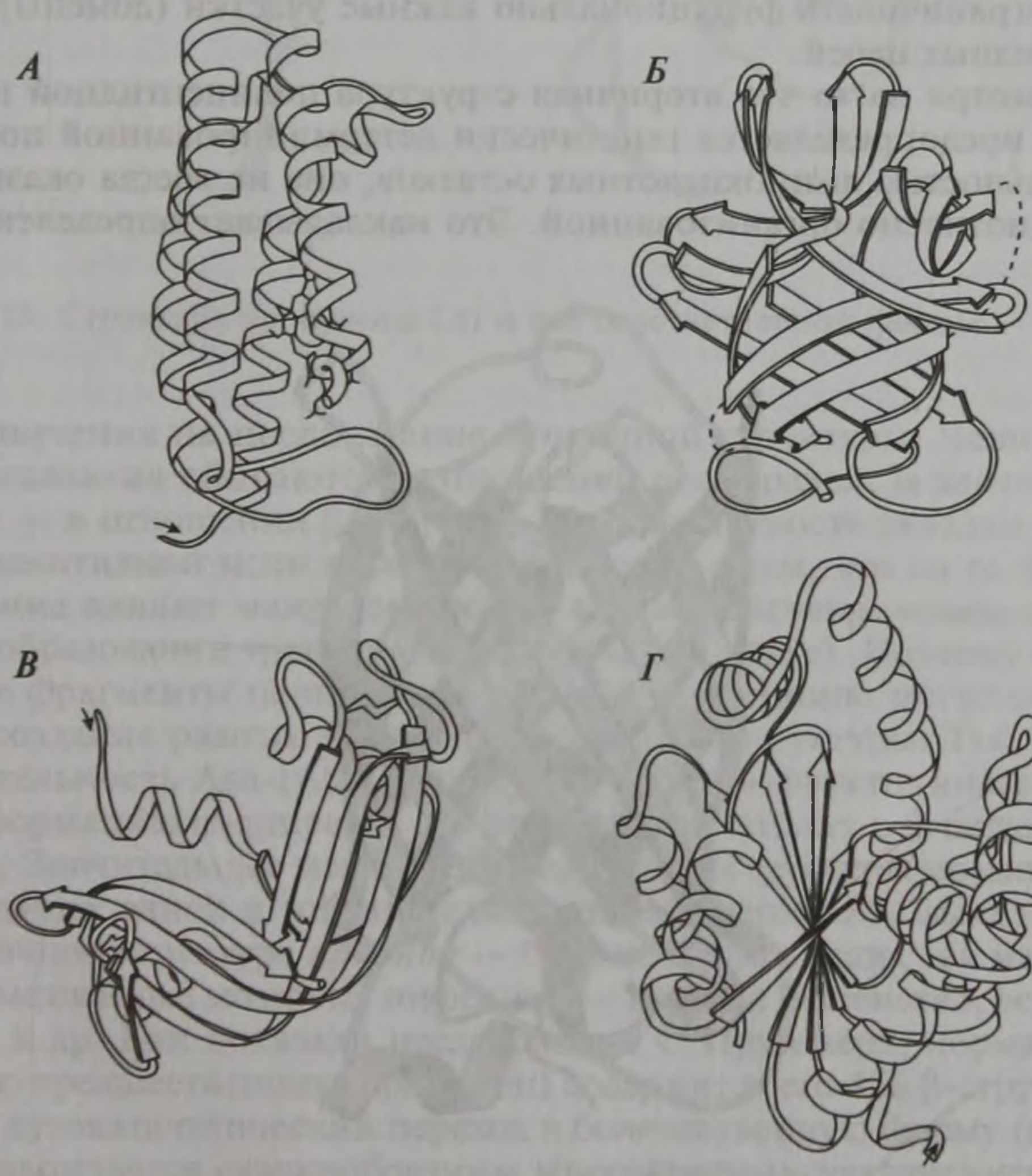


Рис. 11. Белки, относящиеся к различным структурным группам в зависимости от представительства элементов вторичной структуры:

А — цитохром *с* (группа α); *Б* — ингибитор трипсина из соевых бобов (группа β); *В* — рибонуклеаза (группа α + β); *Г* — аденилаткиназа (группа α/β)

тельством вторичных структур. Достаточно причудливое на первый взгляд расположение отдельных элементов вторичной структуры у многих белков, тем не менее, несет в себе определенную функциональную нагрузку. Примером может служить структура зеленого флюоресцирующего белка, у которого 11 β-слоев образуют полый цилиндр (бочонок), защищающий расположенную внутри хромофорную группу (рис. 12).

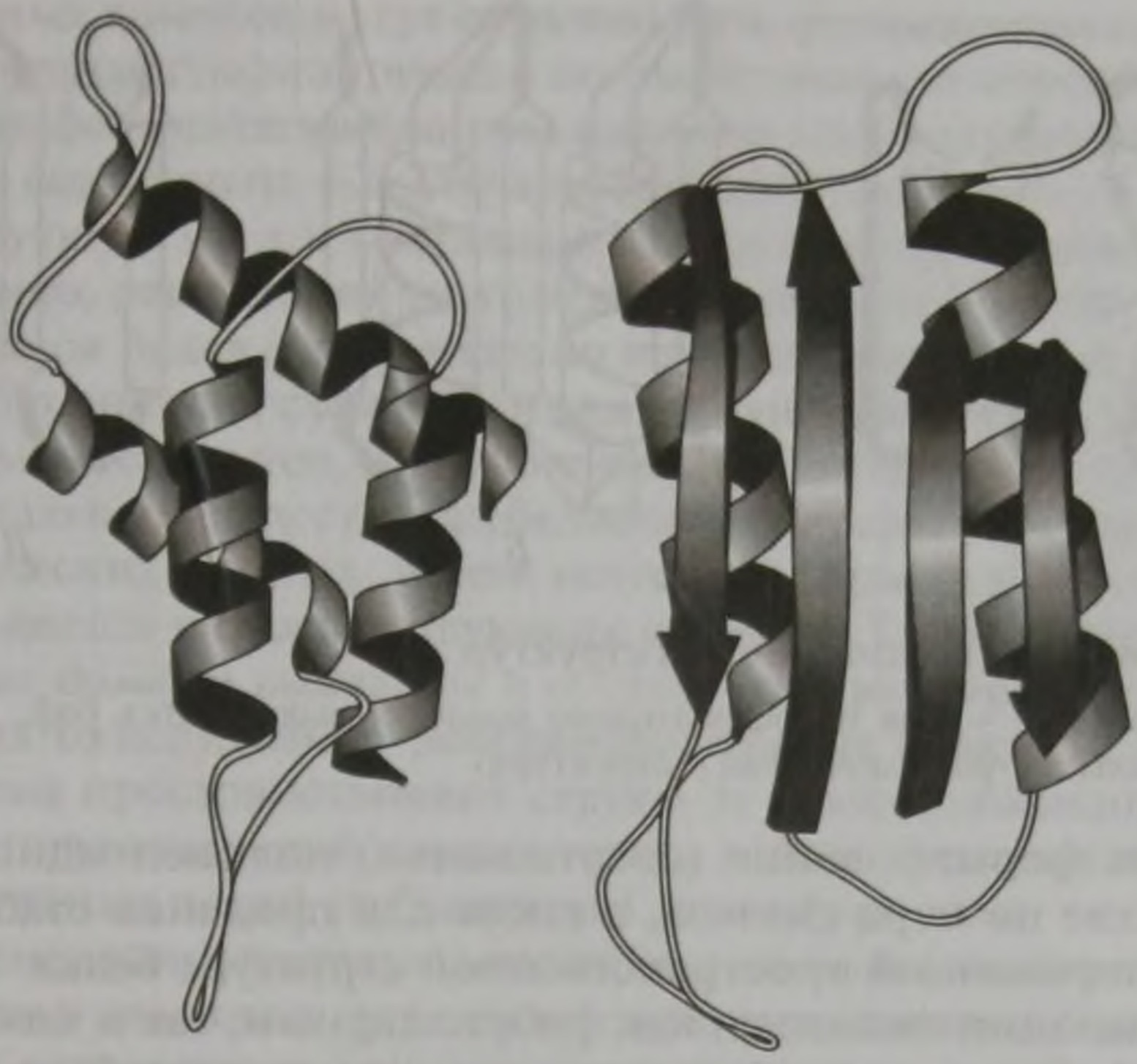
Помимо регулярных вторичных структур в полипептидных цепях белков содержатся значительные участки нерегулярной структуры, образующие различные петли и изгибы. Один из видов такого рода структур получил название β-изгиб. Этот элемент наблюдается при обратном (на 180 °) изменении направления полипептидной цепи и стабилизируется водородной связью между 1-м и 4-м по ходу цепи аминокислотными остатками. Кроме того, петлеобразные структуры могут ограничивать функционально важные участки (домены) полипептидных цепей.

Несмотря на то что вторичная структура полипептидной цепи белков предопределяется генетически детерминированной последовательностью аминокислотных остатков, она не всегда оказывается однозначно организованной. Это накладывает определенные



Рис. 12. Схематическое изображение трехмерной структуры зеленого флюоресцирующего белка, выполненное с использованием специальной компьютерной программы.

Одиннадцать β-слоев образуют полый цилиндр, внутри которого расположена хромофорная группа



А *Б*

Рис. 13. Структура проприона (*А*) и его болезнетворной формы — приона (*Б*)

ограничения на предсказание вторичной структуры. Чаще такие предсказания сбываются в отношении α-спиралей, и значительно реже — в отношении β-структур. Неоднозначность укладки остова полипептидной цепи может быть связана с тем, что на ее формирование влияют межрадикальные взаимодействия, возникающие при образовании третичной структуры (см. ниже). Поэтому одни и те же фрагменты полипептидной цепи в принципе могут служить для создания разных элементов вторичной структуры. Так, последовательность Asn-$(Ala)_2$-Ile-Arg-Ser в фосфофруктокиназе имеет конформацию α-спирали, а в термолизине входит в β-складчатый слой. Значительные изменения вторичной структуры возможны и в пределах одной и той же полипептидной цепи. Они выявлены во вторичной структуре *прионов* — белков, вызывающих так называемые медленные мозговые инфекции — коровье бешенство, вертячку овец и др. Как показали исследования С. Прузенера, нормальный белок-предшественник (проприон) содержит всего 4 % β-структуры, а его аутокаталитический переход в болезнетворную форму (прион) сопровождается скачкообразным многократным увеличением доли β-структур (рис. 13).

Сверхвторичные структуры представляют собой энергетически предпочитаемые ансамбли вторичных структур. В соответствии с принципами кинетики такие образования, вероятно, предпочти-

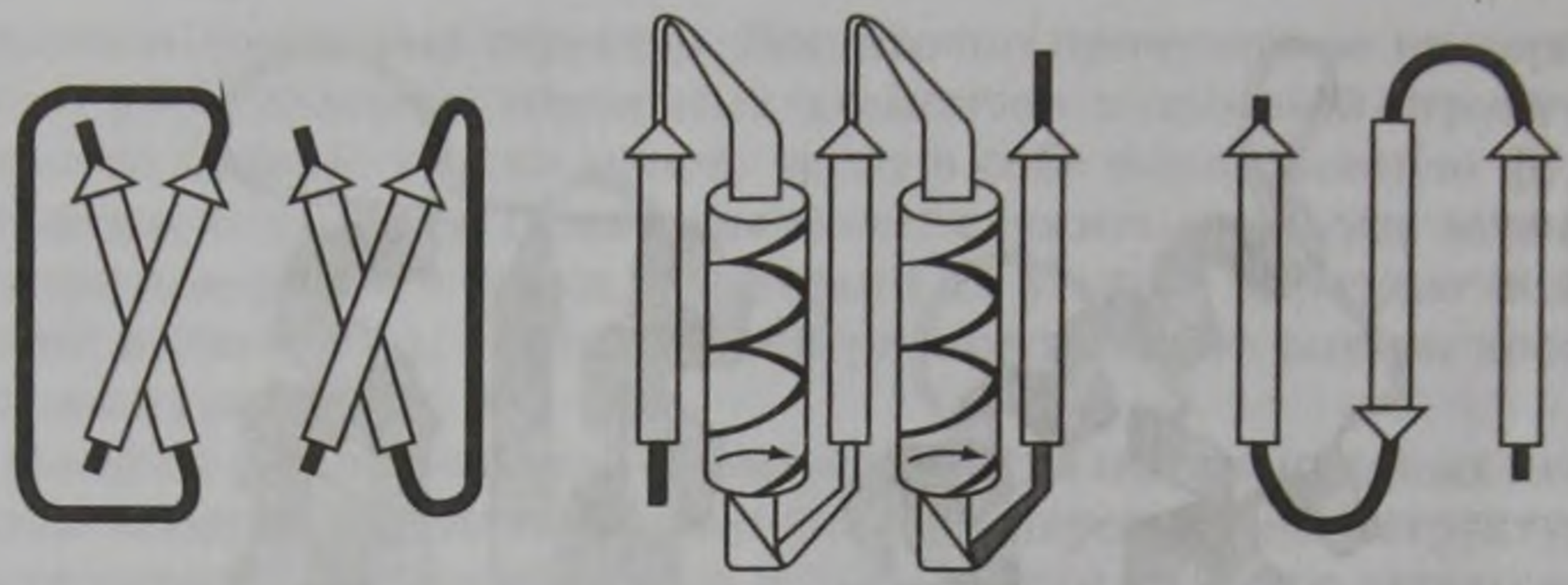


Рис. 14. Примеры сверхвторичных структур:

А — βξβ-звено; *Б* — два последовательно соединенных участка βαβ; *В* — зигзаг (антипараллельная трехцепочечная β-структура)

тельны для формирования (свертывания) полипептидной цепи в пространстве по мере синтеза, а также для придания стабильности уже сформированной пространственной структуре белка. Этот уровень организации свойствен как фибриллярным, так и глобулярным белкам. В фибриллярных белках две α-спирали могут быть скручены относительно друг друга, образуя *левую суперспираль*, период идентичности которой составляет около 14 нм. Такие спирали найдены в α-кератине, тропомиозине и других белках, что, очевидно, придает им определенные функциональные свойства: прочность, эластичность, способность к сокращению. В глобулярных белках чаще представлены сверхвторичные структуры, содержащие β-складчатые листы, между которыми могут располагаться α-спирали (рис. 14). Нередко характерные сверхвторичные структуры содержат в своем составе атомы металлов, что, по-видимому, обеспечивает их дополнительную стабильность. К числу подобных образований можно отнести так называемые «цинковые пальцы» (сверхвторичные структуры с включенными в них атомами цинка), строение которых замечательным образом способствует взаимодействию отдельных видов белков с ДНК (см. гл. 9). Сверхвторичные структуры определенного вида воспроизводятся во многих белках, образуя структурные блоки. Их изучение привело к возникновению понятия блочного принципа структурной организации белков, получившего закономерное развитие в учении о белковых доменах.

Домены. Анализ электронной плотности показывает наличие в белках четко выраженных глобулярных областей — *структурных доменов*. Молекулярная масса элементарных структурных доменов составляет около 20 кДа. Домены часто соединяются между собой короткими участками полипептидной цепи, которые выполняют роль шарниров, обеспечивающих определенную подвижность участков молекул белка относительно друг друга. В пределах одной полипептидной цепи могут встречаться несколько доменов. Один или несколько

структурных доменов могут образовывать *функциональные домены*. Последние идентифицируются по свойствам: по способности осуществлять ферментативную реакцию или присоединять какое-либо вещество (кофермент — в случае ферментов, антиген — в случае иммуноглобулинов и т. д.). «Активный центр» функциональных доменов, как правило, расположен на границе структурных доменов. В качестве примера белка с досконально изученной доменной структурой на рис. 15 приведена схема строения иммуноглобулина. Домены ярко выражены у ферментов, нуждающихся для осуществления катализа в дополнительных веществах небелковой природы (коферментах или простетических группах, ионах металлов), присоединение которых осуществляется в соответствующих доменах. Такие *коферментсвязывающие домены* выявлены в составе многих ферментов, катализирующих окислительно-восстановительные реакции. На рис. 16 изображена пространственная структура альдегидоксидоредуктазы: коферментсвязывающий домен имеется в каждой из двух идентичных полипептидных цепей (субъединиц) этого фермента.

Коферментсвязывающие домены оказываются достаточно консервативными в эволюции, и в целом структурные домены одного плана строения встречаются у многих функционально различных белков.



Рис. 15. Доменная структура иммуноглобулина G.
Молекула состоит из двух тяжелых (H) и двух легких (L) цепей, соединенных дисульфидными мостиками. Домены обозначены в виде кругов: C_H и C_L — константные домены тяжелых и легких цепей соответственно; V_H и V_L — вариабельные домены. Междоменные участки тяжелых цепей обладают определенной подвижностью: в области двух дисульфидных мостиков, соединяющих тяжелые цепи, расположен так называемый шарнирный участок, сообщающий подвижность участкам Fab («лапкам») иммуноглобулина. Вариабельные домены образуют функциональный домен, необходимый для связывания антигена. Углеводные компоненты молекулы не показаны

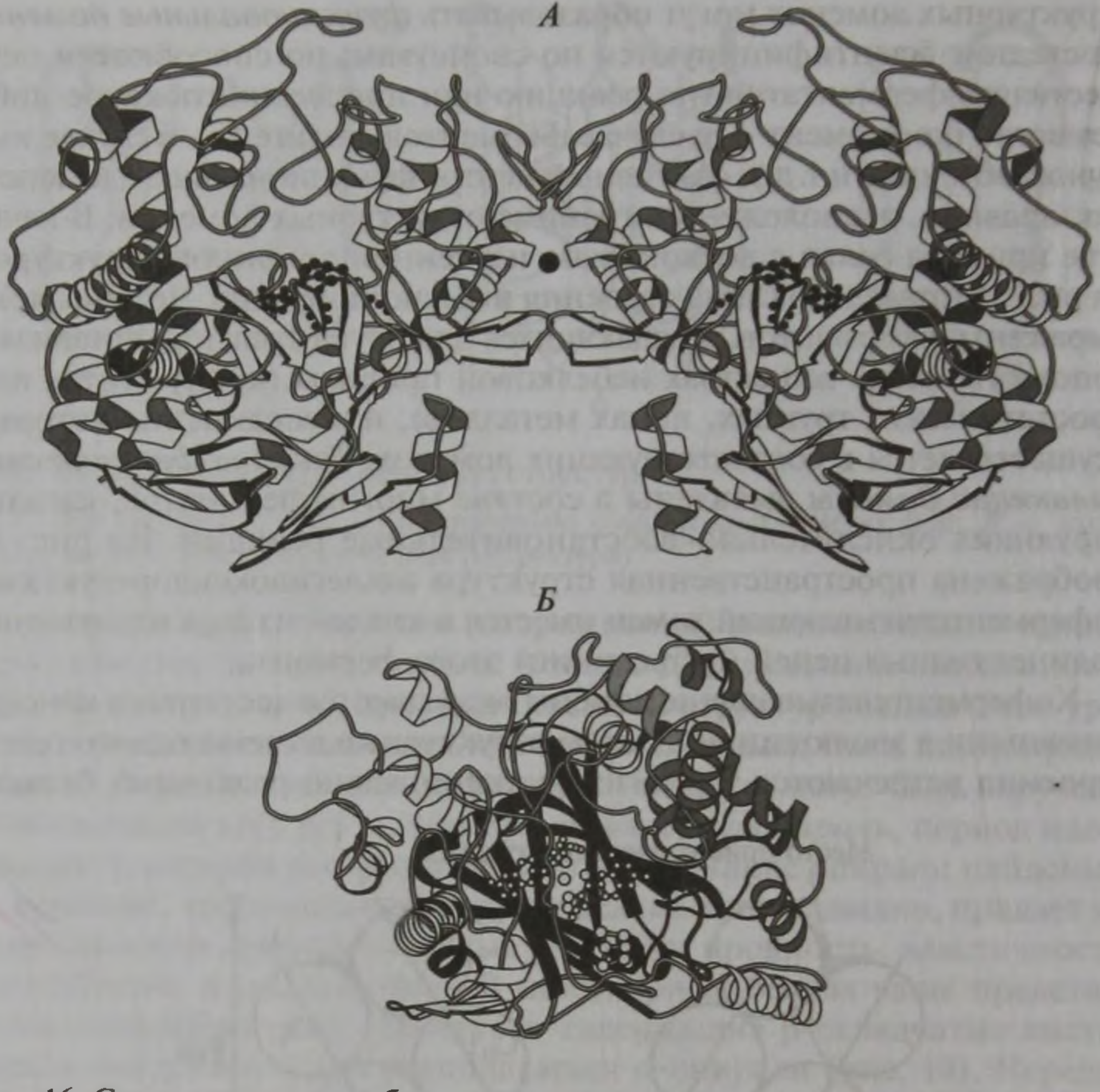


Рис. 16. Схематическое изображение молекулы альдегидоксидоредуктазы.

Молекула этого фермента представляет собой димер, состоящий из двух симметрично расположенных идентичных полипептидных цепей (гомодимер): *А* — общий план строения молекулы, в каждой из полипептидных цепей (субъединиц) которой имеется несколько доменов, образованных определенными участками α-спиралей и β-структур. В центре жирной точкой обозначена ось симметрии. Молекулы коферментов изображены в виде черных шаро-стержневых моделей, расположенных вблизи центра каждой субъединицы; *Б* — одна из субъединиц, состоящая из трех доменов: домен I выделен черным цветом, домен III — серым; в центре (в области домена I) расположены молекулы коферментов

Это относится, в частности, к структурным доменам иммуноглобулинов, NAD-связывающим доменам дегидрогеназ и других ферментов, например триозофосфатизомеразе и пируваткиназе. Присутствие сходных доменов у разных белков легло в основу представлений о модульном (доменном) принципе организации белков. Позднее стало очевидным, что матрицами для синтеза таких структурно подобных (но не идентичных) белков могут служить особым образом процессированные мРНК, возникающие в результате альтернативного сплайсинга (см. гл. 10).

Домены выражены не только у ферментов, связывающих добавочные (небелковые) группы, но и у простых (однокомпонентных) ферментов — протеинов. Разнообразием доменов объясняется способность ферментов выполнять многие каталитические функции. Примером фермента с двумя различными структурными (и одновременно функциональными) доменами может служить эндонуклеаза PI—SceI из клеток дрожжей *Sacharomyces cerevisiae*. Она относится к группе ферментов генной конверсии (homing), участвующей в сайт-специфической интеграции генов в хромосомы. Эндонуклеаза образуется из более крупного белка-предшественника в результате *аутокаталитического протеолиза* (удаления части полипептидной цепи из молекулы предшественника). Таким образом, этот фермент обладает двойной каталитической функцией: вносит разрывы в полинуклеотидные цепи ДНК, действуя как эндонуклеаза (что необходимо для последующей интеграции генов в хромосомы), и осуществляет ограниченный протеолиз своей полипептидной цепи. Двойная функ-

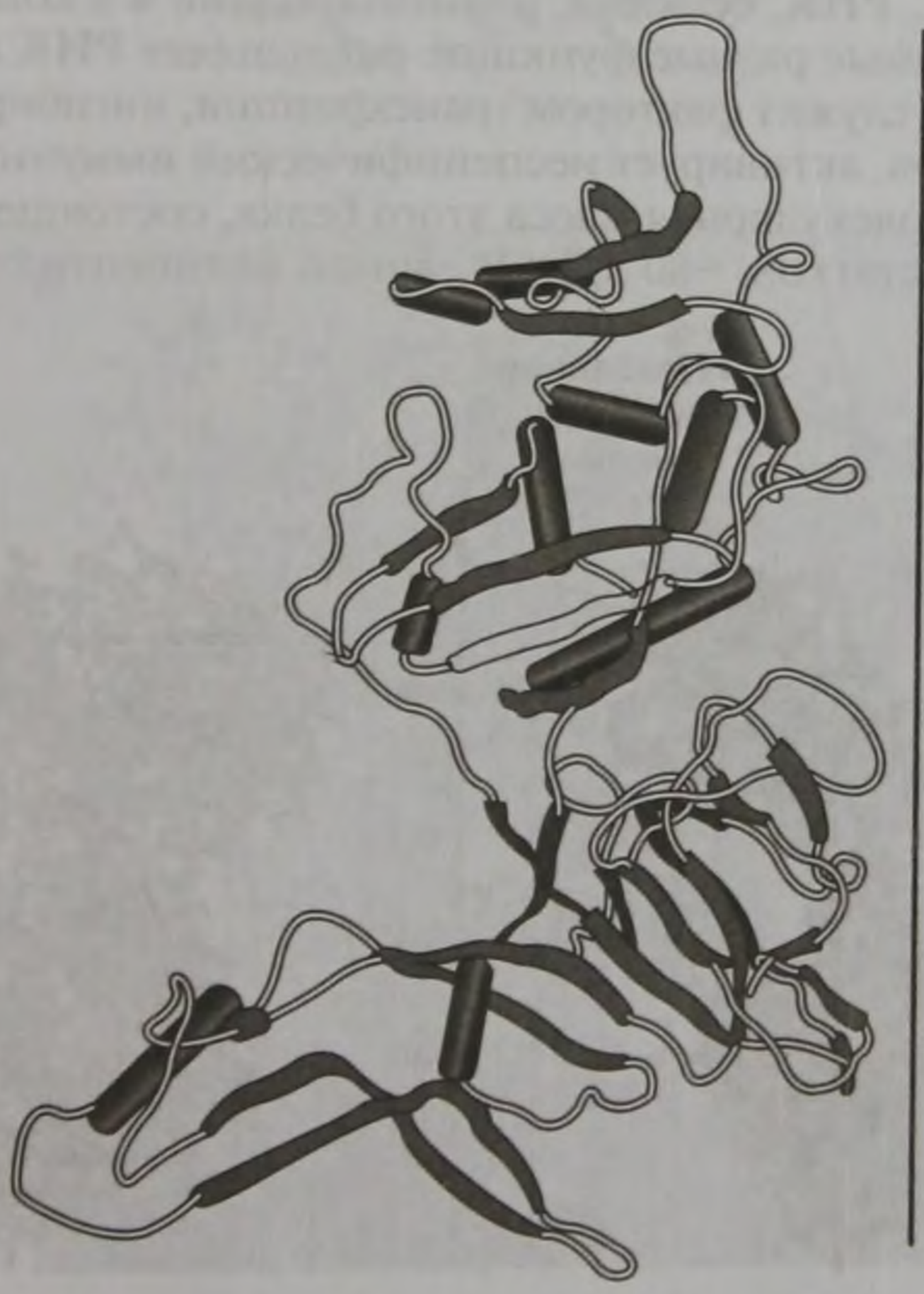


Рис. 17. Доменная структура эндонуклеазы PI-SceI (*Венде В.* и др. Молекулярная биология // 2000. — Т. 34. — Вып. 6. — С. 1054).

Домен I отвечает за протеолитическую активность, домен II — за нуклеазную активность. Оба домена участвуют в связывании ДНК

ция связана с наличием двух доменов, сильно различающихся по структуре: домен I ответствен за протеолиз белка-предшественника, а домен II обладает эндонуклеазной активностью (рис. 17).

Структурно-функциональные исследования некоторых ферментов позволили выявить еще одну особенность строения доменов — наличие в их составе *субдоменов*, каждый из которых способен выполнять определенную каталитическую функцию. Такой сложной субдоменной структурой обладает синтаза высших жирных кислот позвоночных животных: в каждой из двух ее полипептидных цепей содержится по три домена. В составе двух доменов удалось идентифицировать несколько субдоменов, каждый из них катализирует определенную реакцию в серии химических превращений, приводящих к синтезу молекул пальмитиновой и стеариновой кислот. Еще более впечатляющим примером *полифункциональности* белков с различными функциональными доменами является лактоферрин, белок молока млекопитающих из семейства трансферринов — гликопротеинов, переносящих ионы Fe^{3+}. Лактоферрин (рис. 18) способен взаимодействовать с ДНК, РНК, белками, полисахаридами и в комплексе с ними выполняет самые разные функции: расщепляет РНК как фермент рибонуклеаза, служит фактором транскрипции, ингибирует синтез простагландинов, активирует неспецифический иммунный ответ организма и т. д. Молекулярная масса этого белка, состоящего из 673 аминокислотных остатков, ~80 кДа. В единой полипептидной цепи

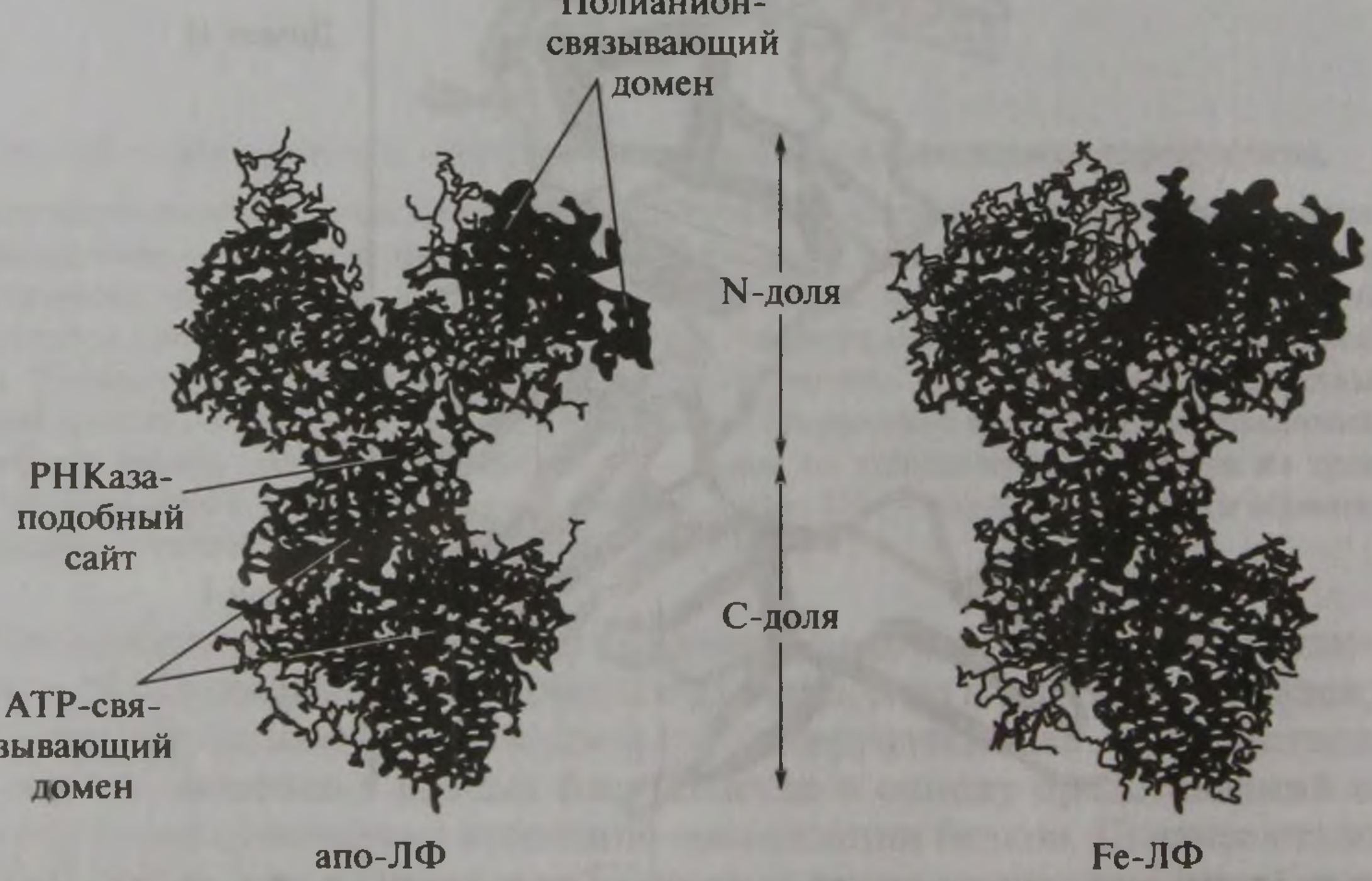
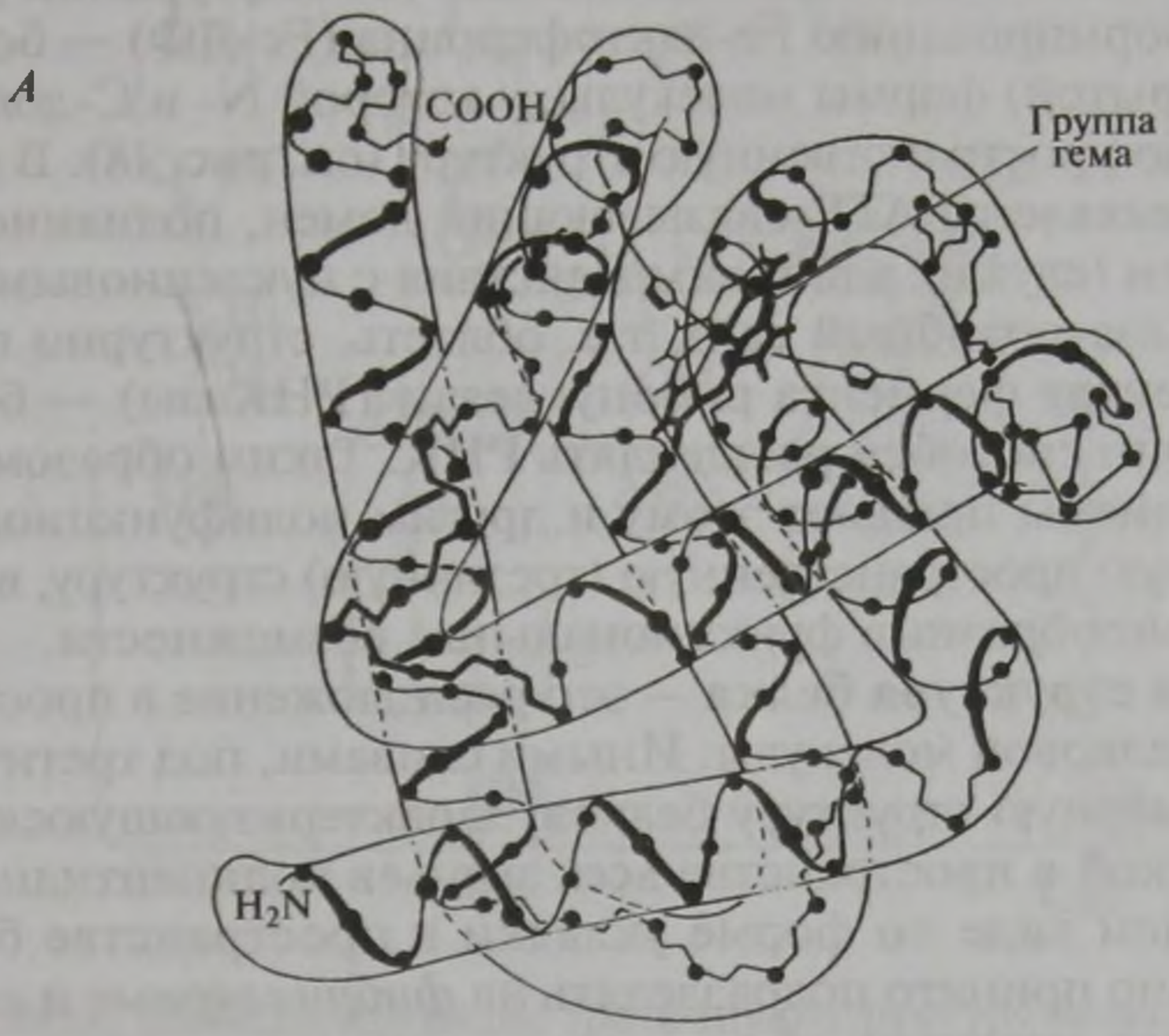


Рис. 18. Пространственная структура полифункционального белка лактоферрина, определенная методом рентгеноструктурного анализа. Пояснения см. в тексте



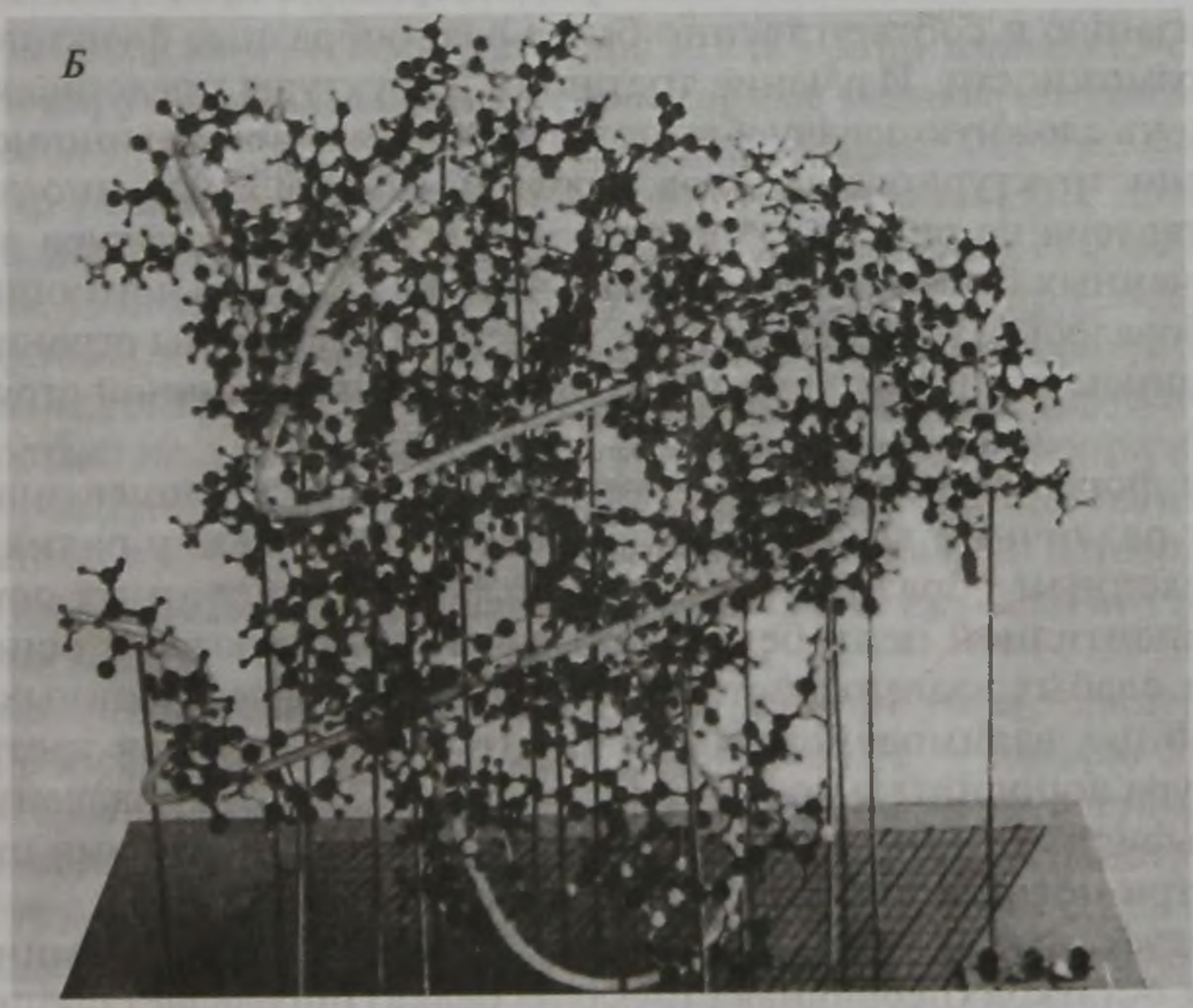


Рис. 19. Третичная структура миоглобина, установленная методом рентгеноструктурного анализа:

А — схематическое изображение молекулы; *Б* — модель, демонстрирующая расположение в пространстве каждого атома молекулы этого белка

лактоферрина методом рентгеноструктурного анализа выявлены два гомологичных домена (N- и С-доли), содержащие железосвязывающие центры. Связывание ионов железа апо-лактоферрином (апо-ЛФ) приводит к формированию Fe-лактоферрина (Fe-ЛФ) — более компактной (закрытой) формы молекулы, у которой N- и С-доли имеют более плотную пространственную структуру (см. рис. 18). В молекуле этого белка выявлены АТР-связывающий домен, полианионсвязывающий домен (служит для взаимодействия с нуклеиновыми кислотами) и РНКаза-подобный сайт, т. е. область, структурно подобная активному центру фермента рибонуклеазы (РНКазы) — благодаря ей лактоферрин способен расщеплять РНК. Таким образом, многочисленные домены придают этому и другим полифункциональным белкам сложную пространственную (третичную) структуру, в которой заложены разнообразные функциональные возможности.

Третичная структура белка — это расположение в пространстве всех атомов белковой молекулы. Иными словами, под третичной понимают трехмерную структуру белков, характеризующуюся определенной укладкой в пространстве всех звеньев полипептидной цепи. В самом общем виде по форме укладки в пространстве белковые молекулы давно принято подразделять на *фибриллярные* и *глобулярные*. Третичную структуру обычно рассматривают в отношении глобулярных белков, имеющих более разнообразную пространственную организацию и соответственно более разнообразные функциональные возможности. Изучение третичной структуры представляет собой очень сложную задачу и выполняется преимущественно методом рентгеноструктурного анализа. Трехмерная структура миоглобина представлена на рис. 19. Уникальная трехмерная структура любого из изученных белков столь сложна, что для ее детального описания потребовалось бы слишком много времени. Поэтому мы ограничимся изложением основных принципов организации третичной структуры белков.

Для формирования третичной структуры важнейшее значение имеют различные виды связей, возникающие между радикалами определенным образом расположенных аминокислотных остатков в полипептидной цепи белка. Эти связи реализуются в основном за счет слабых взаимодействий: водородных связей, ионных и гидрофобных взаимодействий (рис. 20). В ряде случаев третичная структура дополнительно поддерживается прочными ковалентными (дисульфидными) связями, возникающими между остатками цистеина внутри полипептидной цепи.

Следует отметить, что именно порядок расположения аминокислотных остатков (первичная структура белка) определяет специфику укладки в пространстве всех элементов белковой молекулы. В живой клетке на третичную структуру белка огромное влияние оказывает взаимодействие полипептидной цепи с природным растворителем (водой), а также с растворенными в нем макро- и микромолекулами

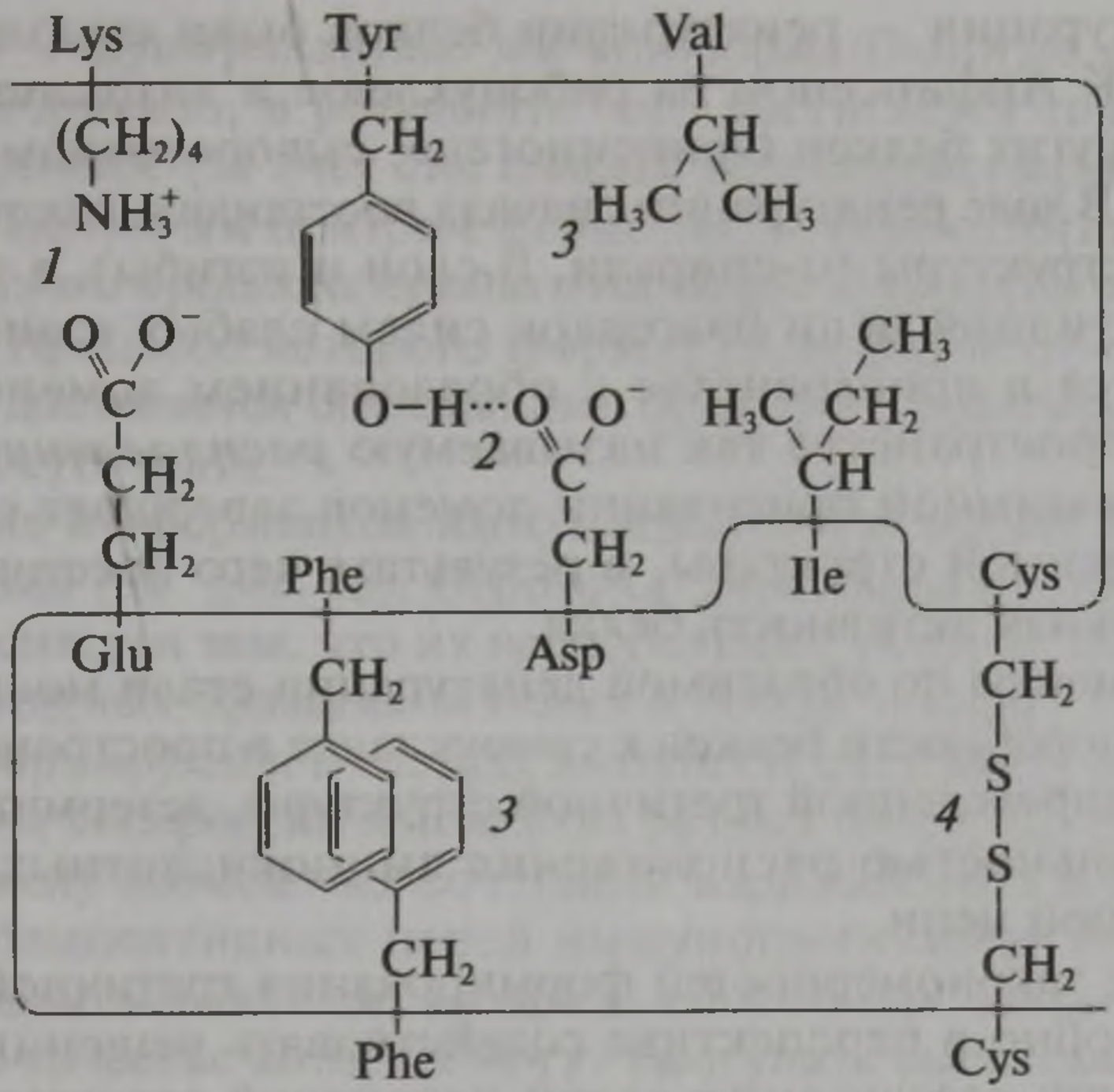


Рис. 20. Связи, стабилизирующие третичную структуру белков:
1 — ионные; *2* — водородные; *3* — гидрофобные; *4* — ковалентные

(другими белками, солями металлов и т. д.). При взаимодействии с водным окружением полярные и неполярные радикалы аминокислот ведут себя по-разному. Полярные радикалы (Ser, Tre, Tyr, Asp, Lys, Glu и др.) гидрофильны, способны взаимодействовать с диполями воды и поэтому обычно расположены на поверхности белковой глобулы. Неполярные радикалы (Ala, Leu, Ile, Phe и др.), напротив, гидрофобны и оказываются внутри молекул белка, образуя гидрофобное ядро. Границы между гидрофобным ядром и наружной поверхностью не всегда выражены, и часть гидрофобных радикалов может выступать на поверхность молекулы, что существенно для контакта между белковыми глобулами при образовании четвертичной структуры белков, а также для взаимодействия с другими внутриклеточными молекулами.

Изменение внешнего окружения (pH, ионной силы, температуры и т. д.) способно вызывать *денатурацию* белков — потерю ими нативной третичной структуры, связанной с нарушением регулярности внутримолекулярных связей, поддерживающих пространственную структуру. При этом наиболее уязвимыми оказываются силы слабых взаимодействий — водородные и ионные связи. Несмотря на то что денатурация не сопровождается разрушением ковалентных (пептидных и дисульфидных) связей, белок утрачивает свои функциональные свойства, что подчеркивает важность его пространственной организации. В определенных условиях белки могут *ренатурировать* — вновь приобрести нормальную глобулярную структуру. Первые эксперимен-

ты по денатурации — ренатурации белков были выполнены в 60-е годы XX в. К. Анфинсеном на рибонуклеазе и лизоциме, а позднее и на ряде других белков (трипсиногене, сывороточном альбумине, α-амилазе). В ходе ренатурации сначала восстанавливаются элементы вторичной структуры (α-спирали, β-слои и изгибы), а затем участки полипептидной цепи благодаря силам слабых взаимодействий организуются в пространстве с образованием доменов, и белок образует в пространстве так называемую *расплавленную глобулу*. Уточнение взаимной ориентации доменов завершает становление пространственной структуры, в результате чего восстанавливается функциональная активность белка.

Эксперименты по обратимой денатурации стали мощным свидетельством способности белков к самоукладке в пространстве — приобретению определенной третичной структуры, детерминированной последовательностью расположения аминокислотных остатков в полипептидной цепи.

Изучение закономерностей формирования третичной структуры белков способно в перспективе содействовать решению альтернативной задачи: составлению такой первичной структуры, которая могла быть ориентирована на создание определенной третичной структуры, т. е. *создание de novo белков, отсутствующих в природе*. Пока в такого рода работах удается создать небольшие полипептиды, состоящие из нескольких десятков аминокислот и напоминающие природные белки (миогемэретрин и цитохром *c*), состоящие из α-спиралей, а также отдельные полипептиды, содержащие не только α-спирали, но и β-структуры (альбебетин). Синтез этих белков на матрицах синтетических мРНК, а также синтез соответствующих генов (кДНК), их клонирование в клетках бактерий (методы генетической инженерии изложены в гл. 14) с последующим выделением и анализом структуры запрограммированных полипептидов дают вполне ожидаемые результаты, что можно расценить как первые шаги истинной белковой инженерии.

Одним из наиболее перспективных направлений белковой инженерии является создание ***каталитически активных антител*** — своеобразной группы белков, получивших название «абзимы».

Известно, что общим свойством ферментов и антител (иммуноглобулинов) является их способность специфически связывать определенные молекулы (лиганды) — субстраты и антигены соответственно. Еще в 1948 г. Л. Полинг, сравнивая строение этих двух фундаментальных классов белков, пришел к выводу, что фермент не так строго комплементарен субстрату, как антитело — антигену. Известно, что активный центр ферментов формируется благодаря уникальному сочетанию определенных аминокислотных остатков (радикалов), нередко достаточно удаленных друг от друга в полипептидной цепи белка, но сближающихся в пространстве при формировании третичной структуры белка. При взаимодействии с субстратом

происходят индуцированные им конформационные изменения в структуре фермента, в результате чего достигается точное и относительно прочное (за счет сил слабых взаимодействий) связывание субстрата с активным центром фермента. В энзимологии эту стадию каталитического процесса принято называть переходным состоянием реакции, в процессе которого формируется фермент-субстратный комплекс и достигается оптимальная ориентация реагирующих групп фермента и субстрата.

В отличие от ферментов иммуноглобулины имеют стандартный план строения (см. рис. 15). Огромное разнообразие антител в организме объясняется тем, что их вариабельные участки, сосредоточенные в N-концевых областях тяжелых и легких полипептидных цепей (V_H и V_L), формируются в клетках иммунной системы (β-лимфоцитах) как продукты экспрессии множества разных генов. Благодаря колоссальному числу возможных сочетаний вариабельных и константных участков полипептидных цепей иммуноглобулинов возникает соответствующее множество антител, способных связывать разные антигены, в качестве которых могут выступать высокомолекулярные соединения, а также небольшие по размерам молекулы (гаптены). Специфическое взаимодействие антитела со своим лигандом, хотя и подобно взаимодействию фермента с субстратом, однако стерически более однозначно и предполагает более прочное связывание лиганда по сравнению со связыванием субстрата в активном центре фермента. Развивая идеи Л. Полинга, в 1969 г. У. Дженкс предположил, что если бы удалось выработать антитела, комплементарные переходному состоянию ферментативной реакции, то они были бы способны катализировать соответствующие химические реакции. Реальные условия для создания каталитически активных антител возникли после того, как были синтезированы отдельные аналоги переходного состояния и разработаны методы получения моноклональных антител (см. гл. 1). Первые антитела-ферменты были получены в 1986 г. Р. Лейнором и П. Шульцем. Ими стали абзимы с эстеразной активностью, полученные иммунизацией животных стабильным аналогом переходного состояния с последующим отбором клонов β-лимфоцитов, способных продуцировать соответствующие антитела. В случае ферментативного катализа, осуществляемого гидролитическими ферментами (эстеразами) при гидролизе сложных эфиров, в которых атом углерода связан с тремя другими атомами, после присоединения молекулы воды образуется нестабильное переходное состояние. При этом атом углерода находится в тетраэдрическом состоянии (рис. 21). В качестве аналога переходного состояния для получения абзима был использован сложный эфир фосфорной кислоты, в котором атом фосфора имитировал тетраэдрический атом углерода. Использование фосфорных эфиров в качестве гаптенов привело к успешному получению антител, обладающих эстеразной активностью. Так, например, были получены антитела к *n*-нитрофенилфосфонилвалериату:

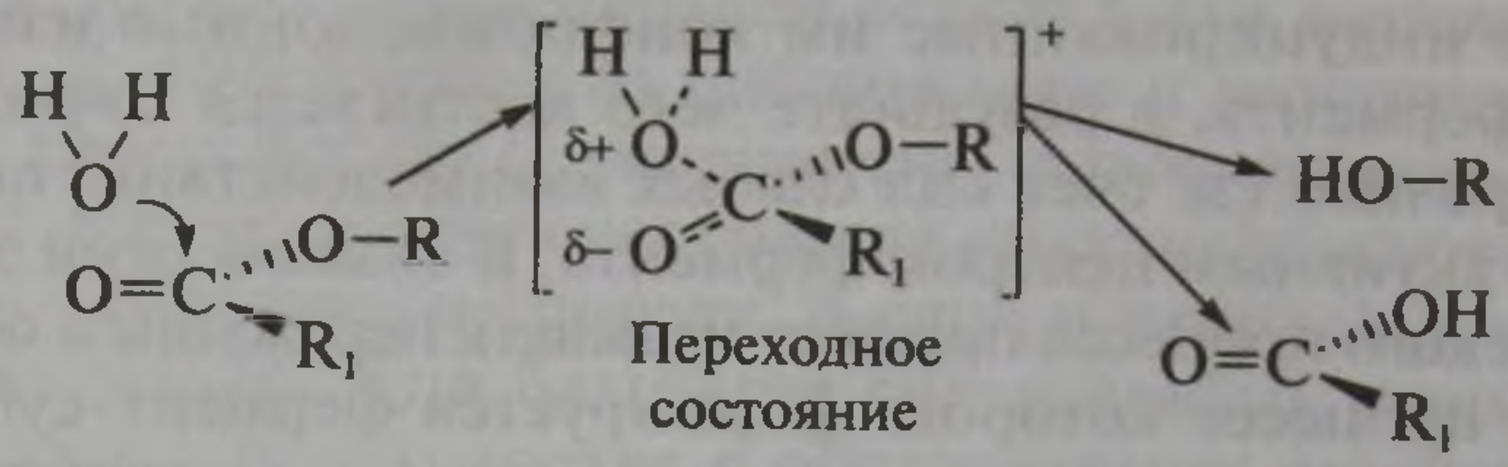
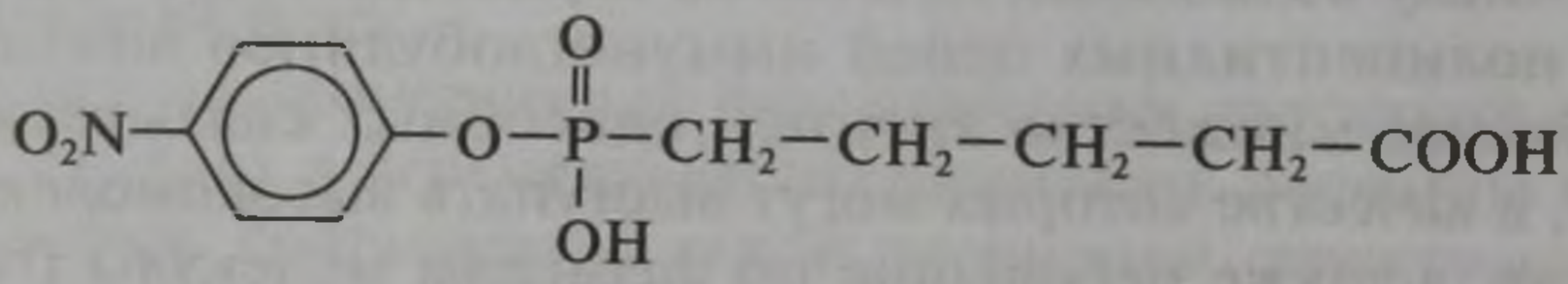
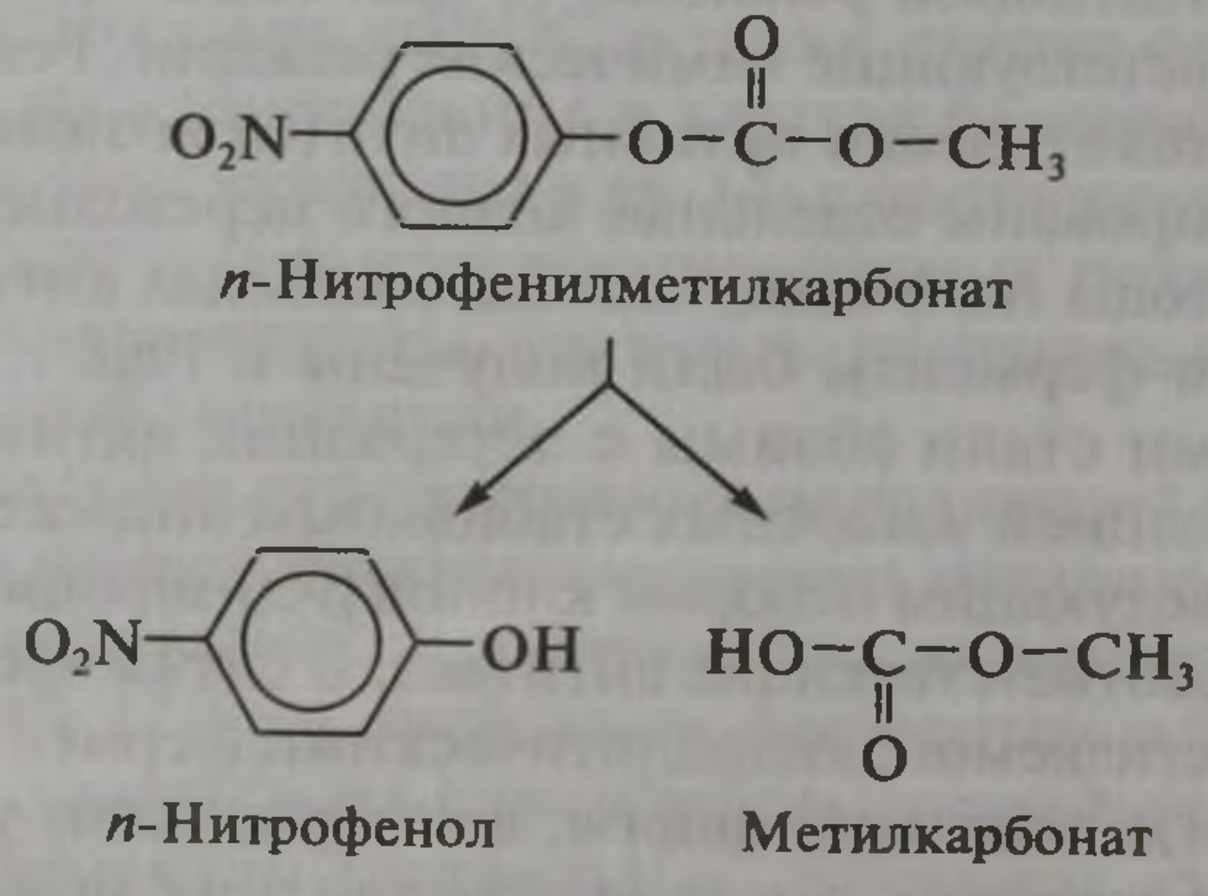


Рис. 21. Реакция гидролиза сложного эфира карбоновой кислоты и структура фосфаната — стабильного аналога переходного состояния



n-Нитрофенилфосфонилвалериат

Полученные антитела оказались способны гидролизовать с высокой скоростью субстрат эстераз — *n*-нитрофенилметилкарбонат, который стерически похож на антиген (гаптен) и при соединении с антителом сразу образует переходное состояние, распадающееся до продуктов реакции:



n-Нитрофенилметилкарбонат

n-Нитрофенол

Метилкарбонат

Дальнейшие работы в области получения каталитически активных антител показали, что абзимы способны ускорять химические реакции в $10^6 - 10^7$ раз, тогда как природные энзимы — в $10^9 - 10^{12}$ раз. Получены абзимы, активные не только в реакциях распада, но и в реакциях синтеза (например, в реакциях образования пептидной связи).

В природе каталитически активные антитела обнаружены у людей, страдающих аутоиммунными заболеваниями. ДНК-гидролизующие

антитела выявлены и в крови больных лейкозом, лучевой болезнью и СПИДом. Выявлены они и у людей, больных системной красной волчанкой, причем в этом случае фракция секреторных IgA-антител обладает протеинкиназной активностью. В крови больных ревматическим артритом, а также людей, страдающих раковыми заболеваниями, обнаружены антитела с амилолитической активностью.

Вопрос о существовании каталитически активных антител в крови здоровых людей остается открытым. Не выявлены они и у больных гриппом, пневмонией и некоторыми видами рака. Значительный интерес в связи с этим представляет обнаружение каталитических антител в молоке человека. Эти антитела обладают протеинкиназной активностью, фосфорилируя казеин и другие белки молока, часть из них способна гидролизовать ДНК и РНК, другие обладают амилолитической активностью. Полагают, что абзимы молока могут играть определенную роль в защите новорожденных от вирусных и бактериальных инфекций.

Практическое значение абзимов может быть весьма многообразным. Оно связано прежде всего с тем, что колоссальное природное многообразие антител ($1 \cdot 10^8$) создает предпосылки для создания биокатализаторов с заданными свойствами, отсутствующих в природе. Большое значение может иметь применение абзимов в биотехнологии и медицине. В перспективе абзимы могут быть использованы для разделения оптических изомеров как природных соединений, так и фармакологических препаратов, что в настоящее время практически недостижимо другими способами.

Формирование пространственной структуры белка в живой клетке представляет собой непростую систему взаимодействий вновь образующейся в процессе трансляции белковой молекулы с предсуществующими в клетке белками. Эти белки могут вступать во взаимодействие с синтезируемой полипептидной цепью и затруднять формирование ее нативной запрограммированной в аминокислотной последовательности структуры. Несмотря на ряд неоспоримых свидетельств определяющего значения первичной структуры для формирования той или иной трехмерной структуры белка (см. выше), в клетках существует группа особых белков-помощников, непосредственно предназначенных для облегчения формирования пространственной структуры вновь синтезируемых полипептидов. Процесс формирования пространственной структуры белка — ***фолдинг*** — обозначает образование определенных складок, изгибов в полипептидной цепи, т.е. процесс становления третичной структуры белка, которую иногда представляют как группу *фолдов* — структур с определенным набором и топологией расположения элементов вторичной структуры. В 1987 г. было предложено называть белки, обслуживающие фолдинг, ***молекулярными шаперонами*** (molecular chaperones). Термин «шапероны» применяется в настоящее время в отношении трех различных групп белков: нуклеоплазминов (они являются ядерными белками,

участвующими в сборке нуклеосом), белков, называемых Hsp 70-Bip (группа белков теплового шока — Heat shock proteins — и их гомологов) и собственно белков-шаперонов, получивших название шапе-ронины и имеющих непосредственное отношение к сворачиванию полипептидной цепи в нативный белок.

Шапероны представляют собой белковые комплексы, встречающиеся в бактериях, цитоплазме эукариотических клеток, матриксе митохондрий и хлоропластов (табл. 2). Они способствуют корректному сворачиванию синтезируемых полипептидных цепей, защищают свернутые в расплавленные глобулы цепи от агрегации. Наиболее исследованы шапероны *E. coli*, называемые GroEL и GroES. GroEL состоит из 14 идентичных субъединиц (протомеров) с молекулярной массой 57 кДа каждый. Он представляет собой цилиндр, сформированный из двух колец, каждое из которых в свою очередь состоит из семи протомеров. Отдельный протомер шаперонина состоит из 3 доменов: эвкаториального, апикального и промежуточного (рис. 22). Апикальные (вершинные) домены каждого кольца образуют входные отверстия в центральную полость цилиндра GroEL. Шаперонин GroES (его также называют *кошаперонином*) состоит из 7 идентичных протомеров с молекулярной массой 10 кДа. Протомеры GroES образуют куполообразную структуру. Образованные частью аминокислотных остатков (Gly 45 — Pro 56) короткие подвижные петли протомеров GroES формируют крышу купола, в центре которой имеется отверстие диаметром около 0,8 нм. GroEL и GroES обладают способностью связывать ATP (14 и 7 молекул соответственно). Места связывания ATP расположены на внутренней поверхности экваториальных доменов GroEL. Подлежащая сворачиванию полипептидная цепь вновь синтезируемого белка связывается с определенным участком внутренней боковой поверхности апикальных доменов. При гидролизе ATP происходит взаимодействие GroEL и GroES — образуется комплекс GroEL/GroES (рис. 23). В этом комплексе создается закрытое пространство (ширина ~7 нм, высота ~6,5 нм), внутри которого полипептидная цепь сворачивается в нативный белок. Процесс сворачивания происходит в течение нескольких циклов диссоциации-реассоциации комплекса шаперонинов, в результате чего, вероятно, происходит постепенное, защищенное от нежелательных контактов формирование нативной третичной структуры белка (рис. 24). Сначала синтезированный полипептид связывается с комплексом шаперонинов в кольце, не занятом GroES (стадия 1). Затем происходят связывание и гидролиз ATP, в результате чего GroES отходит от комплекса (переход от стадии 2 к стадии 3), а затем вновь в комплексе с очередной порцией ATP присоединяется к GroEL (стадия 4). На этой стадии GroES «закрывает» полипептид в полости GroEL, где он сворачивается во время гидролиза ATP (стадии 5 и 6). Отдельный цикл фолдинга занимает 15 — 20 с и завершается гидролизом еще 7 молекул ATP, что приводит

Таблица 2. **Функции шаперонов**

Источник	Белки, подвергающиеся фолдингу	Шаперон	Функция
E. coli	Предшественники секреторных белков; ДНК-репликационный комплекс	CroEL SecB Dna K/Dna J	Задержка свертывания перед транспортом; реорганизация белкового комплекса
Фотосинтезирующие бактерии	Рибулозо-1,5-бисфосфаткарбоксилаза	CroEL/ CroES	Сборка олигомера
Хлоропласты	Рибулозо-1,5-бисфосфаткарбоксилаза	RuSBR	Сборка олигомера
Митохондрии	Митохондриальные белки-предшественники Белки-предшественники в матриксе	Hsp70 Hsp60	Завершение транслокации; стабилизация не полностью свернутых структур в матриксе Стабилизация не полностью свернутых структур и свертывание; реэкспорт предшественников в межмембранное пространство
Эндоплазматический ретикулум	Секреторные белки Субъединицы Т-клеточного рецептора; Запасные белки растений	Bip TRAP, или p28 b70 (Bip)	Завершение транслокации; стабилизация не полностью свернутых структур в канальцах Сборка рецептора Стабилизация вновь синтезируемых белков
Цитозоль	Синтезируемые полипептиды	Hsc70	Стабилизация не полностью свернутых структур

Окончание табл. 2

Источник	Белки, подвергающиеся фолдингу	Шаперон	Функция
Цитозоль	Митохондриальные и секреторные белки-предшественники	Hsc70	Разворачивание перед транслокацией
	Белки-рецепторы стероидных гормонов	Hsp90	Стабилизация неактивной формы рецептора
	Белки, трансформирующие ретровирусы	Hsp90	Стабилизация неактивной формы белка при прохождении через плазматическую мембрану
Ядро	Гистоны	Нуклеоплазмин	Сборка нуклеосом

к освобождению GroES из комплекса с GroEL (переход от стадии 6 к стадии 7) и высвобождению свернутого полипептида через отверстие полости цилиндра GroEL.

Структура комплекса GroEL/GroES детально изучена методом электронной микроскопии. Установлена его ведущая роль в формировании нативной структуры большой полипептидной цепи одного из ключевых ферментов биосинтеза углеводов — рибулозо-

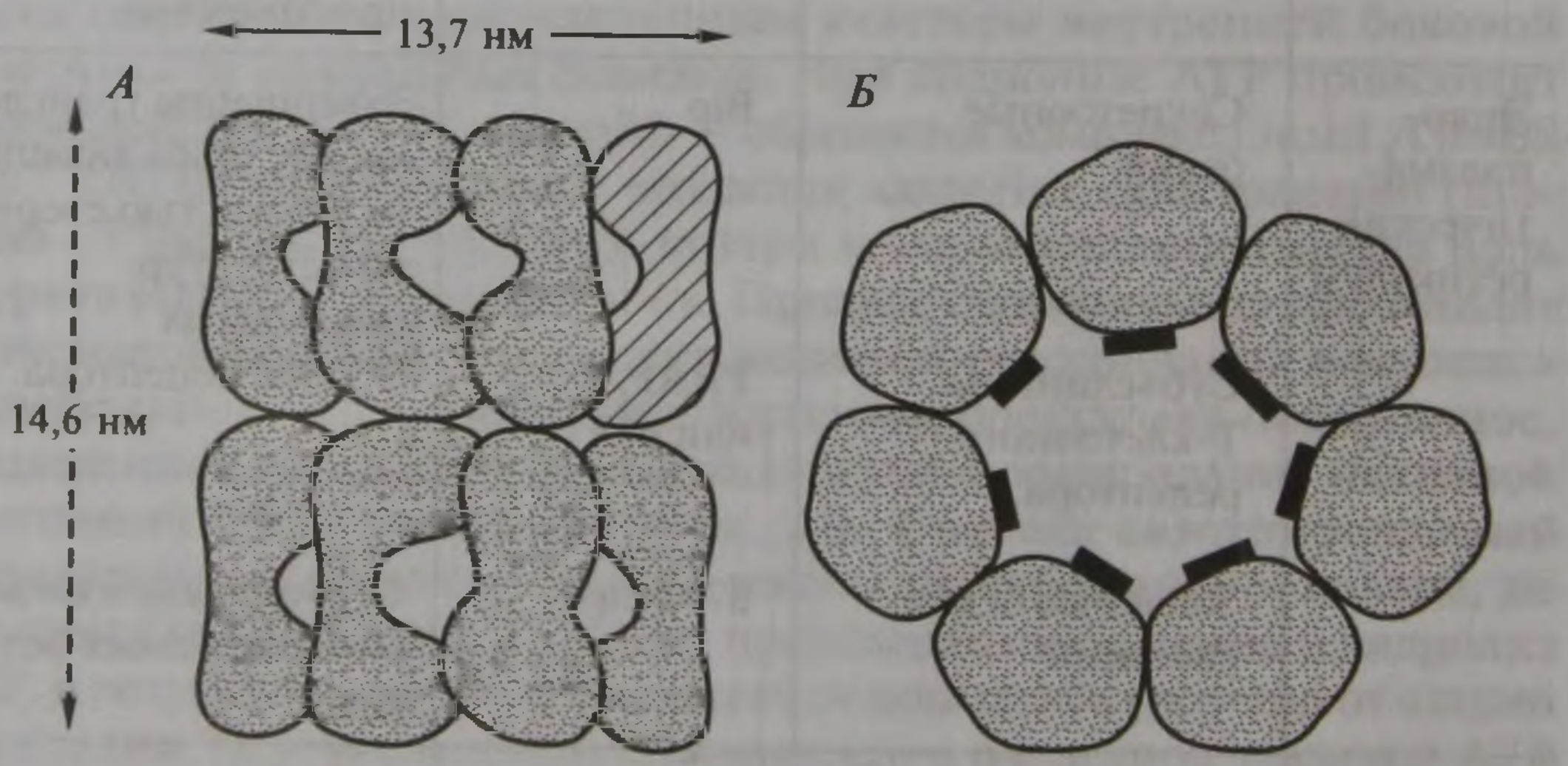


Рис. 22. Структура шаперонинового комплекса GroEL:

А — вид сбоку цилиндра GroEL, состоящего из двух колец; *Б* — вид сверху. Темные прямоугольники — участки связывания полипептида

1,5-бисфосфаткарбоксилазы (Ru Bis Co). Сходный по структуре шаперон выявлен в хлоропластах высших растений — RBP (Ru Bis Co building protein). Как и GroEL, он обладает редким типом 7-лучевой симметрии.

Четвертичная структура характерна для белков, построенных из нескольких полипептидных цепей. Она представляет собой способ взаимного расположения в пространстве отдельных полипептидных цепей в молекуле белка, необходимый для проявления специфических функций. Отдельные полипептидные цепи, входящие в состав белков, обладающих четвертичной структурой, носят название *протомеров*, или *субъединиц*, из которых формируется *эпимолекула* такого *белка-мультимера*.

Четвертичная структура поддерживается исключительно силами слабых взаимодействий, поэтому она представляет собой достаточно лабильное образование, значительно менее прочное, чем третичная и тем более вторичная структура. Межсубъединичные контакты во многом определяются водородными связями, с помощью которых могут, например, контактировать между собой β-складчатые листы протомеров. Существенную роль выполняют и гидрофобные взаимо-

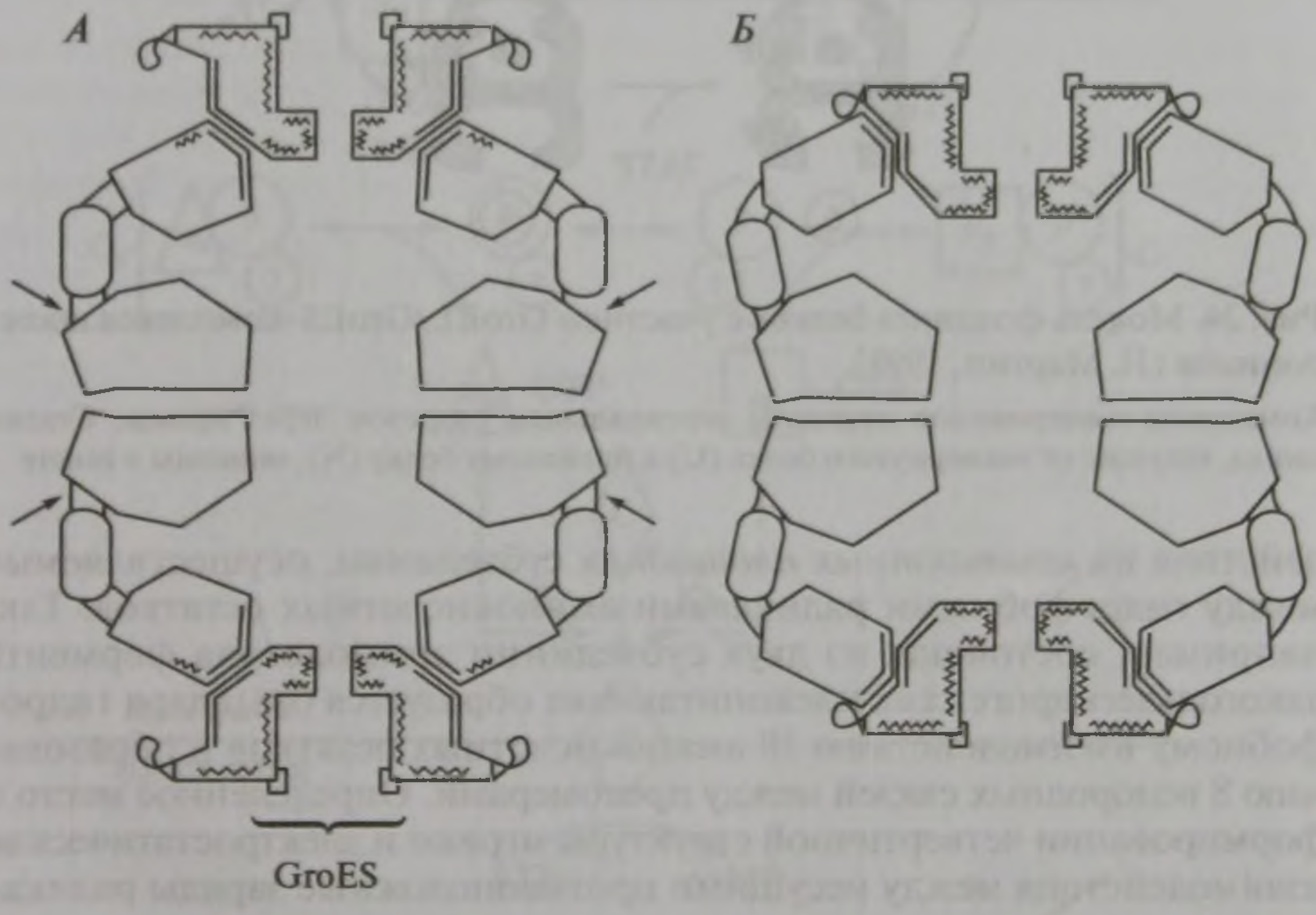


Рис. 23. Предполагаемая модель строения GroEL/GroES-комплекса (*Баширов М.А.* // Биохимия. — М., 1997. — Т. 62. — Вып. 4 — С. 489):

А — для образования комплекса GroES присоединяется к GroEL; *Б* — устойчивый симметричный GroEL/GroES-комплекс. Стрелками обозначены предполагаемые участки GroEL, конформационные перестройки которых приводят к «раскрытию» апикальных доменов

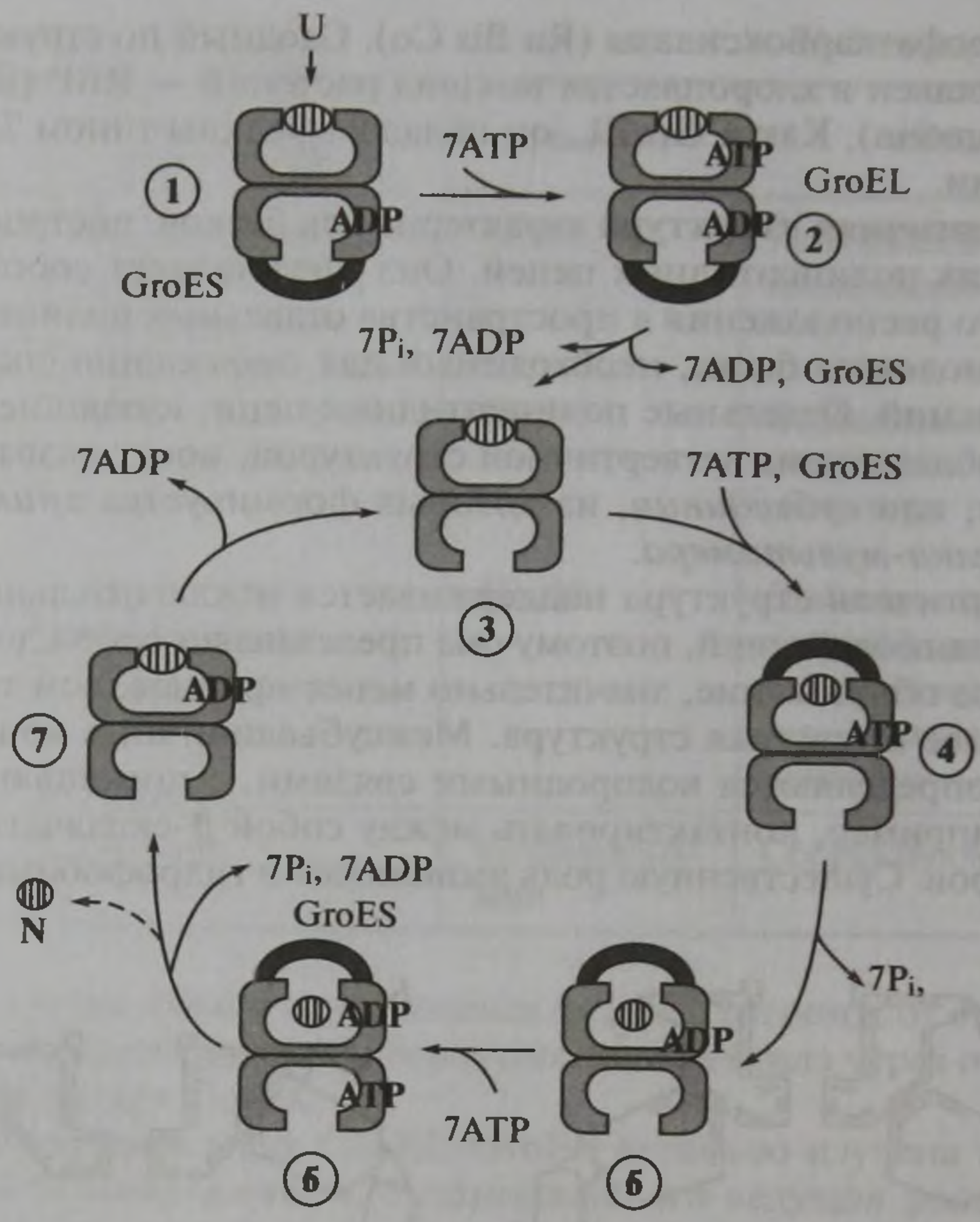


Рис. 24. Модель фолдинга белков с участием GroEL/GroES-комплекса шаперонинов (Й. Мартин, 1998).

Комплексы шаперонинов показаны вертикальным разрезом через кольца. Стадии цикла, ведущие от несвернутого белка (U) к нативному белку (N), описаны в тексте

действия на *контактных площадках* субъединиц, осуществляемые между гидрофобными радикалами аминокислотных остатков. Так, например, состоящая из двух субъединиц эпимолекула фермента алкогольдегидрогеназы млекопитающих образуется благодаря гидрофобному взаимодействию 18 аминокислотных остатков и образованию 8 водородных связей между протомерами. Определенное место в формировании четвертичной структуры играют и электростатические взаимодействия между несущими противоположные заряды радикалами аминокислотных остатков.

Белки, обладающие четвертичной структурой, как правило, содержат четное число субъединиц и представляют собой димеры, тетрамеры, гексамеры, октамеры и т. д. Число димеров и тетрамеров резко преобладает среди изученных белков. Четное число субъединиц, очевидно, отражает общий *принцип симметрии*, свойственный

живой природе. Белки с нечетным числом субъединиц представляют собой относительно редкое явление, и часть из них выполняет очень важные и своеобразные функции у многих организмов. Так, *тримерами* являются некоторые из GTP-связывающих белков (G-белки). Их основная функция состоит в передаче сигналов внутрь клетки или в различных ее компартментах, хотя часть из них выполняют и иные функции (участвуют в биосинтезе белка и других энергоемких процессах). G-белки обладают способностью гидролизовать GTP до GDP и фосфата, переходя из активного состояния (комплекс G-белок — GTP) в неактивное состояние (комплекс G-белок — GDP). GTPразная активность G-белков приводит к высвобождению энергии *макроэргических связей* GTP, которая и используется для передачи сигналов.

К G-белкам относится, например, трансдуцин, который участвует в передаче и усилении светового сигнала, выступая в качестве компонента комплекса белков (трансдуцин-родопсин-фосфодиэстераза циклических нуклеотидов), обеспечивающих восприятие и много-

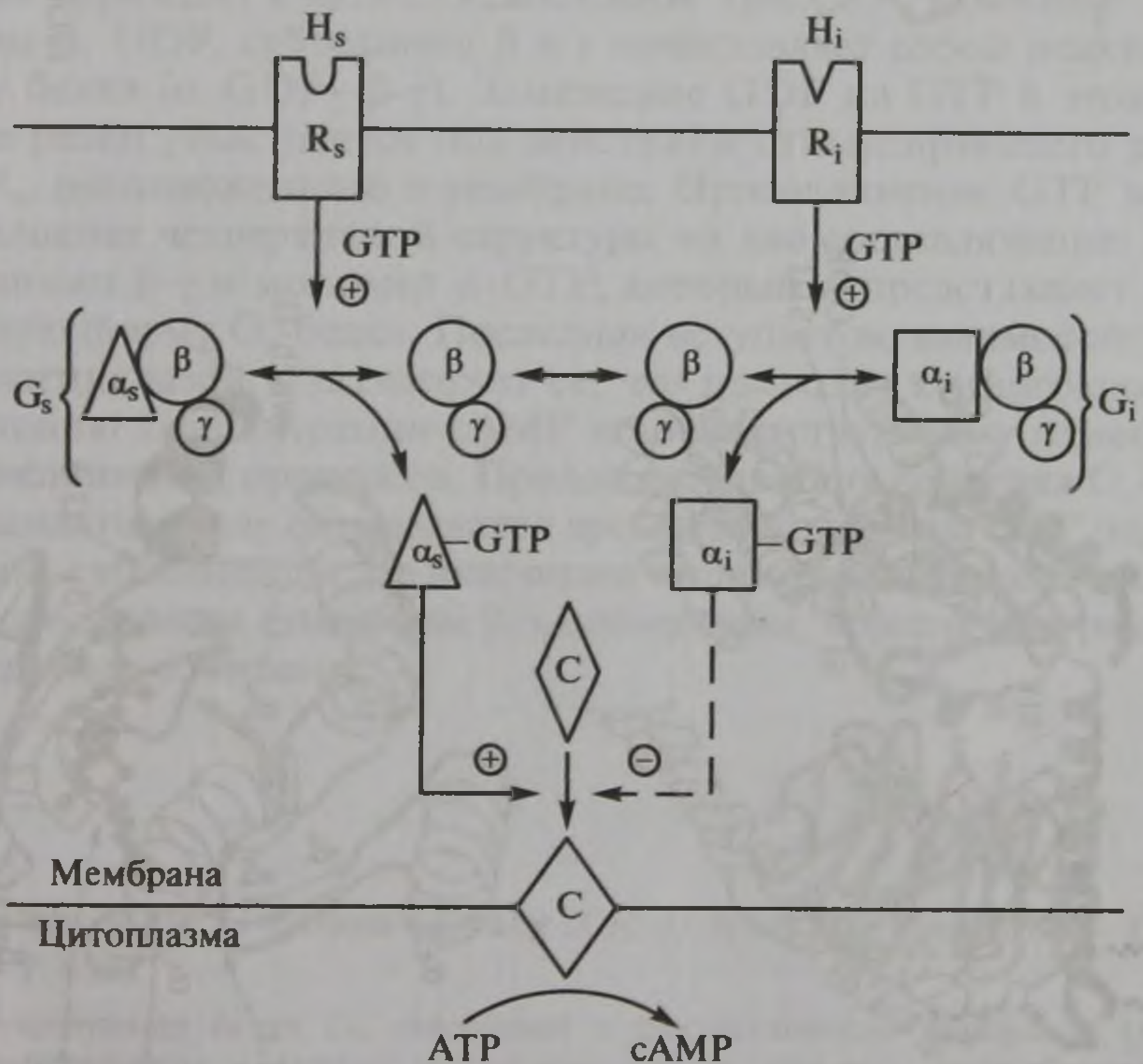
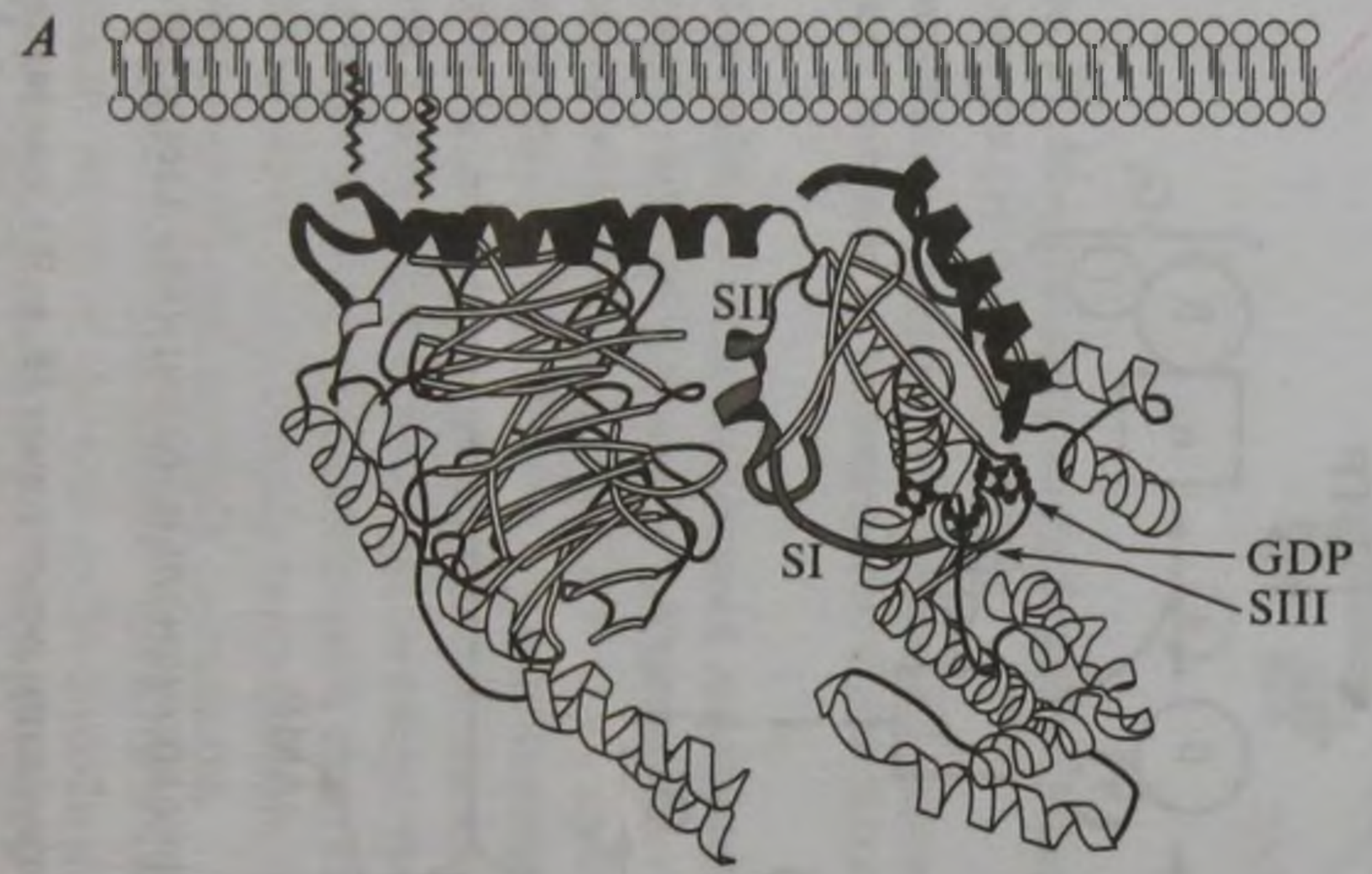
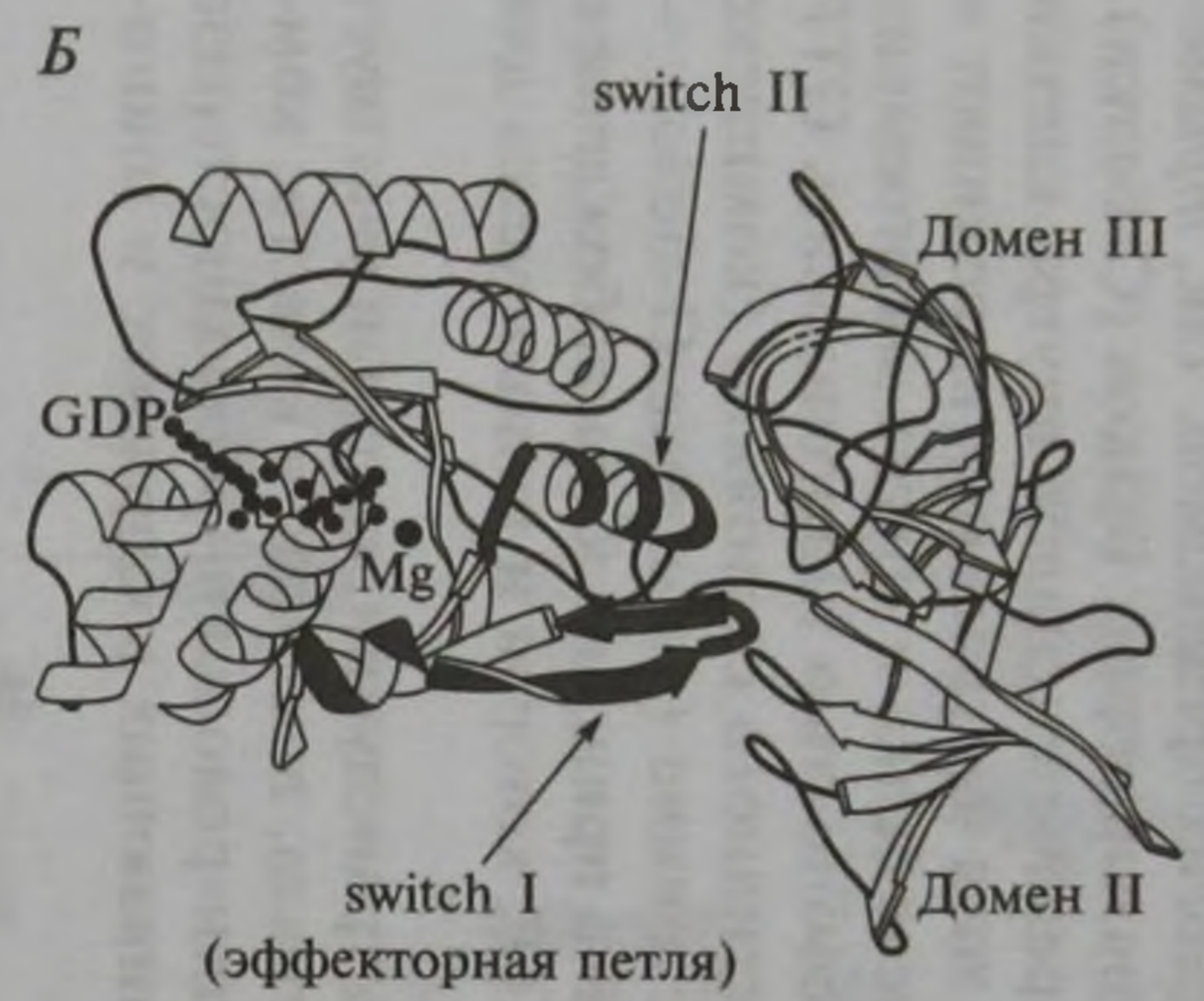
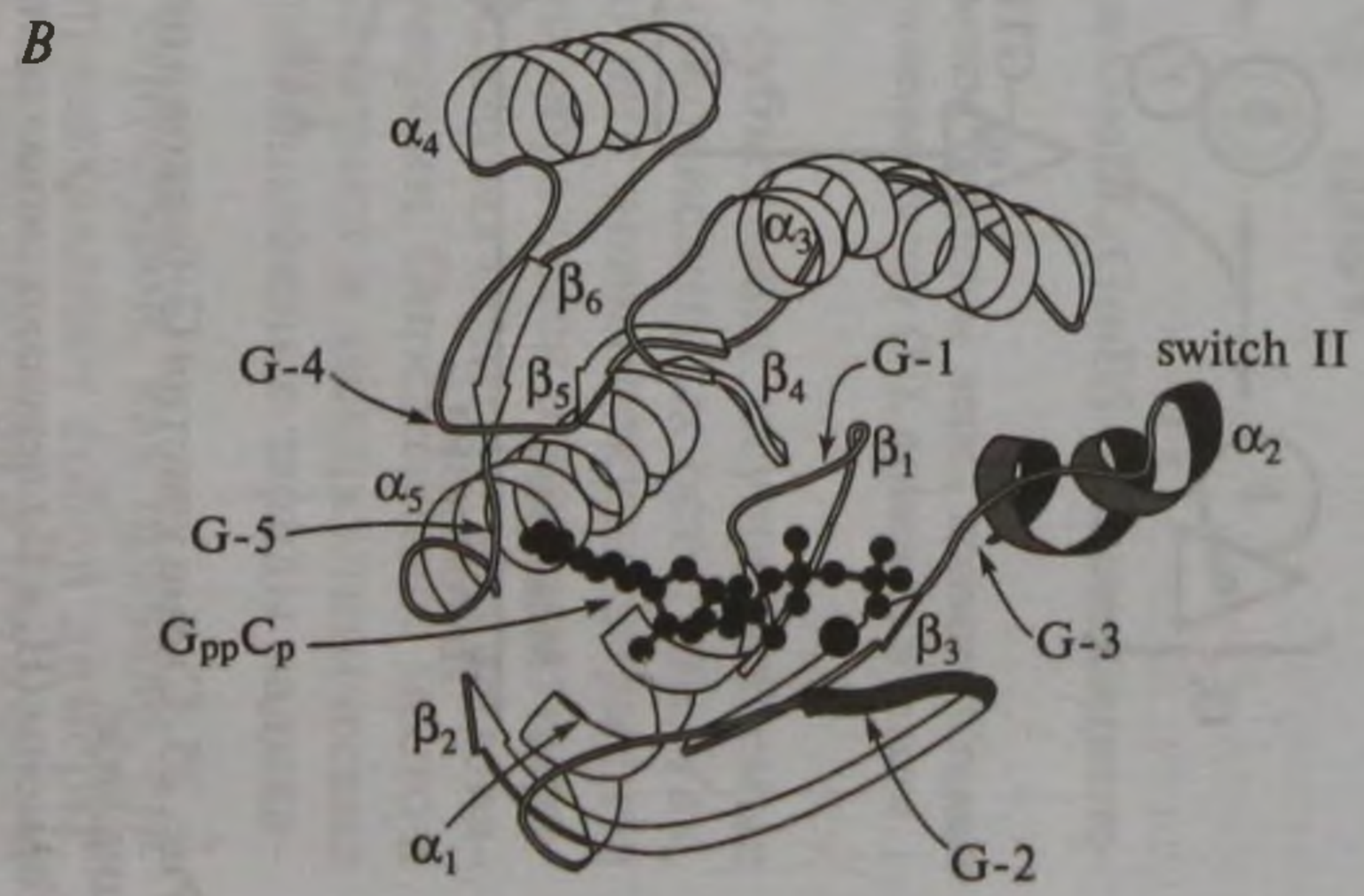


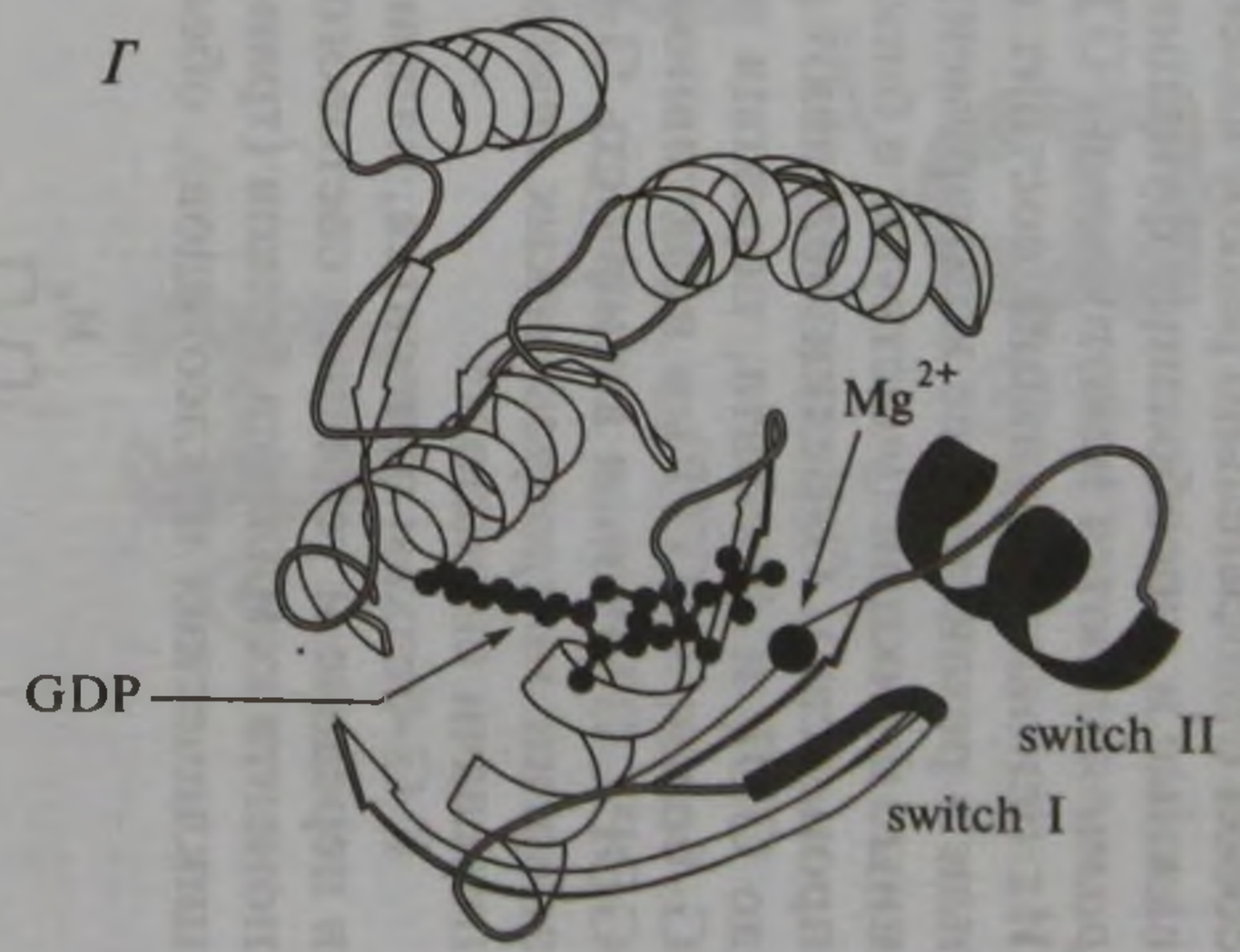
Рис. 25. Схема участия G-белков в передаче гормонального сигнала аденилатциклазе.

Гормоны (H_s и H_i) передают сигнал соответствующим рецепторам (R_s и R_i), с которых он далее поступает на белки G_s и G_i. Комплексы α-субъединиц (α_s и α_i) с GTP либо стимулируют, либо ингибируют активность аденилатциклазы (C). Аденилатциклаза катализирует образование cAMP из ATP


A
SII
SI
GDP
SIII


Б
switch II
Домен III
GDP
Mg
switch I
(эффекторная петля)
Домен II


B
α4
α3
β6
G-4
β5
β4
G-1
switch II
β1
α2
α5
G-5
GppCp
β3
G-3
β2
α1
G-2


Г
Mg2+
GDP
switch II
switch I

кратное ($5 \cdot 10^5$ раз) усиление светового сигнала в сетчатке глаза. Молекула трансдуцина состоит из трех различных субъединиц (T_α, T_β и T_γ), из которых GTPразной активностью обладает субъединица T_α. Сходное строение имеют белки G_s и G_i — компоненты широко распространенной в живых системах аденилатциклазной системы, участвующие, в частности, в передаче гормональных сигналов внутрь клеток-мишеней. Эти белки способны воспринимать сигнал от белка-рецептора гормона и передавать его аденилатциклазе — ферменту, осуществляющему синтез cAMP — вторичного мессенджера (посредника), который в свою очередь регулирует активность целого ряда внутриклеточных ферментов и соответствующих биохимических процессов. Белки G_s и G_i, как и трансдуцин, представляют собой тримеры, состоящие из различных субъединиц (α, β, γ). G_s и G_i называют соответственно стимулирующим и ингибирующим белками, способными альтернативно влиять на активность аденилатциклазы. У G_s-белка субъединица α способна связывать и гидролизовать GTP, что ведет к изменению конформации G-белка и аденилатциклазы, которая переходит в активное состояние (рис. 25). Комплекс субъединицы α, GDP, субъединиц β и γ представляет собой неактивную форму белка (α-GDP—β-γ). Замещение GDP на GTP в этом комплексе резко убыстряется под действием стимулирующего рецептора R_s, расположенного в мембране. Присоединение GTP ведет к диссоциации четвертичной структуры на две составляющие: димер субъединиц β-γ и мономер α-GTP, который и представляет собой активную форму G_s-белка. Последняя вступает во взаимодействие с аденилатциклазой и активирует ее, что приводит к многократному увеличению концентрации cAMP и соответствующему изменению внутриклеточных процессов. Продолжительность действия G_s-белка на аденилатциклазу определяется временем гидролиза GTP под действием α-субъединицы, по истечении которого комплекс Gs-α-GDP вновь соединяется с димером β-γ-субъединиц, и белок возвращается в неактивное состояние.

←

Рис. 26. Строение G-белков (*Sprang S.R.* // Ann. Rev. Biochem. — 1997. — V. 66. — P. 639):

А — гетеротример белка G_i, связанный с плазматической мембраной (поверхности субъединицы γ (слева) и субъединицы α (справа), контактирующие с мембраной, выделены темным цветом; SI, SII, SIII — участки α-субъединицы; молекула GDP связана с α-субъединицей); *Б* — структура фактора элонгации трансляции (EF-Tu) бактерии *Thermus aquaticus*, связывающего комплекс GDP — Mg^{+2}; *В*, *Г* — структура белка Ras (*Г* — комплекс белка Ras с GDP — Mg^{+2}; *В* — комплекс белка Ras с негидролизуемым аналогом GTP ($G_{pp}C_p$); α_1—α_5 — α-спирали; β_1—β_6 — β-структуры; G-1—G-5 — гомологичные области, свойственные многим G-белкам)

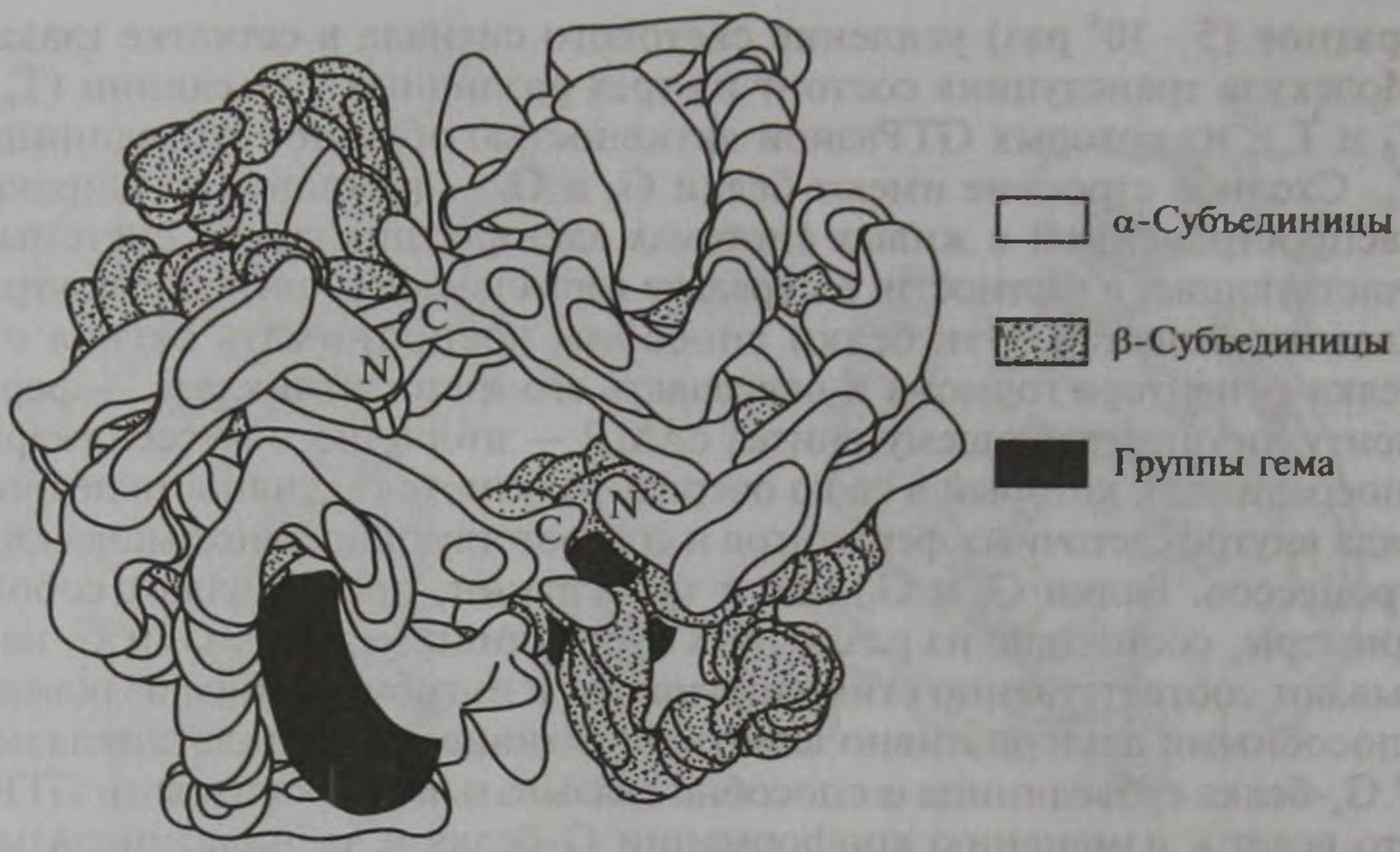


Рис. 27. Модель четвертичной структуры гемоглобина

Пространственная структура тримера Gi-белка представлена на рис. 26. На этом же рисунке приведены структуры двух других G-белков: GTP-связывающего фактора элонгации трансляции (биосинтеза белка) EF-T_u бактерий (см. гл. 11) и белка Ras, участвующего в пролиферации клеток млекопитающих. Последние два белка не имеют четвертичной структуры, однако обладают выраженной гомологией первичной и третичной структур в области связывания GTP и GDP. Важным родственным элементом структур этих белков является участок, содержащий спираль (switch II) и осуществляющий контакт с γ-фосфатной группой GTP. Филогенетическое родство G-белков, выделенных из разных организмов, проявляется в консервативных аминокислотных последовательностях, выявленных в составе этих белков (участки G-1 — G-5 белка Ras на рис. 26) у бактерий, одноклеточных, млекопитающих и человека.

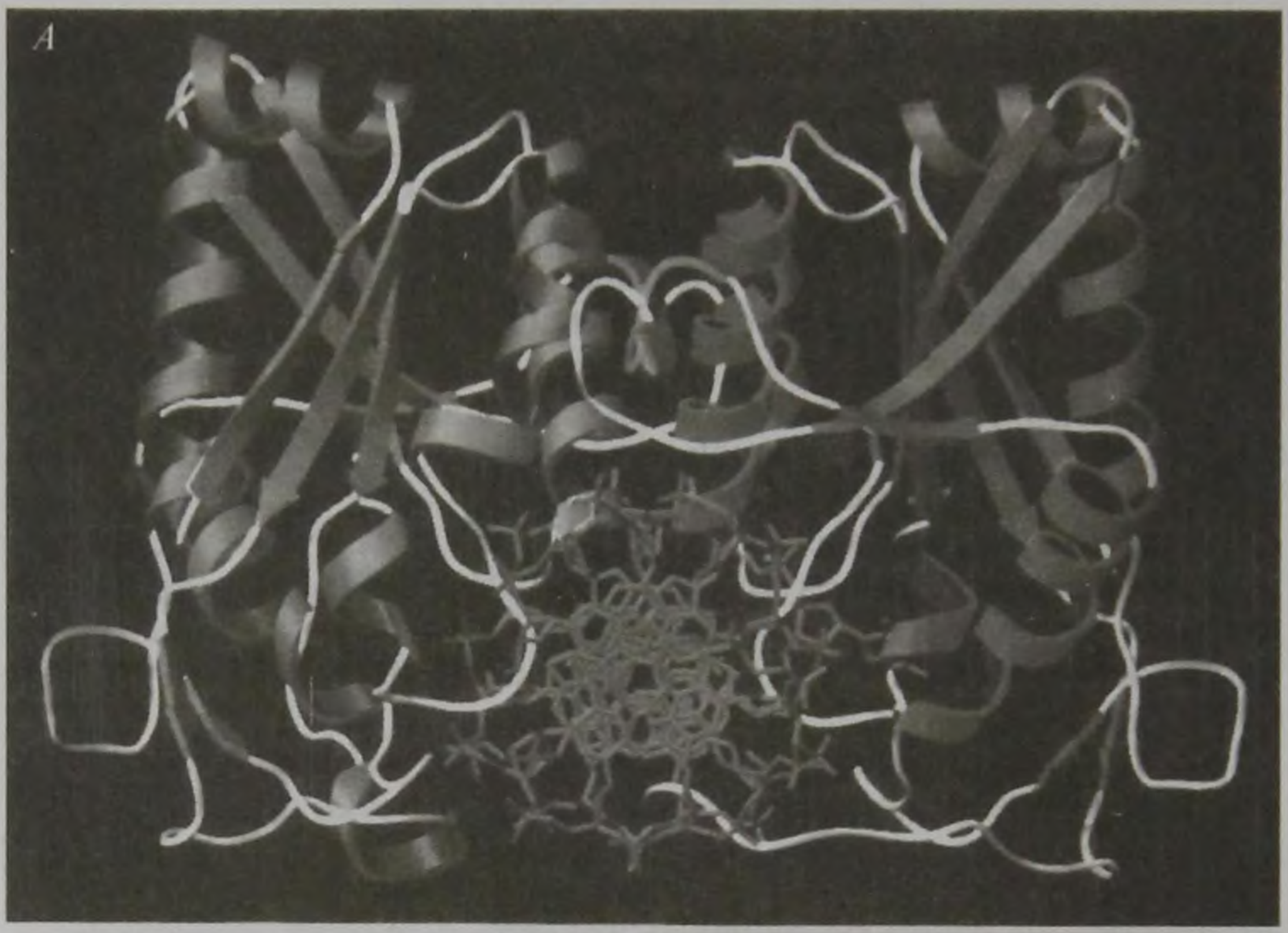
В составе белков с четным числом субъединиц могут содержаться идентичные или различные по структуре протомеры. Так, фермент альдолаза из скелетных мышц кролика состоит из четырех одинаковых субъединиц, а гемоглобины млекопитающих (также тетрамеры по структуре) содержат в своем составе две пары различных субъединиц: две субъединицы типа α и две субъединицы типа β. Таким образом, четвертичная структура гемоглобинов отвечает формуле 2α2β. В обоих вышеотмеченных примерах тетрамерных белков их протомеры пространственно расположены в вершинах тетраэдра (рис. 27). Наиболее яркие примеры четвертичной структуры можно найти среди каталитически активных белков — ферментов, для которых особенно характерна взаимосвязь структуры и функциональных свойств.

Димерами, состоящими из субъединиц одного вида (гомодимерами), являются альдегидоксидоредуктаза (см. рис. 16) и эндонуклеаза рестрикции (рестриктаза) *Bam* HI, схема строения которой представлена на рис. 28. Рестриктаза *Bam* HI, выделенная из *Bacillus amyloliquefacions* и относящаяся к рестриктазам типа II (подробнее о рестриктазах см. гл. 14), представляет собой димер и связывается с ДНК, находящейся в B-форме. После связывания с ДНК молекула этого фермента претерпевает ряд конформационных изменений, в результате которых отдельные участки полипептидных цепей каждой из субъединиц образуют своеобразные структуры («руки»), которые служат для плотного контакта с ДНК (см. рис. 28, *B*). «Рука» одной субъединицы входит в малую борозду ДНК, а «рука» другой субъединицы взаимодействует с сахарофосфатным остовом на другой стороне двойной спирали ДНК. Узнавание зоны расщепления цепей ДНК происходит в основном в зоне большой борозды (большого желоба) молекулы ДНК.

Среди белков, взаимодействующих с ДНК, выявлены молекулы, имеющие своеобразную *тороидальную* (кольцевую) структуру. В целом образование замкнутых структур (колец, трубок или сферических эпимолекул) дополнительно стабилизирует четвертичную структуру за счет увеличения общего числа контактов между протомерами. К таким белкам относятся, в частности, «белки скользящего зажима (обруча)». Так, у *E. coli* выявлен белок кольцевой формы, являющийся интегральной частью белкового комплекса, обеспечивающего репликацию ДНК (см. гл. 7). Белки, расплетающие двойные спирали ДНК в ходе репликации и называемые *хеликазами*, также имеют кольцевую структуру и представляют собой гексамеры. Число субъединиц в мультимерных молекулах тороидальных белков варьирует в широких пределах (от 6 — 7 до 12 — 15).

Еще более крупными молекулами представлены в природе белки *ферритины*, служащие для связывания и депонирования железа у различных видов животных и растений. Так, в эпимолекуле ферритина животных, которая представляет собой сферическую молекулу, содержится 24 субъединицы, между ними располагаются полости (каналы) для проникновения оксидов железа. Внутри центральной полости ферритина насекомых могут быть связаны 3 500 атомов железа на 1 молекулу такого белка-мультимера. Аналогичные каналы выявлены и у ферритинов растений (рис. 29).

Среди ферментов достаточно часто встречаются белки, содержащие различные по структуре и функциям субъединицы. Гетеродимерной структурой обладают протеинкиназы (ферменты, способные фосфорилировать определенные аминокислотные остатки в различных белках), одна из субъединиц которых является собственно каталитической (субъединица типа C), а другая — регуляторной (R). Последняя может связываться с cAMP, что переводит фермент в активное состояние. Гетеромультимером, состоящим из



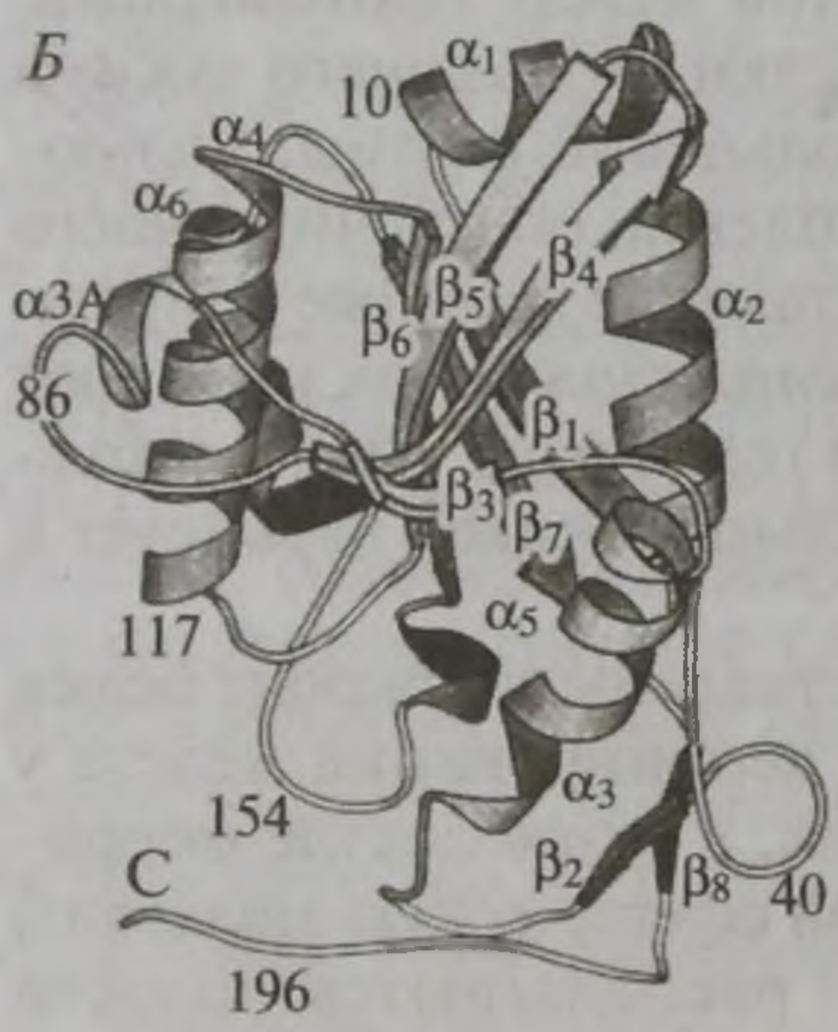




Рис. 28. Структура рестриктазы *Bam* HI, взаимодействующей с ДНК:

А — структура комплекса *Bam* HI-ДНК. Две субъединицы фермента (L и R) расположены соответственно слева и справа от обозначенной внизу молекулы ДНК; *Б* — R-субъединица рестриктазы *Bam* HI (показано расположение α-спиралей (α_1—α_2) и β-слоев (β_1—β_8); цифрами обозначено расположение отдельных аминокислотных остатков начиная с N-конца молекулы, C — карбоксильный конец молекулы); *В* — взаимодействие отдельных структурных элементов R- и L-субъединиц рестриктазы с молекулой ДНК (спирали и петли вторичной структуры субъединицы R выделены темным цветом, соответствующие элементы L-субъединицы — серым)

нескольких различных субъединиц, является и РНК-полимераза (см. гл. 9).

Впечатляющим примером сложной четвертичной структуры может служить молекула аспартаткарбамилтрансферазы *E. coli*. Формула четвертичной структуры этого фермента может быть записана так: $(C_3)_2(R_2)_3$, т.е. в составе данного белка имеются шесть каталитических субъединиц, объединенных в два тримера, и шесть регуляторных субъединиц, образующих три димера. Регуляторные субъединицы могут взаимодействовать с активатором (АТР) или ингибитором (CTP), что влечет за собой изменение конформационного состояния эпимолекулы. В зависимости от того, какой эффектор присоединяется к регуляторным субъединицам, молекула фермента может приобретать рыхлую (релаксированную) форму (R-форма), которой свойственна высокая активность, либо более плотную и менее активную форму (Т-форма).

В целом наличие различных *молекулярных форм* у белков, обладающих четвертичной структурой, представляет собой достаточно распространенное явление, обусловленное относительной подвижностью отдельных субъединиц в эпимолекуле. Существование таких множественных форм имеет фундаментальное значение для регуляции их функциональной активности в клетке, отдельных ее компартментах, а также в тканях и органах высших организмов.

От четвертичной структуры следует отличать *олигомерное состояние белков*. Многие белки обладают способностью образовывать структуры с изменяющимся в определенной пропорции числом полипептидных цепей, однако это не связано с возникновением или исчезновением специфических функций. Так, сывороточный альбумин быка может существовать не только в виде мономера (67 кДа), но и в виде димера, тримера и тетрамера, что, однако, не ведет к появлению каких бы то ни было новых функций. Не обладают, по существу, четвертичными структурами и белковые молекулы, воз-

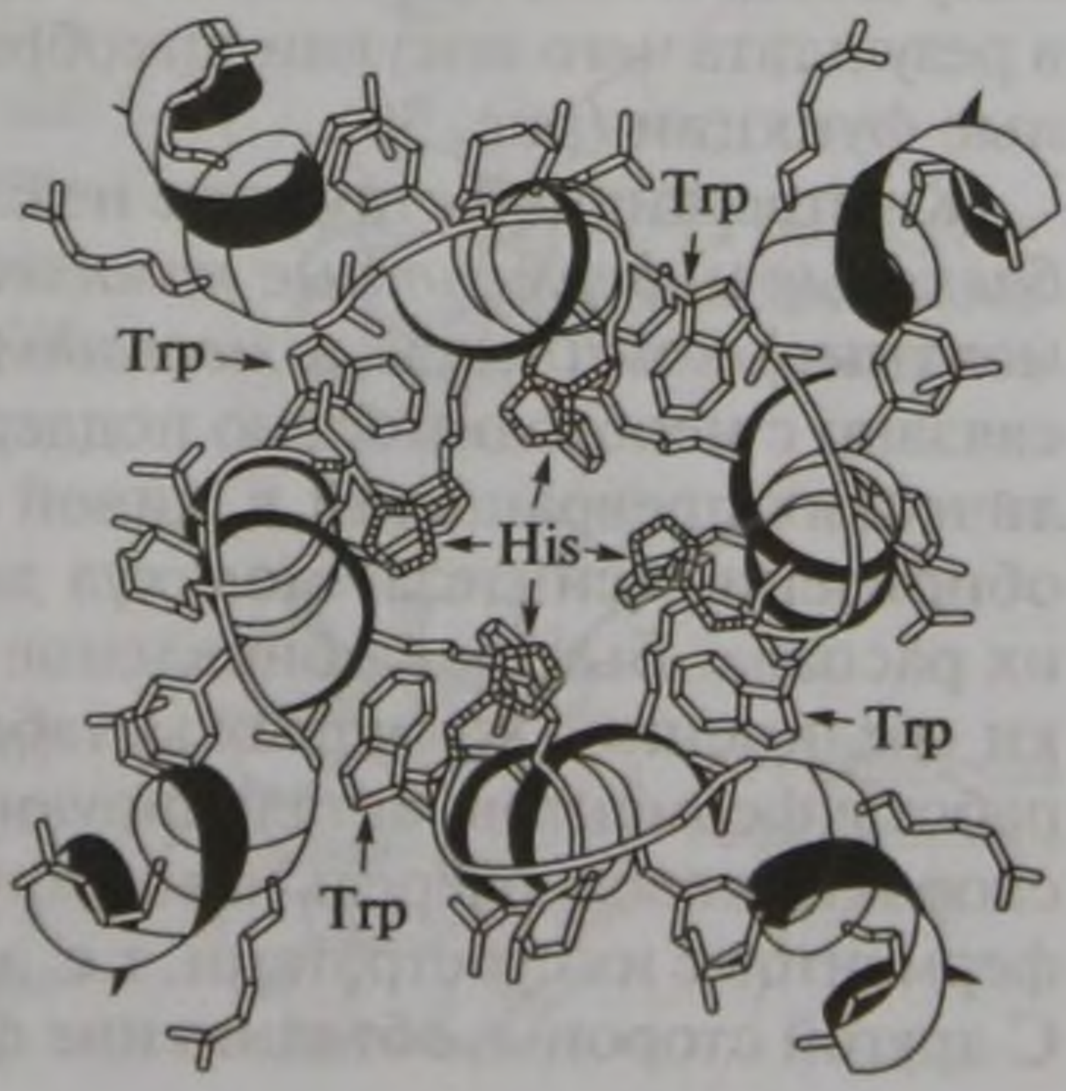
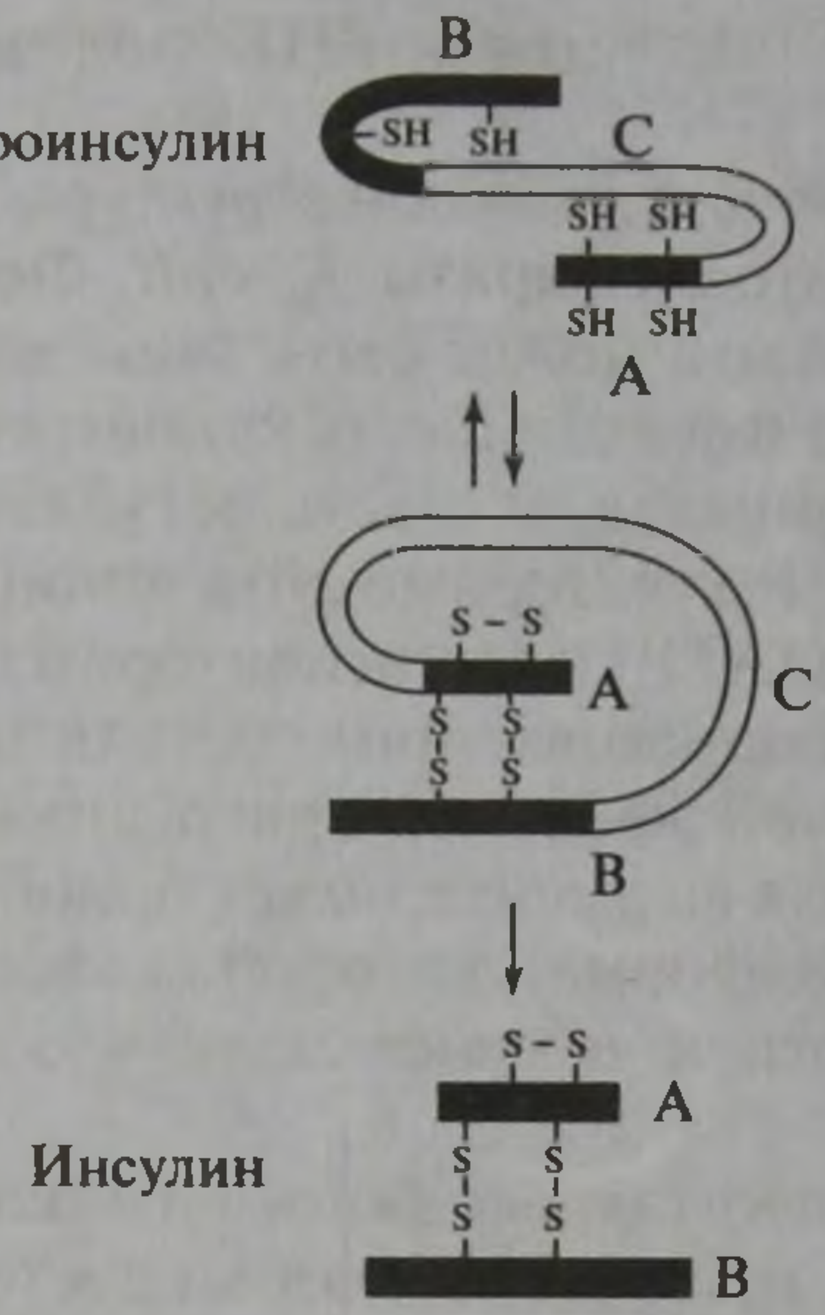


Рис. 29. Компьютерное изображение структуры молекулы апоферритина из семян гороха.

Молекула состоит из четырех субъединиц. Гидрофильные остатки His и Arg, принадлежащие четырем субъединицам, формируют каналы, куда проникают ионы железа. Показаны также остатки Trp, расположенные на границах раздела субъединиц



Сближение А- и В-участков молекулы проинсулина и образование дисульфидных мостиков

Протеолитическое выщепление пептида С приводит к образованию зрелой молекулы гормона

Рис. 30. Образование активной формы инсулина из молекулы предшественника (проинсулина) путем выщепления С-пептида. Пояснения см. в тексте

никающие в результате расщепления (ограниченного протеолиза) белков-предшественников и состоящие в связи с этим из нескольких полипептидных цепей. Примерами таких белков могут служить химотрипсин (возникает из белка-предшественника — *прохимотрипсина* — и состоит из трех полипептидных цепей, ковалентно связанных между собой дисульфидными мостиками) и инсулин. Молекула инсулина состоит из двух полипептидных цепей — А и В, содержащих соответственно 21 и 30 аминокислотных остатков. Эти цепи образуют единую структуру с дисульфидными связями, которая получается после вырезания из проинсулина связующего пептида С, в результате чего инсулин приобретает свойственные ему гормональные функции (рис. 30).

К четвертичной структуре не следует относить и *надмолекулярные белковые* и *ферментные комплексы* — мультиэнзимные (полиферментные) комплексы и *метаболоны*. Образование таких структур связано с необходимостью поддержания высоких скоростей метаболических превращений в живой клетке. Иными словами, скорость обновления (синтеза) молекул должна преобладать над скоростью их распада. Быстрое обновление принципиально важных для клетки соединений (ключевых метаболитов) зависит от эффективности работы ферментов, катализирующих соответствующие реакции, чья скорость в свою очередь лимитируется возможностями столкновения ферментов с их субстратами, т. е. элементарным процессом диффузии. С другой стороны, объединение ферментов в достаточно устойчивые

комплексы способно существенно упростить (ускорить) протекание серии химических реакций за счет адресной передачи метаболитов от одного фермента к другому. Среди такого рода полиферментных комплексов можно различать *абсорбционные* и *интегральные ансамбли* ферментов. Адсорбционные ансамбли образуются за счет сил слабых (не ковалентных) взаимодействий между молекулами ферментов на поверхности мембраны либо какой-то крупной центральной (коровой) молекулы. Интегральные комплексы оказываются встроены во внутриклеточные мембраны. К интегральным комплексам относятся, например, цепочки окислительно-восстановительных ферментов (цитохромов), встроенные во внутренние мембраны митохондрий. Эти ансамбли ферментов осуществляют быструю передачу электронов с окисляемого субстрата на кислород, благодаря чему создаются условия для сопряженного с этим процессом синтеза АТР. Интенсивность работы такой системы ферментов столь высока, что позволяет одной клетке за 1—2 мин полностью возобновлять запасы АТР (10^7 молекул), а весь организм человека способен синтезировать около 1 г АТР в минуту.

К адсорбционным комплексам можно отнести полиферментный комплекс окислительного декарбоксилирования пировиноградной кислоты (ПВК), состоящий из трех ферментов, каждый из которых в этом комплексе представлен определенным числом молекул (рис. 31). Этот мультиэнзимный комплекс осуществляет быстрое превращение ПВК в другой ключевой метаболит — ацетил-СоА. Еще более крупные адсорбционные комплексы носят название *метаболонов* — ансамблей ферментов, катализирующих серию реакций, составляющих определенный *метаболический путь*. В составе известных метаболонов *гликолиза* и *цикла Кребса* входит более 10 различных ферментов, между которыми возможна адресная «эстафетная» передача субстратов.

Один из видов полиферментных комплексов — ***протеасомы***. Это крупные, имеющие определенный план строения комплексы протеолитических ферментов, расщепляющие различные внутриклеточные белки.



Рис. 31. Строение мультиэнзимного комплекса окислительного декарбоксилирования пировиноградной кислоты.

Этот комплекс катализирует превращение пировиноградной кислоты в ацетил-СоА

Они обнаружены у самых разных организмов (от архебактерий до человека) и представляют собой существенный компонент *АТР-зависимого протеолиза*, в ходе которого осуществляется исключительно быстрая деградация прежде всего аномальных белков, возникающих либо вследствие мутаций, либо в результате ошибок белоксинтезирующей системы. Протеасомы расщепляют белки, как правило, до коротких пептидов, которые затем гидролизуются до отдельных аминокислот цитоплазматическими экзопептидазами. Протеасомы, выделенные из разных объектов, имеют молекулярную массу около 2 000 кДа (26S-протеасомы) и весьма сходны по структурно-функциональной организации. По современным данным, протеасома содержит по крайней мере 28 протомеров (отдельных белковых глобул), собранных в стопку из 4 колец по 7 протомеров в каждом кольце, внутри стопки располагается полость, где и происходит деструкция белков (рис. 32). Два внешних кольца формируются α-субъединицами, выполняющими регуляторную функцию, а два внутренних — β-субъединицами, которые обладают протеолитической активностью. Молекулярные массы протомеров варьируют от 22 до 35 кДа, причем различные протомеры не всегда представлены в равной степени. Например, дрожжевые протеасомы содержат 14 различных типов протомеров, протеасомы растений — 12—15, а другие эукариотические протеасомы могут включать до 25 различных белков. В составе некоторых протеасом выявлены РНК, состоящие примерно из 80 нуклеотидных остатков.

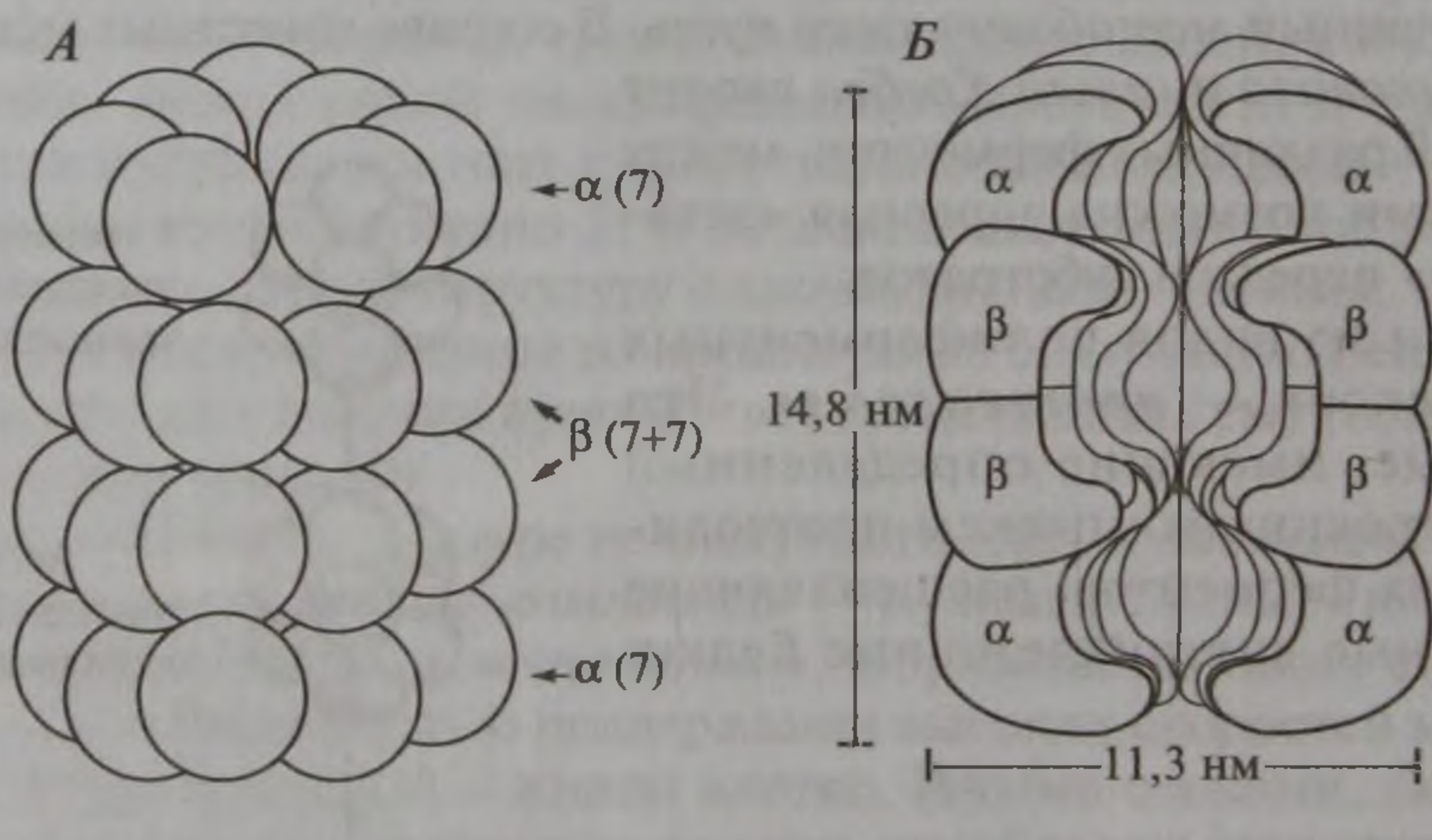


Рис. 32. Схема строения протеасомы из *Thermoplasma acidophillum* (*Coldberg A.L.* // Science. — 1995. — V. 268. — P. 522):

А — общий план строения 20S-протеасомы; *Б* — 20S-протеасома в разрезе. Два внешних кольца состоят из семи α-субъединиц каждое; мол. масса субъединицы — 25,9 кДа. Два внутренних кольца состоят из семи β-субъединиц каждое; мол. масса субъединицы — 22,3 кДа. Кольца образуют центральный канал с тремя камерами, входы в которые располагаются в верхней и нижней частях комплекса между α-субъединицами

Протеасомы присутствуют не только в цитоплазме, но и в ядрах клеток, а также в эндоплазматическом ретикулуме, однако их содержание особенно высоко в цитоплазматическом матриксе (не менее 50 %). Рентгеноструктурный анализ показал, что 20S-протеасома из *Thermoplasma acidophillum* имеет форму бочонка (14,8 · 11,3 нм) с проникающим центральным каналом и тремя большими внутренними камерами. Доступ в полость осуществляется через узкие отверстия в центре каждого кольца. Отверстия внешних колец составляют всего 1,3 нм в диаметре, что чуть больше диаметра α-спиралей и указывает на необходимость разворачивания белковых молекул перед их проникновением в протеасомы. Отдельные протомеры обладают различной специфичностью в отношении пептидных связей расщепляемых ими белков, что способствует быстрой деструкции последних в полости протеасом. В составе протеасом имеются

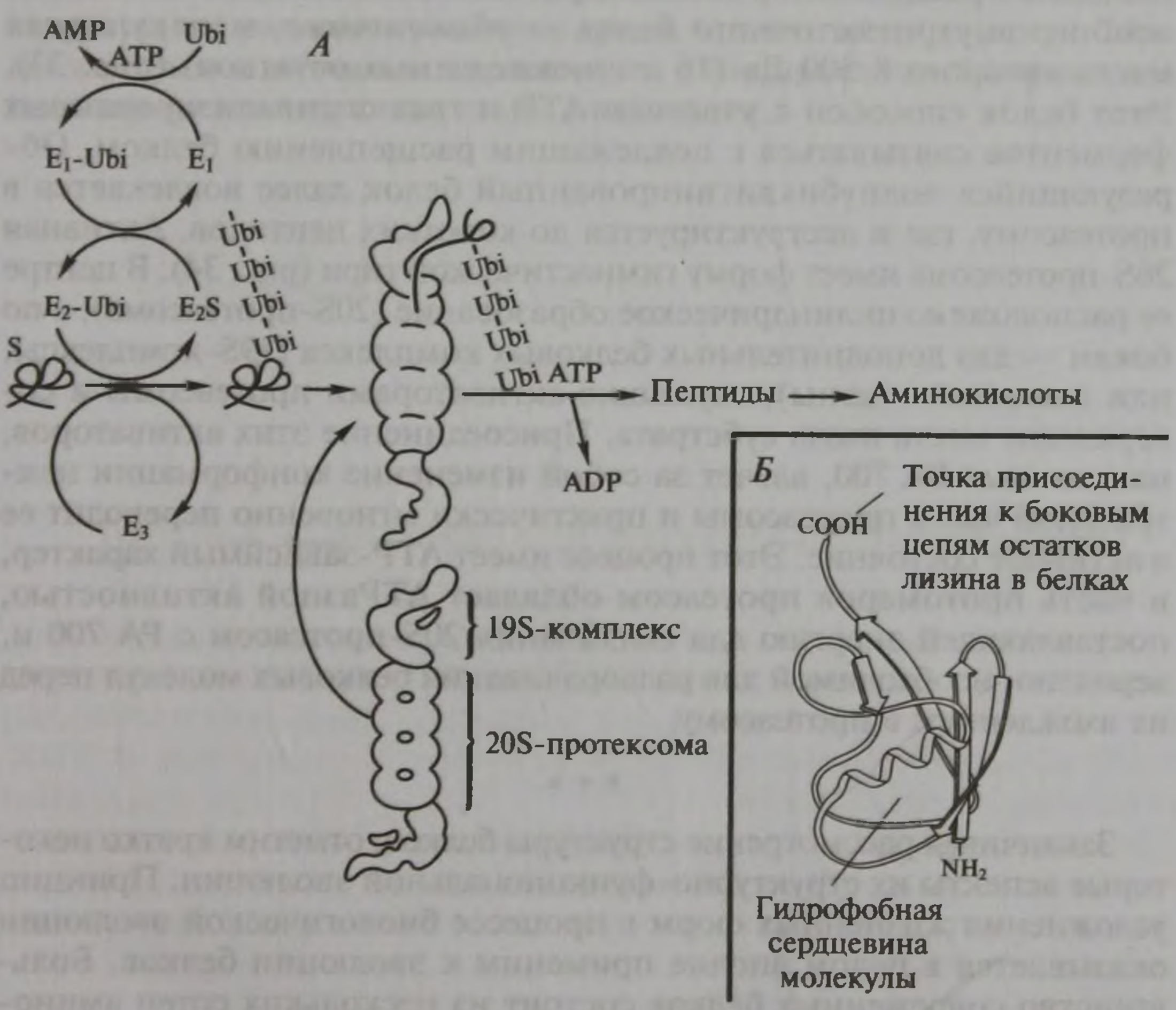


Рис. 33. АТР-зависимый механизм деструкции белка (S) с участием убиквитина и протеасомы. Пояснения см. в тексте:

А — схема процесса с участием ферментов (E_1, E_2, E_3), осуществляющих полиубиквитирование белка; *Б* — трехмерная структура убиквитина; E_1 — убиквитин-активирующий фермент; E_2 — переносчик убиквитина; E_3 — убиквитинпротеинлигаза; Ubi — убиквитин

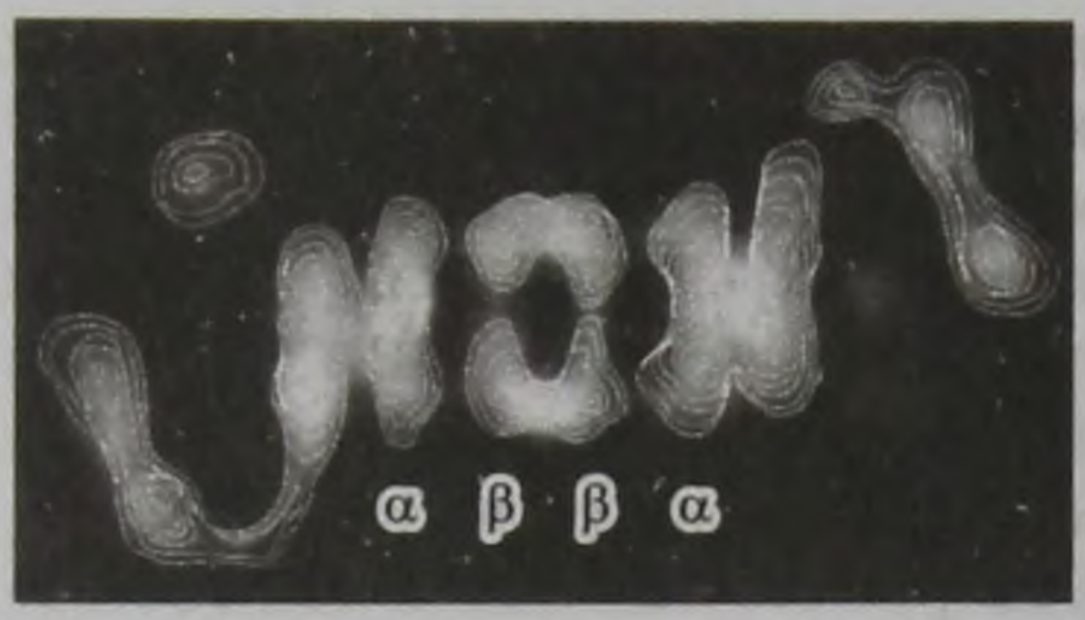


Рис. 34. Электронная микрофотография 26S-протеасомы (Ann. Rev. Biochem. — 1996. — V. 65. — P. 802).

Этот белковый комплекс имеет форму гимнастической гири. α и β — субъединицы 20S-протеасомы. Справа и слева от α-субъединиц протеасомы располагаются белковые комплексы PA 700, выполняющие роль активаторов и имеющие большие V-образные впячивания — места входа субстрата

трипсинподобные и химотрипсинподобные протеиназы, а также постглутамилпротеиназы (о специфичности протеиназ см. выше). Механизм расщепления белков протеасомами связан с действием особого внутриклеточного белка — *убиквитина*, молекулярная масса которого 8 500 Да (76 аминокислотных остатков) (рис. 33). Этот белок способен с участием ATP и трех специализированных ферментов связываться с подлежащим расщеплению белком. Образующийся полиубиквитинированный белок далее вовлекается в протеасому, где и деструктируется до коротких пептидов. Активная 26S-протеасома имеет форму гимнастической гири (рис. 34). В центре ее расположено цилиндрическое образование (20S-протеасома), а по бокам — два дополнительных белковых комплекса (19S-комплексы, или шапочки — кэпы), служащие активаторами протеасомы и содержащие места входа субстрата. Присоединение этих активаторов, называемых PA 700, влечет за собой изменение конформации центральной части протеасомы и практически мгновенно переводит ее в активное состояние. Этот процесс имеет ATP-зависимый характер, и часть протомеров протеасом обладает ATPазной активностью, поставляющей энергию для связывания 20S-протеасом с PA 700 и, вероятно, необходимой для разворачивания белковых молекул перед их вхождением в протеасому.

* * *

Заканчивая рассмотрение структуры белков, отметим кратко некоторые аспекты их структурно-функциональной эволюции. Принцип усложнения жизненных форм в процессе биологической эволюции оказывается в целом вполне применим к эволюции белков. Большинство современных белков состоит из нескольких сотен аминокислотных остатков, и вместе с тем в природе существуют короткие, составленные не более чем из нескольких десятков остатков биологически активные пептиды. Главенствующий и, очевидно, сложившийся на самых ранних этапах эволюции *матричный принцип* биосинтеза белков позволяет считать, что более протяженные полипептиды появились в результате наращивания или слияния более коротких генов.

Увеличение размеров полипептидной цепи открывало возможность для образования серии изгибов (петель) и обособления отдельных фолдов, возникающих в процессе сворачивания (фолдинга) цепей в пространстве. Это постепенно привело к возникновению вторичной, сверхвторичной и доменной структуры глобулярных белков, т.е. к последовательной (плавной) эволюции их структуры. С другой стороны, слияние генов или внедрение чужеродных (вирусных) генов могло приводить к быстрому скачкообразному изменению структуры белков, появлению белковых коньюгатов, к которым, вероятно, принадлежат вышеохарактеризованные полифункциональные белки. Необычные с точки зрения «традиционной» эволюции белковые структуры могли, видимо, возникать в результате внедрения вирусных кДНК в геномы уже достаточно процессированных форм жизни. Возникновение в результате мутаций аллельных вариантов генов, их рекомбинация еще более увеличивали разнообразие белковых структур, не имеющее аналогов среди прочих макромолекул. Молекулярным аспектам кодирования и декодирования белковых тел посвящены последующие главы этой книги.

Глава 3
НУКЛЕИНОВЫЕ КИСЛОТЫ

Нуклеиновые кислоты играют основную роль в хранении и реализации генетической информации. Различают два вида нуклеиновых кислот: дезоксирибонуклеиновые кислоты (ДНК), которые обеспечивают хранение генетической информации, и рибонуклеиновые кислоты (РНК), принимающие участие в ее реализации. Хромосома прокариотической клетки представляет собой одну длинную двуцепочечную молекулу ДНК, собранную в компактное образование — *нуклеоид*. Эукариотические клетки содержат большое число молекул ДНК, которая неравномерно распределена по нескольким хромосомам в виде комплексов с многочисленными белками. Вирусы также содержат в качестве генетического материала нуклеиновые кислоты: у одних это ДНК, у других РНК.

В реализации генетической информации принимают участие различные виды РНК.

Первичная структура нуклеиновых кислот

Нуклеиновые кислоты являются биологическими полимерами. Мономерными звеньями ДНК и РНК являются нуклеотиды (рис. 35) — фосфорные эфиры нуклеозидов, которые, в свою очередь, построены из пентозы и гетероциклического основания. В РНК углеводный компонент — D-рибоза, в ДНК — D-2-дезоксирибоза. Связь между углеводным остатком и гетероциклическим основанием в нуклеотиде осуществляется с помощью N-гликозидной связи. В качестве гетероциклических оснований ДНК содержит два пурина: аденин (A) и гуанин (G) и два пиримидина: тимин (T) и цитозин (C). В РНК вместо тимина содержится урацил (U).

Мономерные остатки в нуклеиновых кислотах связаны между собой фосфодиэфирными связями. Как в ДНК, так и в РНК, эта связь осуществляется только за счет 3′-OH одного нуклеотидного остатка и 5′-OH другого. Такую межнуклеотидную связь называют 3′,5′-фосфодиэфирной. Цепи ДНК и РНК обладают определенной полярностью, или направлением, поскольку все межнуклеотидные

фосфодиэфирные связи ориентированы вдоль цепи одинаково. Благодаря полярности каждая полинуклеотидная цепь имеет 5′-конец и 3′-конец (рис. 36). Из-за наличия 2′-ОН группы у рибозы межнуклеотидные связи в РНК значительно мобильнее, чем в ДНК; они легко гидролизуются в присутствии щелочи.

После окончания цикла синтеза ДНК некоторые пуриновые и пиримидиновые основания могут подвергаться химической модификации. Модификация пуринов и пиримидинов характерна и для РНК (рис. 37). В ней присутствуют 2′-О-метилнуклеотиды, а также может наблюдаться иной способ связывания между урацилом и рибозой (так называемая псевдоуридиловая кислота). Модифицированные нуклеотиды достаточно часто встречаются в тРНК. Как и у ДНК, модификация оснований и рибозных остатков в РНК происходит после завершения синтеза РНК, а не на стадии биосинтетических предшественников.

Определение нуклеотидной последовательности ДНК и РНК. Исходную молекулу ДНК предварительно фрагментируют с помощью рестриктаз, причем осуществляют независимое расщепление двумя и более рестриктазами, в результате чего образуются перекрывающиеся фрагменты. Это позволяет после определения нуклеотидной последовательности соответствующих фрагментов реконструировать первичную структуру всей ДНК.

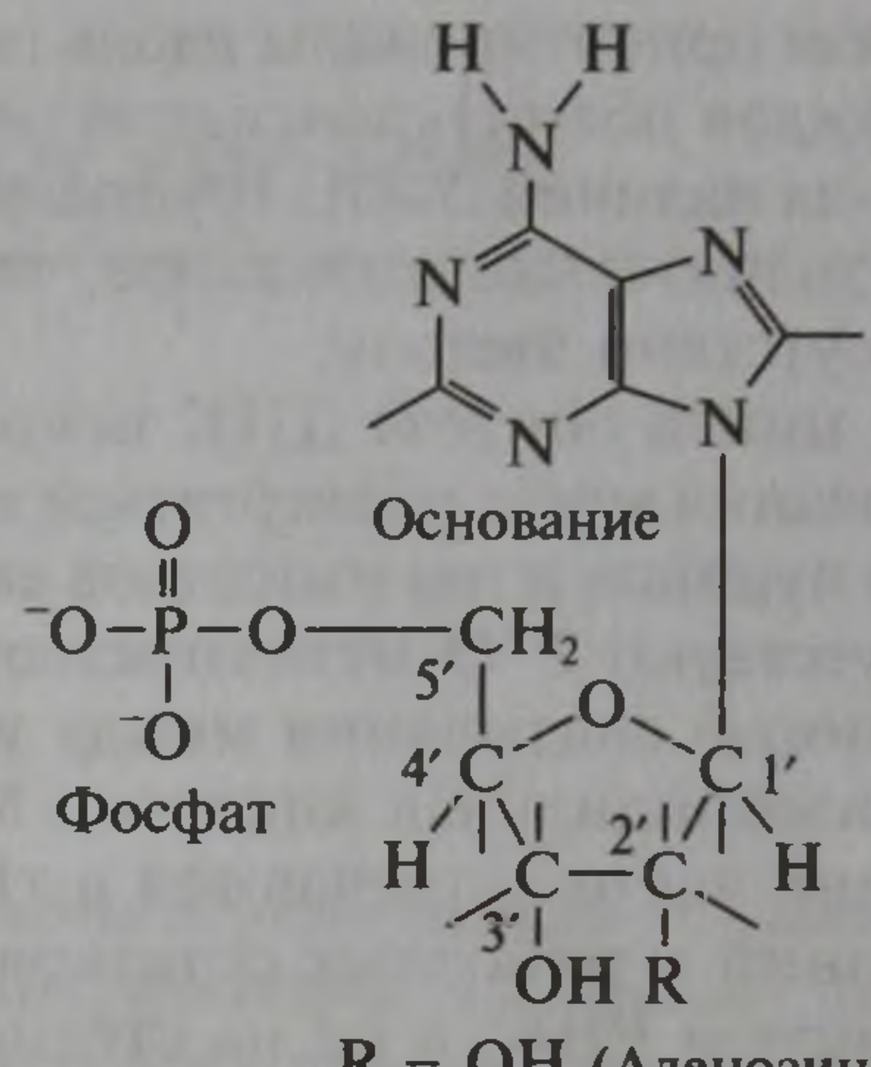
Существуют два принципиально различных подхода к определению первичной структуры ДНК. В основе одного лежат химические реакции, в которых используются непосредственно фрагменты очищенной ДНК, во втором случае используют ДНК-копии очищенных сегментов, полученных ферментативным путем (см. с. 377). Оба метода полностью автоматизированы. При секвенировании ДНК по Сангеру (второй метод) вводят флуоресцентные метки, причем для каждого из четырех анализируемых нуклеотидов используются флуоресцирующие агенты с различными спектральными характеристиками, что позволяет осуществить сканирование гелей при разных длинах волн и передавать данные непосредственно в компьютер.

Полная нуклеотидная последовательность многих видов РНК (длиной до 250—300 нуклеотидов) может быть прямо определена с помощью метода Максама—Гилберта (см. с. 377). Наряду с химическим расщеплением секвенирование РНК можно проводить путем гидролиза ее цепи специфическими РНКазами.

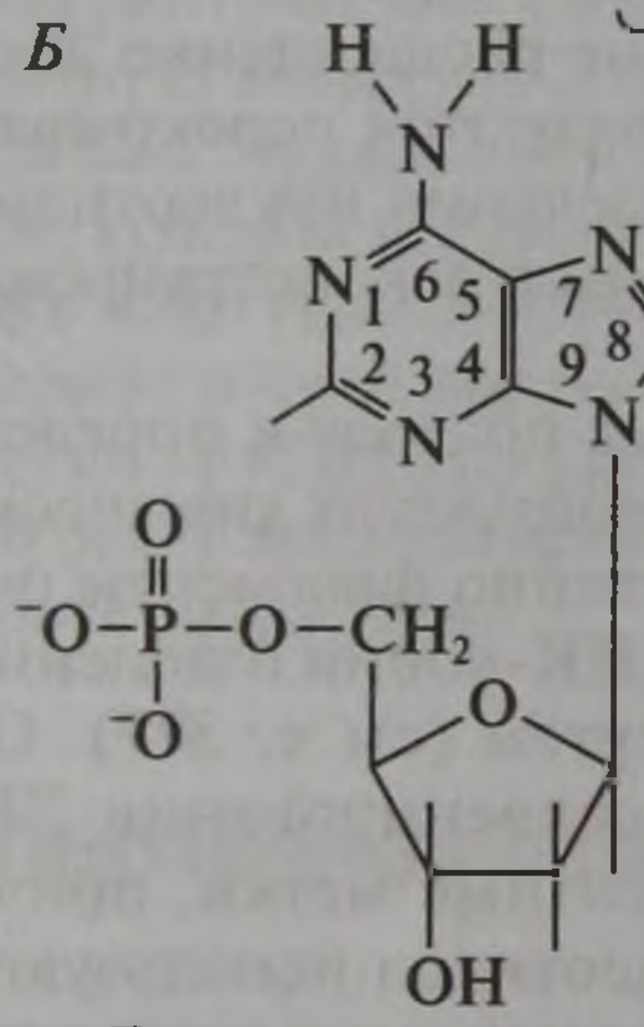
В настоящее время для определения нуклеотидной последовательности достаточно протяженной полинуклеотидной цепи РНК (например, мРНК) секвенируют ген (т.е. ДНК), в котором эта РНК закодирована. Для этой цели выделяют необходимую РНК и с помощью обратной транскриптазы синтезируют кДНК (см. с. 361), которую и используют в анализ.

Компьютерный анализ нуклеотидных последовательностей. На первый взгляд длинные ряды оснований кажутся совершенно

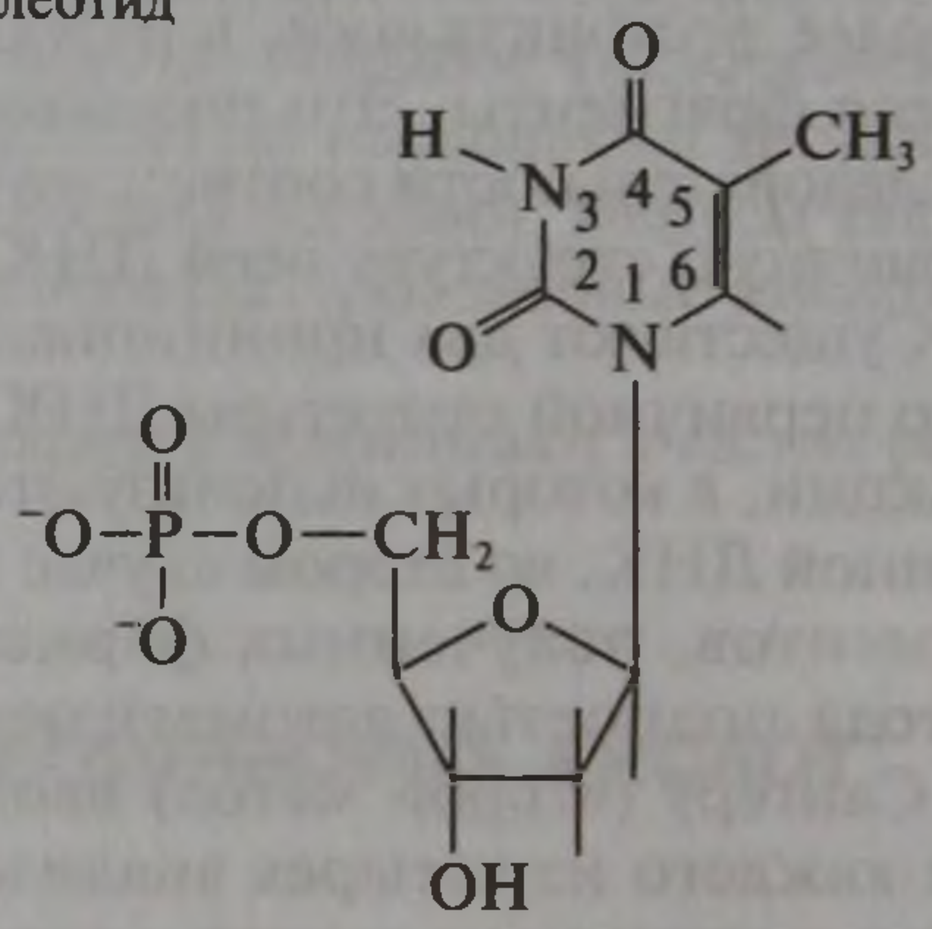
А



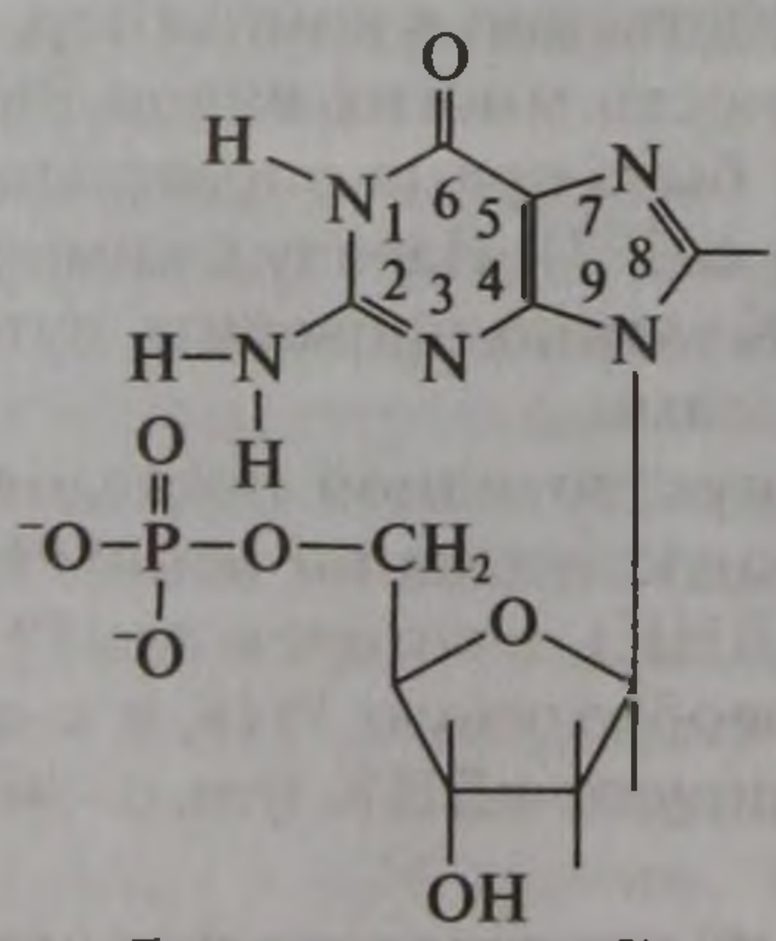
Б



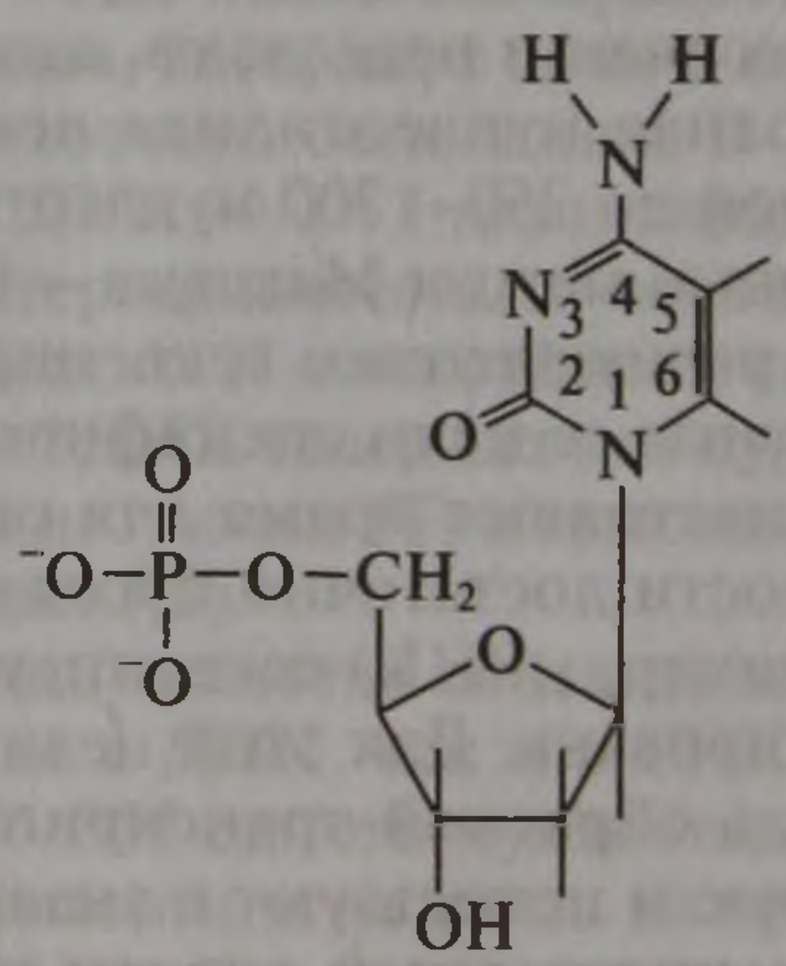
Дезоксиаденозин-5′-монофосфат (dAMP)



Дезокситимидин-5′-монофосфат (dTMP)



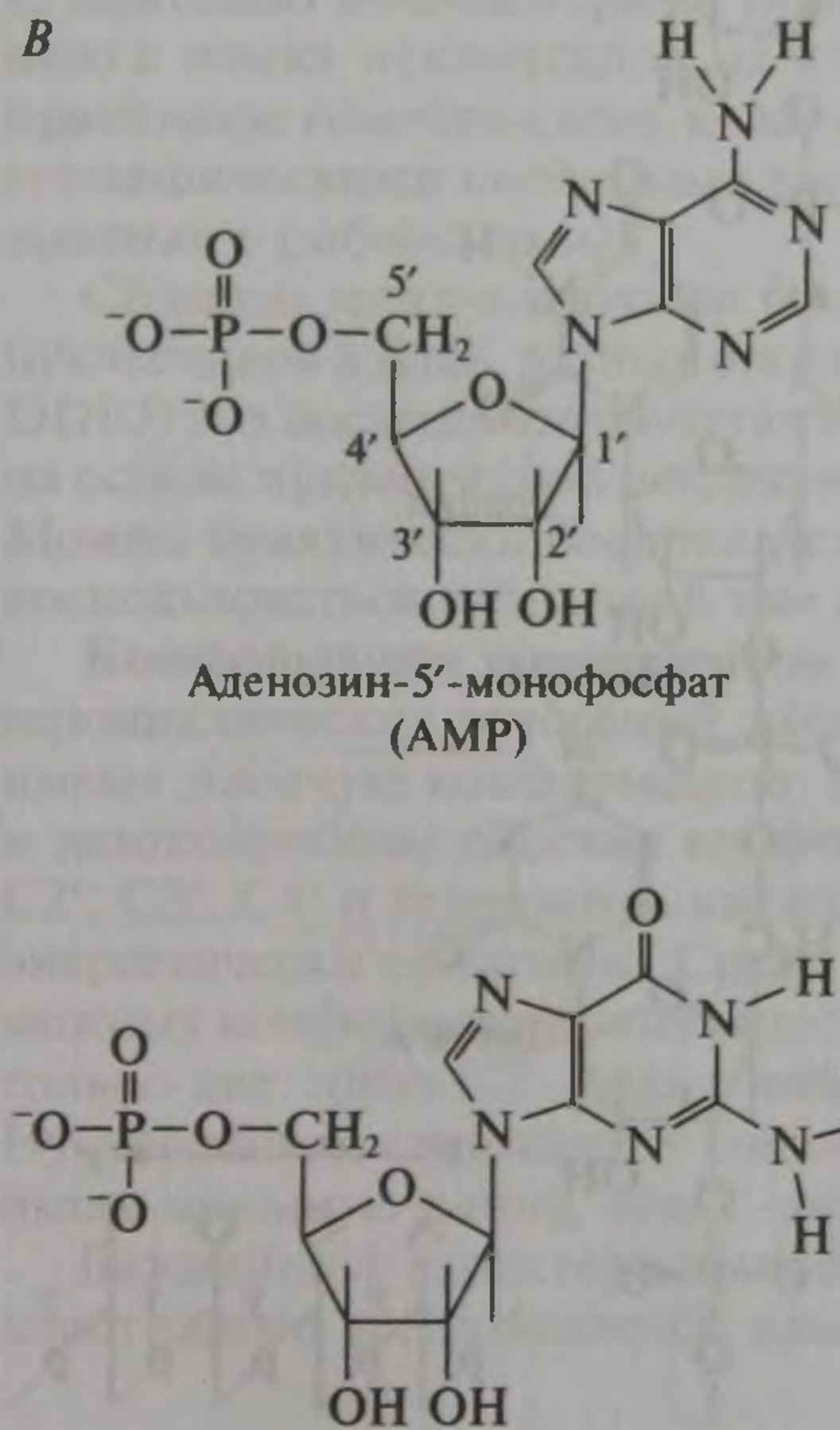
Дезоксигуанозин-5′-монофосфат (dGMP)



Дезоксицитидин-5′-монофосфат (dCMP)

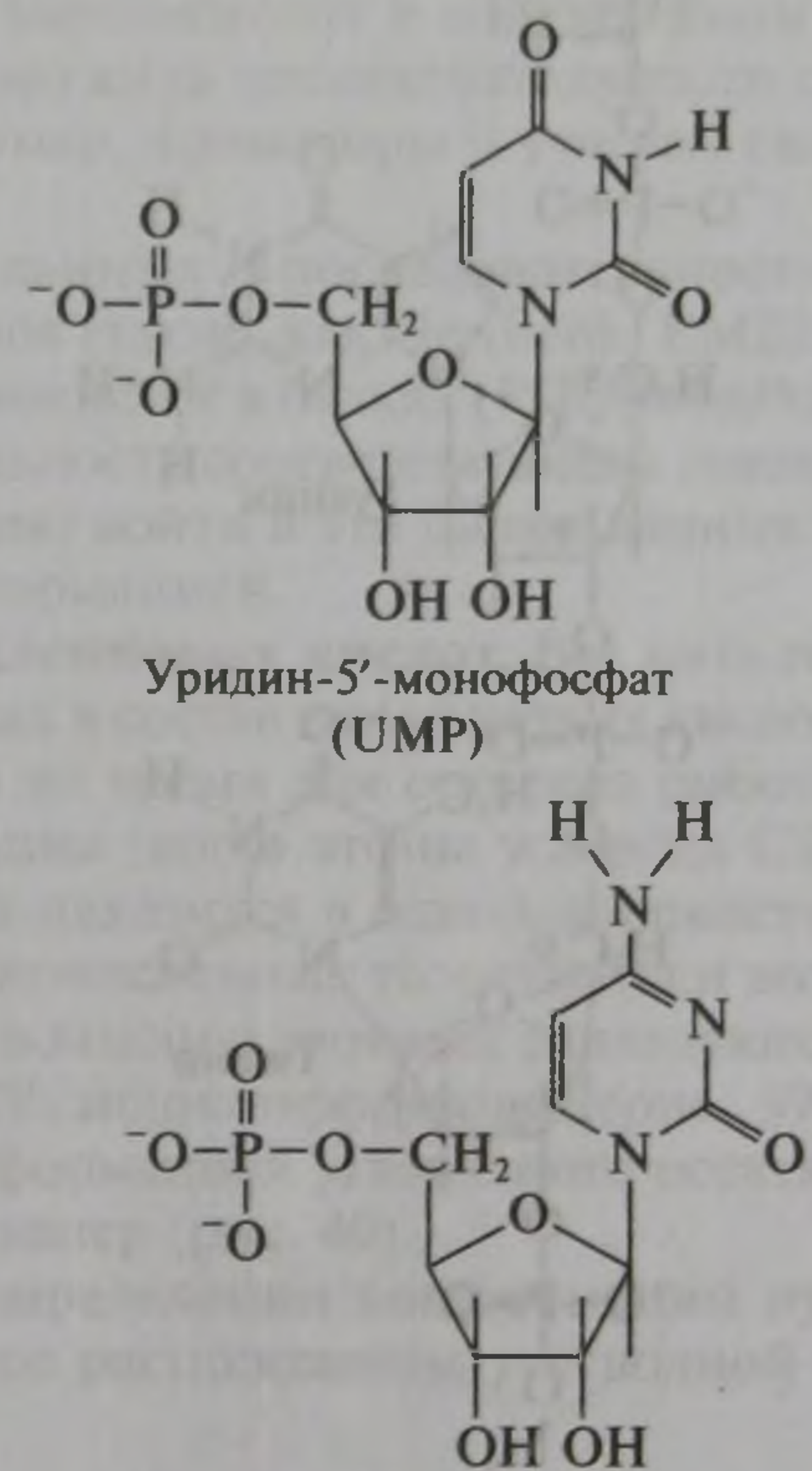
загадочными и не поддающимися расшифровке (рис. 38). Проблему расшифровки решают с помощью высокоэффективных компьютерных программ, которые пригодны для поиска как структурных, так и биологических особенностей этих последовательностей. С помощью программ можно вычислить нуклеотидный состав, выявить участки, богатые пуринами, пиримидинами или определенными сочетаниями оснований, определить частоту встречаемости различных динуклеотидов. Так, одной из характеристик первичной структуры ДНК является частота встречаемости в определенном соседстве любых двух оснований. Частота встречаемости 5′-CG-3′ и 5′-GC-3′ в ДНК прокариот почти одинакова и близка к случайной; то же самое можно сказать и о динуклеотидах 5′-GA-3′ и 5′-AG-3′. Однако в ДНК вирусов, животных и растений частоты встречаемости 5′-CG-3′ составляют от 1/2 до 1/5 частот 5′-GC-3′. Таким образом, последо-

B



OH OH

Аденозин-5′-монофосфат
(AMP)



OH OH

Уридин-5′-монофосфат
(UMP)

OH OH

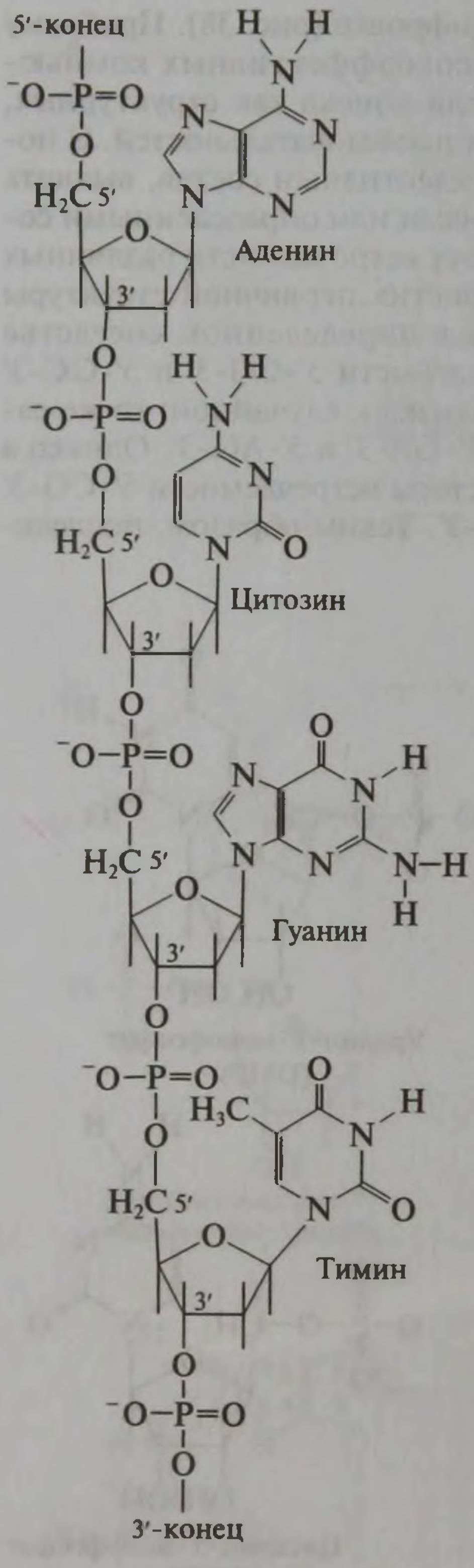
Гуанозин-5′-монофосфат
(GMP)

OH OH

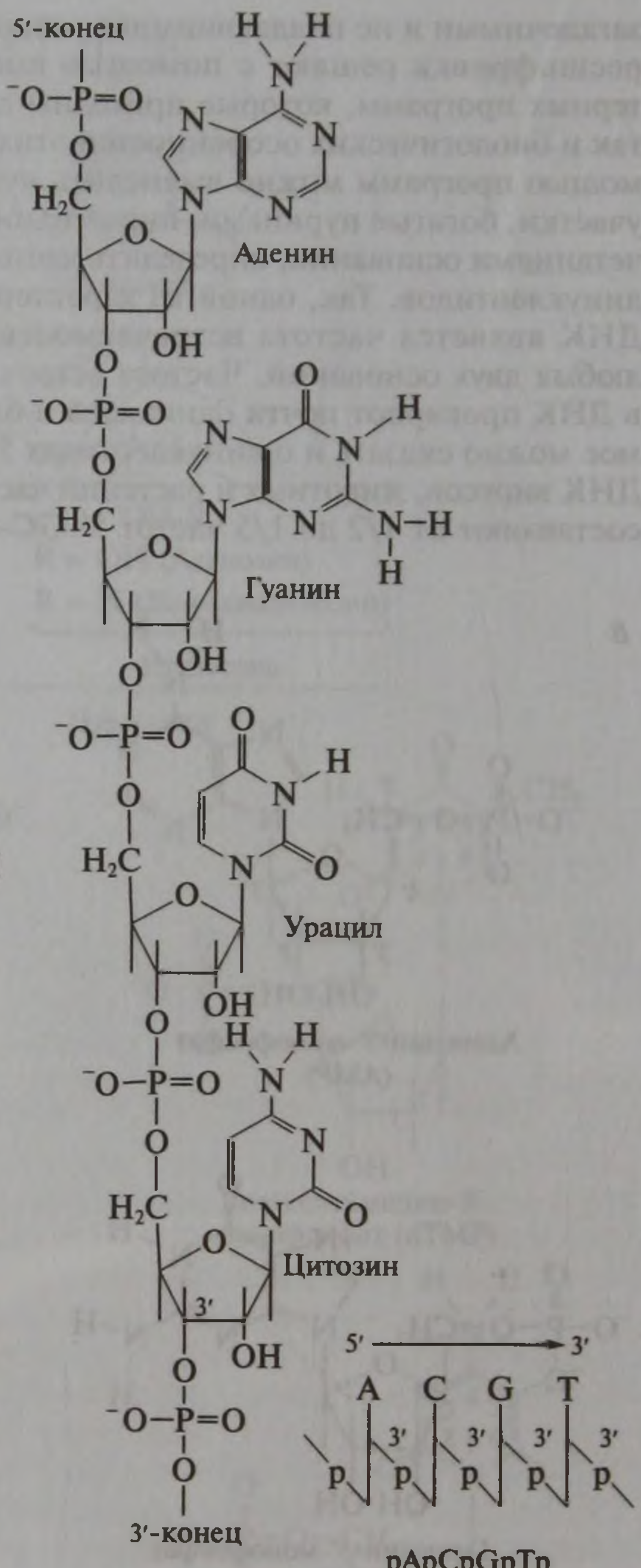
Цитидин-5′-монофосфат
(CMP)

Рис. 35. Структурные формулы нуклеотидов — компонентов нуклеиновых кислот:

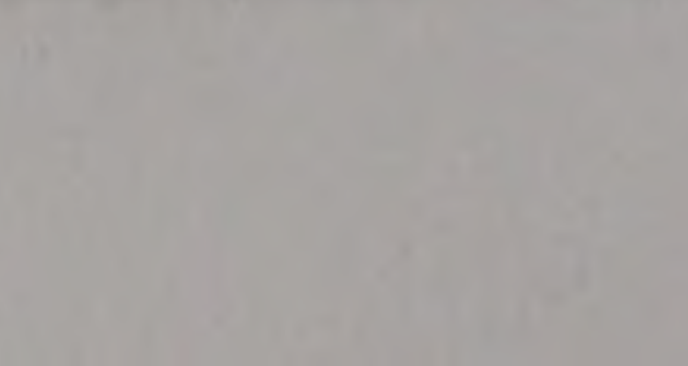
А —химическое строение нуклеотида; *Б* — дезоксирибонуклеотиды, встречающиеся в ДНК; *В* — рибонуклеотиды, встречающиеся в РНК


5′-конец
Аденин
Цитозин
Гуанин
Тимин
3′-конец

А


5′-конец
Аденин
Гуанин
Урацил
Цитозин
3′-конец

Б


5′
3′
A C G T
pApCpGpTp
ACGT

В

вательность 5′-CG-3′ встречается в ДНК высших эукариот довольно редко: это связано со способностью данного динуклеотида служить мишенью при метилировании и с его ролью в регуляции экспрессии генов. Соответствующие программы позволяют выявить специфические субпоследовательности в пределах определенного сегмента, что часто используется при нахождении сайтов для рестриктирующих эндонуклеаз. Имеющиеся программы осуществляют также поиск характерных особенностей последовательностей, включая прямые и обратные повторы. Можно получить данные о гомологии или частичной гомологии различных последовательностей. Это может касаться структуры одного и того же гена у двух разных организмов или родственных генов одного организма. Такой сравнительный анализ широко используется при изучении эволюции на молекулярном уровне. Разработаны программы, позволяющие предсказать стабильную вторичную структуру одноцепочечных РНК или ДНК. С помощью компьютерных программ можно перевести информацию с языка нуклеотидов на язык аминокислот в соответствии с правилами генетического кода; обнаружить последовательности со специфическими свойствами (например, промоторы и участки связывания с рибосомами).

Созданы международные банки данных о последовательностях нуклеотидов в ДНК разных организмов (такие, как GenBank, EMBL/DDBJ) и о последовательностях аминокислот в белках (PIR/SwissPot, на основе нуклеотидной последовательности соответствующих генов). Можно практически беспрепятственно войти в эти банки данных и воспользоваться собранной там информацией.

Конформации компонентов нуклеиновых кислот. Все пять гетероциклических оснований, входящих в состав нуклеиновых кислот, имеют плоскую конформацию. В то же время для остатков рибозы и дезоксирибозы плоская конформация (когда атомы углерода C1′, C2′, C3′, C4′ и гетероатом кислорода находятся в одной плоскости) энергетически невыгодна. Среди многочисленных теоретически возможных конформаций этих остатков в полинуклеотидах реализуются только две: либо C2′-эндо-, либо C3′-эндоконформации (рис. 39). Нуклеотидная единица с 3′-эндоконформацией углеводного остатка имеет меньшую длину, чем 2′-эндоизомер (рис. 40).

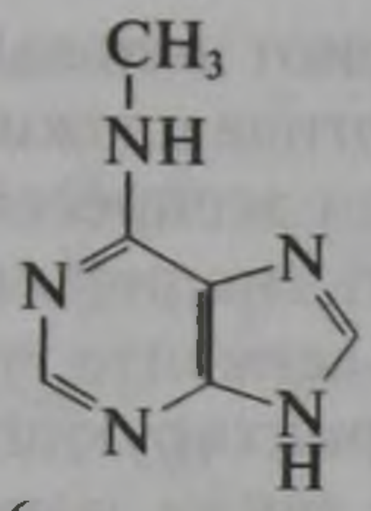
Важнейшей характеристикой в определении конформаций нуклеотидного звена является взаимное расположение углеводной и

←

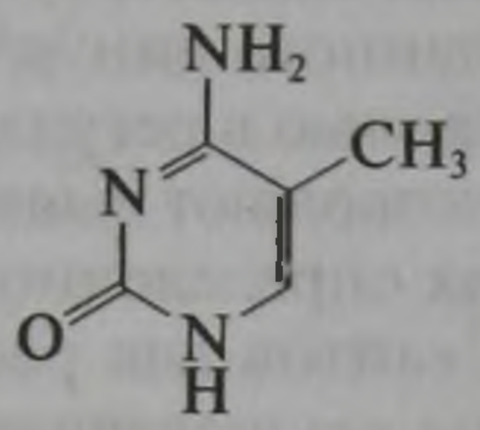
Рис. 36. Связь между нуклеотидами в полинуклеотидной цепи:

А — полидезоксирибонуклеотид; *Б* — полирибонуклеотид; *В* — способы схематического изображения нуклеотидной последовательности (нуклеотидную последовательность принято изображать слева направо, от 5′- к 3′-концу)

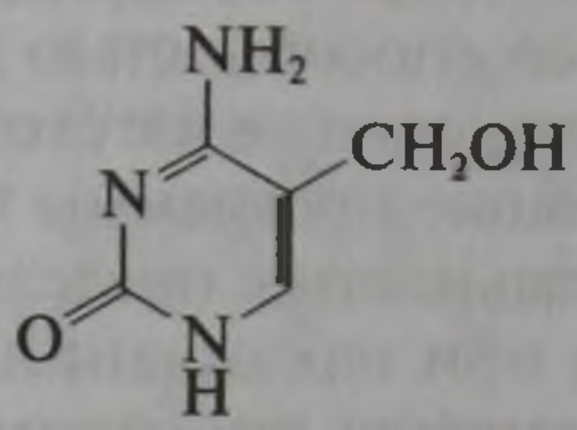
А



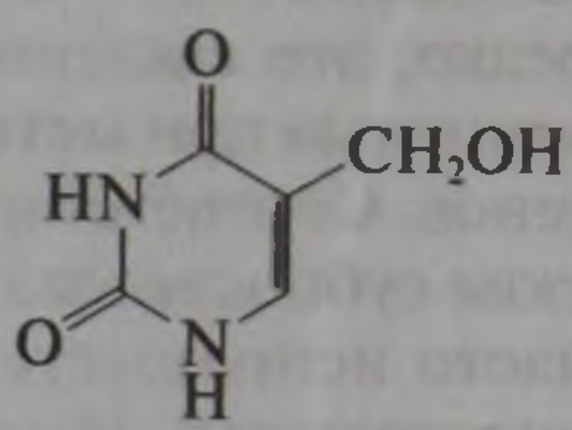
N^6-Метиладенин



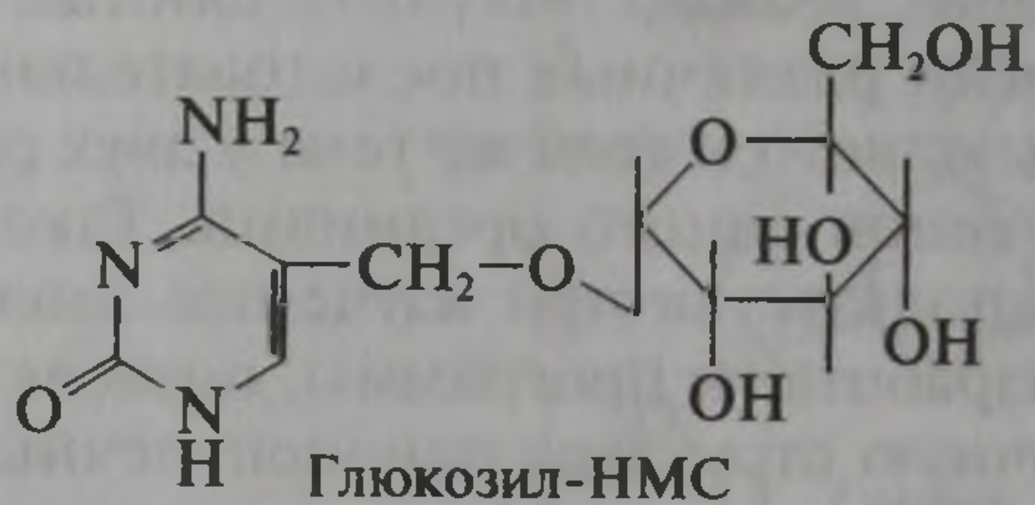
5-Метилцитозин



5-Гидроксиметил-цитозин (HMC)

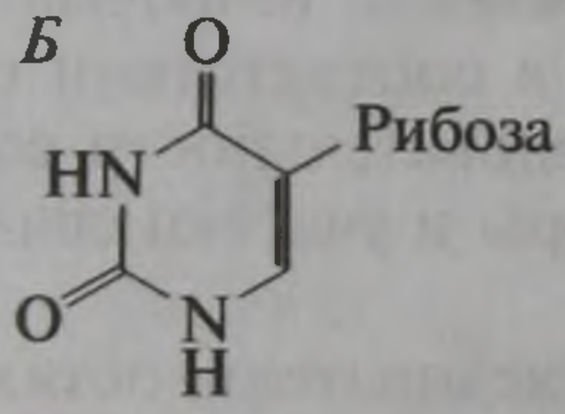


5′-Гидрокси-метилурацил

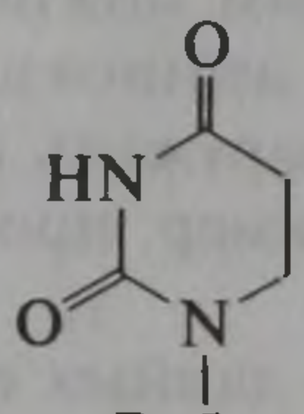


Глюкозил-HMC

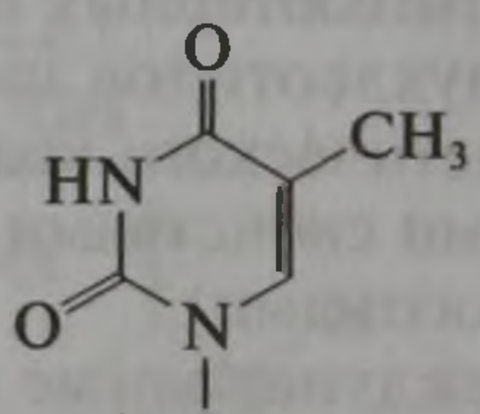
Б



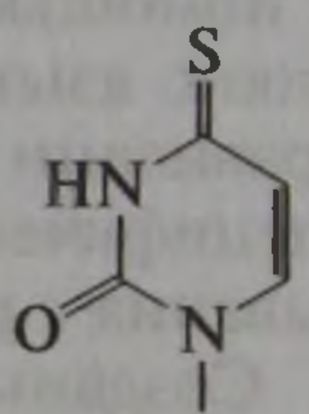
Псевдоуридин (ψ)



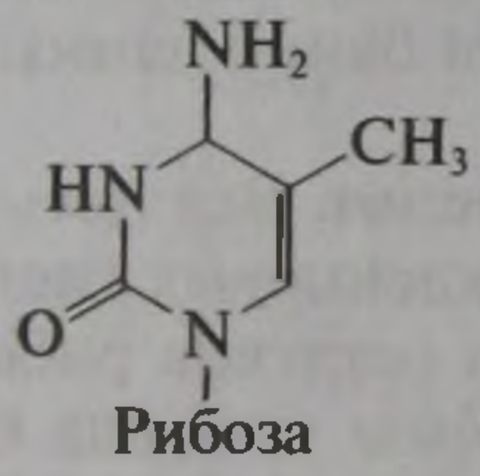
5,6-Дигидроуридин (DHU)



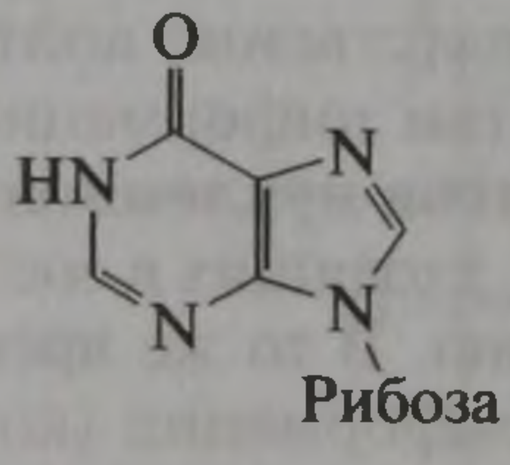
5-Метилуридин-(риботимидин) (rT)



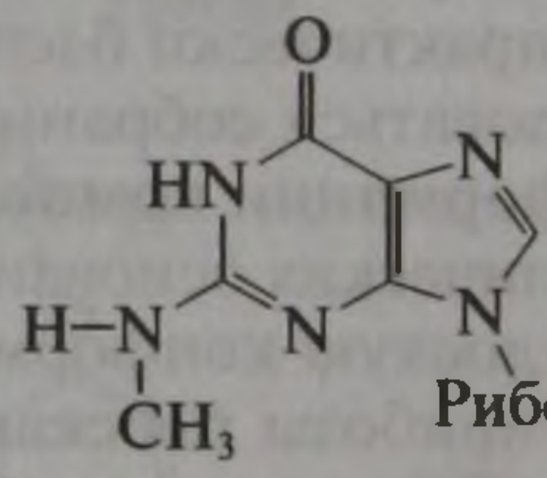
4-Тиоуридин (4-thio U)



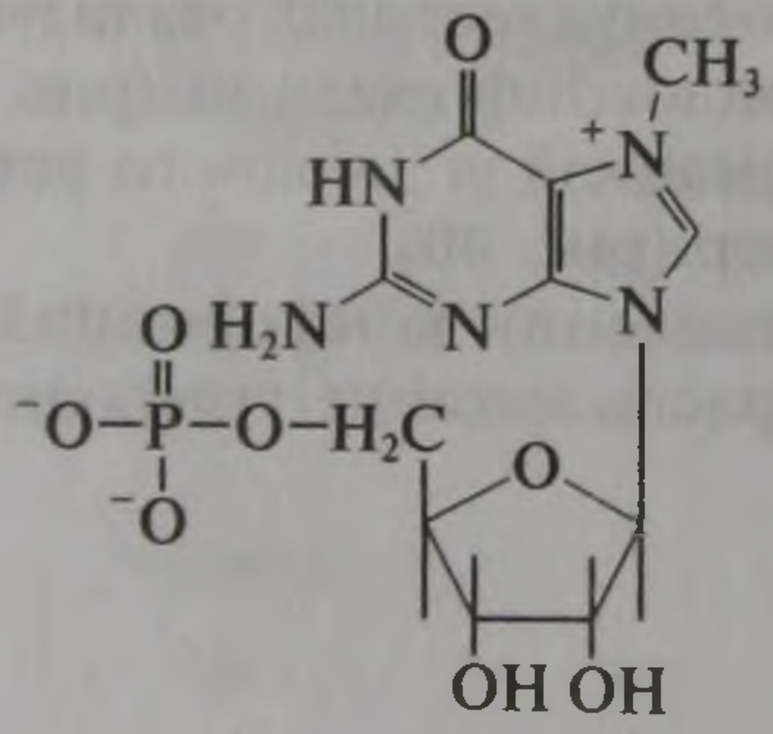
5-Метилцитидин (5-meC)



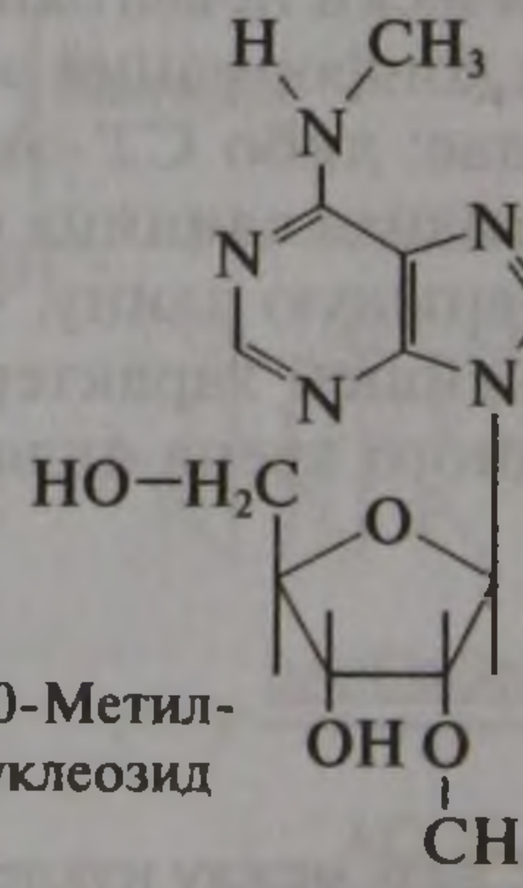
Инозин (I)



^{2}N-Метилгуанозин (^{2}meG)



^{7}N-Метилгуаниловая кислота



2′-0-Метил-нуклеозид

^{6}N-Метиладениловая кислота

Рис. 37. Структурные формулы минорных пуриновых и пиримидиновых оснований, обнаруженных в составе ДНК (*А*) и РНК (*Б*)

60

pCGGAGTTAGGGGCGGGATGGGCGGAGTTAGGGGCGGGACTATGGTTGCTGACTAATTGAG

HOGCCTCAATCCCCGCCCTACCCGCCTCAATCCCCGCCCTGATACCAACGACTGATTAACTC

120

ATGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGTTGCTGACTAATTGAGATGCA

TACGTACGAAACGTATGAAGACGGACGACCCCTCGGACCAACGACTGATTAACTCTACGT

180

TGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCCTAACTGACACA

ACGAAACGTATGAAGACGGACGACCCCTCGGACCCCTGAAAGGTGTGGGATTGACTGTGT

Рис. 38. Первичная структура ДНК вируса SV40.

Показана последовательность нуклеотидов с 1-го по 180-й н.п. (всего 5 226 н.п.)

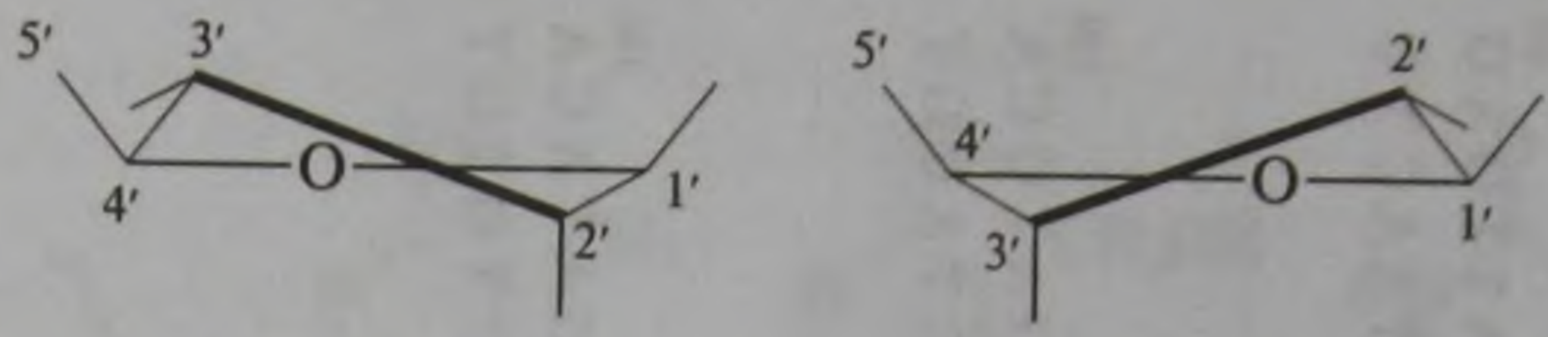


Рис. 39. Две конформации углеводных остатков, встречающиеся в ДНК и РНК.

C1′, 0 и C4′ расположены в одной плоскости, C2′ или C3′ находятся в эндоконформации, когда они выведены над этой плоскостью, т.е. в направлении связи C4′ — C5′

гетероциклической частей, которое определяется углом вращения вокруг N-гликозидной связи. Теоретический анализ конформации нуклеозидов и нуклеотидов и рентгеноструктурные исследования показали, что здесь существуют две области разрешенных конформаций, называемых *син-* и *анти*-конформациями (рис. 41).

В свободных нуклеозидах и нуклеотидах переход от C2′-эндо- к C3′-эндо- и между *син-* и *анти*-конформациями происходит достаточно легко. Если же нуклеотидный остаток встроен в полинуклеотидную цепь, то его конформация при данных внешних условиях оказывается фиксированной: в обычных формах полинуклеотидов мономерные звенья имеют исключительно *анти*-конформацию.

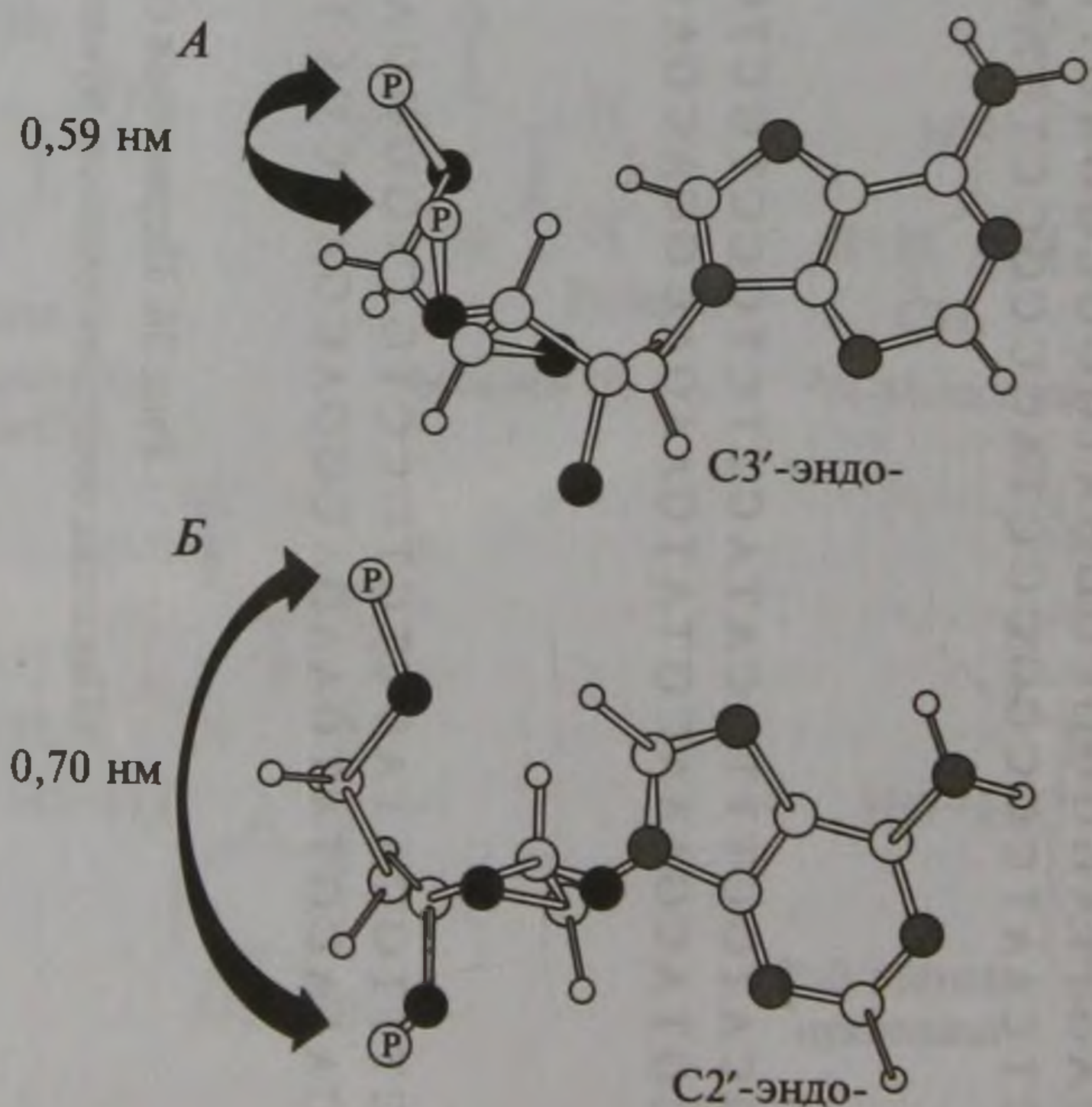


Рис. 40. Возможные конформации нуклеотидных единиц:

А — C3′-эндоконформация; *Б* — C2′-эндоконформация

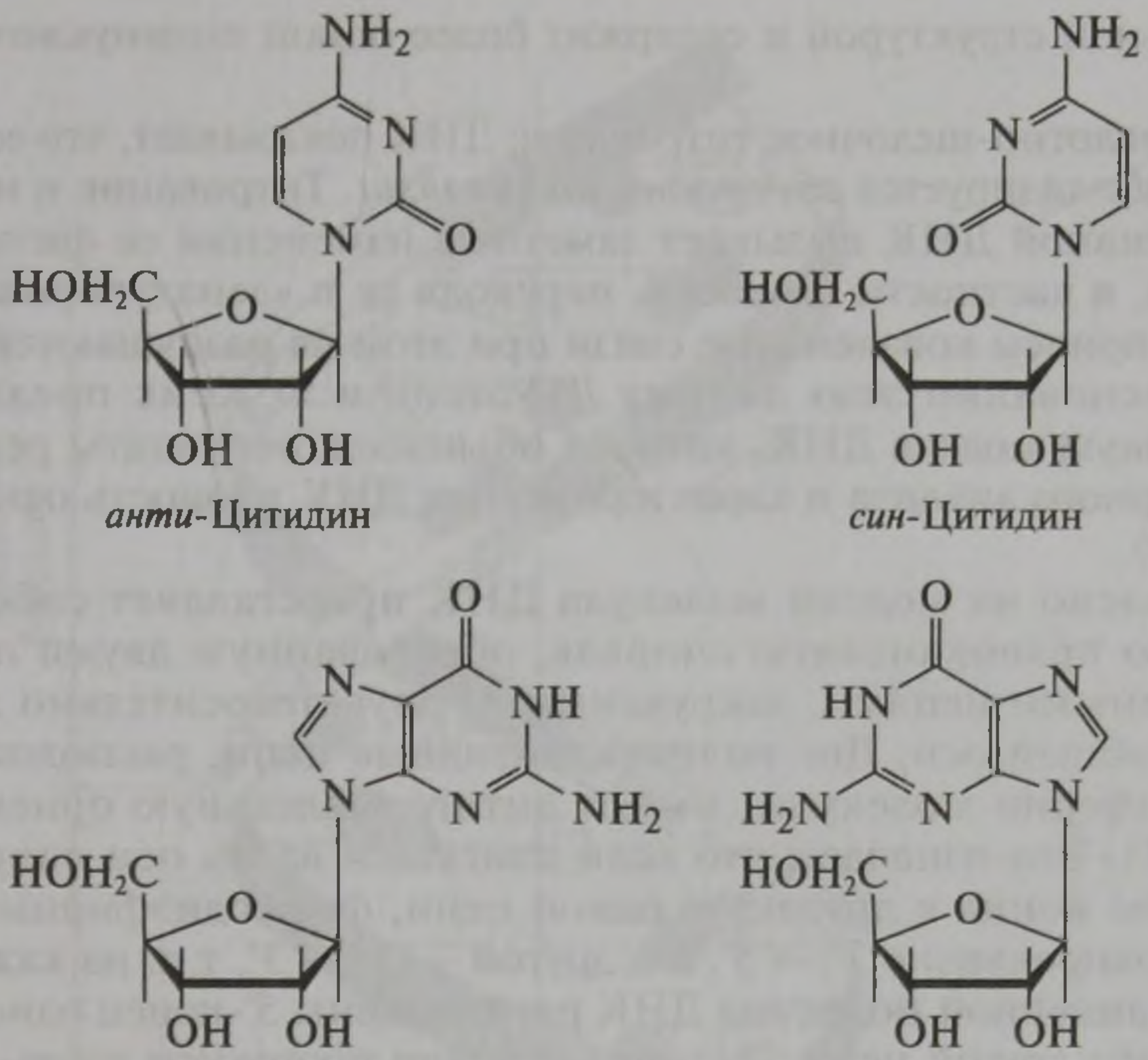


Рис. 41. *Анти-* и *син*-конформации в нуклеотидах

В полирибонуклеотидах углевод чаще всего находится в 3′-эндо-конформации, тогда как в полидезоксирибонуклеотидах он может быть как в 3′-эндо-, так и в 2′-эндоконформации. Это приводит к тому, что число конформаций, возможных для ДНК, больше, чем для РНК.

Макромолекулярная структура ДНК

Двойная спираль ДНК (модель Уотсона — Крика). В 1953 г. Дж. Уотсон и Ф. Крик предложили модель структуры ДНК. При построении модели ученые основывались на четырех группах данных:

1. ДНК представляет собой полимер, состоящий из нуклеотидов, соединенных 3′-5′-*фосфодиэфирными связями*.

2. Состав нуклеотидов ДНК подчиняется правилам Чаргаффа: в любой ДНК содержание пуриновых нуклеотидов (A+G) всегда равно содержанию пиримидиновых нуклеотидов (T+C); число остатков A всегда равно числу остатков T, число остатков G — числу остатков C.

3. Рентгенограммы волокон ДНК, впервые полученные М. Уилкинсом и Р. Франклин, указывают на то, что молекула обладает

спиральной структурой и содержит более одной полинуклеотидной цепи.

4. Кислотно-щелочное титрование ДНК показывает, что ее структура стабилизируется *водородными связями*. Титрование и нагревание нативной ДНК вызывает заметные изменения ее физических свойств, в частности вязкости, переводя ее в «денатурированную» форму, причем ковалентные связи при этом не разрушаются.

На основании этих данных Д. Уотсон и Ф. Крик предложили трехмерную модель ДНК, которая объясняла результаты рентгеноструктурного анализа и характерную для ДНК парность оснований (рис. 42).

Согласно их модели молекула ДНК представляет собой правильную правовинтовую спираль, образованную двумя полинуклеотидными цепями, закрученными друг относительно друга и вокруг общей оси. Две полинуклеотидные цепи, расположенные по периферии молекулы, имеют антипараллельную ориентацию (рис. 43). Это означает, что если двигаться вдоль оси спирали от одного ее конца к другому в одной цепи, фосфодиэфирные связи имеют направление 3′ → 5′, а в другой — 5′ → 3′, т.е. на каждом из концов линейной молекулы ДНК расположены 5′-конец одной цепи и 3′-конец другой цепи. Диаметр спирали постоянен вдоль всей ее длины и равен 2,0 нм.

Гидрофильные пентозофосфатные остовы цепей расположены на внешней стороне двойной спирали. Гидрофобные пуриновые и пиримидиновые основания обеих цепей уложены стопкой с интервалом 0,34 нм и направлены внутрь спирали; плоскости колец гетероциклических оснований перпендикулярны главной оси спирали. Длина витка спирали (полный оборот спирали), который соответствует ее периоду идентичности, составляет 3,40 нм. На один виток спирали приходится 10 нуклеотидных остатков в одной цепи.

Двойная спираль стабилизируется с помощью водородных связей между пуринами одной цепи ДНК и пиримидинами другой, а именно, между А и Т, G и С. Основания, образующие пары, в которых они сочетаются водородными связями, получили название *комплементарных пар*.

В образовании водородных связей участвуют аминогруппы А, Ц, Г и NH-группы гетероциклов Т и Г (группы — доноры). Группами-акцепторами являются карбонильные группы Т, Ц, Г и атомы азота гетероциклов Т и Г. Пара А — Т может образовывать два, а пара Г — Ц три линейных и поэтому особенно устойчивых «мостика» (водородные связи). Последовательность оснований в одной цепи определяет их последовательность в другой. *Комплементарность последовательности оснований в двух полинуклеотидных цепях — ключевое свойство ДНК.*

Позднее было установлено, что модель Д. Уотсона и Ф. Крика описывает структуру одной из нескольких форм двойной спирали,

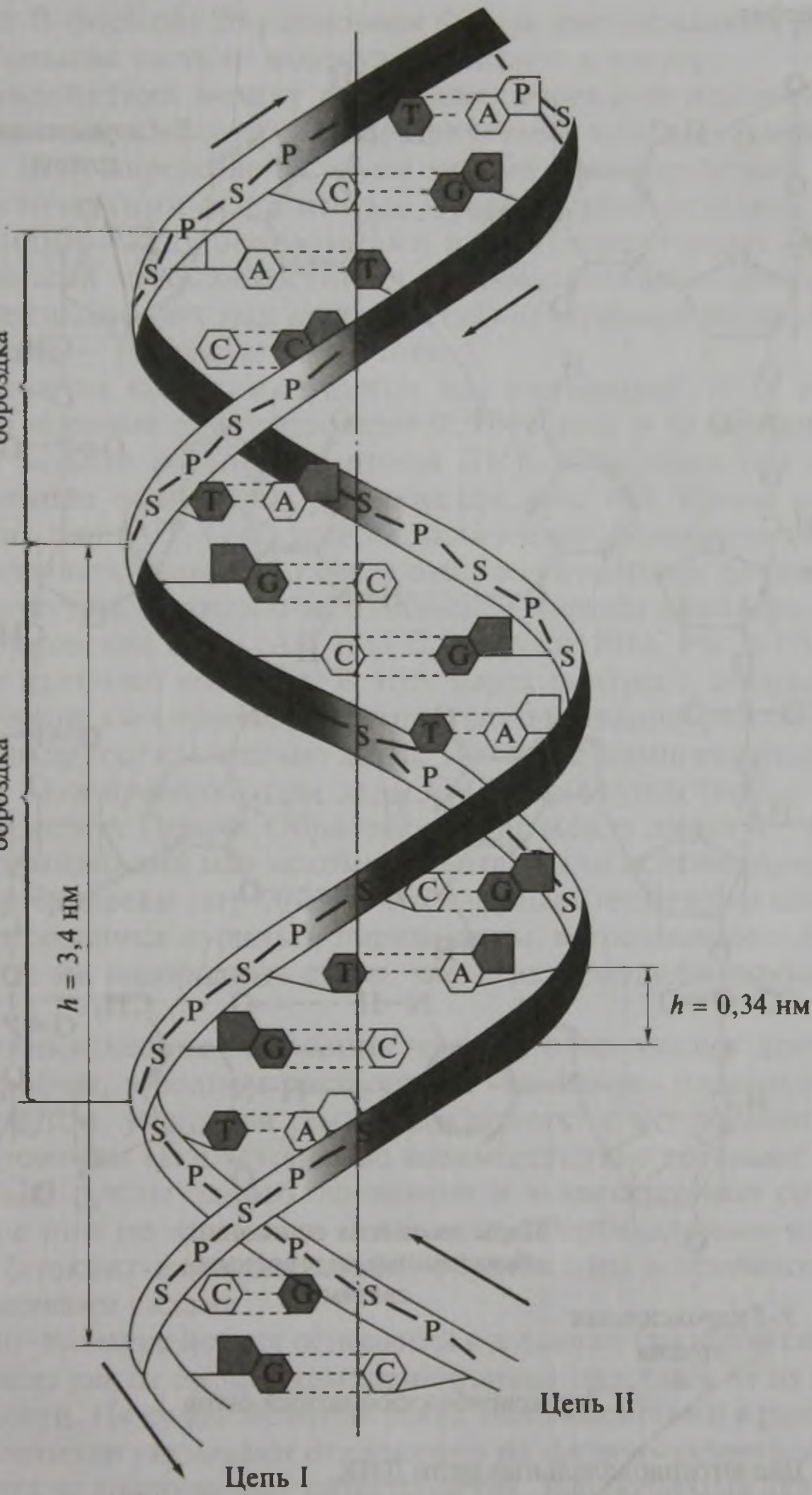


Рис. 42. B-форма двойной спирали ДНК:
P — фосфат; S — дезоксирибоза

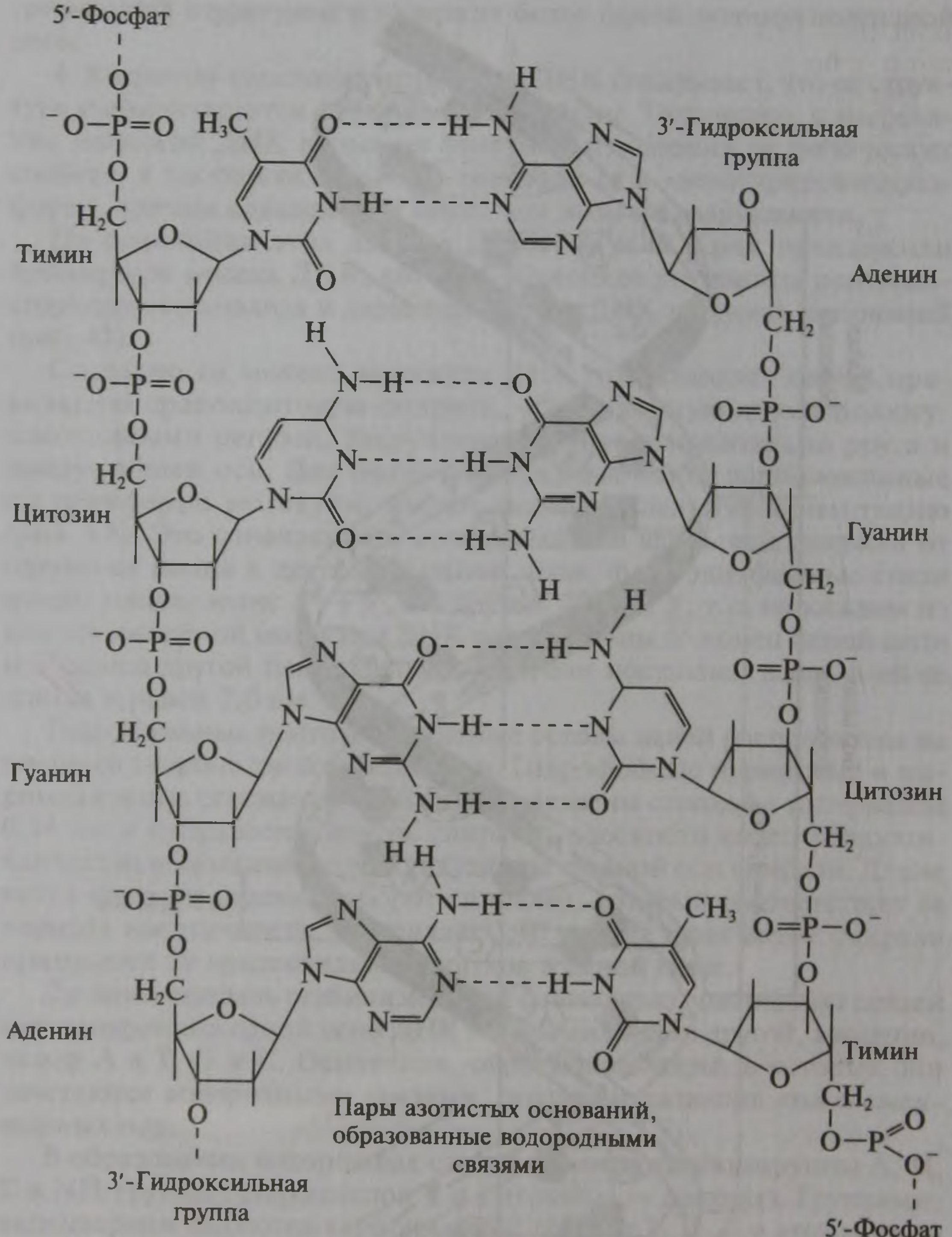


Рис. 43. Две антипараллельные цепи ДНК.

Сахарофосфатный остов расположен снаружи и заряжен отрицательно. Основания повернуты внутрь спирали и располагаются стопкой, друг над другом. Левая цепь сверху вниз имеет направление 5′→3′, а правая — 3′→5′. Закручивание двух цепей в спираль не показано

названной В-формой. Это основная форма двуспиральной ДНК, в которой большая часть ее молекул существует в клетке.

Взаимодействия между гетероциклическими основаниями в нуклеиновых кислотах. Специфическая макромолекулярная структура ДНК определяется двумя типами взаимодействий между гетероциклическими основаниями нуклеотидных остатков: взаимодействиями между основаниями в комплементарных парах и вертикальными межплоскостными взаимодействиями оснований, расположенными друг над другом (стэкинг-взаимодействиями, от англ. *stacking* — укладываться в стопку).

Образование комплементарных пар оснований (A·T- и G·C-пар) было впервые постулировано Д. Уотсоном и Ф. Криком при создании модели двойной спирали ДНК. Обе пары близки по форме и имеют одинаковые размеры (см. рис. 42). Кроме уотсон-криковских пар (A·T, G·C) гетероциклические основания способны образовывать множество связанных водородными связями пар другой структуры. Квантово-механические расчеты показывают, что уотсон-криковские пары (A·T и G·C, в случае РНК это A·U) энергетически наиболее выгодны. В этих парах центры с повышенной и пониженной *электронной* плотностью оснований расположены оптимально друг относительно друга. Такой вид комплементарности называют *электронной*, при этом G·C-пары существенно более стабильны, чем A·T-пары. Образование пар между двумя пуринами, двумя пиримидинами или некомплементарными основаниями (A·C или G·T) стерически затруднено, что нарушает геометрию спирали. Модифицированные пурины и пиримидины, встречающиеся в ДНК, образуют те же водородные связи, что и их немодифицированные аналоги.

Гетероциклические основания нуклеиновых кислот достаточно гидрофобны, в водном растворе им «выгоднее» расположиться друг над другом, уменьшая тем самым контакт с молекулами воды. При образовании таких стопок во взаимодействие вступают функциональные группы одного основания и π-электронные системы соседнего с ним по вертикали основания. «Вертикальные» взаимодействия (стэкинг-взаимодействия) обусловлены в основном ван-дер-ваальсовыми силами.

Стэкинг-взаимодействия оснований в дуплексе (двойной спирали ДНК) зависят как от состава комплементарных пар, так и от их последовательности. На существование таких взаимодействий в растворах полинуклеотидов указывают отклонения их физико-химических характеристик от аналогичных характеристик, наблюдаемых для смеси мономерных компонентов, состав которых идентичен нуклеотидному составу полинуклеотидов. Так, коэффициент молярной экстинкции (поглощение в УФ-свете) полинуклеотидов меньше, чем сумма коэффициентов молярной экстинкции входящих в него мономеров (*гипохромный эффект*).

Полиморфизм двойной спирали

Правые спирали образуют два семейства: А-семейство (конформация сахара С3′-эндо-) и В-семейство (конформация сахара С2′-эндо-). Структуры в пределах каждого из семейств в зависимости от условий (концентрации соли, температуры) могут иметь разное число пар, приходящихся на виток спирали, разный наклон пар оснований к оси спирали и т. д.

А-семейство ДНК. Еще до открытия двойной спирали Р. Франклин получила экспериментальные свидетельства существования весьма упорядоченной структуры в ориентированных вытягиванием и подсушенных волокнах ДНК. Эта структура получила название А-форма ДНК. Этой форме долго не придавали особого значения, так как она возникала при малой влажности, т. е. не при физиологических условиях.

С3′-эндоконформация сахара приводит к уменьшению расстояния между фосфатными группами и, следовательно, к уменьшению расстояния между нуклеотидными парами вдоль оси спирали. Это соответственно ведет к увеличению количества нуклеотидов на виток спирали (11 нуклеотидных остатков). При этом пары оснований в А-форме образуют с осью спирали угол около 20° и очень сильно отодвинуты от оси спирали к периферии молекулы: сдвиг достигает 0,4—0,5 нм, т. е. почти половину радиуса (рис. 44). Вследствие этого

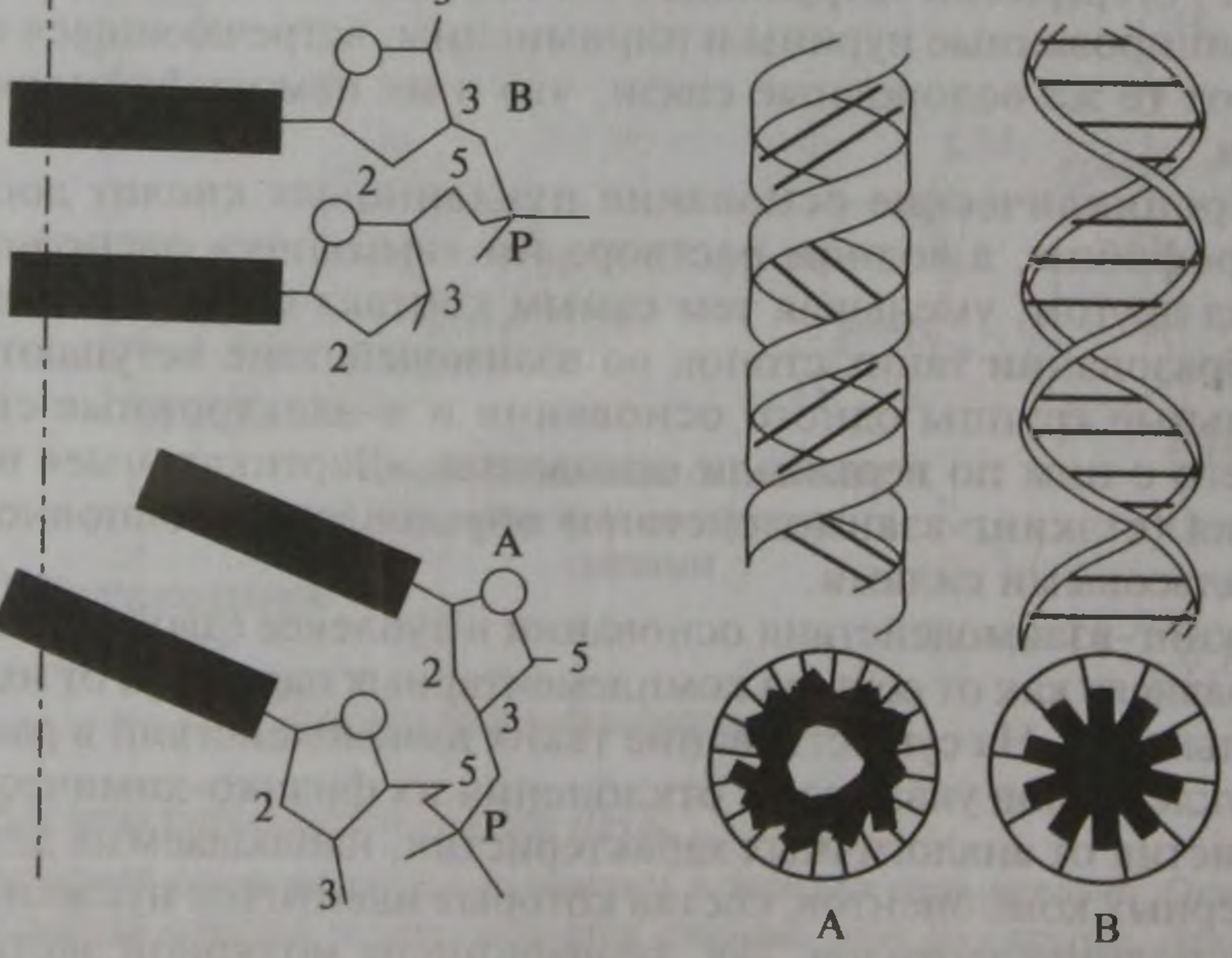


Рис. 44. Пространственная структура соседних звеньев полинуклеотидной цепи в А- и В-формах ДНК и двойной спирали ДНК.

Вертикальная линия слева — ось спирали; ■ — азотистые основания

А-ДНК при взгляде сверху (вдоль оси спирали) выглядит как труба. В А-ДНК в стэкинге участвуют основания, принадлежащие разным цепям, т. е. имеет место как одно-, так и двуцепочечный стэкинг. Это обусловлено величиной угла наклона плоскостей оснований к оси спирали (положительная величина) и углом поворота оснований вокруг оси спирали. Для сравнения в В-ДНК только одноцепочечный стэкинг.

РНК, как правило, представляет собой одну полинуклеотидную цепь, но, навиваясь сама на себя, она может образовать короткий участок двойной спирали. Оказалось, что двойная спираль из цепей РНК имеет структуры А-формы при 100 %-й влажности (в этих условиях ДНК всегда существует в В-форме). Это открытие поставило А-конформацию в ряд биологически значимых. Сейчас уже получены подтверждения биологической роли А-ДНК.

В неблагоприятных условиях некоторые бактерии превращаются в споры и находятся в таком неактивном состоянии сколь угодно долго, пока не появятся подходящие условия для размножения. Их ДНК окружена споровыми белками и находится в А-форме. Именно в этом состоянии ДНК примерно в 10 раз более устойчива к действию ультрафиолетового излучения, чем та же ДНК в В-форме в активно размножающихся бактериях. В случае со спорами бактерий споровые белки способствуют превращению ДНК бактерий из В-формы в А-форму, чтобы защитить ДНК в неблагоприятных условиях.

В этом же плане большой интерес представляют гибридные спирали ДНК-РНК. Они образуются при транскрипции последовательности ДНК в комплементарную ей последовательность РНК, при этом примерно 40 н. п. в области активного центра РНК-полимеразы представляют такой гибрид. Гибридная спираль ДНК-РНК формируется и тогда, когда обратная транскриптаза осуществляет синтез ДНК на РНК-матрице, а также в ходе репликации при образовании коротких молекул РНК-праймеров. Отмечено, что обе цепи полинуклеотидного комплекса — цепь ДНК и цепь РНК — построены из нуклеотидов с 3′-эндоконформацией сахара. Структура таких гибридных спиралей (гетеродуплексов) напоминает А-ДНК (см. рис. 44), при этом В-ДНК должна перейти в А-форму до образования гибрида с РНК.

В-семейство ДНК. Для ДНК В-семейства характерно структурное разнообразие. ДНК со случайными последовательностями могут находиться в В-, С-, D- и других конформационных состояниях. На структуру ДНК влияют тип и концентрация катионов, а также температура. Характеристика отмеченных конформационных состояний представлена в табл. 3

Между этими различными формами осуществляются взаимные переходы, что обуславливается степенью обводненности биомолекулы и различиями в ионной силе окружающей среды. Считают, что в биологическом плане А-форма ДНК участвует в транскрипции и передаче информации от ДНК к РНК (см. гибридные молекулы

Таблица 3. **Характеристика некоторых конформационных состояний ДНК**

Показатель	A-форма	B-форма	C-форма	D-форма	Z-форма
Число пар нуклеотидных остатков на виток	11	10	9,3	8,0	12
Угол наклона плоскостей оснований к оси спирали, град	20	–2	–6	–6	–7
Угол поворота оснований вокруг оси спирали, град	32,7	36	38,6	45,0	–30
Расстояние комплементарных пар от оси спирали, нм	0,425	0,063	0,213	0,143	—
Расстояние между нуклеотидными остатками по высоте спирали, нм	0,256	0,338	0,332	0,304	0,340
Угол между плоскостями комплементраных оснований, град	8	5	5	—	—
Конформация дезоксирибозы	C3′-эндо	C2′-эндо	C2′-эндо	C2′-эндо	C2′-эндо-C3′-эндо

Примечание. A-форма получена из водно-солевых растворов, содержащих ионы щелочных металлов, кроме Li^+; рентгеноструктурный анализ — при относительной влажности 75 %. B-форма получена из водно-солевых растворов, содержащих Li^+; рентгеноструктурный анализ — при относительной влажности 66 %. C-форма получена в тех же условиях, что B-форма, но в растворах с иным соотношением катионов; рентгеноструктурный анализ — при относительной влажности менее 66 %. D-форма выделена из фага T2; содержит остатки глюкозилированного оксиметилцитозина; рентгеноструктурный анализ — при относительной влажности 60 %. Z-форма — монокристаллы поли-(dG—dC) в слабом солевом растворе.

ДНК-РНК); а B-форма ДНК — в репликативных процессах, в умножении количества информации. C-форма наиболее подходяща для упаковки ДНК в составе надмолекулярных структур хроматина, т. е. в хранении информации (рис. 45).

Z-форма ДНК. Левая спираль, получившая название Z-формы ДНК, была обнаружена у полинуклеотида с чередующейся последовательностью (dG—dC). В растворе с низкой ионной силой этот

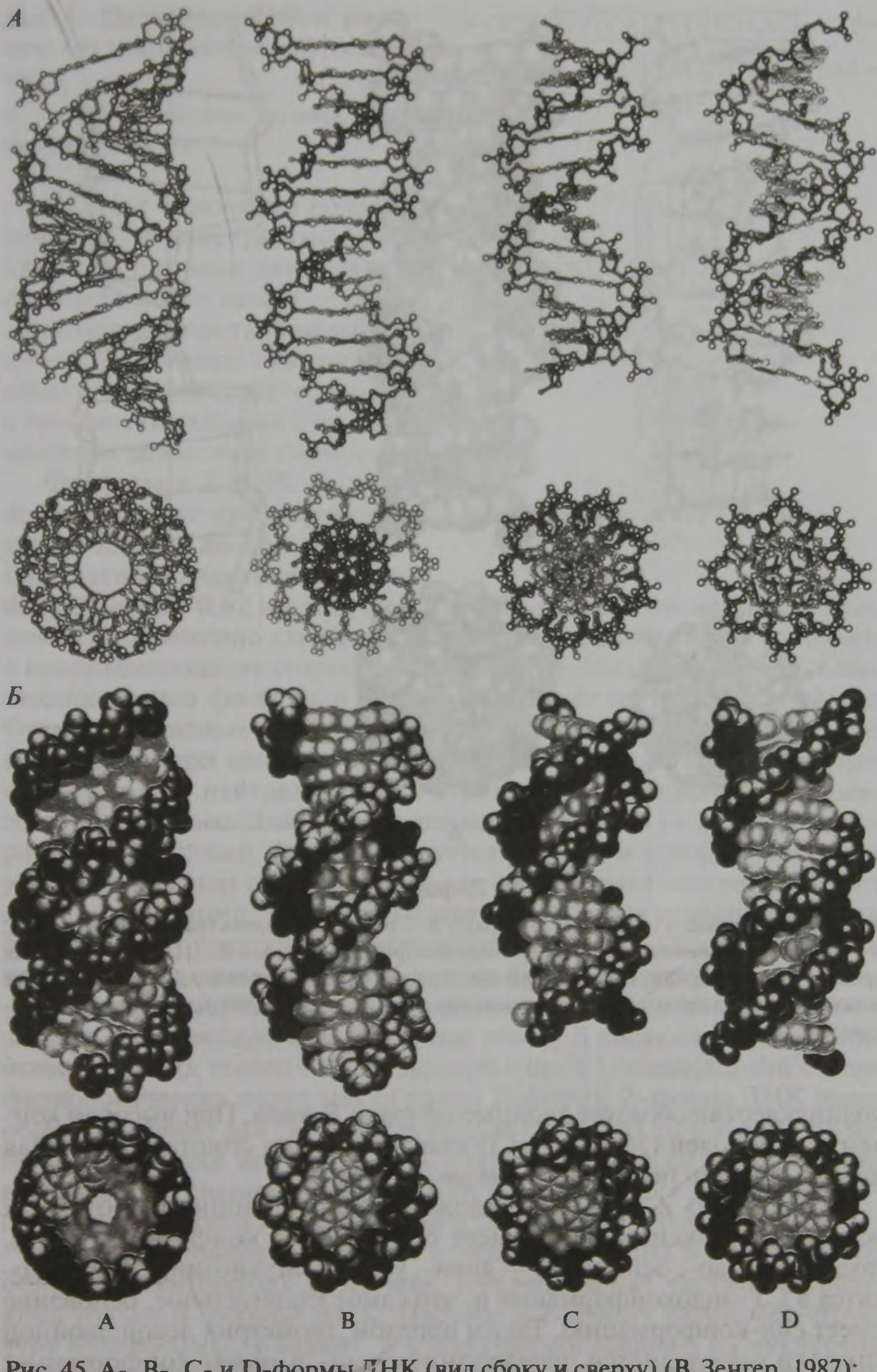
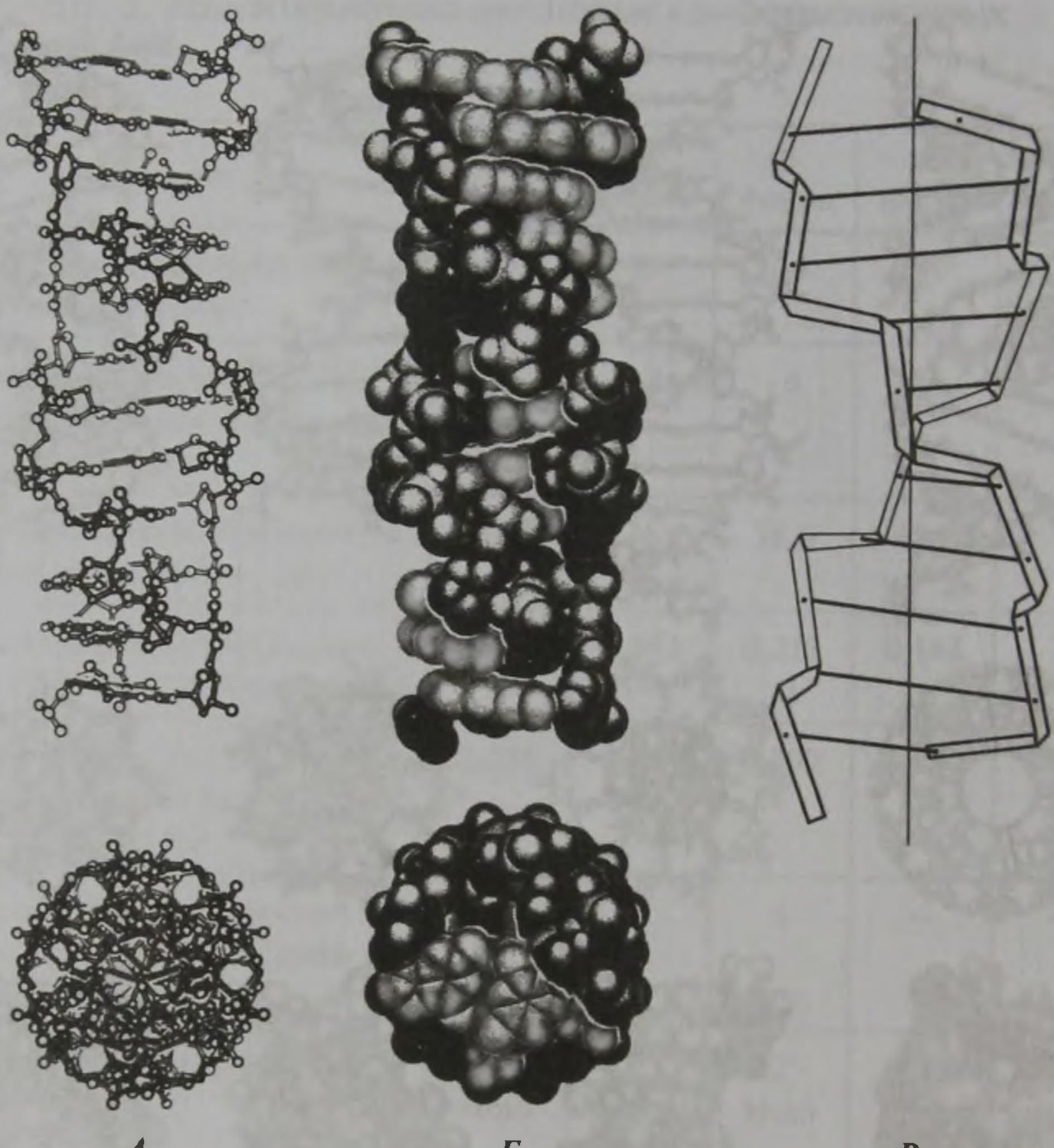


Рис. 45. A-, B-, C- и D-формы ДНК (вид сбоку и сверху) (В. Зенгер, 1987):
А — скелетная модель; *Б* — объемная модель



А *Б* *В*

Рис. 46. Z-форма ДНК:

А — скелетная модель (вид сбоку и сверху); *Б* — объемная модель (вид сбоку и сверху); *В* — зигзагообразная линия, соединяющая фосфатные группы в Z-ДНК. Вертикальная прямая — ось спирали, горизонтальные отрезки — пары оснований и сахара. После остатков гуанина зигзагообразная линия идет вертикально, после остатков цитозина — горизонтально

полинуклеотид образует двойные спирали B-типа. При высокой концентрации солей ($MgCl_2$, NaCl) или добавлении спирта эта двойная правая спираль переходит в *левую* Z-форму.

Особенность Z-ДНК — чередование конформаций нуклеотидных остатков: дезоксицитидин имеет стандартную конформацию, т. е. сахар C2′-эндо-, основание — анти-, у дезоксигуанозина сахар находится в C3′-эндоконформации и, что самое удивительное, основание имеет *син*-конформацию. Таким образом, геометрия левой двойной спирали определяется чередованием *син*- и *анти*-конформаций нуклеотидов.

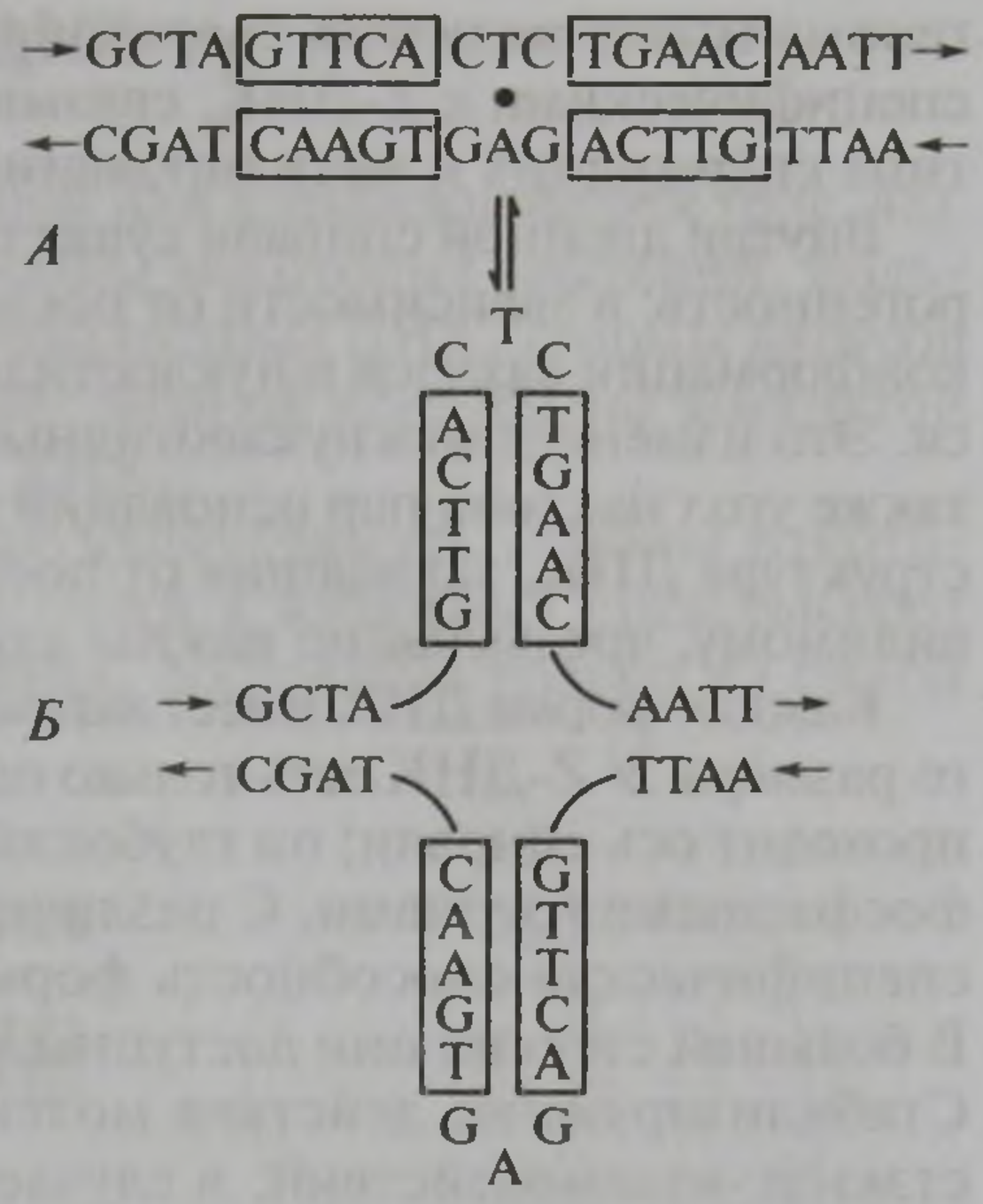


Рис. 47. Палиндром ДНК в линейном (*А*) и шпилечном (*Б*) состояниях:

• — ось симметрии палиндромных последовательностей

Стэкинг оснований обладает новыми, характерными лишь для этой спирали свойствами: он имеет место между остатками цитозина противоположных цепей, а остатки гуанина вообще не взаимодействуют друг с другом, контактируя с атомами соседних дезоксицитидинов.

Фосфаты в Z-ДНК не эквивалентны друг другу: расстояние фосфата до оси спирали (фосфатный радиус спирали) у d(CpG) равно 0,62 нм и d(GpC) — 0,76 нм, т.е. он меняется, принимая попеременно два разных значения; соседние сахара «смотрят» в противоположные стороны, отчего линия, последовательно соединяющая атомы фосфора в цепи, приобретает зигзагообразный вид. Отсюда и название Z-ДНК (рис. 46). B-форма и Z-форма переходят друг в друга при изменении ионной силы раствора или концентрации катионов, нейтрализующих отрицательный заряд на фосфодиэфирном каркасс. Для осуществления перехода B → Z не требуется расхождения цепей. Он инициируется разрывом водородных связей у нескольких пар оснований, после чего гуанин «закрепляется» в *син*-конформации, водородные связи восстанавливаются и основания вновь образуют уотсон-криковские пары. Область перехода перемещается вдоль спирали в виде небольшой петли.

Наличие в ДНК эукариот последовательности (dm^5C-dG) связывают с регуляцией транскрипции генов. В свою очередь, именно полинуклеотид такого состава поли(dG-dm^5C) · поли(dG-dm^5C) при физиологических значениях образует Z-форму. Z-форма ДНК имеет особое значение для сверхспирализации; переход B → Z даже на небольшом участке молекулы ДНК оказывает очень сильное влияние на топологию сверхспиральной ДНК (увеличивается кручение *Tw* — с. 97), и это используется клеткой в процессе экспрессии генов.

Z-ДНК с помощью антител к ней обнаружена в междисковых областях политенных хромосом. Нуклеосомные частицы образуются только B-формой ДНК. Переход в Z-форму разрушает структуру нуклеосомы, а значит, и структуру состоящего из нуклеосом хроматина. Возможно, Z-форма выполняет какую-то регуляторную роль, тем более что переход B → Z обратим. Регуляция экспрессии

генома может включать сверхспирализацию, связывание с белками, специфическими к Z-ДНК, связывание с определенными катионами типа спермидина и метилирование дезоксицитидина.

Внутри двойной спирали существует конформационная микрогетерогенность: в зависимости от последовательности нуклеотидных пар конформации сахаров в нуклеотидных остатках несколько отличаются. Это изменяет межнуклеотидные расстояния вдоль оси спирали, а также угол наклона пар оснований к этой оси. Различия во вторичной структуре ДНК, зависящие от последовательности нуклеотидов, по-видимому, чрезвычайно важны для ее функционирования.

Каждая форма ДНК имеет *малый* и *большой* желобки определенного размера. У Z-ДНК есть только один малый желобок, через который проходит ось спирали; он глубокий и узкий, с двух сторон ограничен фосфатными группами. С различными размерами желобков связана специфическая способность форм ДНК к комплексообразованию. В большей степени они доступны молекулам воды и ионам металлов. Стабилизирующее действие молекул воды направлено на усиление стэкинг-взаимодействий; в случае ослабления последних молекулы воды конкурентно взаимодействуют с протон-донорными и протон-акцепторными центрами оснований, дестабилизируя двойную спираль. Гидратация ДНК играет важную роль в превращении А-формы в В-форму ДНК, и наоборот. Катионы связываются в основном с пентозофосфатным остовом, располагаясь преимущественно в малом желобке двойной спирали. Все это подчеркивает динамичность структуры ДНК и возможность конформационных переходов в ней под действием агентов окружающей среды.

Биспиральные структуры в молекулах ДНК могут возникать и в пределах одной полидезоксирибонуклеотидной цепи. Для этого в последовательности ДНК должны присутствовать инвертированные, или «*палиндромные*», повторяющиеся последовательности длиной от нескольких до многих тысяч пар оснований.

Палиндром («обратно бегущие» последовательности нуклеотидных звеньев; от греч. «палин» — обратно, «дроме» — бегу) представляет собой двуцепочечную последовательность ДНК, которая одинаково читается в обоих цепях, если каждую цепь читать в соответствии с ее ориентацией.

В качестве примера можно привести последовательность $\begin{matrix}5'GGT:ACC3'\\3'CCA:TGG5'\end{matrix}$.
Осью симметрии является центральная точка, относительно которой последовательность остается одинаковой при чтении в противоположных направлениях).

Благодаря наличию собственной симметрии второго порядка палиндромы могут образовывать шпилечные структуры простой крестообразной формы (рис. 47). Такие последовательности способствуют специфическим нуклеиново-белковым взаимодействиям и, возможно, участвуют в процессах регуляции.

Разнообразие форм ДНК

ДНК может находиться в линейной или кольцевой форме (рис. 48). Бактериальные плазмиды, хромосомы некоторых бактерий, большинство митохондриальных и хлоропластных ДНК, геномы вирусов млекопитающих представлены единственной ковалентно замкнутой кольцевой дуплексной молекулой ДНК.

Не все ДНК in vivo являются двуцепочечными. Геномы некоторых мелких вирусов бактерий, растений и животных представляют собой ковалентно замкнутые кольца, состоящие только из одной

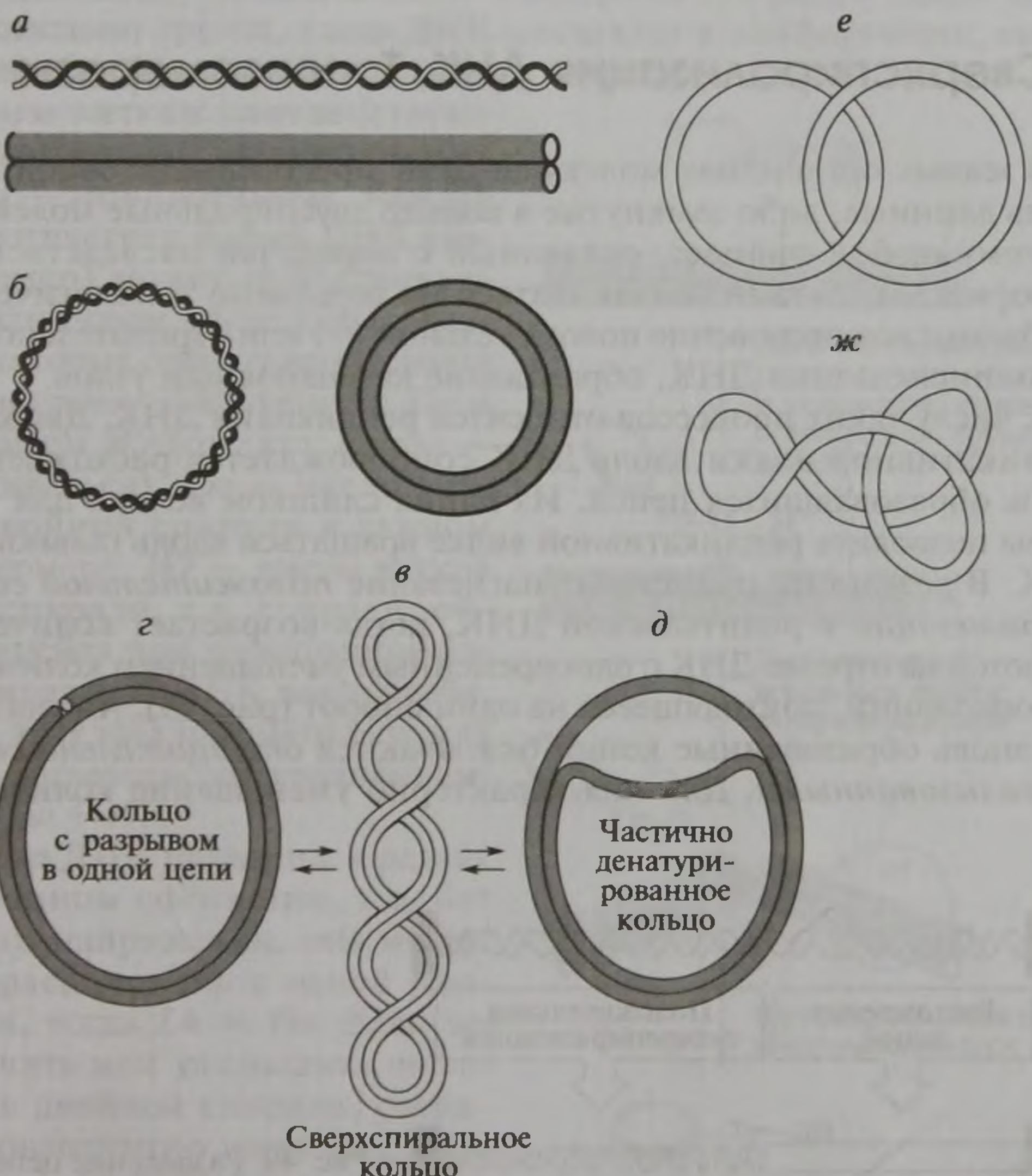


Рис. 48. Схематическое изображение ДНК — линейной (*а*), кольцевой двуцепочечной (*б*) и сверхспиральной кольцевой (*в*); релаксированных кольцевых форм, полученных либо в результате разрыва одной из двух цепей (*г*), либо в результате локального расхождения двух цепей (*д*); катенана (*е*) и узла (*ж*)

цепи. Все известные одноцепочечные кольцевые ДНК относительно малы: ДНК фагов φX174 и M13 содержат 5 300 и 6 000 нуклеотидов и имеют длину 1,5 — 2 мкм. Однако для репликации любой из этих вирусных ДНК совершенно необходимо превращение одноцепочечного кольца в соответствующее двуцепочечное, из которого затем образуются одноцепочечные кольцевые ДНК вирусного потомства. Более того, экспрессия генетической информации в таких геномах всегда осуществляется в фазе двуцепочечной ДНК, поскольку именно она является субстратом для транскрипции последовательностей ДНК в РНК. Третичная структура и линейных и кольцевых форм ДНК характеризуется спирализацией и *супер*(*сверх*)*спирализацией*.

Сверхспирализация ДНК. Топоизомеразы

В живых организмах молекулы ДНК представляют собой либо очень длинные, либо замкнутые в кольцо двуспиральные молекулы, поэтому любой процесс, связанный с передачей наследственной информации, должен наталкиваться на серьезные топологические проблемы: возникновение положительной (+) или отрицательной (–) сверхспирализации ДНК, образование катенанов или узлов.

К числу таких процессов относится репликация ДНК. Движение репликативной вилки вдоль ДНК сопровождается расхождением вновь образовавшихся цепей. Их длина слишком велика для того, чтобы позволить репликативной вилке вращаться вдоль главной оси ДНК. В результате происходит нагнетание ***положительной сверхспирализации*** в родительской ДНК, когда возрастает количество оборотов на отрезке ДНК с одновременным уменьшением количества пар оснований, приходящееся на один оборот (рис. 49). Аналогично обе вновь образованные копии оказываются ***отрицательно сверхспирализованными***. Для них характерно уменьшение количества

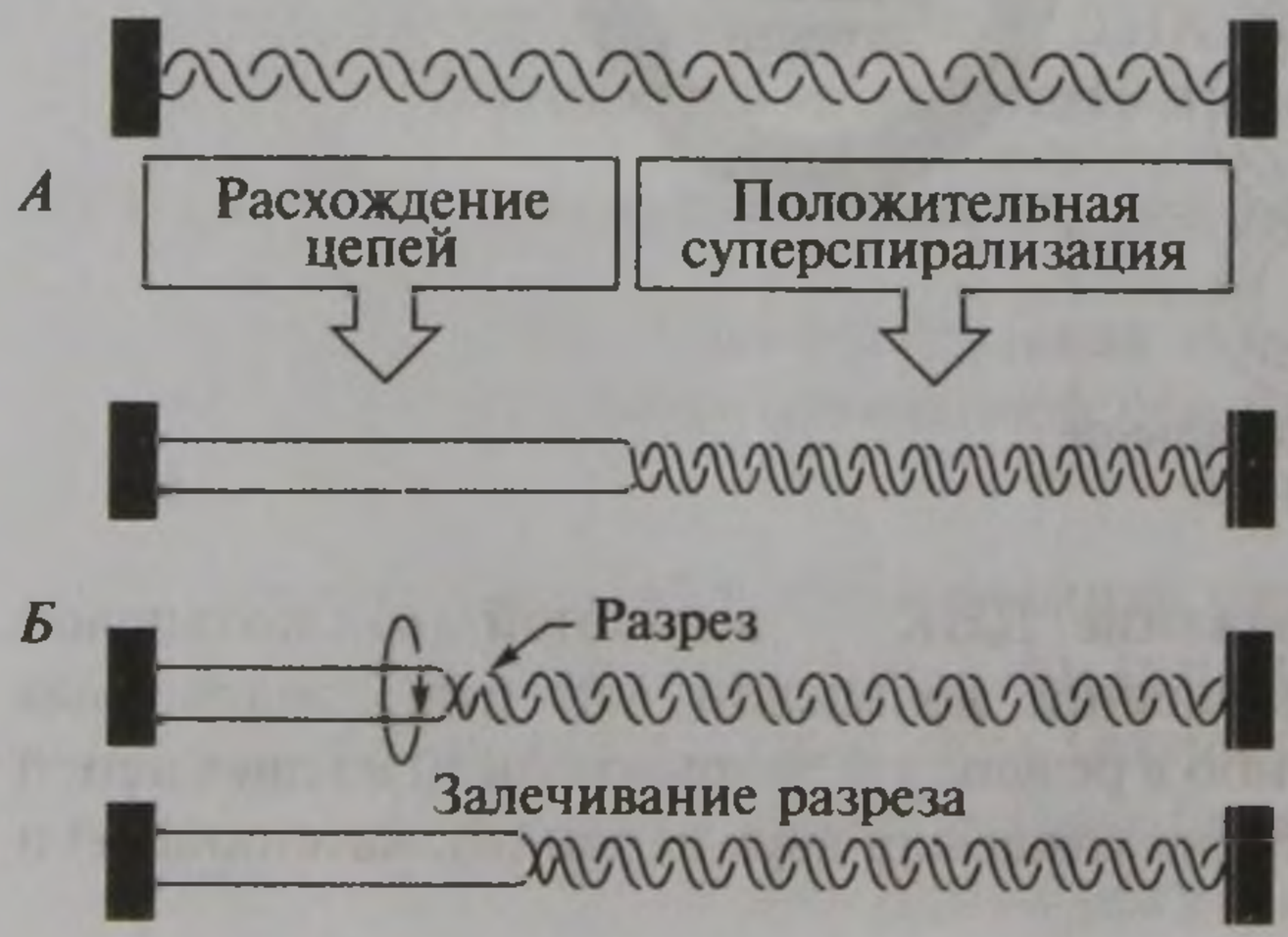


Рис. 49. Разделение цепей дуплексной ДНК посредством компенсирующей положительной суперспирализации в другой области (*А*) либо путем разрезания, вращения и воссоединения (*Б*)

оборотов на отрезке ДНК с одновременным увеличением количества пар оснований на один оборот. К моменту окончания репликации, когда репликативные вилки встречаются друг с другом, образуются узлы или (в случае кольцевых молекул) катенаны. Нагнетание сверхспирализации происходит и в процессе транскрипции (здесь она не играет принципиальной роли). «Перепутывание» сверхдлинных нитей совершается в процессе конденсации и сегрегации хромосом, организации хроматина.

Кольцевая ДНК, не содержащая разрывов ни в одной из цепей, может существовать в различных топологических формах. Эти формы характеризуются порядком зацепления *Lk* — целочисленной величиной, обозначающей количество оборотов одной цепи относительно другой, когда ДНК находится в конформации, соответствующей плоскому кольцу. Правым виткам соответствуют положительные значения *Lk*, левым — отрицательные. Каждая топологическая форма ДНК (топоизомер) может существовать в бесконечном числе конформаций, которые определяются двумя параметрами — кручением *Tw* и райзингом *Wr* (от англ. *writhe* — скручиваться): *Tw* — число витков двойной спирали в данном конформере; *Wr* — число витков суперспирали, т. е. спирали, образующейся при перекручивании двуспиральной ДНК вокруг себя. Связь этих трех параметров между собой определяется уравнением: $Lk = Tw + Wr$.

Когда ДНК находится в релаксированном состоянии, т. е. без всякого напряжения, она может быть расположена в одной плоскости, тогда $Lk = Tw$. Если же увеличить или уменьшить число витков двойной спирали, сохранив ковалентную непрерывность обеих цепей, то такое изменение будет компенсироваться образованием витков сверхспирали (*Wr*). *Wr* может быть как положительной, так и отрицательной величиной.

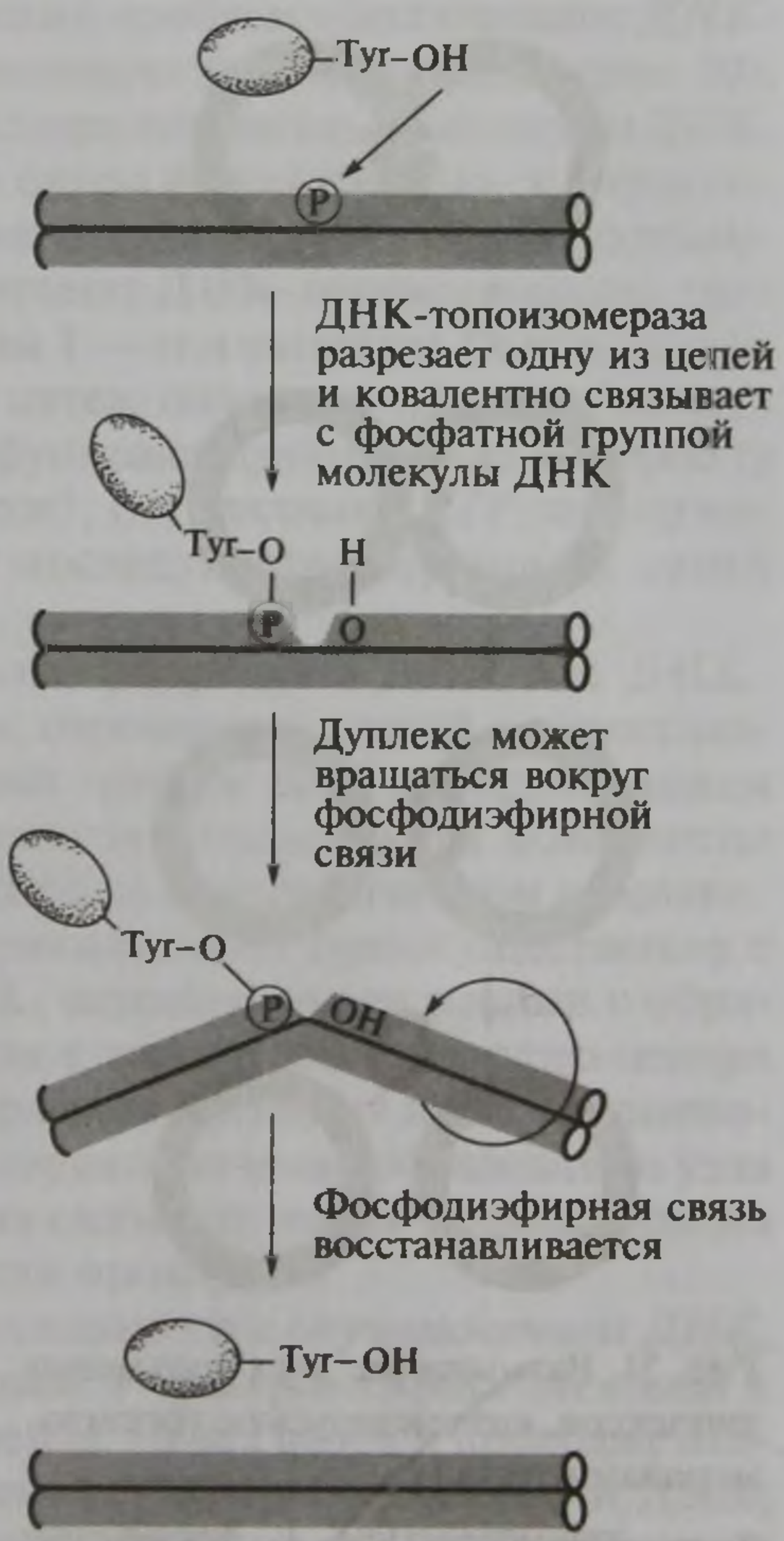


Рис. 50. Реакции, катализируемые топоизомеразами типа I.
Ⓟ — фосфат

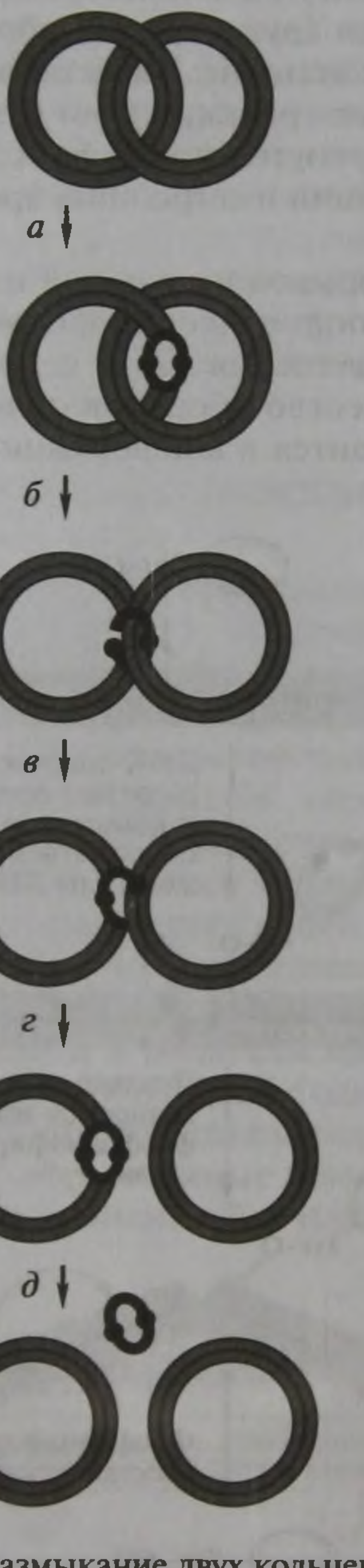


Рис. 51. Размыкание двух кольцевых дуплексов, катализируемое топоизомеразами типа II:

а — взаимодействие с ферментом; *б* — разрыв ДНК; *в* — протягивание ДНК через разрыв; *г* — воссоединение разорванной ДНК; *д* — освобождение ДНК

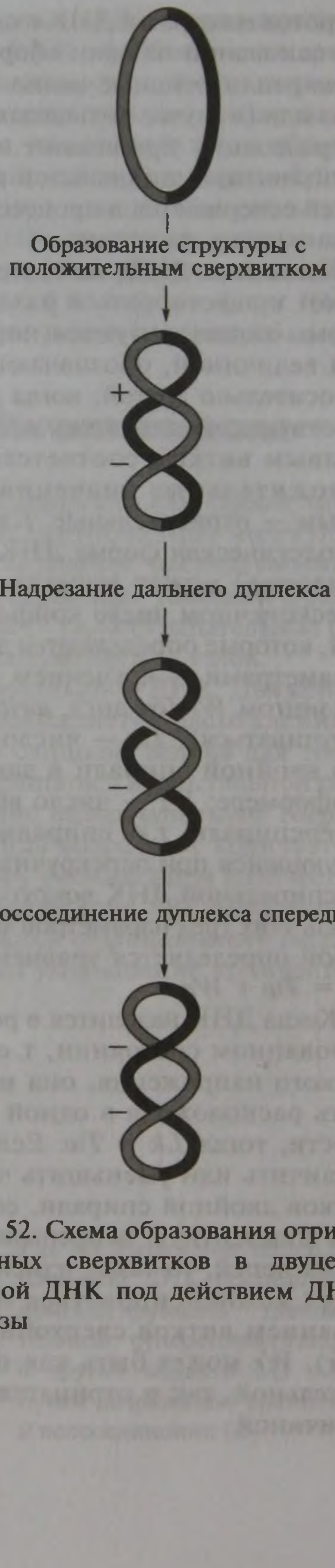


Рис. 52. Схема образования отрицательных сверхвитков в двуцепочечной ДНК под действием ДНК-гиразы

Кольцевые ковалентно замкнутые ДНК, выделенные из клеток, всегда имеют отрицательную суперспирализацию. Степень суперспирализации ДНК удобно характеризовать плотностью спирализации σ, которая равна *Wr/Tw* и не зависит от длины ДНК. Для многих природных сверхспирализованных ДНК σ = –0,05.

Сверхспирализованная ДНК обладает значительным запасом энергии по сравнению с ее релаксированной формой. Следовательно, локальное расплетание двойной спирали ДНК с отрицательными сверхвитками будет приводить к сбросу напряжения сверхспирализации и поэтому энергетически выгодно. Это отчетливо проявляется в том, что отрицательная сверхспирализация заметно стимулирует переход ДНК из правой B-формы в левую Z-форму. Отрицательная сверхспирализация ДНК облегчает связывание с ней белков, раскручивающих ее двойную спираль.

Решение названных топологических проблем обеспечивают *ДНК-топоизомеразы* — ферменты, изменяющие топологию ДНК (рис. 50). Топоизомеразы релаксируют сверхспирализованные молекулы ДНК, снимая их внутреннее напряжение путем внесения одно- и двуцепочечных разрывов с последующим их восстановлением (лигированием). По механизму действия различают ДНК-топоизомеразы двух типов — I и II. ДНК-топоизомеразы I — мономерные белки, релаксируют ДНК без затраты энергии путем внесения одноцепочечных разрывов. ДНК-топоизомеразы II функционируют в виде димеров (у эукариот) и тетрамеров (у прокариот), осуществляя АТР-зависимое расщепление обеих цепей ДНК с последующим переносом цепей через разрыв и его лигированием.

Для внесения одноцепочечного разрыва в ДНК все ДНК-топоизомеразы используют остаток тирозина, который осуществляет нуклеофильную атаку фосфатной группы ДНК с образованием фосфотирозина. В результате ферменты оказываются ковалентно связанными с 5′- или 3′-концами ДНК в одноцепочечном разрыве.

Топоизомеразы подтипа I-5′ связываются преимущественно с *одноцепочечными участками* ДНК, осуществляя ее разрыв с образованием 5′-фосфодиэфирной связи с тирозином активного центра фермента. Через образовавшийся разрыв проходит либо одноцепочечная ДНК (снятие витка), либо двуцепочечная (образование узла или катенана). Ферменты этого типа снимают только отрицательную суперспирализацию. Характерны для прокариот.

Топоизомеразы подтипа I-3′ связываются с *двуцепочечной ДНК*, разрывают одну цепь с образованием 3′-фосфодиэфирной связи с тирозином активного центра фермента. Через разрыв проходит вторая цепь дуплекса, уменьшая плотность суперспирализации ДНК; знак суперспирализации существенной роли не играет. Характерны для эукариот.

Топоизомеразы типа II связываются с двуцепочечной ДНК и осуществляют разрыв обеих цепей с образованием двух 5′-фосфоди-

эфирных связей с тирозинами активного центра. В образовавшуюся щель проходит вторая двуцепочечная ДНК, и результатом является изменение числа положительных или отрицательных супервитков на 2 (в отличие от ферментов типа I, изменяющих число супервитков на единицу за шаг).

Этот фермент способен катенировать, декатенировать и развязывать узлы интактной ДНК. Для многократного повторения цикла требуется АТР (рис. 51).

Топоизомераза II *E. coli* (ДНК-гираза) относится к ферментам типа II, но не требует АТР. Она индуцирует образование отрицательных супервитков в релаксированных кольцевых ДНК. Гираза взаимодействует с ДНК таким образом, что последняя наматывается вокруг белка. При этом возникает положительная сверхспирализация в тех местах молекулы ДНК, которые связаны с белком. Затем фермент разрывает обе нити ДНК и переносит двойную нить с внутренней стороны на внешнюю, после чего скрепляет оба разрыва, превращая положительную петлю в отрицательную (рис. 52). Это особенно важно для инициации репликации, а также необходимо для ее элонгации и терминации.

ДНК-топоизомераза II является жизненно важным ферментом любого эукариотического организма. Выяснена роль ДНК-топоизомеразы в формировании высших уровней структуры хроматина, а именно, участии фермента в образовании петель хроматина во время конденсации хромосом.

Структура и функции РНК

Содержание РНК в любых клетках в 5 — 10 раз превышает содержание ДНК. Основная роль РНК состоит в трансляции генетической информации с образованием белков, а также в осуществлении некоторых специализированных эндонуклеазных функций, возможно регулирующих различные этапы экспрессии генов. Геномы некоторых вирусов (ретровирусов и множества вирусов животных, растений и насекомых) представлены одно- и двуцепочечной молекулой РНК.

Виды РНК. Во всех клетках присутствуют следующие виды РНК: рибосомная (рРНК), транспортная (тРНК) и информационная, или матричная (мРНК). Большинство клеток содержат много малых цитоплазмических РНК (мцРНК), а в клетках эукариот присутствуют малые ядерные РНК (мяРНК) (табл. 4).

Около 80 — 85 % массы клеточных РНК составляют три (прокариоты) или четыре (эукариоты) вида рРНК, около 10 % — почти 100 видов тРНК. На долю нескольких тысяч различных матричных РНК приходится менее 5 % клеточной РНК, а на долю малых ядерной и цитоплазматической РНК, число видов которых пока неизвестно, — менее 2 % от общего количества РНК.

Таблица 4. **Основные виды РНК**

Вид РНК	Приблизительное число разных видов в клетках	Приблизительная длина (число нуклеотидов)	Распространенность*
Транспортная РНК (тРНК)	80—100	75—90	П, Э
Рибосомная 5S РНК (рРНК)	1—2	120	П, Э
Рибосомная 5,8S РНК (рРНК)	1	158	Э
Рибосомная 16S РНК (рРНК)	1	1 600	П
Рибосомная 23S РНК (рРНК)	1	3 200	П
Рибосомная 18S РНК (рРНК)	1	1 900	Э
Рибосомная 28S РНК (рРНК)	1	5 000	Э
Матричная РНК (мРНК)	Тысячи	Варьирует	П, Э
Гетерогенная ядерная РНК (гяРНК)	Тысячи	Варьирует	Э
Малая цитоплазматическая (мцРНК)	Десятки	90—330	П, Э
Малая ядерная РНК (мяРНК)	Десятки	58—220	Э
Малая интерферирующая РНК (миРНК)	Десятки	20—30	Э
Транспортно-матричная РНК (тмРНК)	Десятки	350—400	П

* П — прокариоты, Э — эукариоты

Макромолекулярная структура РНК

Двутяжевые РНК, построенные из двух комплементарных полинуклеотидных цепей, обнаруживаются только в составе некоторых вирусов. Кроме того, известны *однотяжевые РНК* (вироиды), состоящие, по сути дела, из двух комплементарных нуклеотидных последовательностей и поэтому существующие в виде двуспиральных молекул (см. рис. 60, с. 117).

В двойной спирали РНК возможна только одна, C3′-эндо-конформация остатка рибозы. Поэтому двуспиральные РНК, а также комплементарные двутяжевые полирибонуклеотидные комплексы [типа поли(A)·поли(U)] существуют только в правой A-форме, в основных чертах подобной A-форме ДНК.

Общие черты вторичной структуры однотяжевых РНК. Макромолекулы большинства природных РНК построены из одной полирибонуклеотидной цепи. Основной элемент их вторичной структуры — сравнительно короткие двойные спирали, образованные комплементарными участками одной и той же цепи и перемежающиеся ее однотяжевыми сегментами. Полирибонуклеотидные цепи в таких двуспиральных структурах (спирализация сама на себя) *антипараллельны*, а сами двойные спирали, находящиеся в A-форме, не идеальны: в них имеются дефекты в виде неспаренных нуклеотидных остатков или не вписывающихся в двойную спираль однотяжевых петель. Наряду с классическими уотсон-криковскими парами (AU и GC) в двутяжевых участках РНК часто встречается пара GU. Таким образом, стабильность двутяжевых районов поддерживается комплементарными и межплоскостными взаимодействиями оснований. В однотяжевых участках наблюдаются сильные *стэкинг-взаимодействия* оснований, вследствие чего они стремятся принять конформацию однотяжевой спирали.

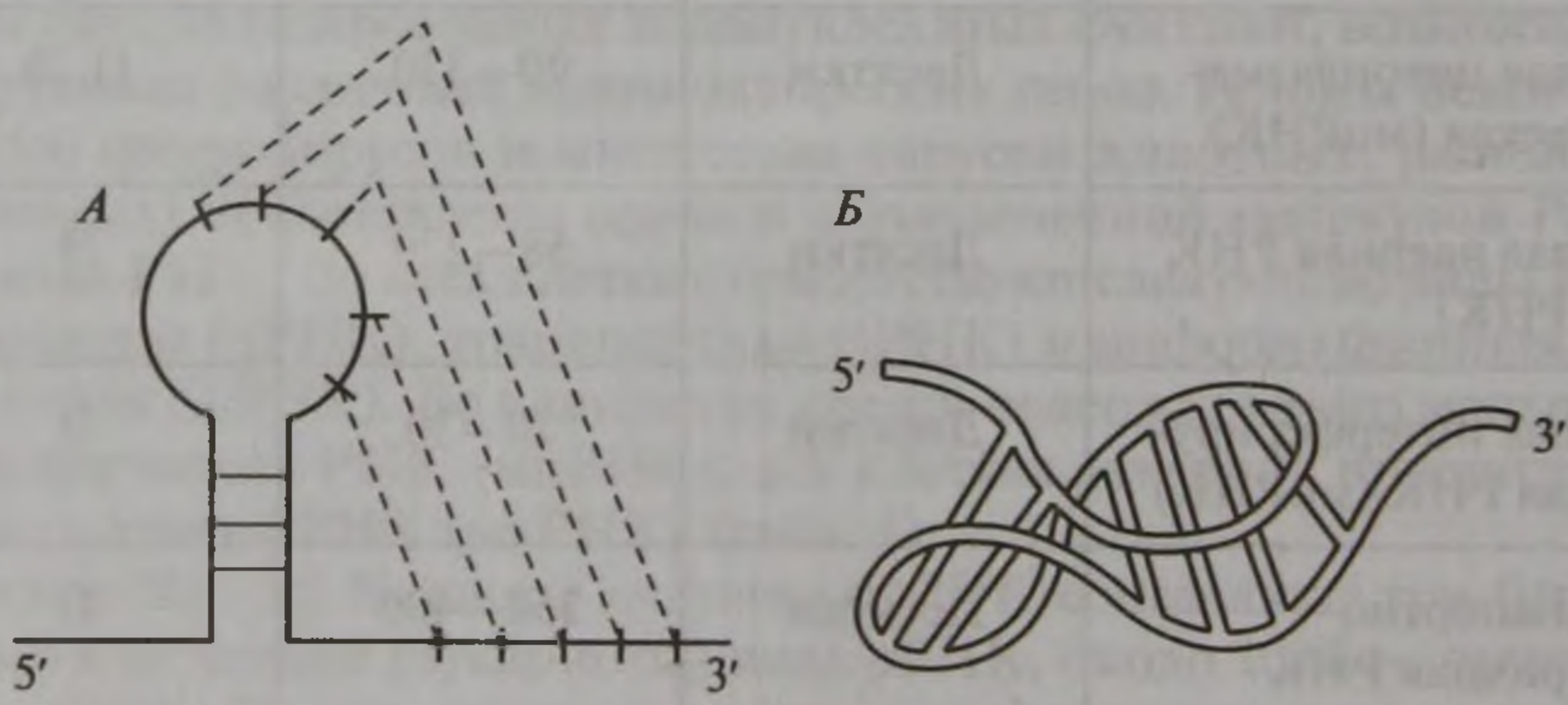


Рис. 53. Схема образования «псевдоузла» (*А*) при спаривании комплементарных участков РНК в вершине шпильки и в другом однотяжевом участке; пространственная модель «псевдоузла» (*Б*)

Третичная структура однотяжевых РНК. При исследованиях макромолекулярной организации однотяжевых РНК было установлено, что в физиологических условиях они характеризуются компактной и упорядоченной третичной структурой, которая возникает за счет взаимодействия элементов их вторичной структуры.

1. Элементы вторичной структуры РНК располагаются друг относительно друга так, чтобы обеспечить максимальный стэкинг оснований в макромолекуле в целом.

2. Контакты между отдельными элементами вторичной структуры осуществляются за счет нескольких типов так называемых «третичных» внутримолекулярных взаимодействий:

• за счет образования дополнительных, зачастую не-уотсон-криковских пар оснований между нуклеотидными остатками удаленных друг от друга (в первичной и вторичной структурах) однотяжевых участков и триплетов оснований между нуклеотидными остатками однотяжевых и двутяжевых элементов. При анализе вторичных структур вирусных и рибосомных РНК было обнаружено, что однотяжевые участки в вершине петель в принципе могут давать двуспиральные комплексы с межпетельными однотяжевыми сегментами с образованием «псевдоузлов» (рис. 53, 54);

• за счет дополнительных («третичных») стэкинг-взаимодействий после интеркаляции (внедрения) оснований одного участка между двумя соседними основаниями другого однотяжевого участка (рис. 55);

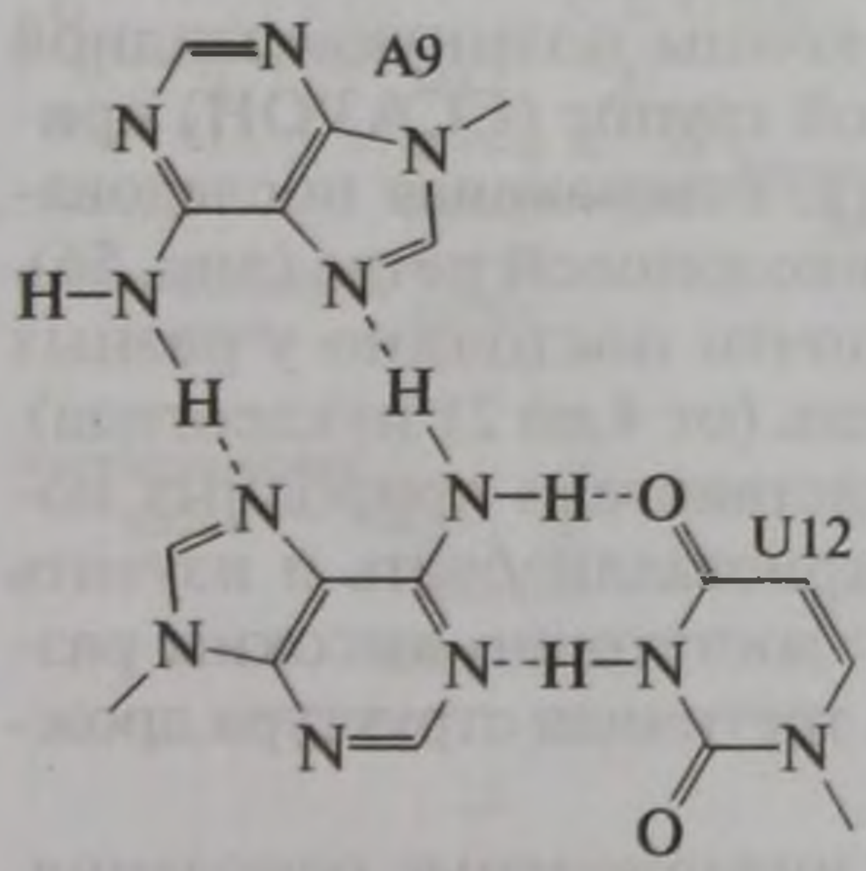


Рис. 54. Образование A·U·A-триплета в тРНК

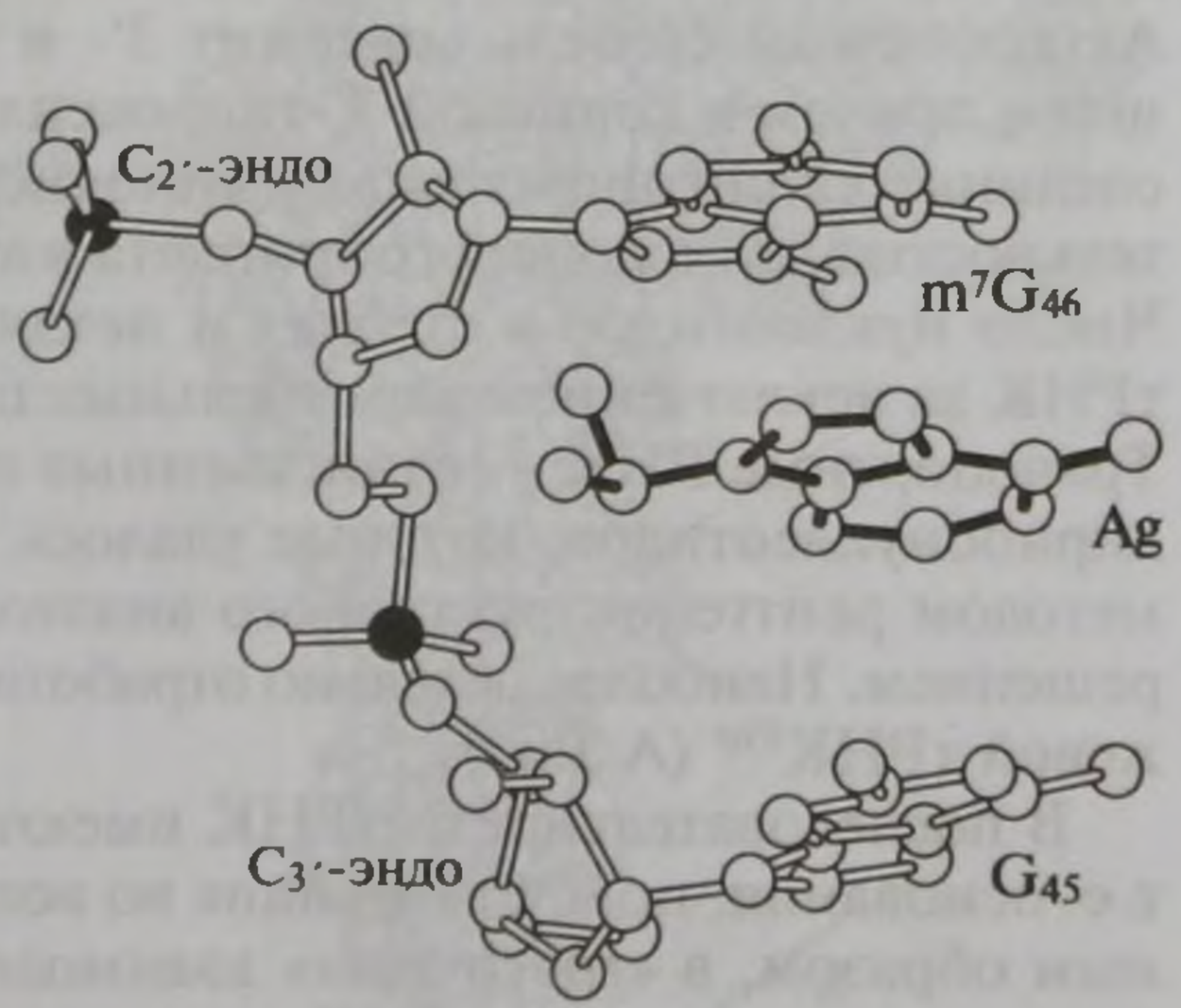


Рис. 55. Интеркаляция A_9 между остатками m^7G_{46} и G_{45} в тРНК.

Этот особый вид стэкинг-взаимодействия сопряжен с расхождением остатков гуанина, при этом конформация рибозы изменяется от $C_{3'}$-эндо- до $C_{2'}$-эндо-

• за счет образования дополнительных водородных связей между 2′OH-группами остатков рибозы и основаниями, а также другими группами пентозофосфатного остова.

3. Третичная структура РНК стабилизирована ионами двухвалентных металлов, которые связываются не только с фосфатными группами, но и с основаниями.

Аналогичные принципы лежат в основе организации структуры всех однотяжевых РНК. Строить конкретные пространственные структуры однотяжевых РНК крайне затруднительно из-за множества возможных вариантов укладки в них элементов вторичной структуры.

Транспортные РНК

Главной функцией транспортных РНК (тРНК) является акцептирование аминокислот и перенос их в белоксинтезирующий аппарат клетки. Они выступают в роли затравки (праймера) в процессе обратной транскрипции (см. с. 237).

Последовательность тРНК включает 70—90 нуклеотидов и около 10 % минорных компонентов. Она образует вторичную структуру, известную под названием «клеверный лист». Эта структура состоит из 4 или 5 двуцепочечных спиральных стеблей и трех петель. Каждый стебель содержит 4—7 уотсон-криковских пар, образующих двойные спирали. Различают акцепторный, антикодоновый, дигидроуридиловый (D), псевдоуридиловый (TψC) и добавочный стебли. Акцепторный стебель содержит 3′- и 5′-концы полинуклеотидной цепи, причем к концевой 3′-гидроксильной группе (CCA3′OH) присоединяется специфическая аминокислота, отвечающая последовательности антикодонового триплета в антикодоновой петле (рис. 56). Число нуклеотидов в стеблях и петлях почти постоянно у разных тРНК за исключением вариабельных петель (от 4 до 21 нуклеотида). Транспортные РНК — единственные представители природных полирибонуклеотидов, которые удалось закристаллизовать и изучить методом рентгеноструктурного анализа с достаточно высоким разрешением. Наиболее детально отработана третичная структура дрожжевой $тРНК^{Phe}$ (А. Рич).

В последовательности тРНК имеются инвариантные основания, т. е. основания, присутствующие во всех тРНК. Они участвуют, главным образом, в «третичных» взаимодействиях, которые возникают тогда, когда молекула тРНК сворачивается в нативную Г-образную структуру (см. рис. 56). Г-форма состоит из двух почти перпендикулярных друг другу спиралей А-РНК, длина которых составляет около 7 нм, толщина — 2 нм. Одну спираль образуют уложенные друг за другом антикодоновый и дигидроуридиловый стебли, другую — акцепторный и псевдоуридиловый стебли.

Третичные взаимодействия включают стэкинг (у тРНК^{Phe} из 76 оснований в стэкинг-взаимодействиях участвует 71 основание); интеркаляцию и необычное спаривание оснований, при котором образуются не только пары, но и триплеты (см. рис. 56). Большая часть стабилизирующих третичную структуру водородных связей образуется между инвариантными и полувариантными основаниями.

Особым видом стэкинга является интеркаляция (в отдельных местах основание из одной цепи встраивается между двумя основаниями другой цепи) (см. рис. 56). Такое взаимодействие осущест-

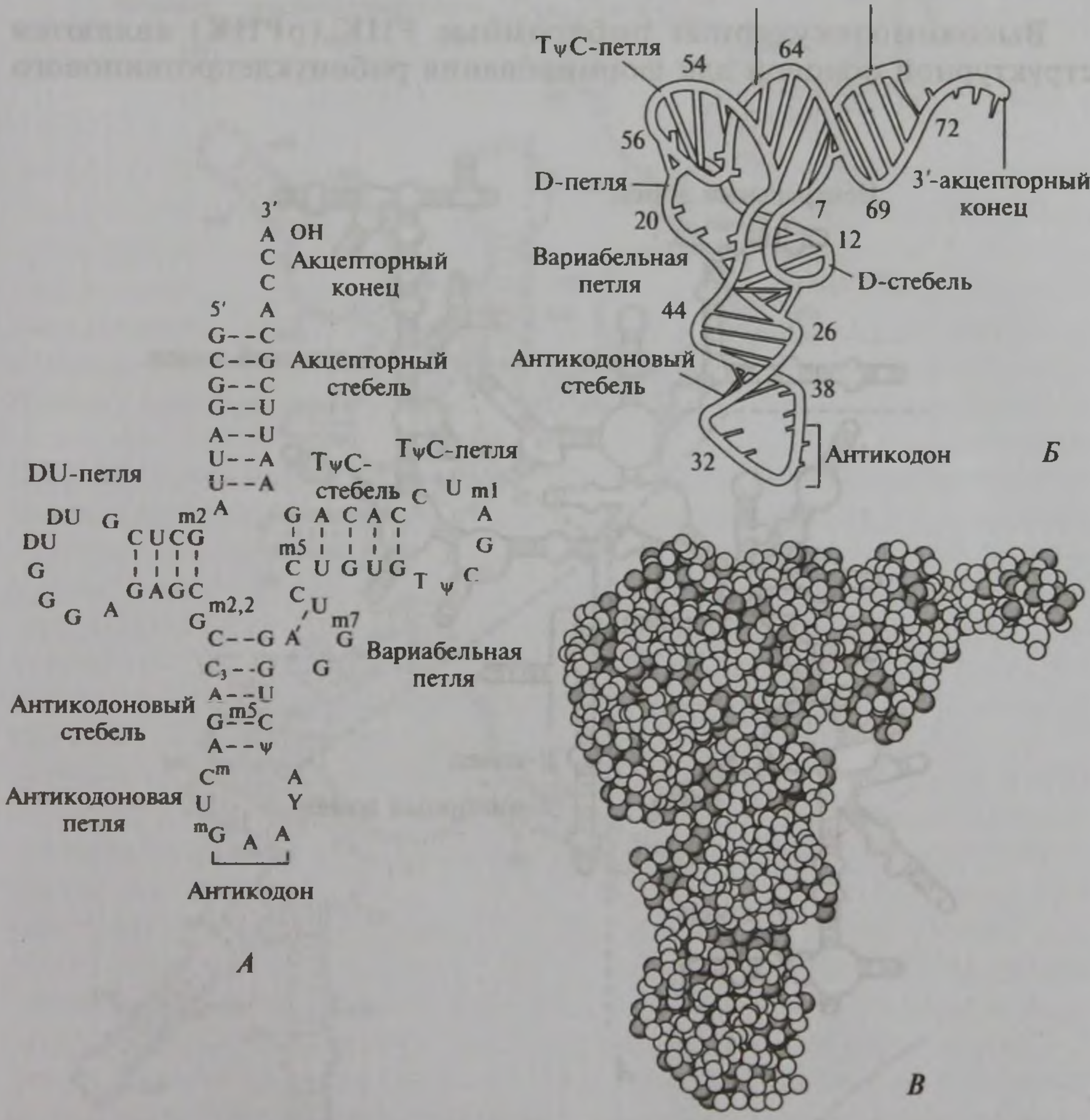


Рис. 56. Структура дрожжевой тРНК^{Phe}:

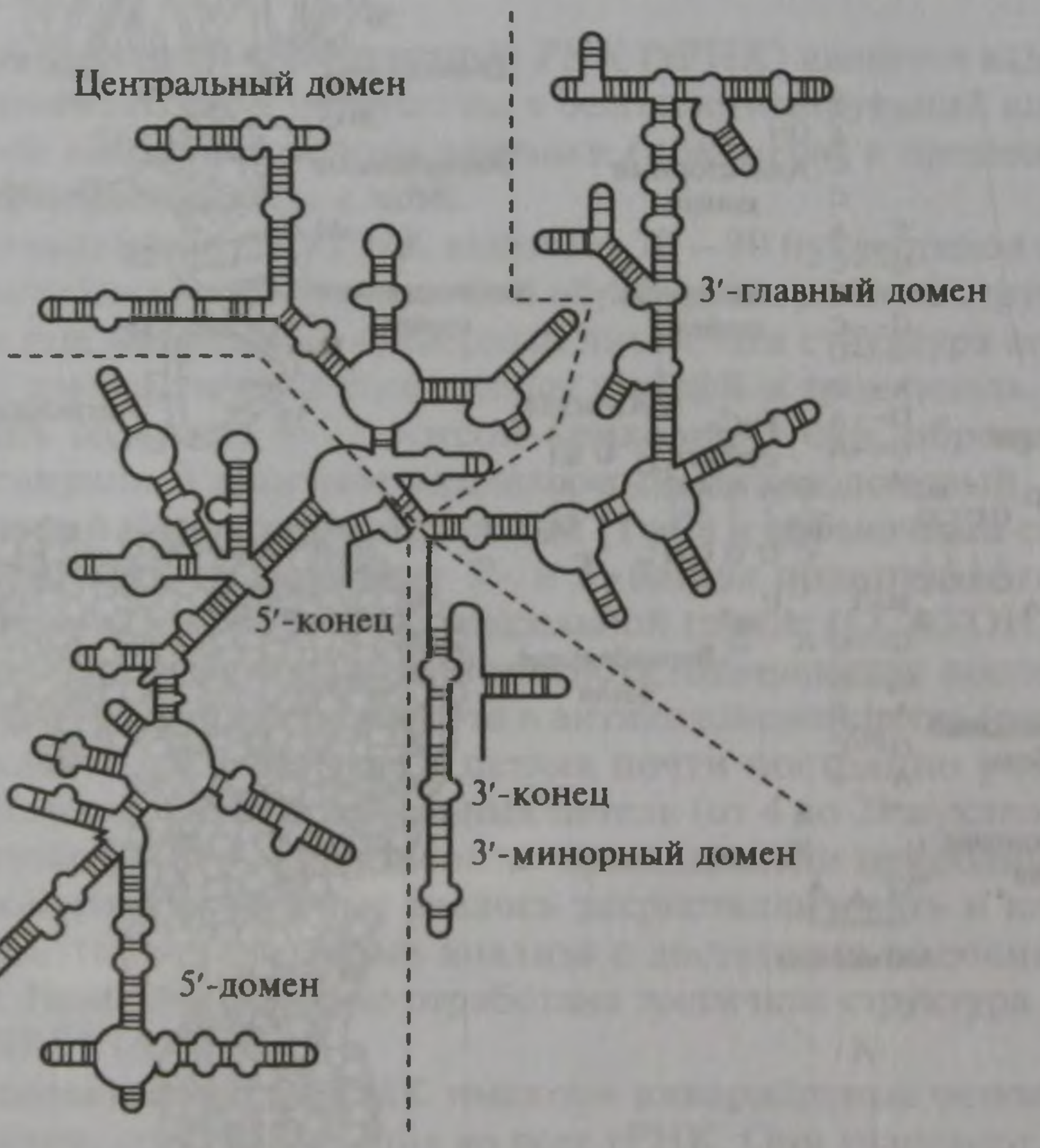
А — первичная и вторичная структуры; *Б* — третичная структура (схематическое изображение). Пентозофосфатный остов изображен в виде непрерывной ленты; пары оснований — в виде длинных полосок; неспаренные основания — в виде коротких полосок. Молекула имеет Г-образную форму. Арабскими цифрами указаны номера нуклеотидных остатков; *В* — пространственная модель

вляется там, где встречаются сразу три цепи. Пентозофосфатный остов в месте вставки основания перестраивается путем изменения конформации сахара C3′-эндо-→C2′-эндо-, в результате чего увеличивается расстояние между фосфатами (от 0,59 нм до 0,70 нм), и это способствует интеркаляции.

Антикодон имеет жесткую архитектуру, которая позволяет ему быстро считывать матричную РНК.

Рибосомные РНК

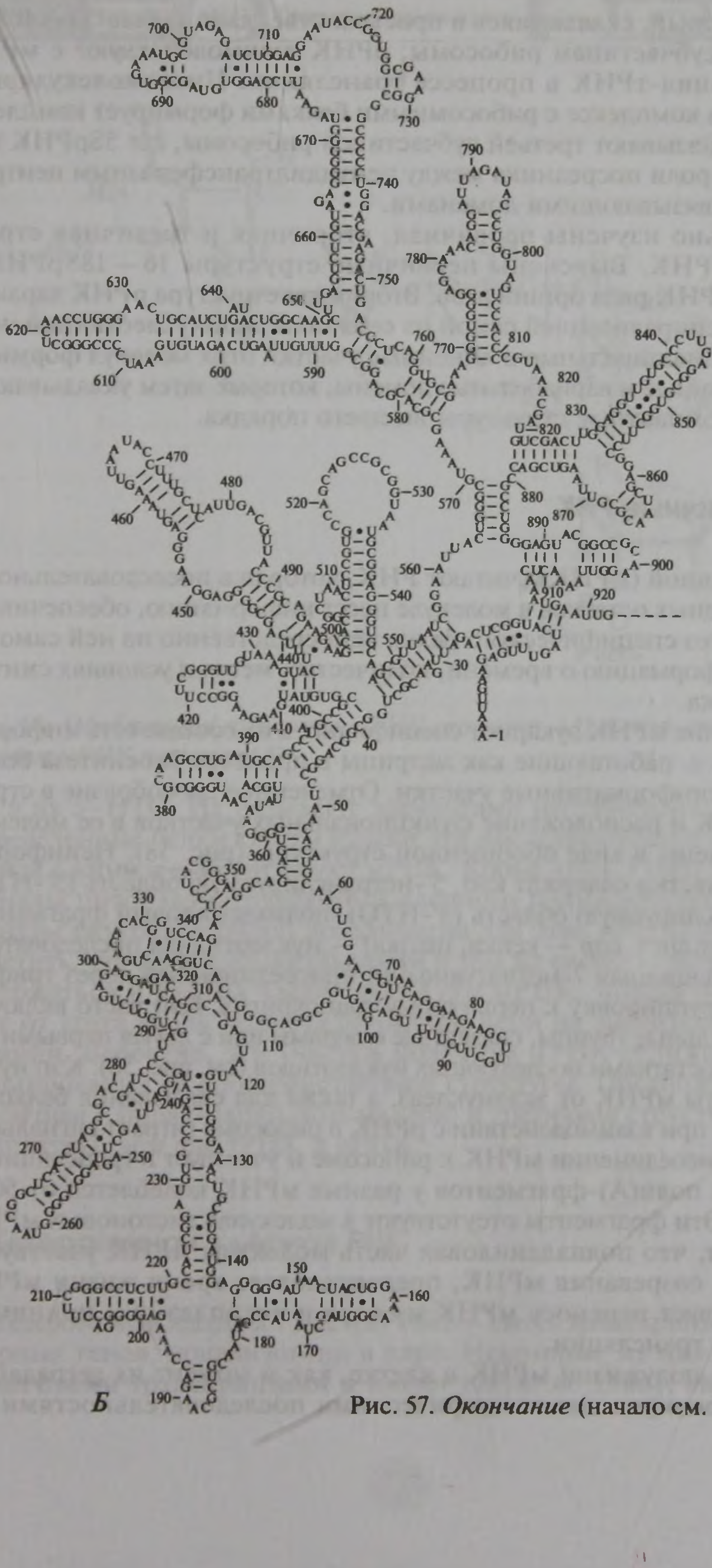
Высокомолекулярные рибосомные РНК (рРНК) являются структурной основой для формирования рибонуклепротеинового



А

Рис. 57. Структуры 16S рРНК *E. coli* (окончание см. на с. 107):

А — вторичная структура, содержащая 4 основных домена: центральный, 3′-главный, 5′- и 3′-минорный (их границы указаны пунктирными линиями); *Б* — первичная и вторичная структуры 5′-домена 16S рРНК. В стабилизации дуплексной структуры рРНК участвуют G·U-пары. Несколько оснований образуют петлю, если в противоположной цепи отсутствуют комплементарные основания



Б

Рис. 57. *Окончание* (начало см. на с. 106)

тяжа, который, складываясь в пространстве, дает начало 30—40S- и 50—60S-субчастицам рибосомы; рРНК взаимодействуют с мРНК и аминоацил-тРНК в процессе трансляции. Низкомолекулярная 5SpРНК в комплексе с рибосомными белками формирует комплекс, который называют третьей субчастицей рибосомы, где 5SpРНК выступает в роли посредника между пептидилтрансферазным центром и EF-G-связывающими доменами.

Детально изучены первичная, вторичная и третичная структуры 5SpРНК. Выяснены первичные структуры 16—18SpРНК и 23—25SpРНК ряда организмов. Вторичная структура рРНК характеризуется спирализацией самой на себя полирибонуклеотидной цепи (рис. 57). Биспиральные и линейные участки этих молекул формируют постоянные и вариабельные домены, которые затем укладываются в более компактные структуры высшего порядка.

Матричные РНК

Матричной (мРНК) считают РНК, которая в последовательности нуклеотидных остатков в молекуле несет информацию, обеспечивающую синтез специфического белка непосредственно на ней самой, а также информацию о времени, количестве, месте и условиях синтеза этого белка.

Строение мРНК эукариот специфично: в их составе есть информативные, т.е. работающие как матрицы в процессе биосинтеза белка, зоны и неинформативные участки. Отмеченное своеобразие в строении мРНК и расположение функциональных участков в ее молекуле представлены в виде обобщенной структуры (рис. 58). Неинформативные участки содержат кэп, 5′-нетранслируемую область (5′-НТО), 3′-нетранслируемую область (3′-НТО), полиаденилоый фрагмент.

Кэп (от англ. *cap* — кепка, шапка) — нуклеотидная последовательность, содержащая 7-метилгуанозин, присоединенный через трифосфатную группировку к первичному транскрипту. Кэп часто включает 2′-О-метильные группы, связанные с первым или с двумя первыми рибозными остатками последующих нуклеотидов (см. рис. 58). Кэп нужен для защиты мРНК от экзонуклеаз, а также для связывания белковых факторов при взаимодействии с рРНК в рибосоме (играет сигнальную роль в присоединении мРНК к рибосоме и участвует в трансляции).

Длина поли(А)-фрагментов у разных мРНК колеблется от 50 до 400 н.п. Эти фрагменты отсутствуют в молекулах гистоновых мРНК. Полагают, что полиаденилоая часть молекулы мРНК участвует в процессе созревания мРНК, предопределяет время жизни мРНК, способствует переносу мРНК из ядра в цитоплазму и принимает участие в трансляции.

Время полужизни мРНК в клетке, как и момент их деградации, запрограммированы специфическими последовательностями ча-

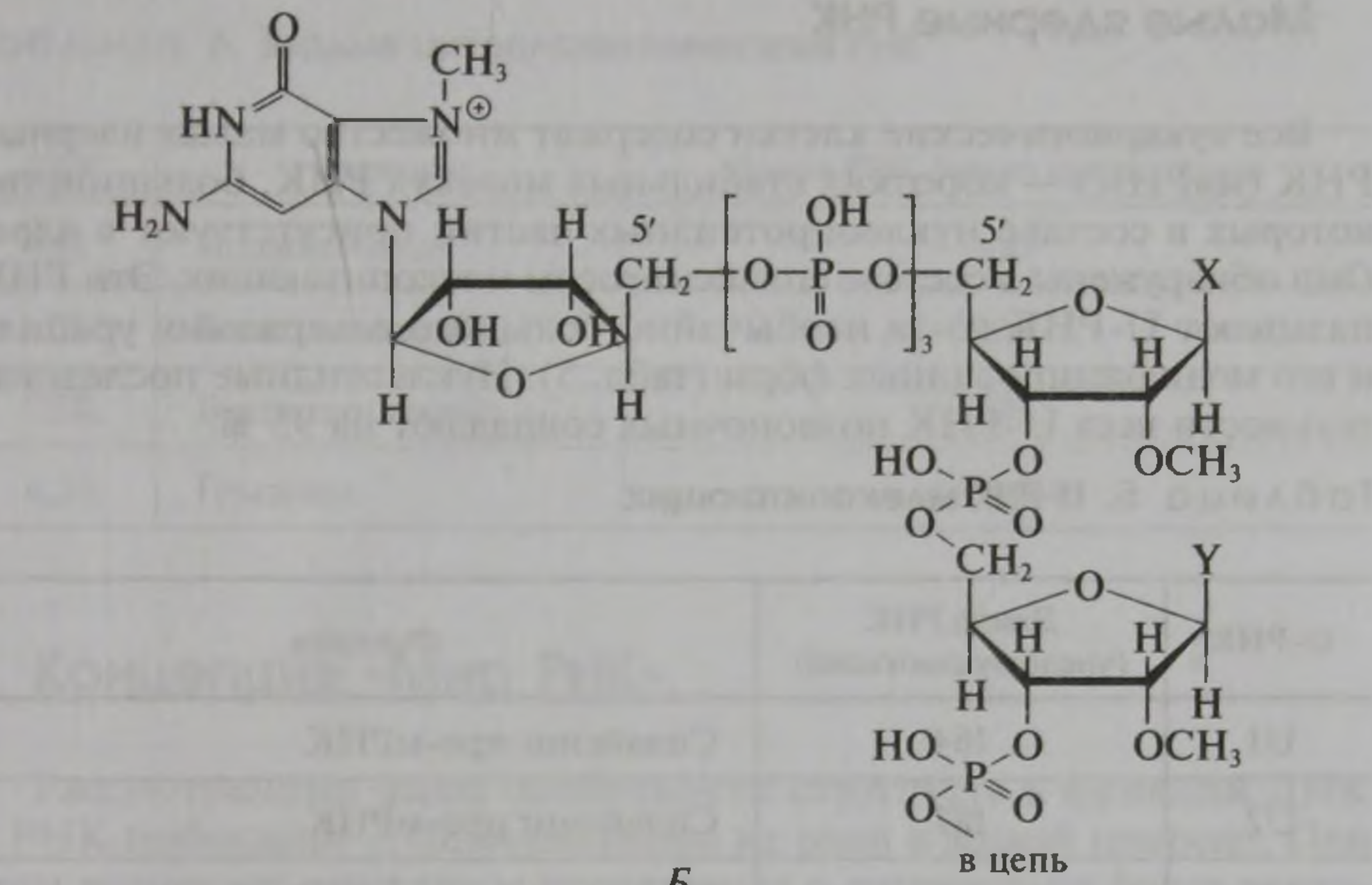


Рис. 58. Обобщенная структура мРНК эукариот (*А*) и кэп-структура на 5′-конце мРНК эукариот (*Б*)

сто в их 3′-НТО. Их называют элементами нестабильности мРНК. Определенные белки клетки узнают эти последовательности, связываются с ними и стабилизируют мРНК. Кроме этого 3′-НТО содержат сигналы цитоплазматического и ядерного полиаденилирования и сигналы внутриклеточной локализации для этих мРНК. Интересно отметить, что информация об одном и том же свойстве мРНК может содержаться в разных частях молекулы мРНК, иногда отстоящих друг от друга вдоль полинуклеотидной цепи на значительном расстоянии. Вероятно, эти участки молекулы мРНК сближаются при формировании ее пространственной структуры. Предполагают, что вторичная (спирализация сама на себя) и третичная структуры менее компактны, чем для тРНК и рРНК.

Гетерогенная ядерная РНК

Гетерогенная ядерная РНК (гяРНК) — смесь транскриптов многих ядерных генов; локализована в ядре. Некоторые из них являются первичными транскриптами и имеют такую же длину, как и гены,

с которых они скопированы, другие — частично процессированы и утратили ряд интронов.

Малые ядерные РНК

Все эукариотические клетки содержат множество малых ядерных РНК (мяРНК) — коротких стабильных молекул РНК, большинство которых в составе нуклеопротеидных частиц присутствуют в ядре. Они обнаружены в составе сплайсингосом млекопитающих. Эти РНК называют U-РНК из-за необычайно большого содержания урацила и его модифицированных форм (табл. 5). Нуклеотидные последовательности всех U-РНК позвоночных совпадают на 95 %.

Таблица 5. **U-РНК млекопитающих**

U-РНК	Длина РНК (число нуклеотидов)	Функция
U1	164	Сплайсинг про-мРНК
U2	187	Сплайсинг про-мРНК
U3	217	Процессинг про-рРНК в ядрышке
U4	145	Сплайсинг про-мРНК
U5	116	Сплайсинг про-мРНК
U6	106	Сплайсинг про-мРНК
U7	65	Образование 3′-концов гистоновых мРНК
U11	131	Полиаденилирование про-мРНК

Малые цитоплазматические РНК

Функции малых цитоплазматических РНК (мцРНК), за исключением 7SL-РНК сигнал-распознающих частиц, не установлены. Известно, что большинство этих РНК ассоциированы с крупными семействами последовательностей, содержащими как гены, так и псевдогены.

В переносе новосинтезированных секретируемых и связанных с мембранами полипептидов через липидный слой эндоплазматического ретикулума участвуют сигналраспознающие частицы, основным

компонентом которых является 7SL-РНК. Нуклеотидные последовательности 7SL-РНК у грызунов и приматов практически одинаковы, а у дрозофилы эта РНК на 64 % гомологична соответствующей РНК человека (табл. 6).

Таблица 6. **Малые цитоплазматические РНК**

РНК	Организм	Длина РНК (число нуклеотидов)
7SL	Позвоночные	300
7S	Беспозвоночные	254
7SK	Беспозвоночные	330
4,5S	Грызуны	90—94

Концепция «Мир РНК»

Рассмотренные выше особенности структуры и функций ДНК и РНК побуждают к сопоставлению их роли в живой природе. При этом возникает отчетливое впечатление о значительно более разнообразных функциональных возможностях рибонуклеиновых кислот по сравнению с ДНК, существование которой связано исключительно с необходимостью сохранения и передачи из поколения в поколение наследственных признаков. Современные знания о структурном и функциональном разнообразии РНК уже не укладываются в те канонические представления об их роли в реализации генетической информации, которые возникли в самом начале развития молекулярной биологии. Представление о том, что РНК служит только инструментом трансформации генетической программы (генотипа), заложенной в структуре ДНК, в конкретный фенотип, формируемый разнообразием белков (ДНК → РНК → белок), было в значительной мере разрушено после открытия *обратной транскрипции*. Оказалось, что РНК может служить матрицей не только для воспроизведения своей собственной структуры в РНК-содержащих вирусных геномах, но и для биосинтеза ДНК у высших организмов. Уже одно это открытие формально поставило РНК в центр основного постулата молекулярной генетики, так как показало, что поток генетической информации распространяется от РНК не в одном, а в двух направлениях: не только к белку, но и к ДНК. Все более глубокое проникновение в механизмы основных молекулярно-генетических процессов (репликацию, транскрипцию и трансляцию) способствовало возникновению понятия о *неканонических функциях* РНК, осознанию *полифункциональности* рибонуклеиновых кислот (рис. 59).

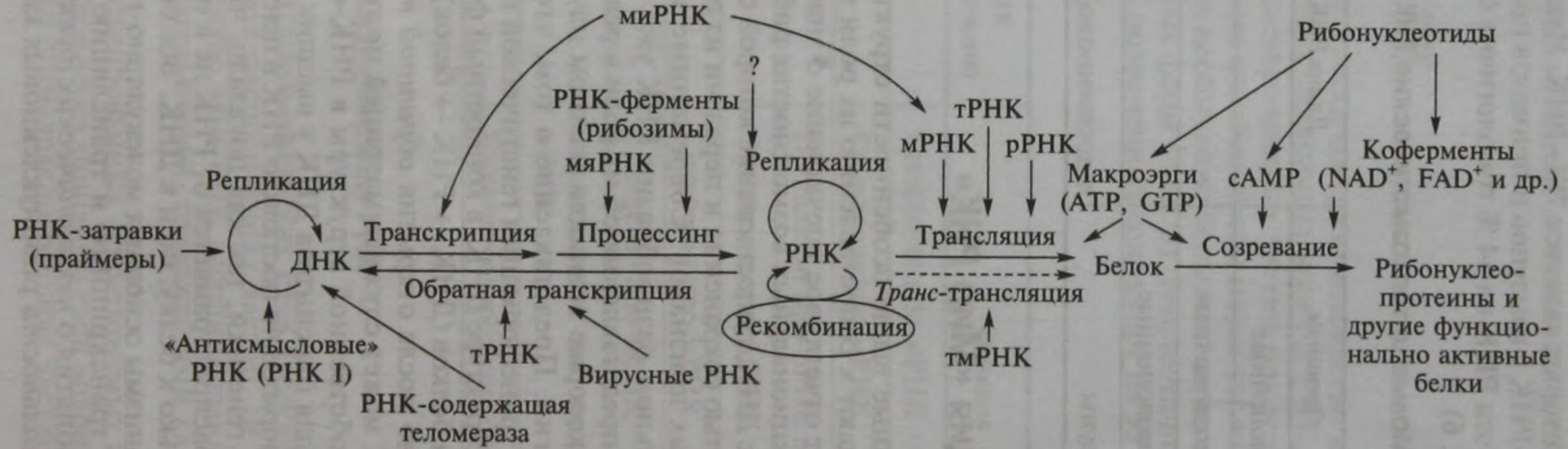


Рис. 59. Роль различных РНК в основных молекулярно-биологических процессах. Пояснения см. в тексте

Для того чтобы оценить значение РНК в природе в целом, следует попытаться ответить на вопрос: *что могут РНК*?

• К достаточно давно определенным каноническим функциям РНК относятся: способность выполнять роль мессенджера при передаче наследственной информации о структуре белка от ДНК к белоксинтезирующему аппарату клеток (мРНК), участвовать в формировании структуры рибосом (рРНК), обеспечивать специфическое акцептирование и перенос аминокислот к рибосомам (тРНК).

Вместе с тем РНК свойственны особые неканонические функции, реализуемые на разных этапах программы жизни тех или иных организмов. В последующих главах неоднократно будут приведены примеры своеобразных структур и функций различных видов РНК, сейчас же ограничимся только перечислением функциональных возможностей этой весьма полиморфной группы молекул, вовлеченных в реализацию большинства фундаментальных биохимических процессов в клетке.

• РНК участвуют в *репликации ДНК*, выступая в роли затравок (праймеров), необходимых для инициации синтеза комплементарных цепей ДНК. Особая «антисмысловая» РНК (РНК I) выполняет роль регулятора инициации репликации ДНК в точках начала репликации, обладая возможностью связывать праймеры и тем самым останавливать биосинтез ДНК.

• РНК выполняет роль *матричной молекулы* в процессах обратной транскрипции (биосинтезе ДНК на матрице РНК) и своей собственной репликации у РНК-содержащих вирусов и фагов. В процессе обратной транскрипции роль затравки, необходимой для синтеза комплементарной цепи ДНК, выполняет тРНК. Матричные свойства РНК реализуются в процессе наращивания теломерных повторов в молекулах ДНК: РНК-матрица является важнейшим компонентом *теломераз* — ферментов, осуществляющих синтез теломерных участков ДНК в хромосомах.

• В процессе *транскрипции* (биосинтезе РНК на матрице ДНК) большое значение имеет способность РНК образовывать разнообразные элементы вторичной структуры (шпильки), которые влияют как на инициацию, так и на терминацию синтеза РНК. РНК активно участвует в процессе своего собственного созревания — процессинге первичных транскриптов (про-РНК). У примитивных одноклеточных организмов выявлена способность РНК к *аутосплайсингу* — вырезанию некодирующих участков (интронов) и сшиванию кодирующих фрагментов (экзонов) без участия белков-ферментов. У организмов, утративших способность к аутосплайсингу, в сплайсировании РНК тем не менее принимают участие особые молекулы — малые ядерные РНК (мяРНК), необходимые для безошибочного вычленения интронов из молекул РНК-предшественников.

• В *биосинтезе белка* (трансляции) РНК безусловно играет определяющую роль. Различные по структуре рРНК формируют

основу субчастиц рибосомы и определяют взаимодействие субчастиц при сборке полной рибосомы. Присоединение мРНК к рибосоме детерминируется комплементарным взаимодействием определенных участков мРНК и рРНК. Активация аминокислот, их специфическое акцептирование и доставка к рибосомам осуществляется тРНК. Кодон-антикодоновое взаимодействие между мРНК и тРНК обеспечивает перевод нуклеотидной последовательности информационных макромолекул в аминокислотную последовательность синтезируемых белков. Сама реакция образования пептидной связи (транспептидирование) и продвижение рибосомы по мРНК (транслокация) также, по всей видимости, связаны с функционированием рРНК. Пространственная структура мРНК непосредственно влияет на скорость трансляции, а ее способность взаимодействовать с разнообразными регуляторными белками, особенно характерная для высших эукариот, является основой для тонкой регуляции биосинтеза белка. Особый вид РНК (тмРНК), сочетающий в себе свойства тРНК и мРНК одновременно, способен обеспечивать *транс-трансляцию* — биосинтез полипептидной цепи белка с использованием различных матричных последовательностей.

• Посттрансляционные модификации синтезированных в ходе трансляции полипептидов, в результате которых образуются функционально активные молекулы, также нередко сопряжены с присоединением к ним значительных по размерам молекул РНК. Таким путем возникают *РНК-содержащие ферменты* — рибонуклеопротеины (РНКаза Р, теломеразы и др.).

• И наконец, нельзя обойти вниманием тот факт, что многие катализаторы белковой природы (ферменты), катализирующие различные биохимические превращения в клетке, функционируют благодаря содержанию в них *коферментов рибонуклеотидной природы* (NAD, FAD, ATP и др.).

После открытия Т. Чеком с соавторами аутосплайсинга РНК у тетрахимены стало ясно, что сами молекулы РНК могут обладать каталитическими свойствами, вполне сопоставимыми со свойствами ферментов-протеинов. Именно открытие рибозимов (РНК-ферментов) привело к созданию концепции «мира РНК» — мира, который, вероятно, возник и существовал задолго до оформления ныне существующего «ДНК-белкового мира». Вскоре после открытия рибозимов Т. Чеком в одной из своих работ Ф. Крик писал: «Эти эксперименты (по каталитической РНК) поддерживают гипотезу, что биохимия РНК предшествовала традиционной биохимии, основанной на нуклеиновых кислотах и белках».

Существует несколько аргументов в пользу того, что РНК представляет собой первичную молекулу — носитель жизни. Известна *способность РНК нести генетическую информацию*. Это в полной мере свойственно ныне существующим РНК-содержащим вирусам. Доказано также, что вирусные РНК способны к *рекомбинации*, в

которую могут вовлекаться как вирусные, так и клеточные РНК. Широко известны ставшие уже классическими результаты опытов Г. Урея и С. Миллера, воспроизводящих первичную (абиотическую) среду Земли. Так, из газообразных веществ (аммиака, углекислого газа, метана и водорода) при воздействии электрического разряда и УФ-облучения можно получить элементарные соединения (формальдегид, синильную кислоту, мочевину, отдельные аминокислоты и др.), из которых далее образуются пуриновые основания (A и G) — исходные молекулы, необходимые для образования нуклеотидов. Сами нуклеотиды уже являются убиквистическими молекулами, способными существовать в виде различных жизненных форм (коферментов, энергоносителей и др.). В результате использования энергии, выделявшейся при нагревании сухих органических остатков или под действием каталитической активности неорганических полифосфатов, нуклеотиды могли взаимодействовать друг с другом с образованием полимерных молекул. Случайное объединение нуклеотидов в полимерные цепи имело решающее значение, ибо привело к возникновению матричных молекул, пригодных для комплементарного копирования. Дальнейшая эволюция этих молекул могла привести к отбору каталитически активных РНК. Таким образом мог возникнуть «мир РНК», где РНК выступает как самодостаточная молекула, сочетающая в себе генотип и фенотип одновременно и способная к эволюционному развитию благодаря рекомбинации и каталитическим способностям. Ключевым ферментом этого мира должен быть фермент РНК-репликаза, способный осуществлять аутокаталитическую репликацию РНК. Поиски такого фермента в настоящее время ведутся среди интронов группы I — участков молекул определенных видов РНК, служащих источником для возникновения и ныне существующих рибозимов.

В ходе дальнейшей биологической эволюции и особенно в связи с возникновением клеточных форм жизни часть функций РНК, вероятно, перешла к ДНК, а другая часть — к белкам. Существующие РНК обладают высоким разнообразием жизненных форм (фенотипов), превосходя в этом отношении ДНК, и сохраняют способность к хранению и передаче генетических признаков, чем принципиально отличаются от белковых молекул, многие из которых они «приспособили» для обеспечения своего собственного существования. В последующих главах мы продемонстрируем примеры межмолекулярных (белково-нуклеиновых взаимодействий), характерных для современного этапа существования главных классов биополимеров.

Глава 4
СТРУКТУРА ГЕНОМА ВИРУСОВ И ФАГОВ

Вирусы — это внеклеточная форма жизни, обладающая собственным геномом и способная к воспроизведению только в клетках живых организмов.

Вирион (или вирусная частица) состоит из одной или нескольких молекул ДНК или РНК, заключенных в белковую оболочку (капсид), иногда содержащую также липидные и углеводные компоненты.

Вирусы размножаются только после инфицирования живых клеток. Различные вирусы проникают в животные и растительные клетки, а также бактерии (вирусы бактерий называются *бактериофагами*). Вирусы являются внутриклеточными паразитами на генетическом уровне и используют для своего размножения белоксинтезирующий аппарат клетки-хозяина.

Жизненный цикл вируса начинается с проникновения внутрь клетки. Для этого он связывается со специфическими рецепторами на ее поверхности и либо вводит свою нуклеиновую кислоту внутрь клетки, оставляя белки вириона на ее поверхности, либо проникает целиком в результате эндоцитоза. В последнем случае после проникновения вируса внутрь клетки следует его «раздевание» — освобождение геномных нуклеиновых кислот от белков оболочки. В результате этой процедуры вирусный геном становится доступным для ферментных систем клетки, обеспечивающих экспрессию генов вируса. Именно после проникновения вирусной геномной нуклеиновой кислоты в клетку заключенная в ней генетическая информация расшифровывается генетическими системами хозяина и используется для синтеза компонентов вирусных частиц. По сравнению с геномами других организмов вирусный геном относительно мал и кодирует лишь ограниченное число белков, в основном белки капсида и один или несколько белков, участвующих в репликации и экспрессии вирусного генома. Необходимые метаболиты и энергия поставляются хозяйской клеткой.

Геном вирусов, заключенный внутри вирионов, может быть представлен ДНК или РНК, последние могут быть одно- и двуцепочечными, кольцевыми и линейными. Значения молекулярных масс ДНК вирусов находятся в пределах $1 \cdot 10^6 — 200 \cdot 10^6$; РНК — $1 \cdot 10^6 — 15 \cdot 10^6$ дальтона.

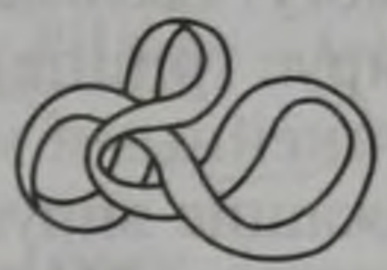
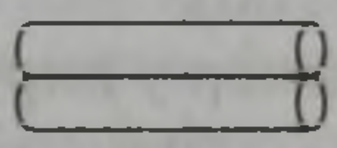
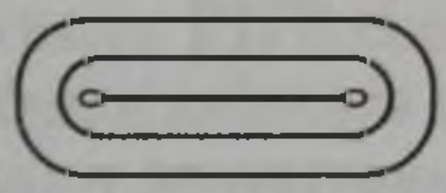
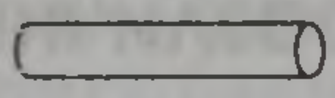
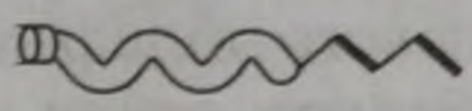
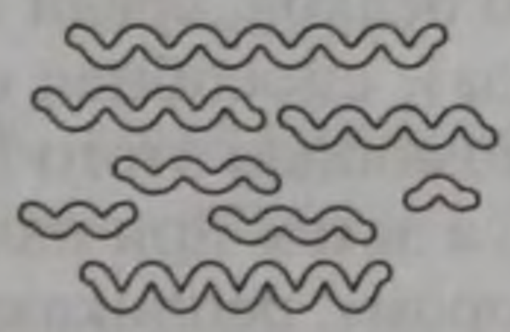
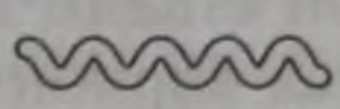
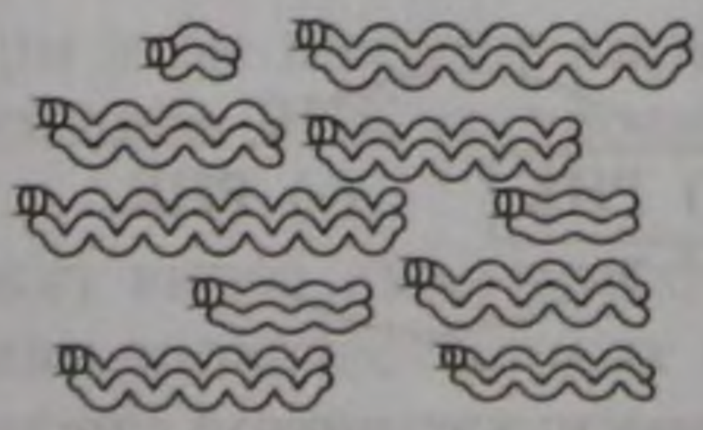
Вирус (примерный размер генома в тыс. н.п. или тыс. н.)	*ДНК-содержащие вирусы*	**Структура генома**
Вирус гепатита В(3)		Частично одноцепочечная кольцевая ДНК
Паповавирусы (5)		Сверхспиральная замкнутая кольцевая ДНК
Аденовирус (36) Вирус герпеса (200)		Линейный дуплекс
Вирус осповакцины (200)		Линейный дуплекс с ковалентно сшитыми концами
Парвовирус (2)		Одноцепочечная линейная ДНК
Фаг φX174 (5,3) Фаг M13 (6,0)		Одноцепочечное кольцо
	РНК-содержащие вирусы	
Ретровирус (10)		Одноцепочечный диплоидный геном
Вирус гриппа (16)		8 разных одноцепочечных молекул
Полиовирус (7) ВТМ (7000)		Одноцепочечная молекула
Реовирус (30)		10 разных двуцепочечных молекул

Рис. 60. Структура генома вирусов

По сложности строения генома вирусы широко варьируют — от фага Q_β (РНК-содержащий вирус бактерий), имеющего 4 гена, до вируса оспы (ДНК-содержащий вирус), геном которого насчитывает около 250 генов. Кроме того, все гены вирусов могут быть заключены в одной молекуле нуклеиновой кислоты или распределены по нескольким молекулам, которые вместе и составляют геном такого вируса. Например, у реовирусов геном представлен двуцепочечной РНК и состоит из 10 молекул (или сегментов). Геномы вирусов, содержащих одноцепочечную РНК, также могут быть либо цельными (ретровирус), либо сегментированными (вирус гриппа). Геном РНК-содержащих вирусов представлен только линейными молекулами РНК.

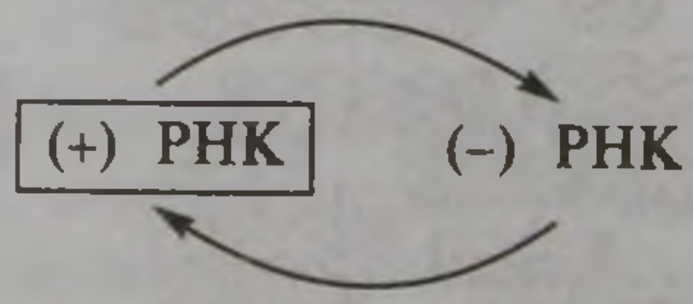
Все известные ДНК-содержащие вирусы позвоночных имеют геном, заключенный в одной молекуле ДНК, линейной или кольцевой, одно- или двуцепочечной (рис. 60).

У некоторых вирусов, например у вируса гепатита В, геном представлен кольцевой молекулой двуцепочечной ДНК, в обеих цепях которой в разных местах обнаружены одноцепочечные участки.

Типы генетического материала и механизм его репликации у различных вирусов

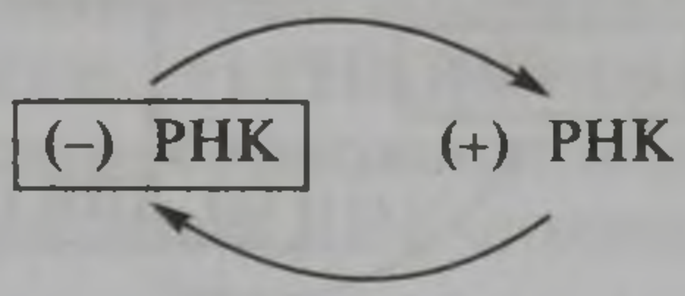
РНК-содержащие вирусы (РНК → РНК). Геномы почти всех известных РНК-содержащих вирусов — это линейные молекулы, их удобно разделить на три группы.

Первая группа — это однонитевые геномы положительной полярности, т.е. с нуклеотидной последовательностью, соответствующей таковой у мРНК. Такие геномы обозначают как (+) РНК. Однонитевый (+) РНК-геном характерен для фага Q_β, вирусов табачной мозаики, полиомиелита, клещевого энцефалита. Вирусные (+) РНК-геномы кодируют несколько белков, среди которых РНК-зависимая РНК-полимераза (репликаза), способная синтезировать молекулы РНК без участия ДНК. С помощью этого фермента синтезируются сначала (–) нити РНК фага, а затем при наличии особого белка, называемого «хозяйским фактором», репликаза осуществляет синтез (+) нити РНК. На заключительной стадии из накопившихся вирусных белков и (+) РНК формируются вирионы. Упрощенная схема этого процесса такова:



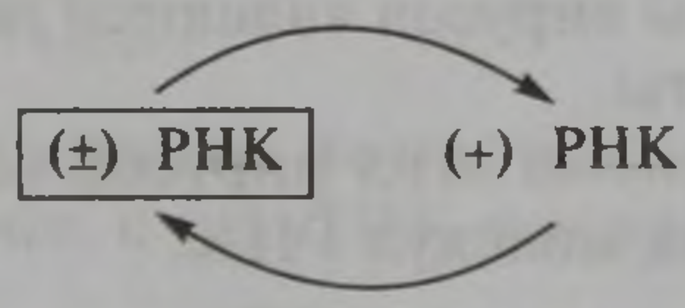
(В рамку заключена та форма нуклеиновой кислоты, которая присутствует в вирионе.)

Вторая группа — это однонитевые геномы с негативной полярностью, т.е. (–) РНК-геномы. К вирусам с негативным РНК-геномом относятся вирусы гриппа, кори, бешенства, желтой карликовости картофеля и др. Поскольку (–) РНК не может выполнять функции мРНК, для образования «своих» мРНК вирус внедряет в клетку не только геном, но и фермент, умеющий снимать с этого генома комплементарные копии по схеме:



Этот вирусный фермент (РНК-зависимая РНК-полимераза, синтезированная в предыдущем цикле размножения) упакован в вирионе в удобной для доставки в клетку форме. Инфекционный процесс начинается с того, что вирусный фермент копирует вирусный геном, образуя (+) РНК, которая выступает в качестве матрицы для синтеза вирусных белков, в том числе РНК-зависимой РНК-полимеразы, которая входит в состав образующихся вирионов.

Третью группу составляют двунитевые геномы, (±) РНК-геномы. Известные двунитевые геномы всегда сегментированы (т.е. состоят из нескольких разных молекул). Сюда относятся реовирусы. Их размножение проходит по варианту, близкому к предыдущему. Вместе с вирусной РНК в клетку попадает и вирусная РНК-зависимая РНК-полимераза, которая обеспечивает синтез молекул (+) РНК. В свою очередь (+) РНК обеспечивает производство вирусных белков на рибосомах хозяйской клетки и служит матрицей для синтеза новых (–) РНК-цепочек вирусной РНК-полимеразой:

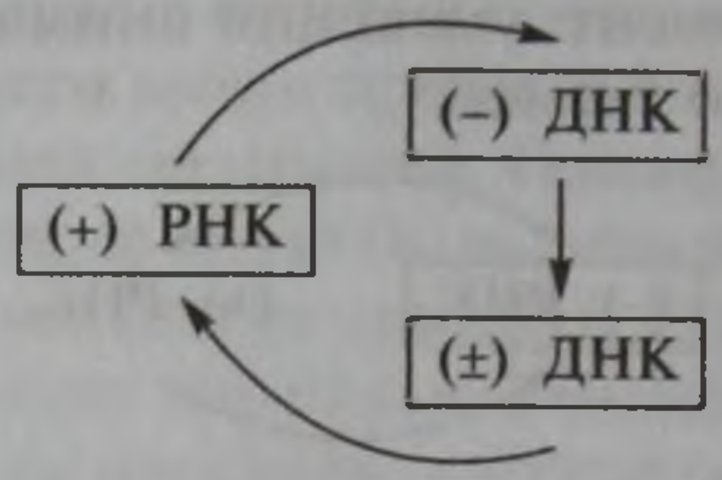


Цепочки (+) и (–) РНК, комплексируясь друг с другом, образуют двунитевый (±) РНК-геном, который упаковывается в белковую оболочку.

РНК-содержащие вирусы (РНК → ДНК → РНК). Сюда относятся вирусы, у которых цикл репликации генома можно разбить на две главные реакции: синтез РНК на матрице ДНК и синтез ДНК на матрице РНК. При этом в состав вирусной частицы в качестве генома может входить либо РНК (ретровирусы), либо ДНК (ретроидные вирусы).

В эту группу входят вирус иммунодефицита человека (ВИЧ) и некоторые возбудители злокачественных новообразований. Вирусная частица содержит две молекулы геномной одноцепочечной (+) РНК. В вирусном геноме закодирован необычный фермент (обрат-

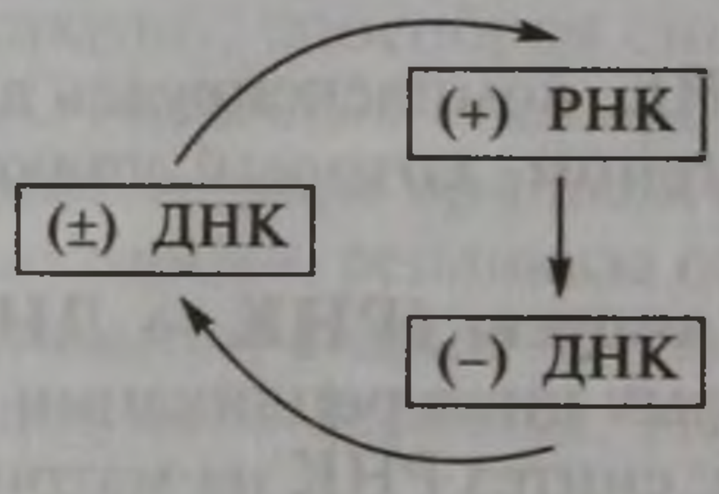
ная транскриптаза, или ревертаза), который обладает свойствами как РНК-зависимой, так и ДНК-зависимой ДНК-полимеразы. Этот фермент попадает в заражаемую клетку вместе с вирусной РНК и обеспечивает синтез ее ДНК-копии сначала в одноцепочечной форме [(–) ДНК], а затем в двуцепочечной [(±) ДНК]:



Вирусный геном в форме нормального дуплекса ДНК (так называемая провирусная ДНК) встраивается в хромосому клетки хозяина. В результате двуцепочечная ДНК вируса представляет собой в сущности дополнительный набор генов клетки, который реплицируется вместе с ДНК хозяина при делении. Для образования новых ретровирусных частиц провирусные гены (гены вируса в хромосомах хозяина) транскрибируются в ядре клетки транскрипционным аппаратом хозяина в (+) РНК-транскрипты. Одни из них становятся геномом нового «потомства» ретровирусов, а другие подвергаются процессингу в мРНК и используются для трансляции белков, необходимых для сборки вирусных частиц. К этой же группе относятся онкогенные ретровирусы: вирус саркомы Рауса (птиц), вирус лейкоза мышей и т. д. Развитие опухоли определяет один ген ретровируса — онкоген.

ДНК-содержащие вирусы. *Первая группа* — вирусы с двуцепочечной ДНК, репликация которых осуществляется по схеме: ДНК → РНК → ДНК. Они получили название *ретроидные вирусы*. Представителями этой группы вирусов являются вирус гепатита В и вирус мозаики цветной капусты.

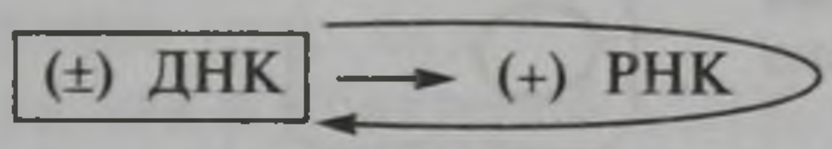
Репликация ДНК-генома этих вирусов осуществляется при посредстве промежуточных молекул РНК:



Молекулы РНК образуются в результате транскрипции вирусных ДНК в клеточном ядре хозяйским ферментом ДНК-зависимой РНК-полимеразой. Транскрибируется только одна из нитей вирусной ДНК. Синтез ДНК на РНК-матрице происходит в результате реакции, катализируемой обратной транскриптазой; сначала синтезируется (–) нить ДНК, а затем на вновь синтезированной (–) нити ДНК тот же фермент строит (+) нить.

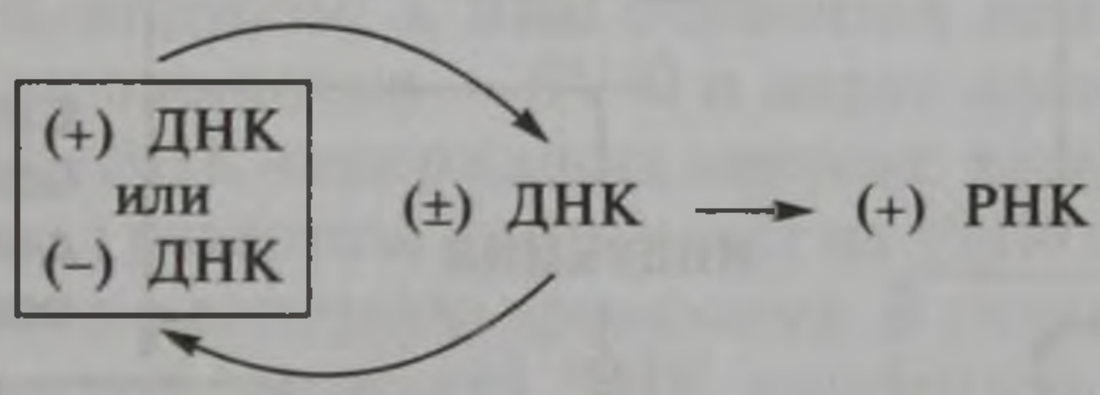
В целом общая схема репликации генома ретроидных вирусов поразительно похожа на схему репликации генома ретровирусов. По-видимому, данное сходство имеет под собой и эволюционную основу, так как первичная структура обратных транскриптаз этих вирусов выявляет определенное сходство между собой.

Вторая группа — вирусы с двуцепочечной ДНК, репликация которых осуществляется по схеме ДНК→ДНК. В зараженной клетке ДНК-зависимая РНК-полимераза транскрибирует с генома этих вирусов молекулы мРНК (т.е. (+) РНК), которые принимают участие в синтезе вирусных белков, а размножение вирусного генома осуществляет фермент ДНК-зависимая ДНК-полимераза:



В одних случаях производством как мРНК, так и ДНК занимаются клеточные ферменты; в других случаях вирусы используют собственные ферменты. Бывает, что те и другие ферменты обслуживают процесс репликации и транскрипции. К этой группе относятся вирусы герпеса, оспы и др.

Третья группа — вирусы с одноцепочечной ДНК, либо с негативной, либо с позитивной полярностью. Попав в клетку, вирусный геном сначала превращается в двуцепочечную форму, это превращение обеспечивает клеточная ДНК-зависимая ДНК-полимераза:



Транскрипция и репликация на последующих этапах происходит так же, как и для вирусов, с (±) ДНК-геномом.

Типы взаимодействия вируса с клеткой-хозяином

Существуют два типа взаимодействия вируса и клетки, где главное различие — степень автономии вируса от клетки-хозяина. Это детально исследовано на примере бактериофага λ.

Когда бактериофаг λ заражает подходящую клетку кишечной палочки, он немедленно реплицируется и образует несколько сотен дочерних фаговых частиц, которые выходят наружу в момент лизиса клетки. Лизис клеточной стенки осуществляет кодируемый фагом фермент лизоцим. Это так называемый *литический* путь инфек-

ции — он заканчивается гибелью клетки-хозяина. В данном случае вирус не подчиняется клеточному контролю.

Альтернативным является *лизогенный* путь инфекции, при котором линейные инфицирующие молекулы ДНК фага λ замыкаются в кольцо и включаются (интегрируют) в кольцевую хромосому клетки-хозяина. Эта интеграция осуществляется путем сайт-специфической генетической рекомбинации, катализируемой особым белком фага λ, так называемой *интегразой*. Клетка, содержащая ДНК фага λ в своей хромосоме, называется *лизогенной* клеткой, а встроенная ДНК

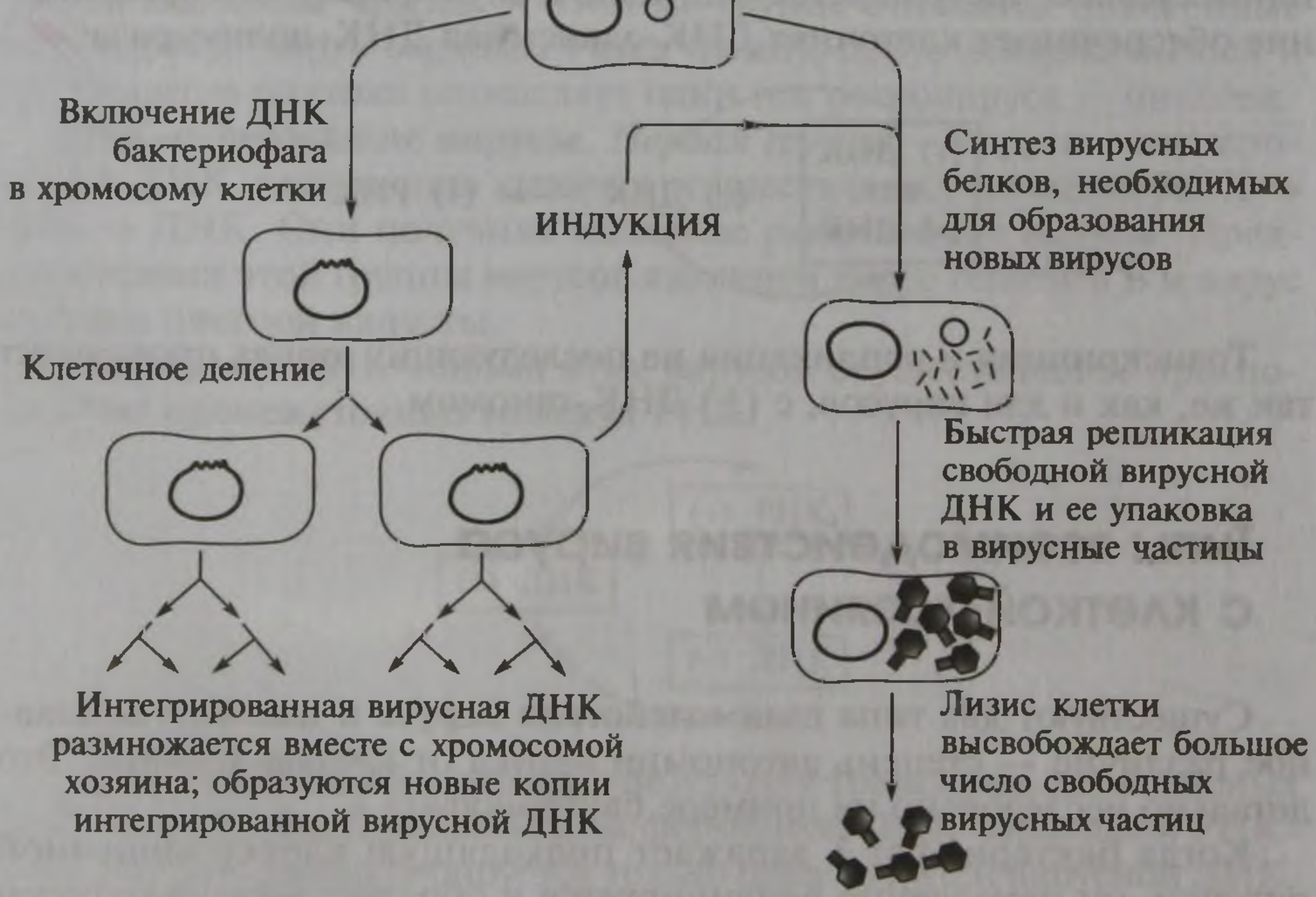


Рис. 61. Жизненный цикл бактериофага λ

фага λ — *профагом*. Клетки и их хромосомы делятся, а вместе с хромосомами в каждую дочернюю клетку попадают и фаговые гены. При определенных обстоятельствах, а именно при воздействии какого-нибудь повреждающего внешнего фактора (например, ультрафиолетовых лучей, ионизирующего излучения, химических реагентов), активируются так называемые «SOS»-функции, предназначенные для устранения нарушений. В лизогенной клетке это приводит к выщеплению (эксцизии) профага из клеточной ДНК и началу обычного, свойственного фагу цикла размножения, т. е. отщепившийся профаг вступает в литический цикл репликации. Интегрированный профаг, таким образом, не погибает вместе с поврежденной клеткой-хозяином, а у него есть шанс перейти в соседнюю неповрежденную клетку кишечной палочки (рис. 61).

В животных клетках, как и у бактерий, ДНК-вирусы могут размножаться литическим путем, ведущим к гибели клетки. Такие животные клетки называют *пермиссивными*. Клетки, в которых размножение вирусов блокируется, называются *непермиссивными*; вирусная хромосома в таких случаях либо включается в геном клетки-хозяина и затем размножается вместе с ним, либо образует плазмиду — кольцевую молекулу ДНК, репликация которой регулируется и не ведет к гибели клетки. Иногда это вызывает в непермиссивных клетках определенное генетическое изменение, в результате которого начинается их неконтролируемый рост, т. е. нормальные клетки превращаются в раковые. Соответствующий ДНК-вирус называют в таких случаях опухолевым ДНК-вирусом; к ним относятся наиболее изученные представители паповавирусов — SV40 и вирус полиомы.

В ретровирусах (РНК-содержащих вирусах: вирус саркомы Рауса, вирус СПИДа и т. д.) ревертаза превращает их РНК в форму ДНК, которая встраивается в клеточную хромосому. В результате транскрипции интегрированной вирусной ДНК клеточной ДНК-зависимой РНК-полимеразой образуется большое количество молекул вирусной РНК, идентичных исходному инфицирующему геному. Завершается процесс трансляцией этих молекул РНК с образованием капсидных и мембранных белков, а также обратной транскриптазы; наконец, происходит сборка новых вирусных частиц, включающих вирусный РНК-геном и обратную транскриптазу. Вирусные частицы покидают клетку, не убивая ее. Возникает особая форма хронической инфекции, при которой работающий вирусный геном, включенный в состав клеточной хромосомы, передается дочерним клеткам. Такое сосуществование не обходится бесследно для клетки. Свойства ее под влиянием находящихся в ней вирусных белков могут драматически меняться, и в некоторых случаях зараженная клетка становится раковой.

Еще один важный результат был получен при исследовании ретровирусов — создание на их основе рекомбинантных векторов. Эти векторы оказались особенно ценными для получения трансгенных животных.

Характеристика некоторых вирусов

Фаг λ. ДНК фага λ упакована в гексагональную головку, образованную несколькими белками (рис. 62, *А*). К головке прикреплен хвостовой отросток, построенный из других белков. Белки головки и хвостового отростка кодируются геномом фага λ и синтезируются в клетках бактерий, инфицированных фагом. Фаг прикрепляется к клеточной стенке и вводит ДНК в бактериальную клетку подобно шприцу.

ДНК фага λ — линейная двуцепочечная молекула, содержащая 48 502 н. п. Осуществлено картирование практически каждого из его 61 гена, полностью определена нуклеотидная последовательность его генома.

Молекула ДНК фага λ содержит область начала репликации (ori), гены структурных вирусных белков (головки и хвостового отростка) и ферментов, участвующих в репликации ДНК, лизисе инфицированных клеток и установлении лизогении (рис. 62, *В*).

Приблизительно 1/3 генома, представленная непрерывным внутренним сегментом, находящемся на карте между точками 20 и 38 тыс. н. п., необходима для установления лизогенного состояния. У всех векторов, сконструированных на основе ДНК фага λ, многие или даже все гены, требующиеся для установления лизогенного состояния, удаляются, и в этих случаях инфицирование фагом λ кишечной палочки сопровождается их лизисом.

Во время своего жизненного цикла λ-геном может находиться в разных формах. Внутри фаговой частицы ДНК находится в линейной форме: 5′-конец каждой ее цепи представляет собой одноцепочечную последовательность из 12 нуклеотидов. Эти последовательности называются «*липкими концами*», так как они взаимно комплементарны и могут спариваться друг с другом. При попадании в клетки *E. coli* геном фага λ замыкается в кольцо и сшивается с помощью бактериальной ДНК-лигазы (рис. 62, *Б*).

Места спаривания липких концов получили название *cos*-сайтов. Кольцевая форма обязательна для осуществления репликации фаговой ДНК, она также служит предшественником при лизогенной интеграции ДНК фага в геном хозяина. После интеграции геном фага λ снова становится линейным, однако благодаря удаленности сайта интеграции на фаговой ДНК (*att*P) от *cos*-сайта, интегрированная линейная ДНК профага оказывается пермутированной (измененной с перестановкой) относительно ДНК, находящейся в фаговой частице.

На основе фага λ сконструировано множество разных λ-векторов, которые включают вставки разных размеров; максимальный — немногим более 24 тыс. н. п.

Фаг φX174. Геном этого фага представлен кольцевой одноцепочечной ДНК, содержащей 5 386 нуклеотидов. Нуклеотидная после-

довательность ДНК этого мелкого фага была установлена в 1977 г. Ф. Сангером и его коллегами в Кембридже (Англия). Это достижение знаменовало собой начало новой эры в молекулярной биологии и принесло Сангеру вторую Нобелевскую премию. Важным результатом

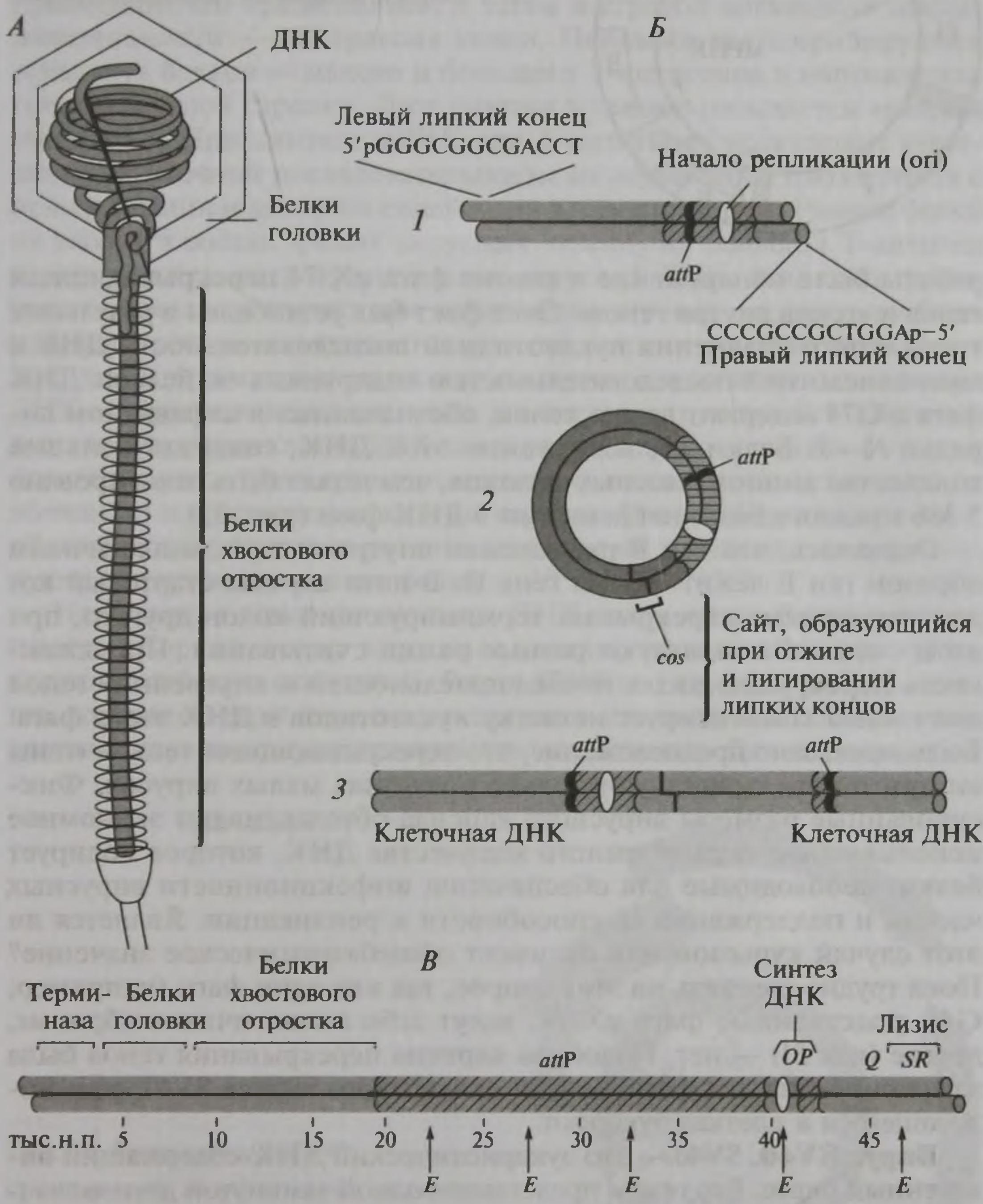


Рис. 62. Бактериофаг λ (М. Сингер, П. Берг, 1998):

А — строение; *Б* — формы генома: *1* — линейная; *2* — кольцевая; *3* — интегрированная; *att*P-сайт интеграции; *В* — структура генома: заштрихованы гены, необязательные для литической инфекции; светлый овал — область начала репликации; *E* — *Eco*RI-сайты (*E* — *Eco*RI-сайты (сайт расщепления ДНК рестриктазой EcoRI))

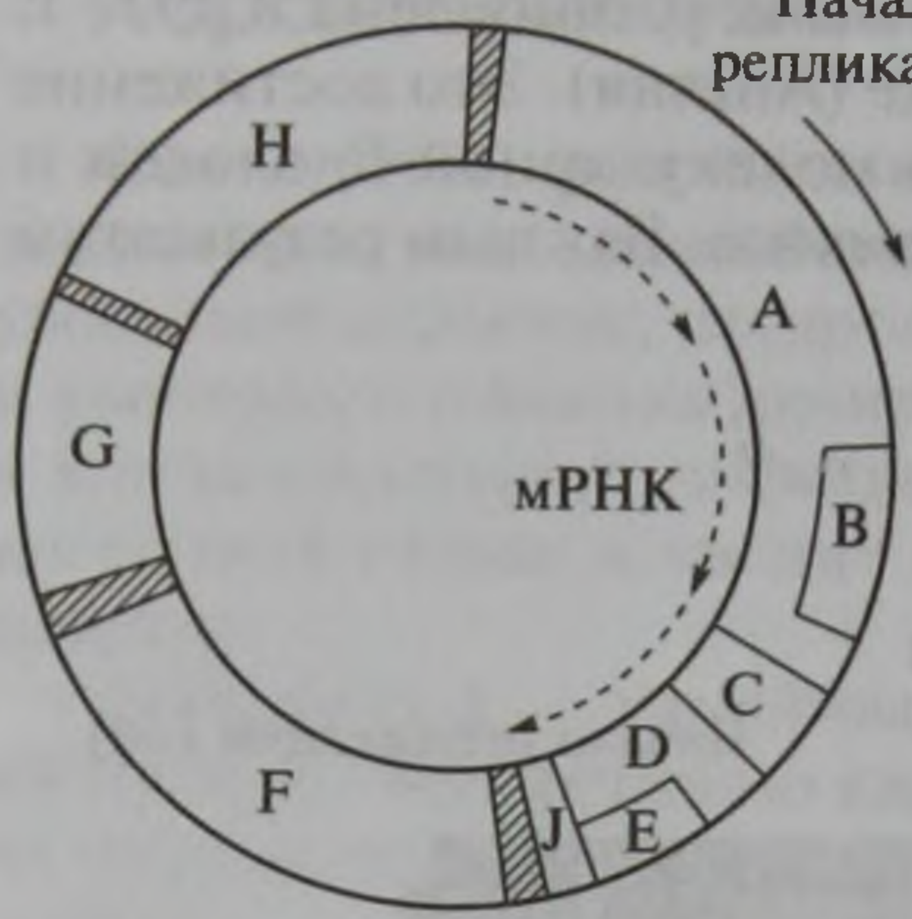


Рис. 63. Генетическая карта фага φX174.

Буквами обозначены гены. Ген B расположен внутри гена A, а ген E внутри гена D. Штриховкой показаны участки, которые не транскрибируются

работы было обнаружение в геноме фага φX174 перекрывающихся генов и «генов внутри генов». Этот факт был установлен в результате тщательного сравнения нуклеотидной последовательности ДНК с аминокислотной последовательностью кодируемых ею белков. ДНК фага φX174 содержит девять генов, обозначенных в алфавитном порядке A — J. Белки же, кодируемые этой ДНК, содержат большее количество аминокислотных остатков, чем может быть закодировано 5 386 нуклеотидами, найденными в ДНК фага (рис. 63).

Оказалось, что ген B расположен внутри гена A, аналогичным образом ген E лежит внутри гена D. В пяти случаях стартовый кодон одного гена перекрывает терминирующий кодон другого, при этом сигналы используют разные рамки считывания. Протяженность перекрывающихся последовательностей и внутренних генов полностью компенсирует нехватку нуклеотидов в ДНК этого фага. Было высказано предположение, что перекрывающиеся гены и «гены внутри генов» существуют только в геномах малых вирусов. Фиксированные размеры вирусного капсида обуславливают экономное использование ограниченного количества ДНК, которое кодирует белки, необходимые для обеспечения инфекционности вирусных частиц и поддержания ее способности к репликации. Является ли этот случай курьезом или он имеет общебиологическое значение? Пока трудно ответить на этот вопрос, так как одни фаги (например, G4), родственные фагу φX174, ведут себя аналогичным образом, другие (как fd) — нет. Подобная картина перекрывания генов была зафиксирована для фага λ и для опухолевого вируса SV40, размножающегося в клетках эукариот.

Вирус SV40. SV40 — это эукариотический ДНК-содержащий онкогенный вирус. Его геном представлен одной замкнутой двуцепочечной молекулой ДНК. ДНК в составе вирусной частицы находится в комплексе с четырьмя клеточными гистонами и образует кольцевую, содержащую нуклеосомы минихромосому, уложенную в компактную структуру и окруженную внешней белковой оболочкой (капсидом). Природным хозяином SV40 является макака-резус.

Расшифрована нуклеотидная последовательность всего генома SV40, состоящая из 5 243 н.п., и установлены участки, кодирующие пять вирусных белков: малый и большой Т-антигены (от англ. *tumour* — опухоль), VP1, VP2 и VP3 (рис. 64).

Вирус SV40, проникнув в клетку, направляется к ядру. Здесь происходит его «раздевание», а затем в строгой временной последовательности — экспрессия генов. Первыми транскрибируются гены двух белков — малого и большого Т-антигенов в направлении против часовой стрелки. Этот участок в геноме называется «ранней областью». При синтезе мРНК для Т-антигенов происходит вырезание вставочной последовательности из первичного транскрипта с использованием аппарата сплайсинга клеточного ядра. Ранние белки не входят в состав зрелых вирусных частиц, но большой Т-антиген играет важную роль в репликации вирусной ДНК, выполняя функцию хеликазы. Репликация начинается в точке ori, идет в обоих направлениях и заканчивается в месте встречи репликативных вилок; затем происходит сшивание дочерних двуцепочечных молекул ДНК с образованием замкнутых кольцевых структур. Вскоре после инициации репликации ДНК замедляется синтез «ранних» белков и активизируется транскрипция «поздних» генов, которая начинается вблизи ori и протекает по часовой стрелке. Три мРНК, по-видимому, образуются в результате различных реакций сплайсинга и транслируются в соответствующие белки — VP1, VP2 и VP3.

Накопление дочерних молекул ДНК, вирионных белков, завершается сборкой новых вирусных частиц, их высвобождением из клетки и гибелью последней. Вирус SV40 не реплицируется в клетках грызунов, его геном встраивается в клеточную ДНК, вызывая злокачественное перерождение клетки-хозяина. Для поддержания

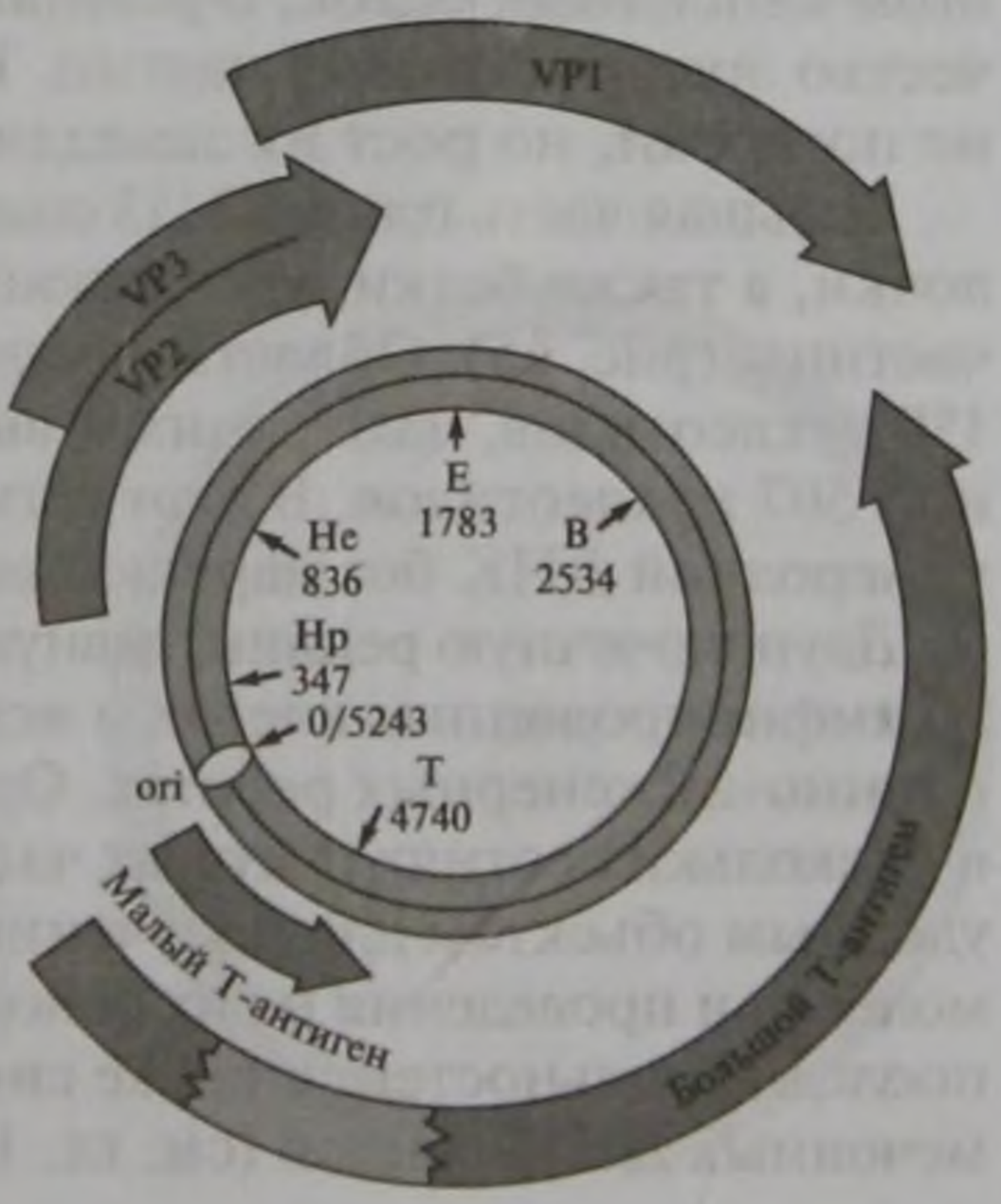


Рис. 64. Геном вируса SV40.

С внутренней стороны карты указано число пар оснований; Hp, T, E, B, He — уникальные рестриктазные сайты. Участки, кодирующие пять вирусных белков, вынесены за пределы карты. Большой Т-антиген кодируется двумя отдельными группами генов

трансформированного состояния необходимы оба вирусных антигена (экспрессия «ранней» области генома вируса), «поздняя» область не участвует в этом процессе.

Таким образом, ограниченное количество генетической информации у вируса SV40 используется с максимальным эффектом (перекрывание генов, сплайсинг и т. д.); кроме того, новосинтезированная ДНК связывается с гистонами клетки-хозяина и образует сверхспирализованный комплекс, который затем упаковывается в капсулу с помощью вирусных белков (VP1, VP2 и VP3).

На основе вируса SV40 конструируют различные типы векторов: это трансдуцирующие, плазмидные, пассивные трансформирующие SV40 векторы.

Фаг M13. Фаг M13 входит в группу нитчатых бактериофагов *E. coli*. Его геном представлен одноцепочечной кольцевой молекулой ДНК, содержащей 6 400 нуклеотидов. Он упакован в напоминающий пробирку капсид, построенный из трех разных кодируемых фагом белков. Фаг M13 инфицирует клетки *E. coli* в том случае, если последние имеют тонкие, напоминающие волоски структуры, называемые F-пилями, которые необходимы для прикрепления фага к клетке. В процессе репликации фага M13 в клетках *E. coli* одноцепочечная ДНК превращается в двуцепочечную кольцевую форму, т. е. синтезируется комплементарная инфицирующей цепи ДНК. Двуцепочечная ДНК детерминирует синтез примерно 300 своих копий, которые затем используются в качестве матриц при образовании одноцепочечных фаговых ДНК, упаковывающихся в фаговые частицы. Одновременно транскрибируются и транслируются фаговые гены, а на внутренней мембране клетки *E. coli* концентрируется основной капсидный белок. Зрелые одноцепочечные молекулы фаговой ДНК проходят через мембрану, захватывая при этом капсидный белок, в результате чего образуется большое количество зрелых фаговых частиц. Инфицированные клетки при этом не погибают, но рост их замедляется.

Большая часть генома M13 содержит гены, кодирующие белки оболочки, а также белки, участвующие в репликации и сборке фаговой частицы (рис. 65). Область начала репликации состоит примерно из 150 нуклеотидов, находящихся внутри некодирующего сегмента длиной 507 нуклеотидов. В этот сегмент может быть помещена вставка чужеродной ДНК без нарушения способности M13 к репликации.

Двуцепочечную репликативную форму ДНК можно легко выделить из инфицированных клеток и использовать затем в качестве вектора в генно-инженерных работах. Одноцепочечная ДНК, содержащаяся в нескольких сотнях фаговых частиц культуральной среды, является удобным объектом для получения одноцепочечных рекомбинантных молекул и проведения на их основе детального анализа нуклеотидных последовательностей, а также синтеза специфических радиоактивно меченных ДНК-зондов (см. гл. 14).

Ретровирусы — это обширная группа вирусов, представители которой различаются как по биологическим свойствам (например, по способности вызывать злокачественные опухоли), так и по морфологии; имеются и заметные вариации в структуре их геномов. Общие черты этих геномов следующие:

1) вирусная РНК — одноцепочечная молекула длиной до десяти тысяч нуклеотидов, представляет собой (+) нить, молекулы РНК кэпированы на 5′-конце и полиаденилированы на 3′-конце;

2) в молекулах РНК есть прямой концевой повтор (R) длиной несколько десятков нуклеотидов (на 3′-конце он предшествует поли(А)-тракту);

3) в вирусной РНК записана информация для синтеза по крайней мере трех групп вирус-специфических белков: структурных белков сердцевины вириона (Gag-белков), ферментативных белков, принимающих участие в обратной транскрипции вирусного генома и в интеграции вирус-специфической ДНК в клеточную хромосому (продуктов гена pol) и белков, входящих в состав наружной липопротеиновой оболочки вириона (Env-белков). У некоторых ретровирусов есть дополнительные гены;

4) вирусная частица содержит две молекулы геномной РНК; таким образом, редкой (если не уникальной) особенностью ретровирусов является диплоидность их генома.

Фундаментальные исследования ретровирусов, вызывающих опухоли у птиц и млекопитающих, указали путь к решению насущных клинических проблем, связанных с раковыми заболеваниями человека и синдромом приобретенного иммунного дефицита — СПИДом. СПИД вызывает один из ретровирусов ВИЧ (вирус иммунодефицита человека), который инфицирует макрофаги и один из типов Т-лимфоцитов, выводя из строя иммунную систему больного. Другие лимфотропные ретровирусы человека также инфицируют Т-клетки, но вызывают лейкозы.



Рис. 65. Структура генома фага М13.

Гены обозначены цифрами: гены 2 и 5 ответственны за репликацию, остальные детерминируют образование капсида и сборку вирусной частицы. Светлый овал — область начала репликации; ■ — некодирующий участок

Структура вируса иммунодефицита человека представлена на рис. 66. Вирусная частица имеет сферическую форму, ее диаметр около 100 нм. Оболочка частицы образована двойным слоем липидов,

Таблица 7. **Белки, закодированные в геноме ВИЧ-1**

Ген	Белок	Функция
	Белки вириона	
	MA (p17)	Матричный белок; связывание с мембраной, сборка вируса
gag	CA (p24)	Белок капсида; сборка вируса
	NC (p9)	Белок нуклеокапсида; РНК-связывание; сборка вируса
	PR (p11)	Протеаза; созревание вируса
pol	RT (p66/p51)	Обратная транскриптаза; репликация вируса
	IN (p32)	Интеграза; репликация вируса
env	SU (gp120)	Вирусный белок оболочки; связывание CD 4-рецептора, инфекционность вируса
	TM (gp41)	Трансмембранный гликопротеин оболочки; инфекционность вируса
	Вспомогательные белки	
vpr	Vp*r*	Вирусный белок R; перемещение предынтеграционного комплекса в ядро
vif	Vif	Фактор вирусной инфекционности; предынфекционные события
vpu	Vpu	Вирусный белок U; созревание вирусной оболочки и выход вируса
	Регуляторные белки	
tat	Tat	Трансактиватор; стимулирует элонгацию транскрипции
rev	Rev	Регулятор экспрессии поздних генов; стимулирует появление несплансированных мРНК в цитоплазме
nef	Nef	Отрицательные регуляции CD4; регулирует сигнальные пути Т-клеток

пронизанных гликопротеинами. Липидная мембрана происходит из внешней мембраны клетки-хозяина. В каждой гликопротеиновой молекуле имеются две субъединицы: gp41, пронизывающая мембрану, и gp120, находящаяся снаружи. Внутри оболочки заключена сердцевина вируса, образованная белками p17 и p24. В сердцевине вирусной частицы находятся вирусная РНК и молекулы обратной транскриптазы. В настоящее время расшифрована структура генома ВИЧ-1. Белки, закодированные в геноме ВИЧ, представлены в табл. 7.

Из данных таблицы видно, что геном ВИЧ содержит 9 генов, но процессинг (клеточными и вирусными ферментами) генерирует 14 белковых продуктов. Ген *gag* кодирует внутренние белки вириона; они синтезируются в виде полипротеина, из которого потом протеаза вырезает конкретные белки. Следующий кодирующий участок, *pol*, детерминирует синтез обратной транскриптазы. Фермент полифункционален: он способен синтезировать ДНК, используя в качестве матрицы либо РНК, либо ДНК; кроме того, он ведет себя подобно РНКазе Н (т.е. разрушает цепь РНК, входящую в состав РНК/ДНК дуплекса) и обладает эндонуклеазной активностью. Ген *pol* кодирует и интегразу, осуществляющую включение ДНК вируса в геном зараженной клетки. Ген *env* кодирует компоненты вирусной оболочки, которая захватывает часть клеточной цитоплазматической мембраны. В геноме ВИЧ представлены также гены вспомогательных и регуляторных белков. Так, белки Tat и Rev совместно обеспечивают значительное увеличение синтеза вирусных белков при переходе провируса из латентной в активную форму.

Репликативный цикл ВИЧ. После взаимодействия вирусной частицы с чувствительной клеткой (Т-лимфоцитом) сердцевина вириона освобождается от липопротеиновой оболочки и оказывается в цитоплазме Т-лимфоцита. Здесь обратная транскриптаза превращает вирусную РНК в линейную двуцепочечную молекулу ДНК. Образовавшаяся молекула вирусной ДНК длиннее молекулы вирусной РНК, которая послужила матрицей для ее синтеза. Действительно, к 5′-концу вирусной ДНК добавляется последовательность U3, а к 3′-концу этой цепи — последовательность U5 (см. рис. 115). В результате на концах молекулы вирус-специфической ДНК появился длинный концевой повтор (ДКП), имеющий структуру U3RU5. На следующем этапе линейная молекула ДНК проникает в ядро, где превращается в кольцевую молекулу (рис. 66) за счет взаимодействия между ее «тупыми» концами. При этом одна форма кольцевой ДНК содержит одну копию ДКП, другая — две тандемные копии ДКП. Одна или несколько копий ДНК интегрируют в геном лимфоцита.

Интеграция ДНК ВИЧ в геном зараженной клетки является ключевой стадией репликативного цикла этого ретровируса. Она осуществляется вирусным ферментом интегразой, кодируемым геном *pol*. Интеграза узнает строго определенные нуклеотидные последовательности, локализованные в U3 и U5 ДКП вирусной ДНК,

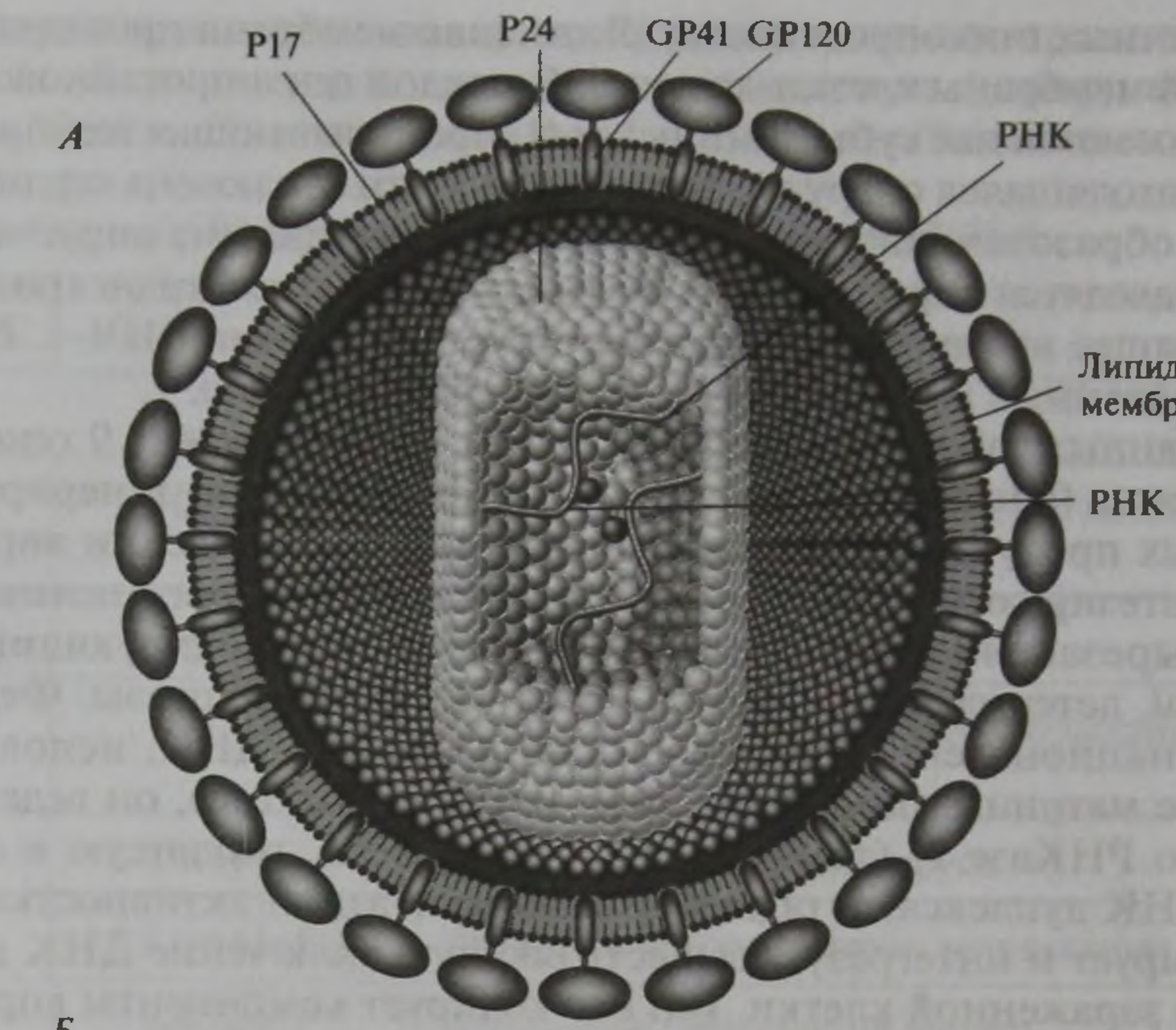


Рис. 66. Вирус иммунодефицита человека:

А — структура вируса: GP — гликопротеины; P — белки сердцевины вируса; • — обратная транскриптаза; *Б* — структура РНК-генома ретровируса (расшифровка генов дана в табл. 7); *В* и *Г* — продукты, образующиеся в результате обратной транскрипции РНК вируса: *В* — линейный дуплекс ДНК с последовательностью U3RU5 (длинные концевые повторы — ДКП); *Г* — два кольцевых дуплекса, производных формы В: *1* — с одним ДКП, *2* — с двумя ДКП

связывается с ними и катализирует реакцию 3′-концевого процессинга. Эта реакция представляет собой эндонуклеазное расщепление вирусной ДНК, в результате которого с 3′-конца каждой цепи удаляется динуклеотид GT (рис. 67). Затем интеграза осуществляет реакцию трансэтерификации, которая включает нуклеофильную атаку 3′-гидроксильной группы процессированных цепей ДНК на межнуклеотидный фосфат клеточной ДНК с образованием ковалентного продукта. Важно отметить, что встраивание вирусной ДНК может осуществляться в любое место клеточной ДНК, т.е. процесс (называемый переносом цепи) не зависит от нуклеотидной последовательности. Обычно реакции трансэтерификации подвергаются межнуклеотидные связи, расположенные в различных цепях ДНК на расстоянии 5 н.п. одна от другой. При завершении интеграции происходят процессинг 5′-концов вирусной ДНК, полимеразная достройка пяти недостающих нуклеотидов и лигирование, осуществляемое клеточными белками.

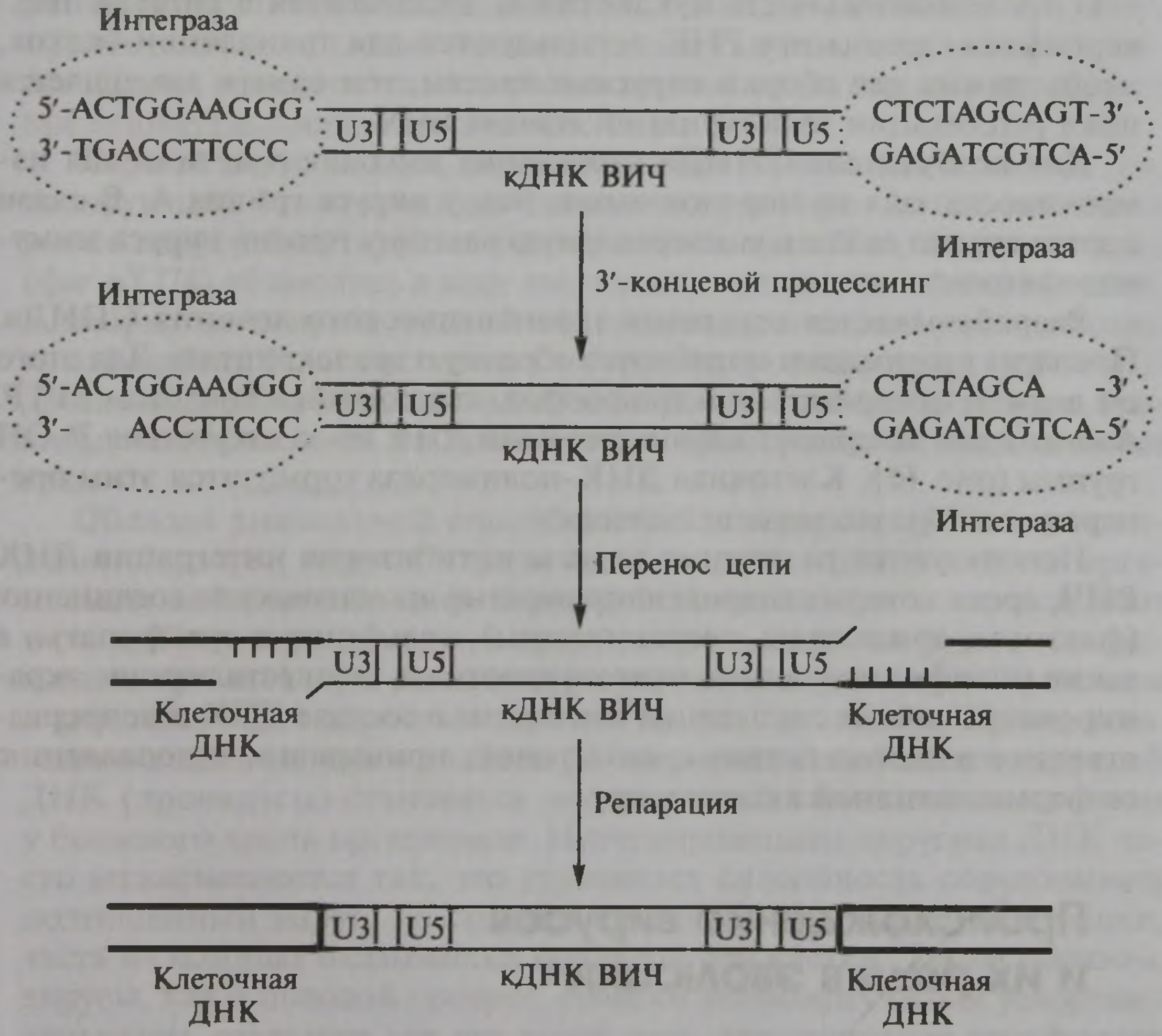


Рис. 67. Механизм интеграции ДНК вируса иммунодефицита в клеточную ДНК

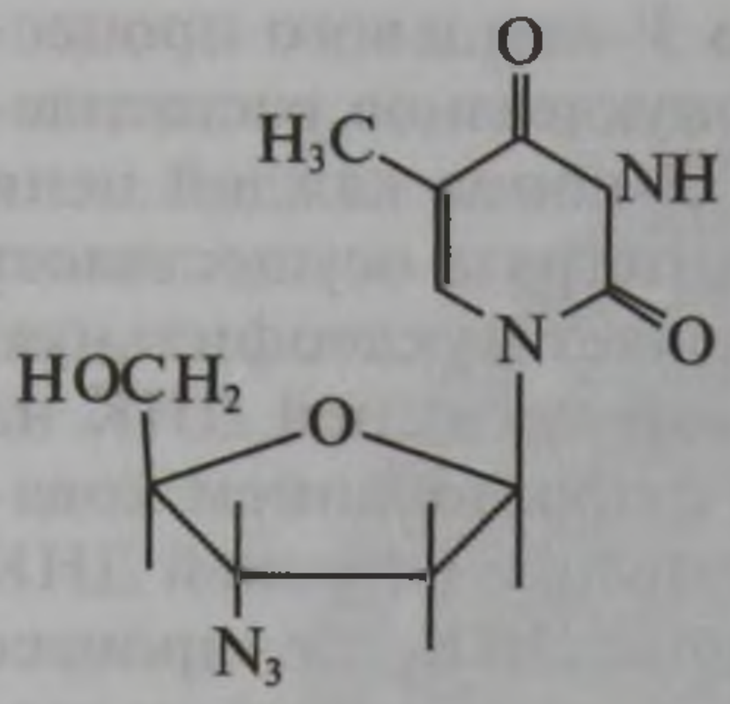


Рис. 68. Азотимидин

Синтез РНК ретровирусов идет тогда, когда геном вируса в виде провирусной ДНК является интегральной частью клеточной хромосомы, и осуществляется в ядре клеточным транскрипционным аппаратом; в качестве основного фермента используется РНК-полимераза II.

Провирус длиннее РНК, поэтому транскрипция начинается и оканчивается внутри последовательностей ДКП: промотор расположен в районе U3, терминатор — между последовательностями R и U5.

Первичные транскрипты провируса подвергаются обычным посттранскрипционным модификациям: кэпированию 5′-конца, полиаденилированию 3′-конца (в районе R есть сигнал полиаденилирования) и сплайсингу. Последняя модификация затрагивает не все РНК вируса. Молекулы РНК, сохранившие всю последовательность нуклеотидов, включаются в вирион, подвергшиеся сплайсингу РНК используются для трансляции белков, необходимых для сборки вирусных частиц, тем самым завершается цикл репликации транскрипции генома ВИЧ.

Для возбудителя СПИДа характерна высокая генетическая изменчивость: она на порядок выше, чем у вируса гриппа А. В связи с этим трудно создать универсальную вакцину против вируса иммунодефицита человека.

Разрабатывается стратегия терапевтического лечения СПИДа. Препарат азотимидин ингибирует обратную транскриптазу. Для этого он должен превратиться в трифосфат, являющийся аналогом дТТР. Азотимидин завершает элонгацию цепи ДНК из-за отсутствия 3′-ОН группы (рис. 68). Клеточная ДНК-полимераза тормозится этим препаратом в гораздо меньшей степени.

Используются различные классы ингибиторов интеграции ДНК ВИЧ, среди которых гидроксилированные ароматические соединения (флавоны, ариламиды, тетрациклины), сульфоны и сульфонаты, а также модифицированные олигонуклеотиды, осуществляющие экранирование сайтов связывания интегразы в составе ДНК и непосредственное взаимодействие с интегразой, приводящее к подавлению ее ферментативной активности.

Происхождение вирусов и их роль в эволюции

Известно, что даже самые крупные вирусы не способны осуществлять биосинтетические процессы самостоятельно — они зависят в

этом отношении от клеток-хозяев. Таким образом, клетки должны были появиться в процессе эволюции раньше вирусов. Предшественниками вирусов, вероятно, были небольшие фрагменты нуклеиновых кислот, которые приобрели способность размножаться независимо от хромосом их клеток-хозяев. Такими независимо размножающимися элементами являются плазмиды. Они встречаются в ДНК- и РНК-форме. Плазмиды не имеют белковой оболочки и не могут свободно переходить из клетки в клетку.

Возможно, что первые РНК-плазмиды напоминали собой вироиды, встречающиеся в некоторых растительных клетках. Эти небольшие кольцевые молекулы РНК (не более 300—400 нуклеотидов) размножаются, хотя не кодируют никаких белков. Вироиды проявляют свойства инфекционных патогенов растений. Они переходят от растения к растению только в том случае, когда и донорная клетка, и клетка-реципиент оказываются поврежденными, когда между ними не существует мембранного барьера, который вироид не способен преодолеть. Под давлением естественного отбора такие независимо реплицирующиеся элементы могли, очевидно, включать в себя те нуклеотидные последовательности клетки-хозяина, которые облегчали их самостоятельное размножение, в том числе и некоторые последовательности, кодирующие белки.

Первый вирус возник, вероятно, когда у плазмиды появился ген, кодирующий белок капсида. Число генов, которое может заключать в себе вирус, лимитируется размерами капсида. Так, мелкие вирусы (фаг φX174) обзавелись в ходе эволюции перекрывающимися генами, в которых часть нуклеотидной последовательности, кодирующей один какой-нибудь белок, используется для кодирования второго белка. У других вирусов в процессе эволюции возникли более крупные капсиды, что могло дать им возможность приобретать новые полезные гены.

Обладая уникальной способностью к переносу ДНК через видовые барьеры, вирусы играли важную роль в эволюции тех организмов, которые они инфицируют. Многие вирусы часто вступают в рекомбинацию как друг с другом, так и с хромосомами клеток-хозяев; при этом они захватывают случайные фрагменты хромосом и переносят их в другие клетки или другие организмы. Кроме того, включившиеся в геном хозяина (интегрированные) копии вирусной ДНК (провирусы) становятся постоянными компонентами генома у большого числа организмов. Интегрированная вирусная ДНК часто видоизменяется так, что утрачивает способность образовывать полноценный вирус, но сохраняет способность кодировать белки, часть из которых оказывается полезной для клетки. Таким образом, вирусы, как и половой процесс, создают возможности для ускорения эволюции, открывая для нее такой путь, как смешение генофондов различных организмов.

Глава 5
ГЕНОМ ПРОКАРИОТ

Основной чертой молекулярной организации прокариот является отсутствие в их клетках ядра, отгороженного ядерной мембраной от цитоплазмы. Отсутствие ядра является лишь внешним проявлением особой организации генома у прокариот.

Геном прокариот построен очень компактно. Количество некодирующих последовательностей нуклеотидов минимально. Многие механизмы регуляции экспрессии генов, использующиеся у эукариот, никогда не встречаются у прокариот. Простота строения генома прокариот объясняется их упрощенным жизненным циклом.

Структура бактериальной хромосомы

Пятьдесят лет назад все хромосомы прокариот считались линейными. Но в 1956 г. Ф. Жакоб и Е. Вильман предложили кольцевую модель бактериальной хромосомы. С тех пор эта модель считалась общепринятой до появления метода прямого анализа физической структуры хромосом — электрофореза в пульсирующем поле. В 1989 г. впервые этим методом была идентифицирована линейная бактериальная хромосома у возбудителя клещевого спирохетоза *Borrelia burgdorferi*. Вскоре было обнаружено, что линейная и кольцевая хромосомы сосуществуют одновременно у *Agrobacterium tumefaciens*, а у грамположительных бактерий рода *Streptomyces*, обладающих одним из самых больших бактериальных геномов (~8 000 тыс. н.п.), имеется одна линейная хромосома. Линейные хромосомы у бактерий часто сосуществуют с линейными плазмидами и широко распространены в природе. Состав геномов некоторых бактерий приведен в табл. 8.

Размер генома у различных бактерий колеблется от 580 тыс. н.п. у *Micoplasma genitalium* до 9 500 тыс. н.п. у *Myxococcus xanthus*. Для кишечной палочки *E. coli* размер генома составляет 4 600 тыс. н.п. (молекулярная масса около $3 \cdot 10^9$ Да, длина молекулы около 1,5 мм). В бактериальных клетках хромосома сильно компактизована. Так, кольцевая молекула ДНК *E. coli* длиной ~1,5 мм заключена в клетку, имеющую форму палочки диаметром 1 мкм и длиной 2 мкм.

Таблица 8. **Состав сложных геномов бактерий**

Бактерии	Геном
Agrobacterium tumefacient C58	1 линейная хромосома; 1 кольцевая хромосома; 2 плазмиды
Bacillus cereus F0836 76	1 кольцевая хромосома; 1 мегаплазмида
Brucelia melitensis	2 кольцевые хромосомы
Leptospira interrogans	1 кольцевая хромосома; 1 мегаплазмида
Rhizobium meliloti	1 кольцевая хромосома; 2 мегаплазмиды
Rhodobacter sphaeroides	2 кольцевые хромосомы
Rhodococcus facians	1 линейная хромосома; линейная плазмида
Streptomyces ambofaciens	1 линейная хромосома

В последние годы наблюдается прогресс в исследовании первичной структуры бактериальных хромосом. На конец 2011 г. была определена полная нуклеотидная последовательность ~200 прокариотических геномов.

Размеры и характеристики некоторых прокариотических геномов с известной нуклеотидной последовательностью приведены в табл. 9.

Структуры, связанные с репликацией. Полная нуклеотидная последовательность бактериальной хромосомы позволяет определить структуры генома, связанные с его фунционированием, т.е. с процессами репликации, транскрипции, трансляции, регуляции и т.д.

Репликация бактериальной хромосомы начинается в точке ori C (origine) и продолжается в обоих направлениях до участка терминации репликации ter C (terminus). У большинства бактерий участки ori C и ter C делят кольцевую хромосому на две почти равные реплихоры. Область начала репликации во многих бактериальных геномах имеет консервативную структуру. Подобные структуры найдены почти во всех секвенированных бактериальных геномах. Для линейных хромосом точка ori C определяется ровно посередине ее. Направление транскрипции большинства генов бактерий с высокой скоростью роста совпадает с направлением репликации. Структура области инициации репликации ДНК архебактерий неизвестна. Отсутствие асимметрии GC в геноме архей указывает на наличие множественных участков инициации репликации.

Открытые рамки считывания и определение функций белка. Открытая рамка считывания (англ. *open reading frame*, ORF) состоит из ряда триплетов, кодирующих аминокислоты; не содержит каких-

Таблица 9. **Характеристики некоторых бактериальных геномов, определенные по их нуклеотидной последовательности**

Видовое название	Размер генома, тыс. н. п.	Кодирующие последовательности	Количество ORF (предположительное число генов)	Идентифицировано генов		Характеристика
				число	%	
Mycoplasma genitalium	580	89,7	479	324	69	Вызывает воспаления мочеполовых путей
Mycoplasma pneumoniae	817	88,7	677	333	49	Возбудитель пневмонии
Borrelia burgdoferii	1 300	93,0	863	499	58	Возбудитель клещевого спирохетоза (болезнь Лайма)
Helicobacter pylory	1 700	91,0	1 590	907	57	Вызывает язвенную болезнь, гастрит
Methanococcus jannaschii	1 700		1 692	776	46	Анаэроб, метаноген, $T_{опт}$ роста = 85 °С
Haemophilus influenzae	1 800		1 740	1 015	58	Вызывает отиты, ОРЗ, менингиты
Synechocystis sp.	3 600	87,0	3 168	1 402	48	Фотосинтезирующая цианобактерия
Bacillus sabtilis	4 200	87,0	4 000	2 320	58	Автотроф
Mycobacterium tuberculosis	4 400	91,0	3 924	~1 600	40	Возбудитель туберкулеза
Escherichia coli K-12	4 600	88,6	4 288	2 656	62	Прототрофная энтеробактерия



Рис. 69. Функциональная характеристика генов *E. coli*, выявленных по полной нуклеотидной последовательности ее генома

либо терминирующих кодонов; потенциально может транслироваться в белок. В установленной нуклеотидной последовательности генома прокариот с помощью компьютерных программ выявлены открытые рамки считывания. Средний размер их в прокариотических геномах соответствует примерно 300 аминокислотным остаткам. Сопоставление состава ORF по функциям (также с помощью компьютерных программ) позволяет идентифицировать компоненты генома, общие для всех организмов, общие для данной группы видов и уникальные, специфичные только для данного организма. Функциональная классификация генов и генных продуктов является основой для сравнения геномов и выявления сходства и различия их метаболизма. В качестве модельного объекта использован геном *E. coli*. Все белки и кодирующие их гены классифицированы на три функциональные группы: энергообмен, информация, коммуникация (внутри- и внеклеточная). Распределение функций генов *E. coli*, выявленных по полной нуклеотидной последовательности генома, представлено на рис. 69. Среди 4 288 ORF *E. coli* не выявлены функции у 40 % ORF. Примерно такая же часть генов с неизвестной функцией отмечена и в других прокариотических геномах.

Минимальный размер генома прокариот. На основании анализа первичной структуры прокариотических геномов был составлен перечень генов, абсолютно необходимых для существования свободно живущих клеток. Этот перечень генов получил название *минимальный набор генов*. Он кодирует 256 белков и включает следующие жизненно важные генетические системы микроорганизмов: почти полный набор генов системы репликации; гены рудиментарной системы репарации и рекомбинации; гены аппарата транскрипции, в которой почти полностью отсутствуют гены регуляции транскрипции; большой набор генов, кодирующих белки, гомологичные шаперонам; гены, контролирующие анаэробный метаболизм, включая гены гликолиза и фосфорилирования субстратов; гены биосинтеза липидов; гены системы транспорта белков; гены, кодирующие ферменты,

используюшие сложные кофакторы; набор генов, обеспечивающий транспорт метаболитов; полный набор генов утилизации нуклеотидов de novo и гены их биосинтеза (табл. 10).

Таблица 10. **Функции белков, соответствующих минимальному набору из 256 генов**

Функция	Число белков
Преобразование энергии	28
Транспорт и метаболизм аминокислот	11
Транспорт и метаболизм нуклеотидов	20
Транспорт и метаболизм углеводов	5
Метаболизм кофакторов	8
Метаболизм липидов	6
Трансляция и биогенез рибосом	94
Репликация, транскрипция, рекомбинация, репарация	35
Структурная функция (белки наружной мембраны)	7
Секреция и адгезия	5
Шапероны	13
Транспорт неорганических ионов	4
Предсказана гипотетическая функция	15
Функция неизвестна	4

Если же из генома прокариотического организма вычесть консервативный минимальный набор генов для микроорганизмов, живущих в сходной среде, то можно получить те гены, которые определяют специфические фенотипические характеристики, составляющие уникальность организма.

Экологическая специфичность на уровне генома. Адаптация к определенным условиям обитания проявляется как в изменении общих характеристик генома (ГЦ-состав, суперспирализация ДНК), так и в наличии генов, продукты которых обеспечивают приспособление микроорганизма к тем или иным условиям. Показателен в этом отношении геном хеликобактерии, обитающей в желудке человека и вызывающей язвенную болезнь. Хеликобактерии способны к обитанию в кислой среде, так как устанавливают положительный внутри-

мембранный потенциал при низких pH и модифицируют микросреду с помощью уреазы. Белки *H. pylori* содержат вдвое больше основных аминокислотных остатков аргинина и лизина, чем белки *E. coli*. С паразитизмом связано наличие у *H. pylori* не менее пяти различных систем адгезии для прикрепления к эпителиальным клеткам желудка. Неожиданным является наличие в геноме *H. pylori* 14 генов (более 1 % генома), относящихся к системе рестрикции — модификации типа II. Вероятно, эта система, кроме расщепления чужеродной ДНК, имеет иную, неизвестную пока биологическую функцию.

Отсутствие определенных генов также указывает на особенности метаболизма, связанные с условиями жизни микроба. Если организм обеспечен ресурсами за счет их адсорбции, то механизмы синтеза (а значит, и гены) могут исчезнуть.

Структура прокариотических генов

Ген — единица наследственной информации, занимающая определенное положение в геноме или хромосоме и контролирующая выполнение определенной функции в организме. По результатам исследования прокариот, главным образом *E. coli*, ген состоит из двух основных элементов: регуляторной части и собственно кодирующей части. Регуляторная часть гена обеспечивает первые этапы реализации генетической информации, заключенной в структурной части гена; структурная часть гена содержит информацию о структуре кодируемого данным геном полипептида. Количество некодирующих последовательностей в структурной части гена у прокариот минимально. 5′-конец прокариотического гена имеет характерную организацию регуляторных элементов, особенно на расстоянии 50—70 н.п. от точки инициации транскрипции. Этот участок гена называют промотором. Он важен для транскрипции гена, но сам в РНК не транскрибируется. Противоположный 3′-конец — терминаторная область, необходимая для терминации транскрипции. В РНК он также не транскрибируется. Транскрипция начинается со стартовой точки (+1).

Промотор имеет две консервативные последовательности: одна состоит из 6 или 7 пар оснований и расположена на расстоянии ~10 оснований от стартовой точки; ее обозначают как –10-последовательность, или бокс Прибнова — в честь ее первооткрывателя. Вторая последовательность, длина которой равна обычно 9 нуклеотидам, расположена на расстоянии ~35 оснований от сайта инициации (рис. 70). Эта 35-последовательность участвует в связывании РНК-полимеразы; в боксе Прибнова РНК-полимераза начинает локальное раскручивание спирали и создает условия для инициации синтеза РНК.

Последовательности ДНК, являющиеся сигналами остановки транскрипции, находятся на 3′-конце гена и называются транс-

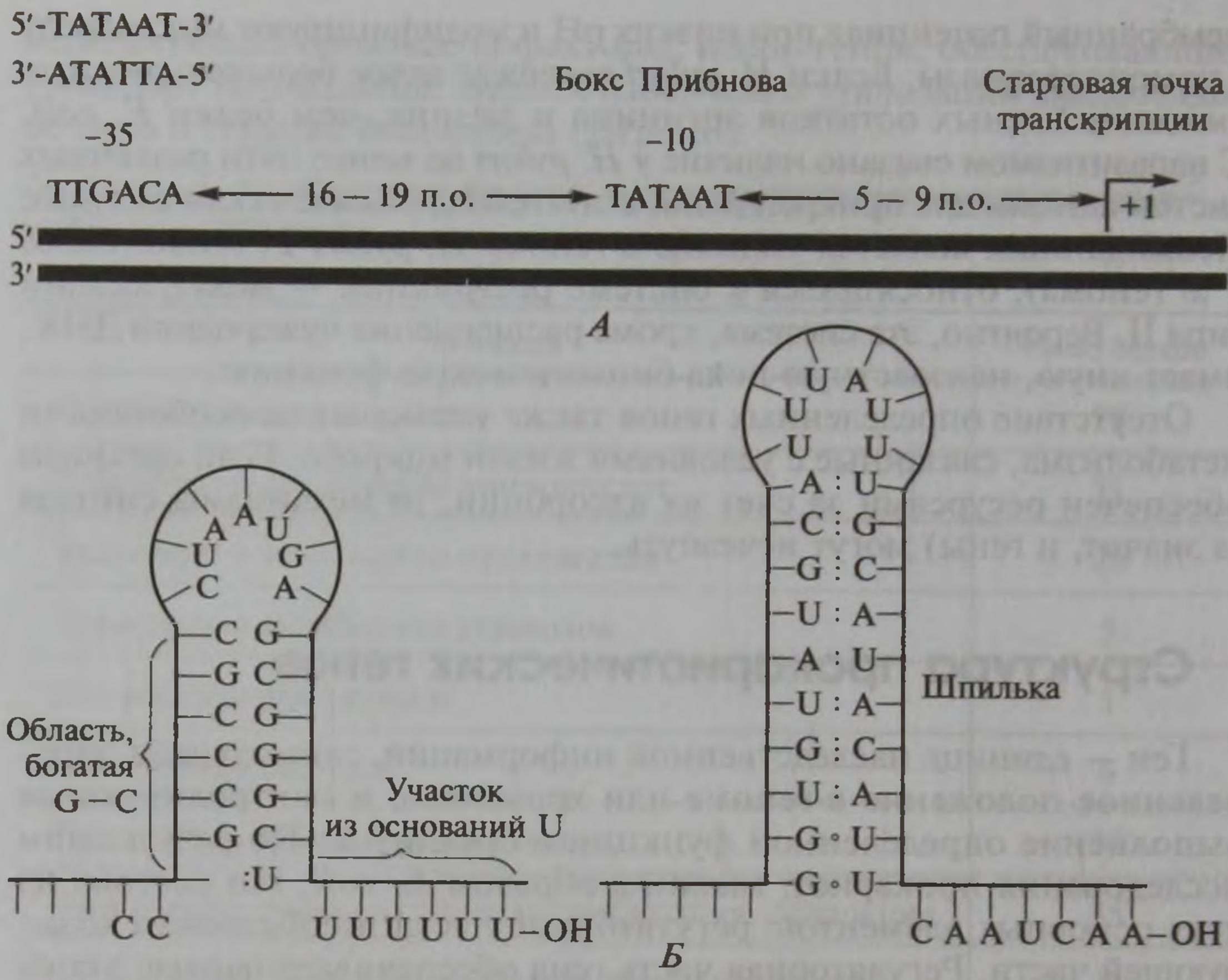


Рис. 70. Структура промотора и терминатора прокариотических генов:

А — консенсусные последовательности элементов промотора *E. coli*; *Б* — структуры терминаторных шпилек в РНК прокариот

крипционными терминаторами. Они содержат последовательности, которые в транскрибируемой РНК формируют структуру шпильки. Последовательность оснований в этом месте способствует образованию стабильных пар G—C за счет водородных связей. Вслед за шпилечной структурой в транскрипте РНК находится несколько остатков U (соответственно в гене последовательность A_{5-6}), дающих слабое связывание РНК с ДНК. Это и облегчает терминацию транскрипции.

Участок ДНК между промотором и терминатором называется единицей транскрипции. Последовательности, кодирующие белки, бывают обычно фланкированы нетранслируемыми сегментами (их обозначают 5′-НТП и 3′-НТП). Гены РНК перемежаются спейсерными (от англ. *spacer* — распорка) последовательностями, которые вырезаются в ходе процессинга рРНК и тРНК (рис. 71). Число генов, выявленное в результате расшифровки нуклеотидной последовательности геномов некоторых прокариот, см. в табл. 10.

Оперонная организация геномов прокариот. У *E. coli* гены, кодирующие белки одного и того же метаболического пути или определяющие близкородственные функции, часто бывают сцеплены.

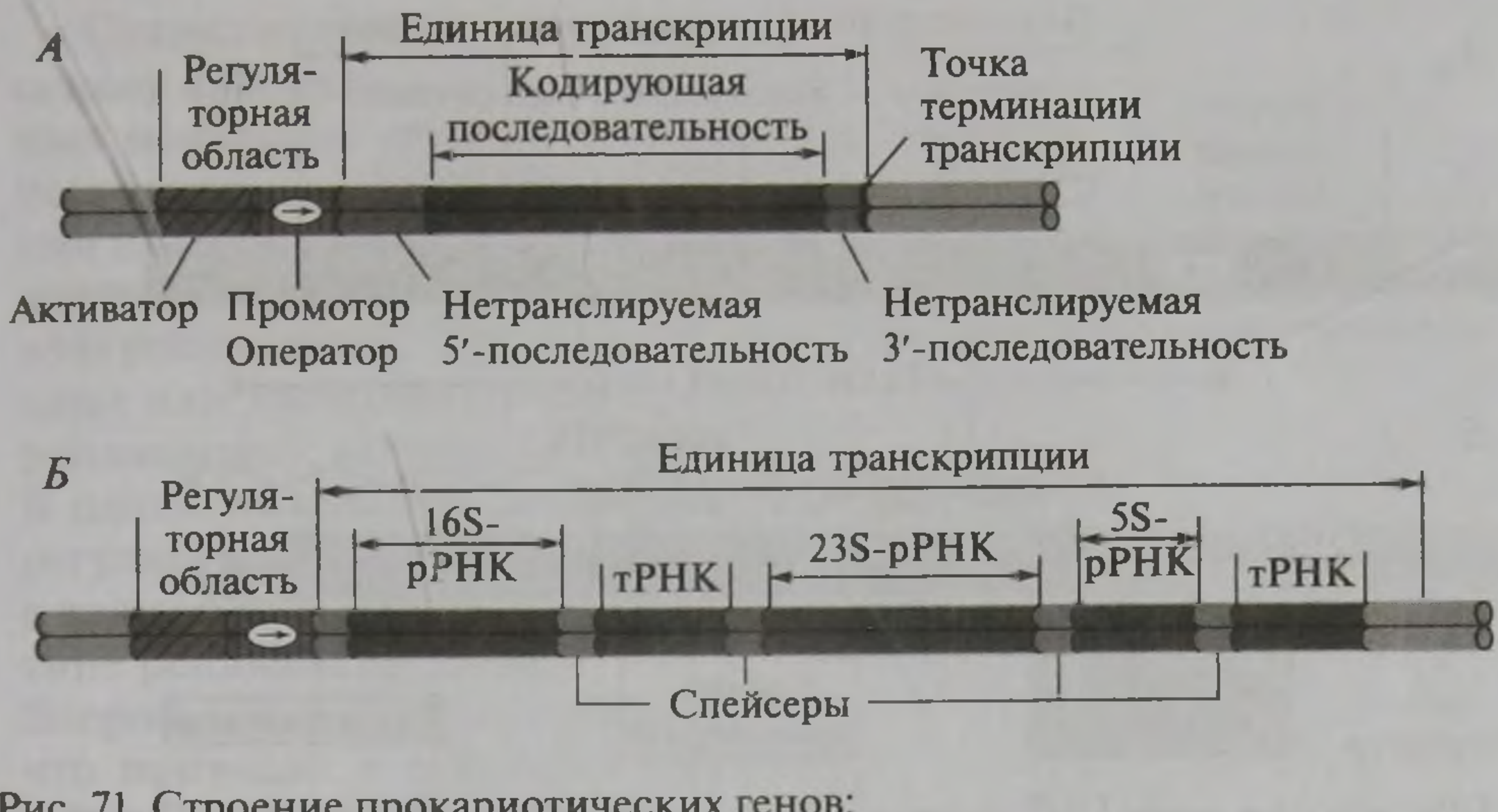


Рис. 71. Строение прокариотических генов:

А — ген, кодирующий один белок; *Б* — гены, кодирующие рРНК и тРНК

Они транскрибируются с промотора, находящегося на 5′-конце такой группы генов (кластера), в виде единственной молекулы РНК, называемой *полицистронным* (или *полигенным*) транскриптом. Группа координированно экспрессирующихся генов называется *опероном*. Три гена, кодирующие ферменты, ответственные за метаболизм галактозы у *E. coli*, организованы в оперон с промотором (Р) и примыкающим к нему регуляторным сегментом-оператором (О) на 5′-конце транскрибируемой последовательности (рис. 72).

При сравнении геномов *E. coli* и *B. subtilis*, имеющих около 1 000 общих генов, отмечено, что 100 из них находятся в составе одних и тех же оперонов, хотя в некоторых случаях порядок генов в оперонах изменен (арабинозный оперон ara ABD у *B. subtilis* и ara BAD у *E. coli*).

Гены, кодирующие несколько родственных функций, не всегда образуют единый оперон. Так, гены, кодирующие рибосомные белки, организованы во множественные опероны, в чей состав иногда входят гены, кодирующие другие белки, участвующие в транскрипции и/или трансляции. Как правило, отдельные опероны, кодирующие родственные функции, имеют одинаковые или сходные регуляторные последовательности и поэтому реагируют на определенный регуляторный сигнал сходным образом. Сохранение генов в составе одного оперона дает определенные преимущества: во-первых, обеспечивает координированную экспрессию генов; во-вторых, такое расположение генов важно для их распространения путем горизонтального переноса. Горизонтальный перенос генов вносит свой вклад в эволюцию прокариотических геномов. Показано, что он происходит как между родственными, так и между неродственными бактериями. Гены, появившиеся в данном геноме в результате горизонтального

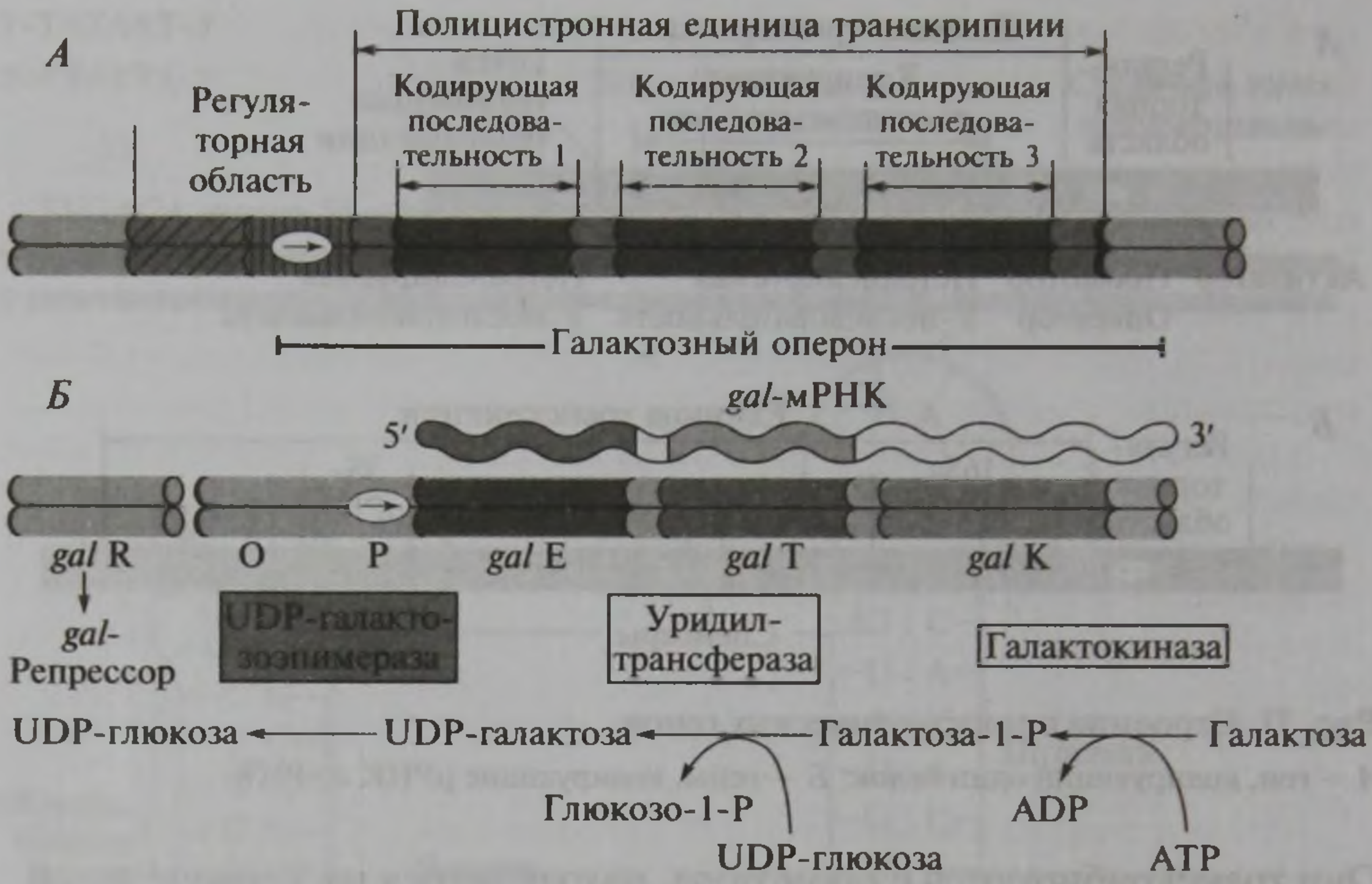


Рис. 72. Оперонная организация генов прокариот:

А — общая схема оперона; *Б* — галактозный оперон *E. coli*. Представлены оператор (O), промотор (P) и три гена оперона — *gal* E, *gal* T, *gal* K. Внизу показаны реакции, катализируемые тремя генными продуктами — галактокиназой, уридилтрансферазой и UDP-галактозоэпимеразой. Все три фермента транслируются единственной полицистронной *gal* мРНК, ген *gal* R, кодирующий репрессор галактозного оперона, не сцеплен с опероном

переноса, называют ксенологами. Отмечен горизонтальный перенос генов между эукариотами и прокариотами: наибольшее число генов эукариотического типа (около 30) обнаружено в геноме хламидии.

Бактериальные плазмиды

Кроме хромосомы у большинства бактерий существуют другие способные к автономной репликации структуры — *плазмиды*. Это двуцепочечные кольцевые ДНК размером от 0,1 до 5 % размера хромосомы, несущие гены, необязательные для клетки-хозяина, или гены, необходимые лишь в определенной среде. Именно такие внехромосомные элементы и содержат гены, которые придают клеткам наследуемую устойчивость к одному или нескольким антибиотикам. Они получили название *факторов резистентности*, или *R-факторов*. Другие плазмиды определяют болезнетворность патогенных бактерий, например патогенных штаммов *E. coli*, возбудителей чумы и столбняка. Третьи — определяют способность почвенных бактерий использовать необычные источники углерода, например углеводороды нефти.

Существуют разные виды классификации плазмид, но чаще всего основу классификации составляет наличие в плазмидах определенных *модульных сегментов ДНК*. Для своей репликации плазмиды используют репликативный аппарат клетки-хозяина, однако репликация плазмид происходит независимо от хромосомы. Каждая плазмида является самостоятельным репликоном, сама контролирует собственную репликацию. Для этой цели она должна иметь по крайней мере один или несколько *репликативных модулей* (область инициации репликации), которые и позволяют ей автономно реплицироваться. В одном из классов плазмид (F-плазмид) репликация и сегрегация регулируются согласованно с репликацией бактериального генома, и в каждой клетке содержится одна или две копии плазмиды. О таком типе репликации говорят как о репликации со строгим контролем. Второй тип репликационного модуля свободен от такого контроля, что приводит к существованию в клетке многих копий плазмид (10—30 копий). Подобный тип репликации получил название *репликации с ослабленным контролем*.

Наличие других модулей, не связанных с репликацией, не является обязательным для каждой определенной плазмиды. Половые факторы, т.е. *конъюгативные* (или *самотрансмиссивные*) плазмиды, подобно F-фактору имеют модули, содержащие гены и регуляторные области (гены *tra*), необходимые для переноса плазмиды из одной клетки в другую. Трансмиссивные плазмиды кодируют специальные ворсинки, половые пили, которые появляются на поверхности клеток, содержащих плазмиды, и способны специфически связываться с поверхностью бесплазмидных клеток. Последующее сокращение пиля притягивает клетки друг к другу, и между ними образуется мостик, через который плазмидная ДНК может передаваться в новую клетку. Способность передаваться в новые клетки — полезное для плазмид качество, но лишь большие плазмиды могут кодировать сложную систему поверхностных изменений клетки, обеспечивающую конъюгацию. *Неконъюгативные* плазмиды (утратившие модуль конъюгации) не способны к самотрансмиссивности, но способны к передаче в присутствии трансмиссивных плазмид, используя их аппарат конъюгации. Такие плазмиды называют *мобилизуемыми*.

Модули другого типа содержат гены, белковые продукты которых (например, β-лактамаза и др.) инактивируют антибиотики. Плазмиды, несущие такие модули, часто называют R-плазмидами (от англ. *resistance* — устойчивость). Одна плазмида обеспечивает устойчивость к нескольким антибиотикам, при этом все или несколько генов резистентности разного типа могут быть сгруппированы в одном модуле.

Имеется еще один тип модулей, встречающийся в плазмидах: *Col-модули*. Они кодируют один из нескольких белков-колицинов (антибактериальных агентов, продуцируемых бактериями). Колицины различаются по структуре и способу действия. Плазмиды, кодирующие определенный колицин, часто содержат гены, обеспечивающие

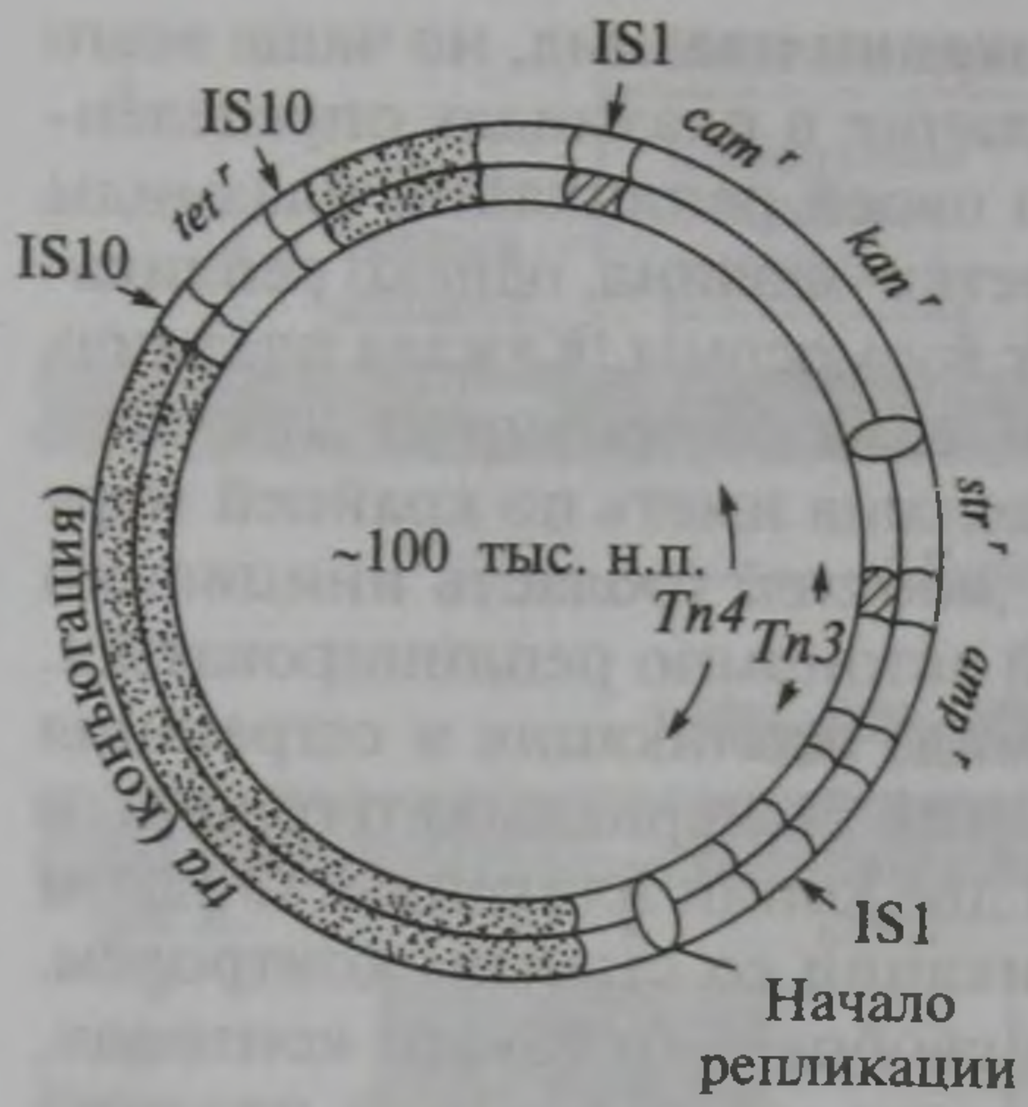


Рис. 73. Структура типичной конъюгативной R-плазмиды.

Светлый овал — область начала репликации (ori). Гены резистентности к антибиотикам: тетрациклину — *tet*^{r}, хлорамфениколу — *cam*^{r}, канамицину — *kan*^{r}, стрептомицину — *str*^{r}, ампициллину — *amp*^{r}. Точками выделены сегменты, необходимые для конъюгации (*tra*), штриховкой — мобильные элементы (см. с. 147)

иммунность к этому колицину, страхуя клетку-продуцент от повреждений, вызываемых ее собственным средством защиты. Некоторые плазмиды содержат также модули, кодирующие системы *рестрикции-модификации*. Примером такого модуля является система *Eco* RI.

Типичная плазмида. Модульное строение типичных R-плазмид схематически представлено на рис. 73. Около 50 % последовательностей из приблизительно 100 тыс. н. п. этих плазмид гомологичны одному из участков F-плазмиды. Эта область включает гены, необходимые для конъюгации (*tra*) и репликации со строгим контролем. В F-подобной области находится дополнительный модуль, содержащий гены, контролирующие устойчивость к тетрациклину. Вторая половина R-плазмиды не родственна F и содержит модули, ответственные за резистентность к антибиотикам (стрептомицину, сульфаниламидам, хлорамфениколу и канамицину).

Плазмиды, выделенные независимо и даже в различных частях света, часто содержат близкородственные модули (одни и те же гены резистентности к антибиотикам). На основании этих данных возникло представление о том, что между геномами происходит обмен некоторыми генетическими модулями в виде интактных сегментов ДНК. Быстрое распространение генов резистентности к антибиотикам объясняется именно обменом модулями между плазмидами. Эти и другие данные, полученные при изучении генетики бактерий, привели к идентификации подвижных элементов, названных *инерционными последовательностями* (IS) и *транспозонами*, которые способны перемещаться не только между плазмидами, но и между плазмидными и клеточными геномами, а также в пределах самого бактериального генома.

Эти последовательности всегда встраиваются в другие молекулы ДНК и никогда не находятся в свободном состоянии, они не могут инициировать собственную репликацию, а реплицируются вместе с молекулой ДНК, в которую они интегрированы.

IS-элементы и транспозоны бактерий

IS-элементы (от англ. *insertion sequences* — последовательности-вставки) — это сегменты ДНК, способные как целое перемещаться из одного участка локализации в другой (рис. 74). IS-элементы содержат лишь те гены, которые необходимы для их собственного перемещения — *транспозиции*. Кроме того, IS-элементы имеют особую последовательность на концах, как правило, *инвертированные повторы*. При встраивании в новую последовательность ДНК IS-элементы вызывают небольшую дупликацию: дуплицированный участок с двух сторон фланкирует встроившийся IS-элемент. Характеристика типичных IS-элементов *E. coli* представлена в табл. 11.

Таблица 11. **Типичные IS-элементы *E. coli***

IS	Суммарная длина, н. п.	Длина концевых инвертированных повторов, н. п.	Длина дуплицированной мишени, н. п.	Обычное местонахождение
IS1	768	23	9	Хромосома *E. coli*, 19 копий; плазмида R6, две копии; *Shigella dysenteria*, > 40 копий
IS2	1 327	41	5	Хромосома *E. coli*, 0 — 12 копий; R6, одна копия; F-плазмида, одна копия
IS4	1 426	18	11	Хромосома *E. coli*, 1 — 2 копии
IS10	1 329	23	9	R-плазмиды

Транспозонами (Tn-элементами) называют сегменты ДНК, обладающие теми же свойствами, что и IS-элементы, но содержащие также гены, не имеющие непосредственного отношения к транспозиции (гены устойчивости к антибиотикам, гены токсинов или гены дополнительных ферментов клеточного метаболизма). В общем, для транспозонов характерны те же гены, которые имеются в плазмидах. Более того, нередко присутствие в составе плазмиды того или иного гена обусловлено наличием в последовательности плазмидной ДНК соответствующего транспозона. Он может быть устроен так же, как IS-элемент, но с дополнительным геном. Однако важно отметить, что часто два IS-элемента, оказавшиеся поблизости друг от друга, способны перемещаться вместе, одновременно перенося заключенный

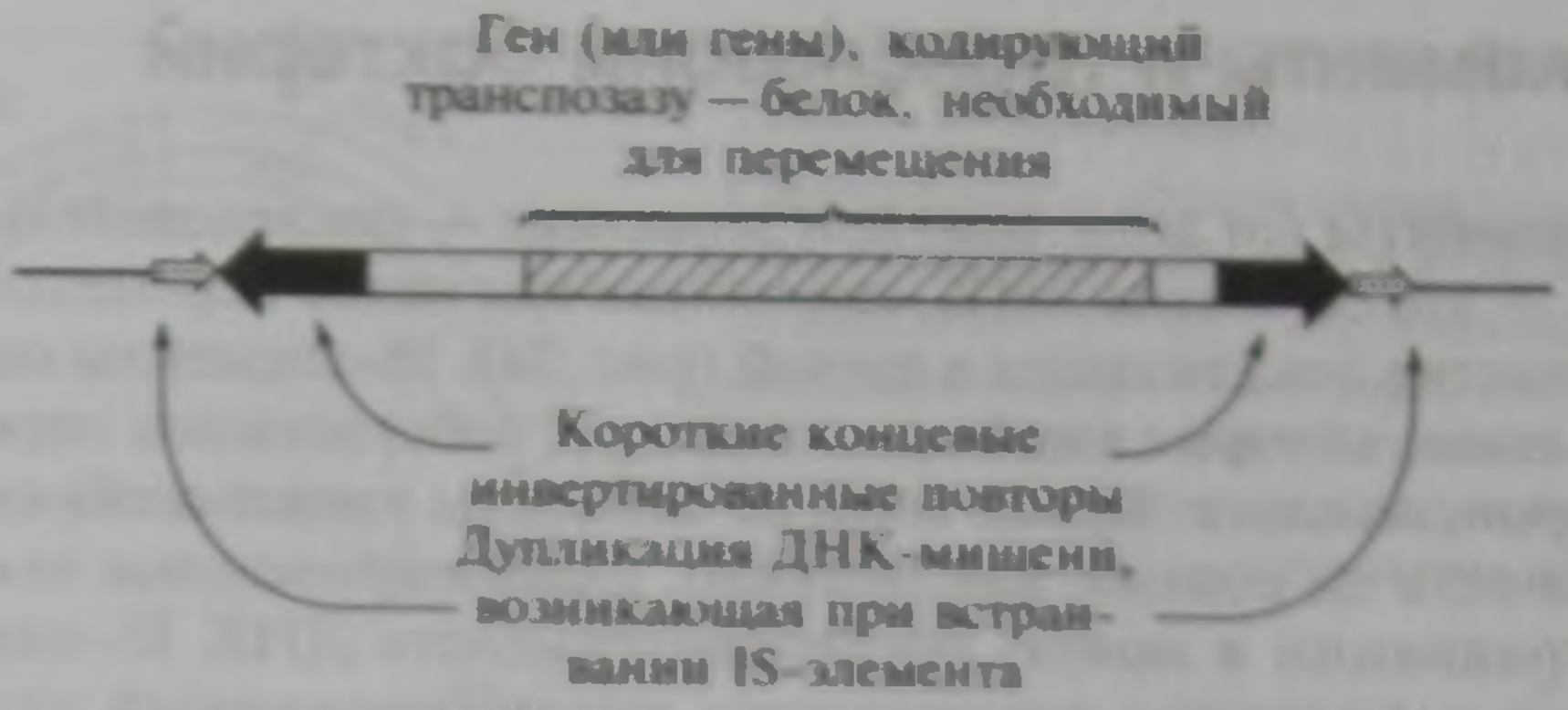


Рис. 74. IS-элемент *E. coli*

между ними сегмент ДНК. Таким образом, транспозон могут образовать два расположенных рядом IS-элемента. Транспозоны различной структуры показаны на конкретных примерах на рис. 75.

Транспозоны и IS-элементы ответственны за целый ряд генетических явлений у бактерий. Встраивание мобильного элемента в какой-либо ген может привести к его инактивации. Кроме того, некоторые IS-элементы и транспозоны вызывают генетическую нестабильность поблизости от места своей локализации: в окрестностях элемента заметно повышается частота делеций и инверсий, причем одна из границ перестройки всегда совпадает с одним из концов IS-элемента, автономного или в составе транспозона. Мобильные элементы способны также вызывать транслокации.

Механизмы перемещения мобильных элементов бактерий. Существует два типа транспозиций из одного генома в другой. В ходе транспозиции первого типа, называемой *кointеграционной*, донорный геном, который несет IS-элемент или транспозон, сливается с реципиентной молекулой ДНК. Репликация за счет клеточного репликативного аппарата приводит к удвоению мобильного элемента. Образуется промежуточный кointеграт с дупликацией мобильного элемента. Чтобы произошло разделение кointеграта на исходные

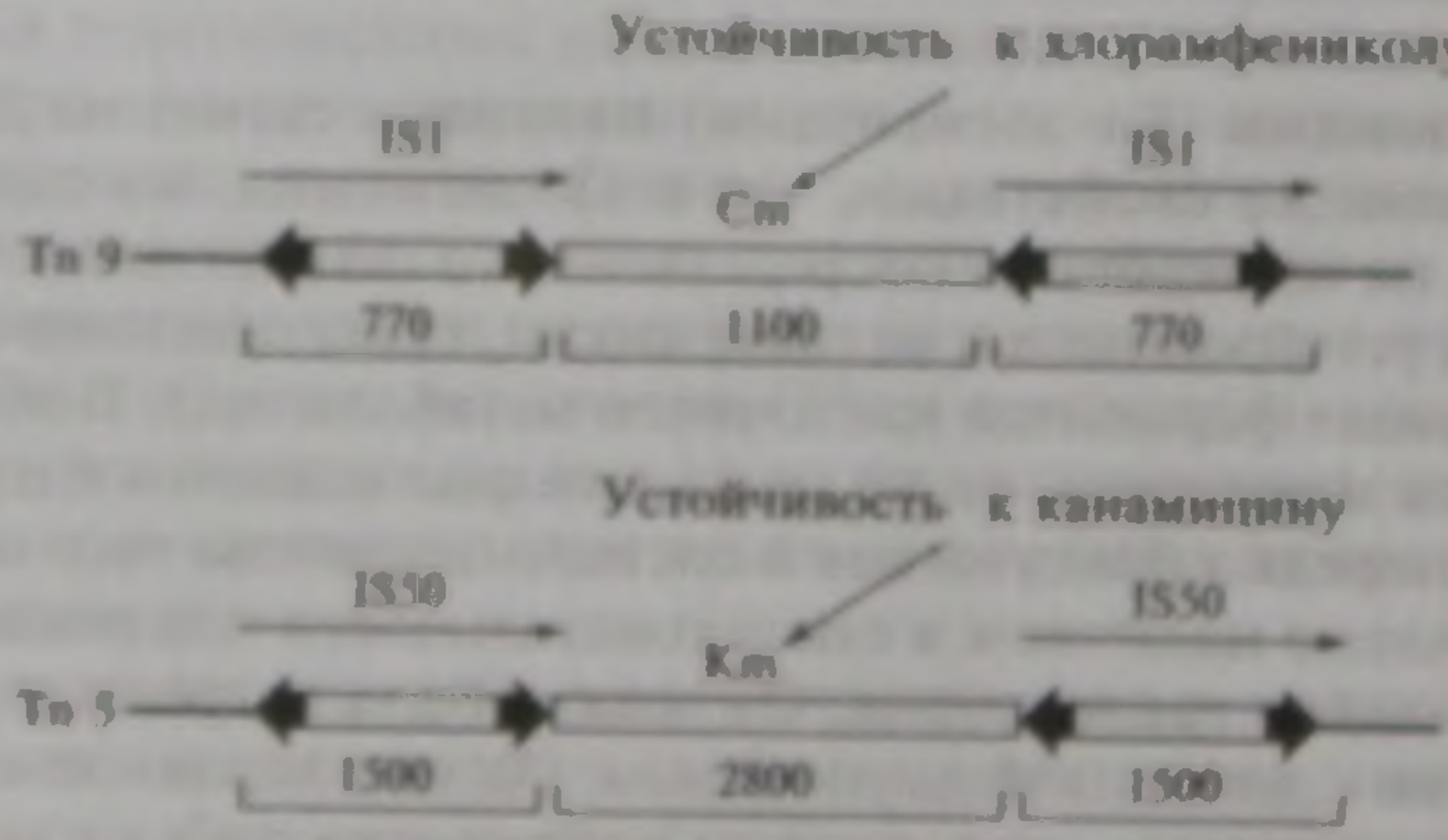


Рис. 75. Некоторые бактериальные транспозоны

репликоны (один из которых приобрел бы новую копию мобильного элемента), необходимо действие продукта гена tnpR, называемого *резолвазой* (от англ. *resolution* — разрешение), которая разрезает коинтеграт на исходные репликаторы (рис. 76).

Транспозиция второго типа называется *простым встраиванием*. Мобильный элемент перемещается в новый геномный локус, при этом никаких других перестроек, кроме дупликации сайта-мишени, не происходит. Этот вид транспозиции иногда называют консервативным (или нерепликативным). Для завершения процесса потребуется лишь ограниченный репаративный синтез ДНК.

Генетическая изменчивость бактерий. Мобильные элементы вызывают генетическую нестабильность поблизости от участка своей локализации, особенно в процессе репликативного механизма транспозиций. В зависимости от того, как внесены разрывы в ДНК-мишень, получится либо делеция, либо инверсия генетического материала между местом расположения транспозона и мишенью его перемещения. В связи с этим интересно отметить, что хромосомы родственных видов бактерий отличаются друг от друга многочисленными перестройками именно этого типа. По-видимому, мобильные



Рис. 76. Типы встраивания мобильного элемента в новый геном

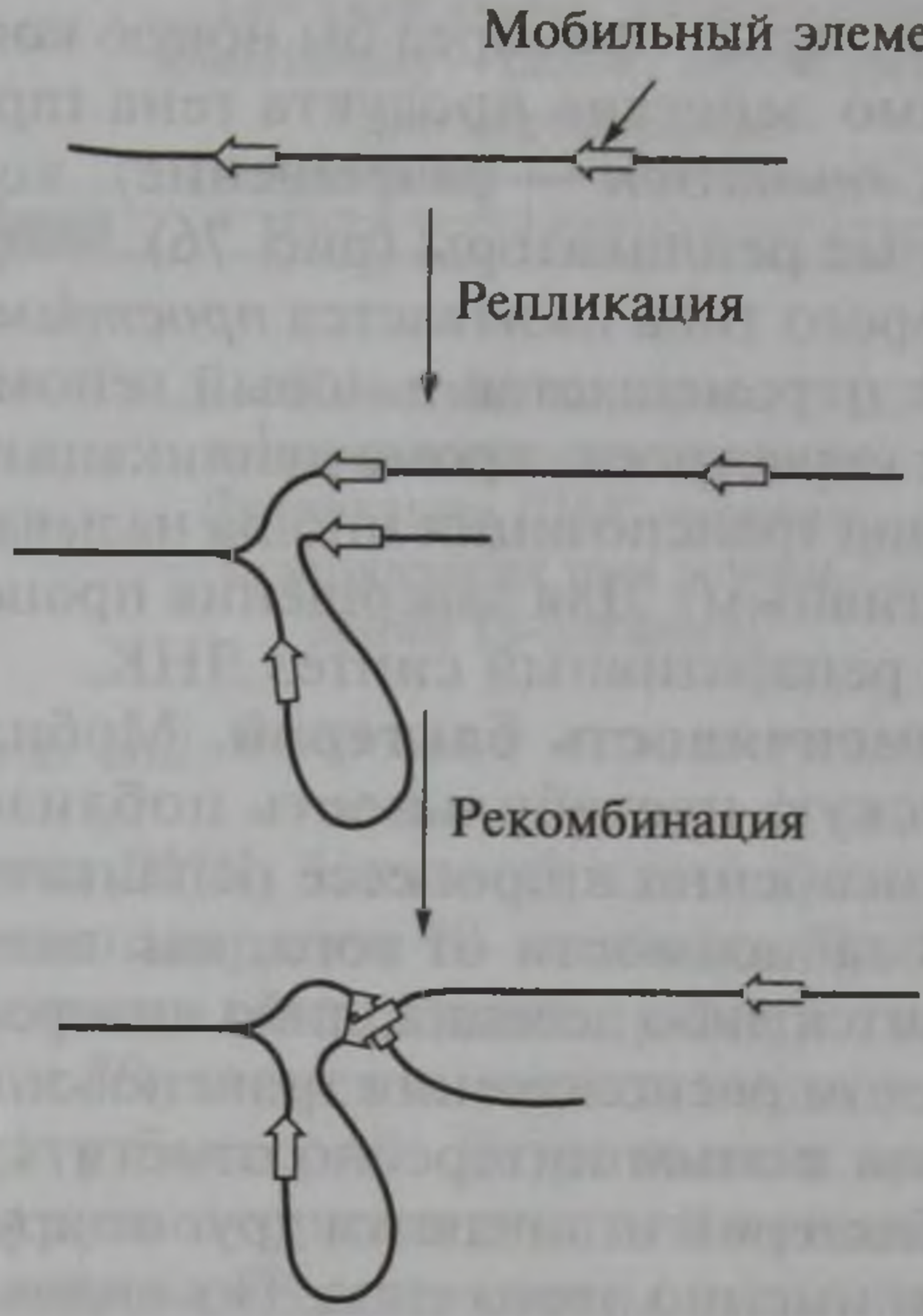


Рис. 77. Механизм образования дупликаций за счет гомологической рекомбинации по мобильным элементам (неравный кроссинговер)

элементы сыграли существенную роль в дивергенции и видообразовании бактерий. Встраивание IS-элементов поблизости от молчащего гена может приводить к его активации за счет транскрипции с промотора IS-элемента, т. е. изменяется регуляция бактериального гена. Очень важно, что мобильные элементы служат подвижными участками гомологии, гомологическая рекомбинация между которыми может приводить к дупликациям генов (рис. 77). Считается, что дупликация — один из основных путей эволюционного возникновения новых функций. Действительно, «лишняя» копия гена выходит из-под давления естественного отбора и получает возможность накапливать изменения. Чаще всего это приведет к утрате какой бы то ни было функции, но иногда может получиться ген с новыми функциями.

Нельзя забывать и тот факт, что клетка может получить селективное преимущество за счет приобретения в составе транспозона гена, который сам по себе способен оказаться выгодным для бактерии в определенных условиях. Действительно, на транспозонах «путешествуют» гены устойчивости к различным бактериальным ядам, в том числе к тяжелым металлам и антибиотикам, гены дополнительных метаболических путей, позволяющие использовать необычный источник углерода, наконец, гены некоторых токсинов, делающие бактерии патогенными и позволяющие им тем самым существенно изменить образ жизни. Сказанное в равной степени относится и к плазмидам, поскольку большинство полезных для клетки-хозяина плазмидных генов находится в составе транспозонов.

Глава 6
СТРУКТУРА ГЕНОМА ЭУКАРИОТ

Для клеток эукариот характерно наличие оформленного *ядра*. Информационной макромолекулой их генома является ДНК, которая неравномерно распределена по нескольким хромосомам в виде комплексов с многочисленными белками. Однако генетическую информацию в клетках содержат не только хромосомы ядра. Жизненно важная генетическая информация заключена и во внехромосомных молекулах ДНК. У эукариот — это ДНК хлоропластов, митохондрий и других пластид. Под геномом эукариотического организма в настоящее время понимают суммарную ДНК гаплоидного набора хромосом и каждого из внехромосомных генетических элементов, содержащуюся в отдельной клетке зародышевой линии многоклеточного организма.

Суммарное количество ДНК в гаплоидном геноме принято обозначать латинским символом C. Оно измеряется в пикограммах (пг), может быть выражено в парах нуклеотидов или дальтонах: 1 пг = $= 0{,}965 \cdot 10^9$ н.п. $= 6{,}1 \cdot 10^{11}$ Да). Размер генома эукариот в среднем на 2—3 порядка выше, чем у прокариот (табл. 12).

Таблица 12. **Средний размер гаплоидного генома у некоторых организмов**

Организм	Средний размер генома, н.п.
Мелкие вирусы	$1{,}0 \cdot 10^4$
Микоплазмы	$1{,}6 \cdot 10^6$
Бактерии	$2{,}0 \cdot 10^6$
Грибы	$4{,}7 \cdot 10^7$
Насекомые	$2{,}3 \cdot 10^8$
Моллюски	$1{,}6 \cdot 10^9$
Костистые рыбы	$1{,}4 \cdot 10^9$

Окончание табл. 12

Организм	Средний размер генома, н. п.
Бесхвостые амфибии	$2{,}7 \cdot 10^9$
Хвостатые амфибии	$3{,}6 \cdot 10^{10}$
Рептилии	$1{,}5 \cdot 10^9$
Птицы	$1{,}2 \cdot 10^9$
Млекопитающие	$2{,}6 \cdot 10^9$
Человек	$3{,}2 \cdot 10^9$
Растения:	
голосеменные	$1{,}6 \cdot 10^{10}$
покрытосеменные	$2{,}7 \cdot 10^{10}$
лилия *Lilium longiflorum*	$1{,}8 \cdot 10^{11}$

Геном эукариот существенно отличается от генома прокариот по ряду признаков, среди которых необходимо отметить его *избыточность*. Эукариотическая клетка содержит во много раз больше генов, чем прокариотическая. Повышенное содержание ДНК в геноме эукариот нельзя объяснить лишь увеличением потребности этих организмов в дополнительной генетической информации в связи с усложнением организации, поскольку большая часть их геномной ДНК, как правило, представлена *некодирующими последовательностями* нуклеотидов. Феномен значительной избыточности генома эукариот в отношении некодирующих последовательностей нуклеотидов известен под названием «парадокса С».

Получить суммарную характеристику степени разнообразия последовательностей ДНК, образующих геном, можно, исследуя кинетику реассоциации денатурированной ДНК.

Кинетика реассоциации денатурированной ДНК и сложность генома эукариот

Водородные связи и межплоскостные взаимодействия, стабилизирующие двойную спираль, достаточно слабы и при повышении температуры (до 100 °С) происходит разделение цепей. Такой процесс называют *плавлением двойной спирали*, или ее *денатурацией*. Поскольку для разрушения двух водородных связей АТ-пар требуется меньше энергии, чем для разрыва трех водородных связей GC-пар,

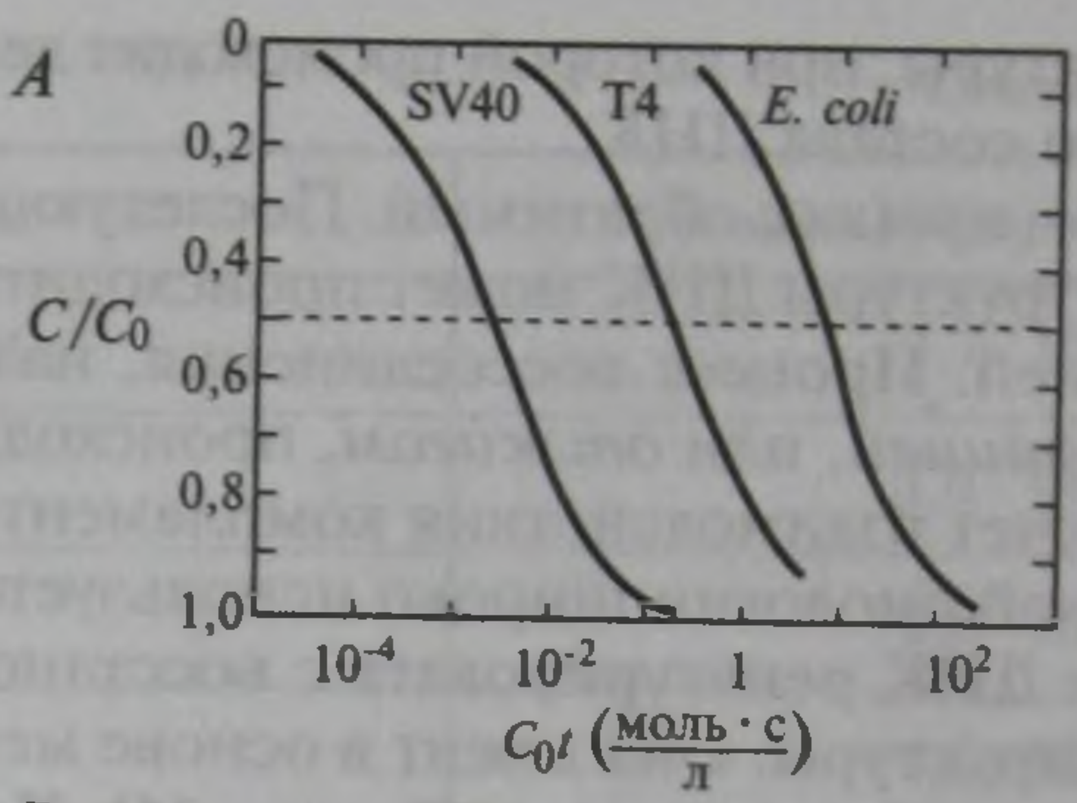
значения температуры, при которой происходит денатурация, зависят от нуклеотидного состава ДНК.

Денатурация — процесс обратимый. Последующее восстановление двуцепочечной структуры ДНК может происходить даже при полном расхождении цепей. Процесс воссоединения, называемый *ренатурацией*, *реассоциацией*, или *отжигом*, происходит при понижении температуры за счет взаимодействия комплементарных оснований.

В молекулярной биологии широко используется способность денатурированных ДНК ренатурировать с восстановлением исходной двуспиральной структуры. Она лежит в основе метода молекулярной гибридизации нуклеиновых кислот (см. гл. 14). Некоторые формальные характеристики эукариотического генома могут быть получены с помощью этого метода. Исследование кинетики реассоциации однонитевых денатурированных фрагментов ДНК путем определения в каждый данный момент реакции количества ренатурировавшей двуспиральной ДНК позволяет получить суммарную валовую характеристику степени разнообразия последовательностей ДНК, образующих геном, или сложность ДНК. Скорость реассоциации ДНК пропорциональна C_0t ($v = kC_0t$), где C_0 — исходная молярная концентрация однонитевой ДНК; t — время реассоциации, с. Значение $C_0t_{1/2}$, когда реассоциирует 50 % ДНК, характеризует степень сложности генома. Сложность ДНК определяется общей длиной (в нуклеотидных парах) разных последовательностей ДНК, содержащихся в исследуемом образце ДНК.

Исследование кинетики процесса реассоциации, разработка математического описания реакции, создание вычислительных программ для обсчета кинетических данных (Бриттен и сотр., 1966 — 1968) выявили чрезвычайную информативность кинетических параметров реакции реассоциации ДНК для понимания сложности генома, особенно генома эукариот. Для получения кинетических параметров реакции реассоциации ДНК последнюю фрагментируют (~1000 н. п.), фрагменты денатурируют при 100 °С в течение 10 — 12 мин, затем образец ДНК охлаждают до 65 °С (на 20 °С ниже температуры плавления). В этих условиях фрагменты восстанавливают двуспиральную структуру. Разделение одноцепочечных и двуцепочечных фрагментов ДНК на колонке с гидроксиапатитом [$3Ca_3(PO_4)_2 \cdot Ca(OH)_2$] с помощью фосфатного буфера (0,4 М — для элюции двуцепочечных фрагментов и 0,12М для элюции одноцепочечных фрагментов), определение их концентрации позволяют построить кривую реассоциации ДНК. Для ее построения на оси абсцисс откладывают величину C_0t в логарифмическом масштабе, на оси ординат — долю реассоциированной ДНК.

Мерой скорости реассоциации является величина $C_0t_{1/2}$, т. е. произведение начальной концентрации одноцепочечной ДНК на время ее полупревращения из одноцепочечной в биспиральное состояние. Для ДНК прокариот ход реассоциации ДНК описывается симметричной



Размер генома (н.п.)	SV40	T4	*E. coli*
	$5{,}2 \cdot 10^3$	$1{,}7 \cdot 10^5$	$4 \cdot 10^6$

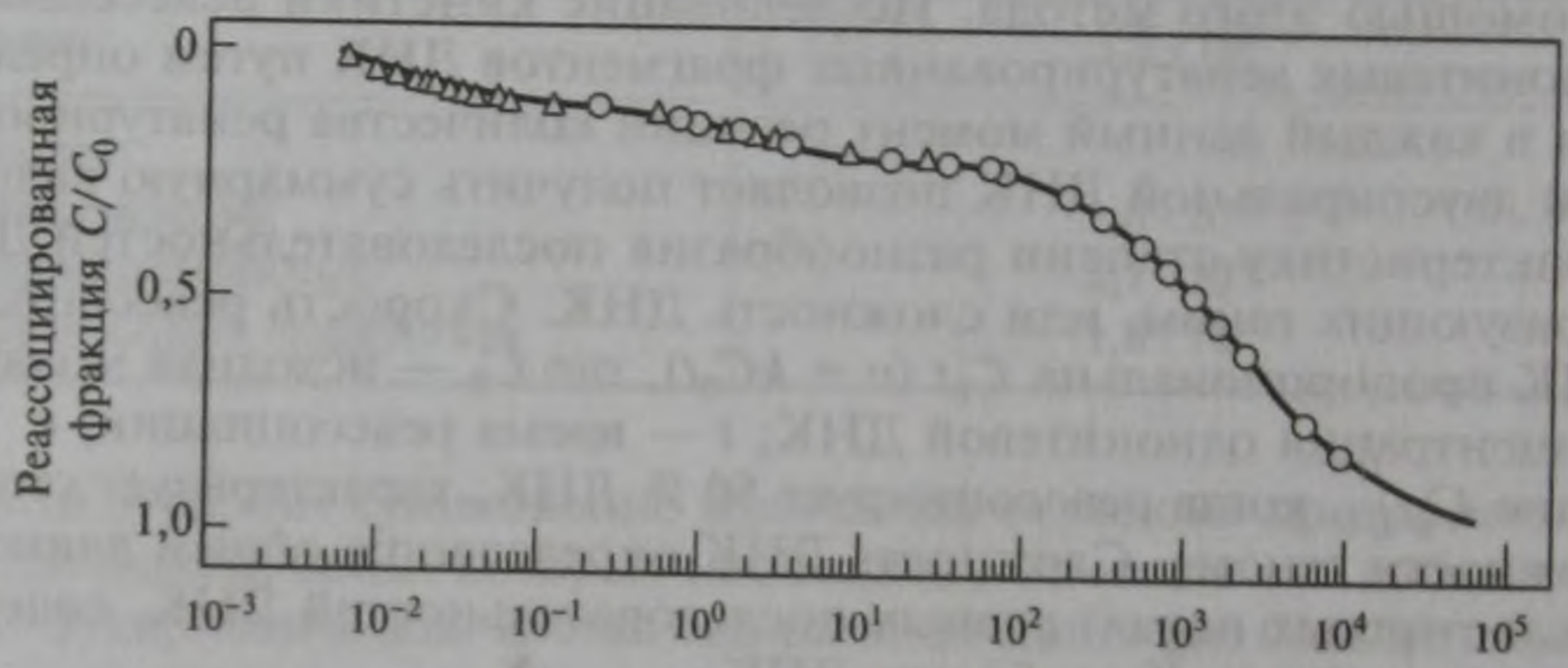


Рис. 78. Кривые кинетики реассоциации двуцепочечных ДНК разного происхождения:

А — ДНК вируса SV40, фага T4 и *E. coli*; *Б* — ДНК мыши; C/C_0 — доля реассоциированной ДНК, где C и C_0 — суммарные концентрации денатурированной ДНК в момент времени t после начала реассоциации и в момент времени t_0. Последовательности ДНК, реассоциирующие при значениях C_0t от 10^{-3} до 10^{-1}, называют высокоповторяющимися последовательностями, от 10^{-1} до 10 — умеренно повторяющимися; оставшаяся фракция ДНК — уникальные последовательности

S-образной кривой с точкой перегиба, соответствующей половине реакции. У эукариот ход реассоциации ДНК растягивается на много порядков значений C_0t (рис. 78). Эти результаты привели к заключению, что существенная часть генома эукариотических организмов состоит из повторяющихся в разной степени последовательностей ДНК. Различают высокоповторяющиеся, умеренно повторяющиеся и уникальные последовательности ДНК эукариот, соответствующие трем выявляемым при ренатурации фракциям. Для ДНК каждого вида величина $C_0t_{1/2}$ пропорциональна величине генома. Сопоставляя значения $C_0t_{1/2}$ ДНК разных видов и ДНК вида с известной величиной генома, можно определить размер генома этих видов.

При изучении кинетики реассоциации генома эукариот впервые были обнаружены диспергированные повторяющиеся последовательности, встроенные в другие последовательности ДНК.

Последовательности нуклеотидов эукариотического генома

Повторяющиеся последовательности. В современной классификации повторов принято различать высокоповторяющиеся последовательности, число которых превышает 10^5 на гаплоидный геном, и умеренно повторяющиеся, представленные $10 — 10^4$ копиями.

Высокоповторяющиеся последовательности включают функционально и структурно обособленную часть генома, представленную *сателлитными ДНК* (*сат ДНК*). Сателлитные ДНК состоят из коротких тандемных повторов длиной 1 — 20 н. п., организованных в длинные блоки. Свое название эти ДНК получили на основании того, что при анализе суммарной эукариотической ДНК центрифугированием в градиенте плотности хлорида цезия (CsCl) они сопровождали основной пик оптической плотности в виде плеча (спутника, сателлита) (рис. 79). Именно гомогенный нуклеотидный состав фракции сателлитных ДНК, определяемый наличием в ней многочисленных коротких повторов, изменил ее плавучую плотность, что легко обнаруживалось при центрифугировании.

Для сателлитов характерен ряд свойств, среди которых наиболее важны: а) быстрая и точная реассоциация в процессе ренатурации ДНК; б) множество копий; в) простая первичная структура; г) гомогенный состав (протяженные кластеры одних и тех же повторяющихся блоков последовательностей); д) пурин-пиримидиновая асимметрия в распределении нуклеотидов по цепям ДНК; е) концентрирование в прицентромерном гетерохроматине; ж) ограниченная репликация (недорепликация) при политенизации хромосом; з) нахождение в составе хромосом в виде тандемно (друг за другом) расположенных кластеров. Содержание сателлитной ДНК в геноме эукариот может достигать 5 — 50 % от суммарного количества ДНК. В составе одного генома можно обнаружить несколько разных сателлитных ДНК.

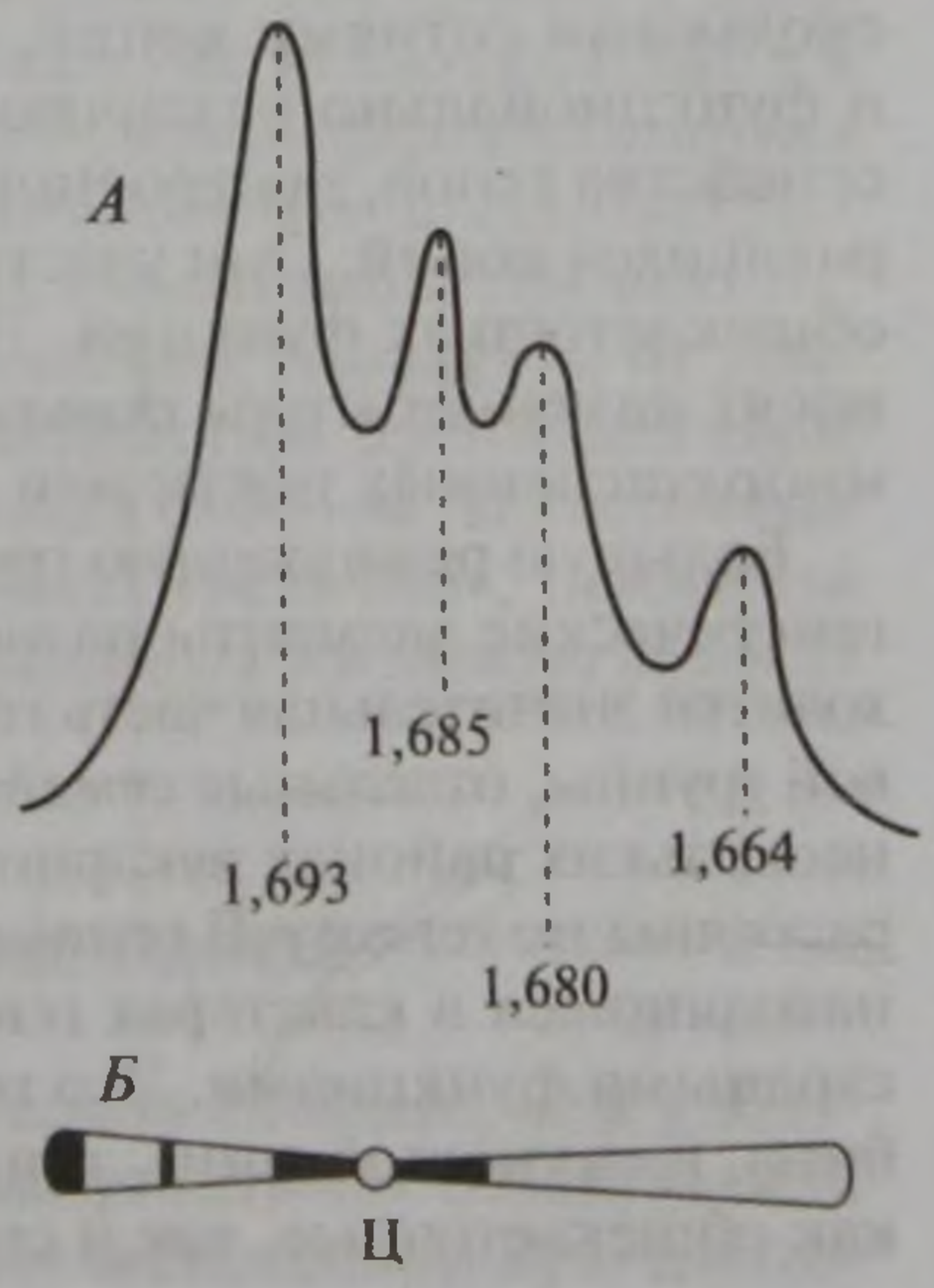


Рис. 79. Сателлитные ДНК:

А — центрифугирование препарата ДНК *Drosophila virilis* в градиенте CsCl, основная масса ДНК имеет плавучую плотность 1,693; *Б* — локализация сателлитных ДНК (темные участки), выявляемая методом гибридизации in situ в хромосомах млекопитающих (схема); Ц — центромера

Определение нуклеотидных последовательностей сателлитных ДНК показало наличие в их составе тандемных повторов из нескольких нуклеотидов. Главная повторяющаяся единица сателлитных ДНК (базовая последовательность), например у *D. melanogaster*, может быть достаточно простой, состоящей из пяти или семи нуклеотидов: (AATAT)n, (AATAG)n, (AATAC)n, (AATAAAC)n и (AATAGAC)n. Встречаются и более сложные повторяющиеся единицы из 359 н. п. Базовая последовательность сателлитной ДНК повторяется многократно на протяжении 10 тыс. н.п. и более. Таким образом, сателлитные ДНК образуют протяженные геномные блоки.

Меченые сателлитные ДНК можно гибридизовать с хромосомными препаратами in situ, а после гибридизации выявлять их локализацию. В хромосомах самых разных эукариот (дрозофила, млекопитающие) сателлитные ДНК преимущественно обнаруживаются в центромерных и теломерных районах хромосом (см. рис. 79). Это так называемые гетерохроматические районы, в которых материал хромосом сохраняется в процессе клеточного цикла в значительно более компактном состоянии в отличие от основной (эухроматической) части хромосом.

Отождествление сателлитной ДНК с гетерохроматином означает, что высокоповторяющаяся ДНК не экспрессируется с образованием полипептидов или РНК. Локализация сателлитной ДНК в области центромеры говорит о том, что она выполняет некоторую структурную функцию в хромосоме при митозе и мейозе. Возможно, функция сателлитной ДНК может быть связана с процессами расхождения хромосом.

Структура и функции мини- и микросателлитной, а также теломерной ДНК представлены на с. 172 — 173.

К ***умеренно повторяющимся последовательностям*** относятся последовательности, представленные в геноме десятками или несколькими сотнями копий. Среди них можно выделить структурно и функционально различающиеся геномные фракции. Существуют семейства генов, построенные из сгруппированных тандемно повторяющихся копий. Они участвуют в осуществлении жизненно важных общеклеточных функций. Это гены гистонов, тРНК и рРНК. Они носят название «гены *домашнего хозяйства*», в геноме находятся в многочисленных почти или совершенно идентичных копиях.

Большую разнородную группу составляют *подвижные* (*мобильные*) генетические элементы разной природы (см. с. 175). На их долю приходится значительная часть генома (10 — 20 %). В отличие от генов первой группы, отдельные семейства которых сгруппированы в одном или нескольких районах эукариотического генома, подвижные элементы рассеяны по геному. В геноме обнаруживаются также рассеянные или находящиеся в кластерах гены, кодирующие гомологичные белки со сходными функциями. Это гены разных видов актина, тубулина, глобина, иммуноглобулина, гены теплового шока и др. Они выполняют как общеклеточные, так и специализированные функции.

Уникальные последовательности. Уникальные (неповторяющиеся) последовательности — наиболее сложный компонент генома эукариот. Содержание их варьирует у разных организмов в пределах 15—93 % от всей ДНК. Процентное содержание уникальных последовательностей у низших эукариот заметно больше, чем у высших. Особенно мало уникальной ДНК у растений.

Во фракцию уникальных последовательностей попадают многие структурные гены, кодирующие белки. Бóльшая часть уникальных последовательностей является некодирующей и обычно не заключает в себе генетической информации в общепринятом значении этого термина, т. е. не кодирует функционально значимые полипептидные цепи или РНК. Хорошо известным примером таких уникальных последовательностей являются интроны, общий размер которых, как правило, на порядок и более превышает суммарный размер экзонов содержащих их генов. Сюда следует отнести и последовательности, расположенные между генами.

Несмотря на широкую распространенность повторяющихся последовательностей в геноме эукариот, биологическое значение этих и других некодирующих элементов генома остается непонятным. Вызывает сомнение правильность активно обсуждаемой в течение двух десятилетий гипотезы об «эгоистичности» избыточной геномной ДНК, о ее ненужности.

Недавно для объяснения основной функциональной роли «избыточной» ДНК в геноме эукариот предложена гипотеза «альтруистической» ДНК (Л. И. Патрушев, 1997). Именно избыточные последовательности ДНК оказывают стабилизирующее влияние на генетическую информацию, заключенную в геноме многоклеточных организмов. Они разбавляют кодирующие последовательности некодирующими таким образом, что последние выступают в роли ловушек мутагенов любого вида. Это разбавление могло произойти в результате эволюционных преобразований генома-предшественника путем включения в него некодирующих избыточных последовательностей при участии разных молекулярных механизмов. Организмы, у которых жизненно важные гены максимально защищены от действия мутагенов, должны обладать эволюционным преимуществом и в первую очередь сохраняться естественным отбором. Избыточные последовательности ведут себя вполне «альтруистично», принимая удар мутагенов на себя и специфически защищая жизненно важные участки ДНК от мутаций.

Структура эукариотических генов

Организация генов у эукариот значительно сложнее, чем у прокариот. Эукариотический ген можно рассматривать как совокупность сегментов ДНК, которые вместе составляют экспрессируемую едини-

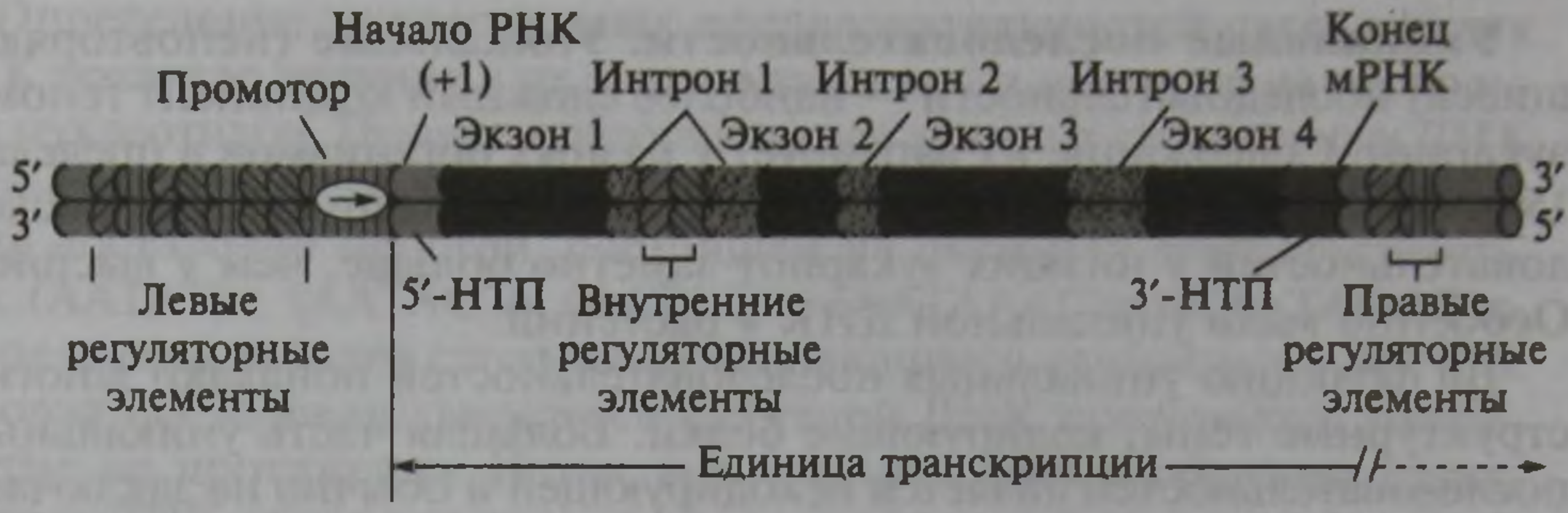


Рис. 80. Структура эукариотического гена, кодирующего белок:

+1 — точка инициации транскрипции; 5′-НТП и 3′-НТП: 5′- и 3′- — нетранслируемые последовательности

цу, ответственную за образование специфического функционального продукта — либо молекулы РНК, либо полипептида.

К сегментам ДНК, составляющим ген, относятся следующие элементы:

1. Единица транскрипции — это участок ДНК, кодирующий первичный транскрипт. Он включает: а) последовательность, которая обнаруживается в зрелых функциональных молекулах РНК; б) интроны (для мРНК); в) промежуточные последовательности — спейсеры (для рРНК). Интроны и спейсеры удаляются в ходе процессинга первичных транскриптов; г) 5′- и 3′-нетранслируемые последовательности (5′-НТП и 3′-НТП).

2. Минимальные последовательности, необходимые для начала транскрипции (промотор) и конца транскрипции (терминатор).

3. Последовательности, регулирующие частоту инициации транскрипции; ответственные за индуцибельность и репрессию транскрипции, а также клеточную, тканевую и временную специфичность транскрипции. Они разнообразны по строению, положению и функциям (рис. 80). К их числу относятся *энхансеры* (от англ. *enhance* — усиливать) и *сайленсеры* (от англ. *silence* — заглушать) — это последовательности ДНК, расположенные в тысячах пар нуклеотидов от промотора эукариотического гена и оказывающие дистанционное влияние на его транскрипцию. В это определение гена не включены последовательности ДНК, которые влияют на пространственную конфигурацию гена в хроматине, ни последовательности, которые регулируют его топологию.

Гены, кодирующие белки

В отличие от прокариотических генов, почти всегда коллинеарных своим РНК, многие гены эукариот имеют *мозаичное строение*. Под мозаичностью в данном случае подразумевается чередование коди-

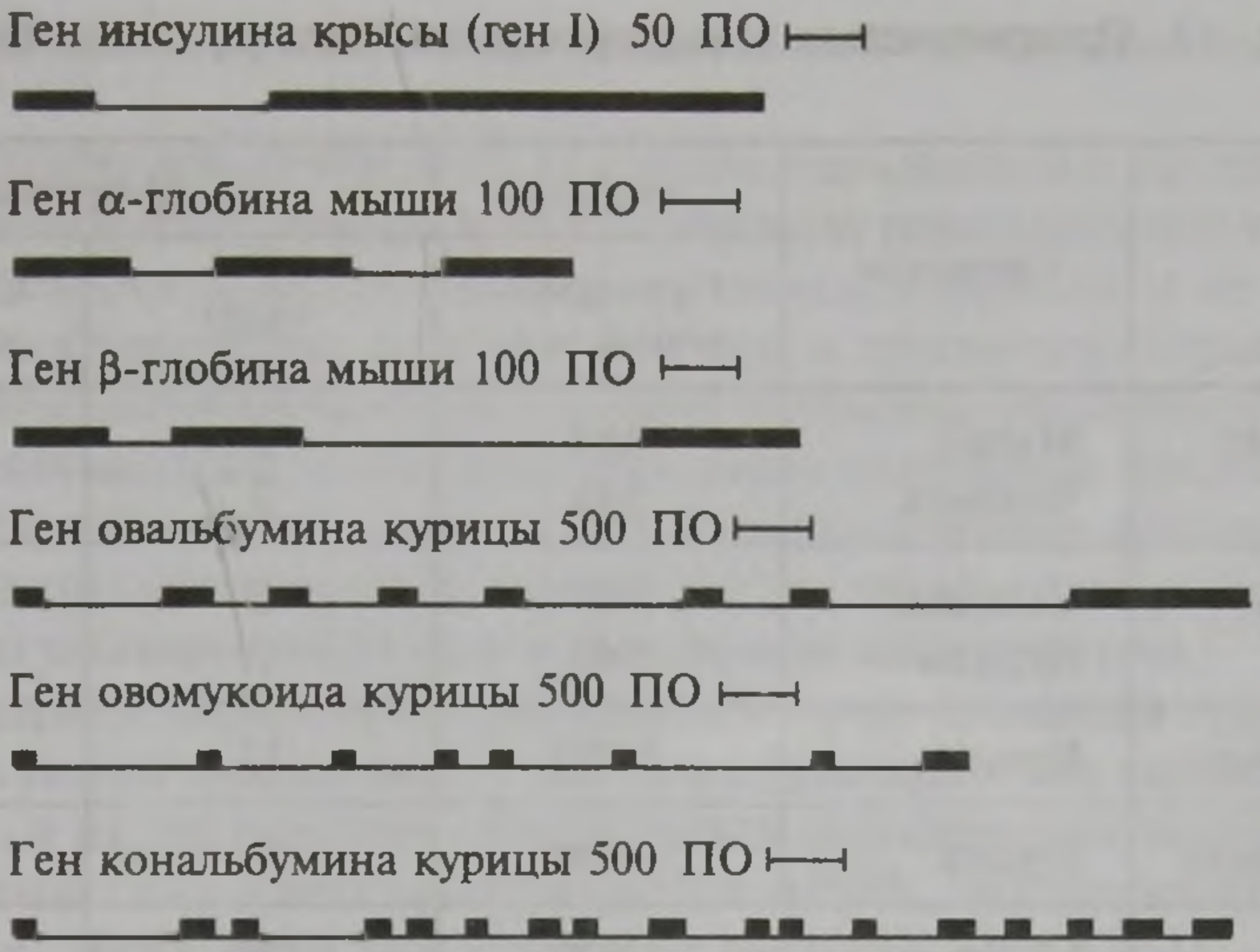


Рис. 81. Карты некоторых прерывистых генов.

Жирными линиями обозначены экзоны, тонкими — интроны

рующих (*экзоны*) и некодирующих (вставочные последовательности, или *интроны*) последовательностей в пределах единицы транскрипции. Интроны чаще всего встречаются в генах, кодирующих белки. За исключением некоторых генов, кодирующих пять гистонов, α- и β-интерфероны, все гены, кодирующие белки позвоночных, содержат интроны. У большей части генов *Drosophila* интроны отсутствуют. Однако некоторые кластеры генов, регулирующие развитие органов тела насекомого, содержат множество интронов. Интроны присутствуют и во многих генах растений.

Размеры, число и местоположение интронов у разных генов различны. Тем не менее сходные гены у организмов разных видов часто имеют одинаковое число интронов в одних и тех же позициях, хотя длина и нуклеотидная последовательность интронов могут заметно различаться. Обычно число интронов на ген возрастает пропорционально длине последовательности, кодирующей белок, а размеры экзонов в среднем составляют около 300 н. п. (рис. 81). В целом общая длина последовательностей интронов превышает суммарную длину экзонов обычно от 2 до 10 раз, а иногда еще больше (табл. 13).

Интроны располагаются в единицах транскрипции не случайным образом. В белок-кодирующих генах они часто находятся между сегментами, которые кодируют отдельные структурные или функциональные домены белка.

Эволюционное возникновение мозаичной (интрон-экзонной) структуры генов эукариот, так же как и консервативный характер наследования размеров и взаимного расположения интронов в генах, не находят в данное время исчерпывающего объяснения из-за кажущегося отсутствия фактора давления естественного отбора на

Талица 13. **Прерывистые гены эукариот, кодирующие белки**

Ген	Организм	Экзоны	Интроны	
		Суммарная длина, н. п.	число	Суммарная длина, н. п.
α-Глобин	Мышь	463	2	256
	Человек	596	2	265
Инсулин	Человек	417	2	964
	Курица	540	2	3 619
Альбумин	Крыса	2 000	13	12 500
α-Амилаза	Крыса	1 600	7	7 400
Дигидро-фолатре-дуктаза	Мышь	1 600	26	40 400
Овальбу-мин	Курица	1 872	7	5 758
Фиброин	Тутовый шелкопряд	18 000	1	970
Фазеолин	Фасоль	1 263	5	515

последовательности нуклеотидов без четких биологических функций. Наибольшее распространение получила концепция В. Гилберта (1977), согласно которой появление интронов, по-видимому, совпавшее по времени с эволюционным возникновением многоклеточных организмов, обеспечило возможность обмена экзонами между неродственными генами. Такой обмен должен сопровождаться образованием новых белков мозаичного строения, составленных из готовых полипептидных функционально значимых модулей (доменов), ранее принадлежавшим другим белкам (гипотеза о перетасовке экзонов в процессе эволюции). Следствием этого, по мнению сторонников данной концепции, были резкое ускорение образования белков и ферментов с новыми функциями, а также глубокие эволюционные преобразования самих организмов, реализующих такие молекулярные механизмы. Эта точка зрения получила еще название «гипотезы позднего возникновения интронов». В соответствии с другой гипотезой (Дж. Дарнелл и В. Дулиттл, 1978) современные интроны представляют собой «эволюционные реликты». Когда-то интроны были частью гигантских генов. Предполагается, что мозаичная структура генов определяет эволюционные возможности генома, тогда как про-

кариоты, утерявшие интроны, представляют собой эволюционный тупик.

Определение нуклеотидных последовательностей в составе гомологичных генов (например, генов глобинов) показало, что наибольшим изменениям в эволюции подвергались интроны, а не экзоны. В интронах обнаружены вставки, делеции и другие перестройки, в то время как последовательности экзонов оказываются значительно более консервативными: в них имеются лишь отдельные нуклеотидные замены. Эти наблюдения можно истолковать в пользу представлений о том, что перекомбинация экзонов осуществлялась по районам интронов, что и сопровождалось изменением их структуры.

Для эукариот характерны моноцистронные единицы транскрипции в отличие от прокариот, у которых обычно гены организованы в опероны, а их экспрессия опосредуется полицистронными мРНК. В эукариотических клетках образуются мРНК, кодирующие только один белок. Даже если ген кодирует не один вид мРНК, как при альтернативном сплайсинге, каждая зрелая мРНК имеет только одну транслируемую кодирующую последовательность.

Регуляторные элементы генов, кодирующих белки

У высших эукариот в пределах 100 — 200 н. п. перед стартом транскрипции выявлена сложная мозаика промоторных элементов, представленных короткими нуклеотидными последовательностями («мотивами»). На расстоянии 27 — 30 н. п. от сайта инициации транскрипции расположен ТАТА-мотив, который строго определяет сайт инициации транскрипции, т. е. 5′-конец транскрипта (рис. 82).

На расстоянии 80 — 50 н. п. имеются последовательность ССААТ, обычно встречающаяся в промоторной зоне разных тканеспецифических генов (глобинов, тиреоглобулина, актина и др.), и последовательность GGGCGGG, усредненный вариант «GC»-мотива. Последний может встречаться несколько раз по длине промоторных районов генов, работающих конститутивно и обеспечивающих общеклеточные функции: перед генами глюкозо-6-фосфатдегидрогеназы

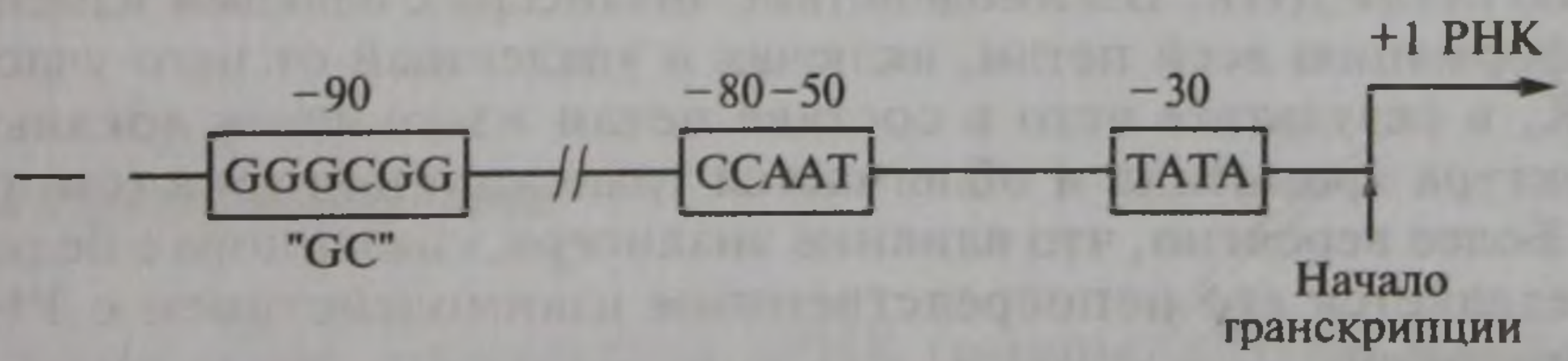


Рис. 82. Общая схема строения промоторов генов, кодирующих белки, у эукариот

и фосфоглицераткиназы вкраплены соответственно девять и два таких мотива.

Эффективность промотора определяется как эффективностью отдельного элемента («мотива») в составе промотора, числом его элементов, так и ближайшим нуклеотидным окружением вокруг него.

Гены, кодирующие адаптивные белки, образование которых резко усиливается под влиянием разных факторов среды (повышение температуры, отравление металлами), содержат в составе промоторов дополнительные характерные короткие нуклеотидные последовательности. Так, в гене белка теплового шока дрозофилы в промоторном районе имеются повторяющиеся мотивы CTCG и GTTC; в генах белка металлотионеина выявлены повторяющиеся девятинуклеотидные «мотивы»: TGCGCTCGG. Если такие мотивы вставить в промоторы других генов, то их экспрессия начнет зависеть от присутствия металлов. Выявлены участки, ответственные за гормонзависимую регуляцию. Таким образом, область промотора эукариотического гена рассматривается как специфический ДНК-остов, на котором собираются белки транскрипции, узнающие свои сайты связывания и взаимодействующие как друг с другом, так и с РНК-полимеразой.

В составе эукариотического гена имеются особые *цис*-действующие элементы регуляции, получившие название усилителя (энхансера), или активатора, транскрипции. Оказалось, что энхансеры могут располагаться как на 5′-, так и на 3′-конце фрагмента ДНК, включающего ген, а также в составе интронов (рис. 83). Они способны действовать и на больших расстояниях (более 1 000 н. п.), и вне зависимости от ориентации по отношению к направлению транскрипции гена.

Энхансеры состоят из «нуклеотидных мотивов». Короткие нуклеотидные блоки могут служить центрами связывания специфических ядерных белков. Сила энхансера, вероятно, может зависеть от числа таких блоков. Обсуждаются два основных механизма действия энхансеров. Считается, что функциональные участки генома, содержащие один или несколько генов, образуют длинные петли, включающие десятки тысяч нуклеотидных пар ДНК. Петли закреплены в матриксе клеточного ядра и сверхспирализованы. В состав матрикса входит топоизомераза II, определяющая топологию петли ДНК. Взаимодействие энхансера с белками изменяет конформацию всей петли, включая и удаленный от него участок ДНК, в результате чего в составе петли изменяется локальная структура хроматина и облегчается транскрипция гена (см. рис. 83). Более вероятно, что влияние энхансера, связанного с белком, определяется его непосредственным взаимодействием с РНК-полимеразой.

Энхансер в зависимости от белкового фактора может вести себя и как негативно действующий «глушитель» (сайленсер) экспрессии

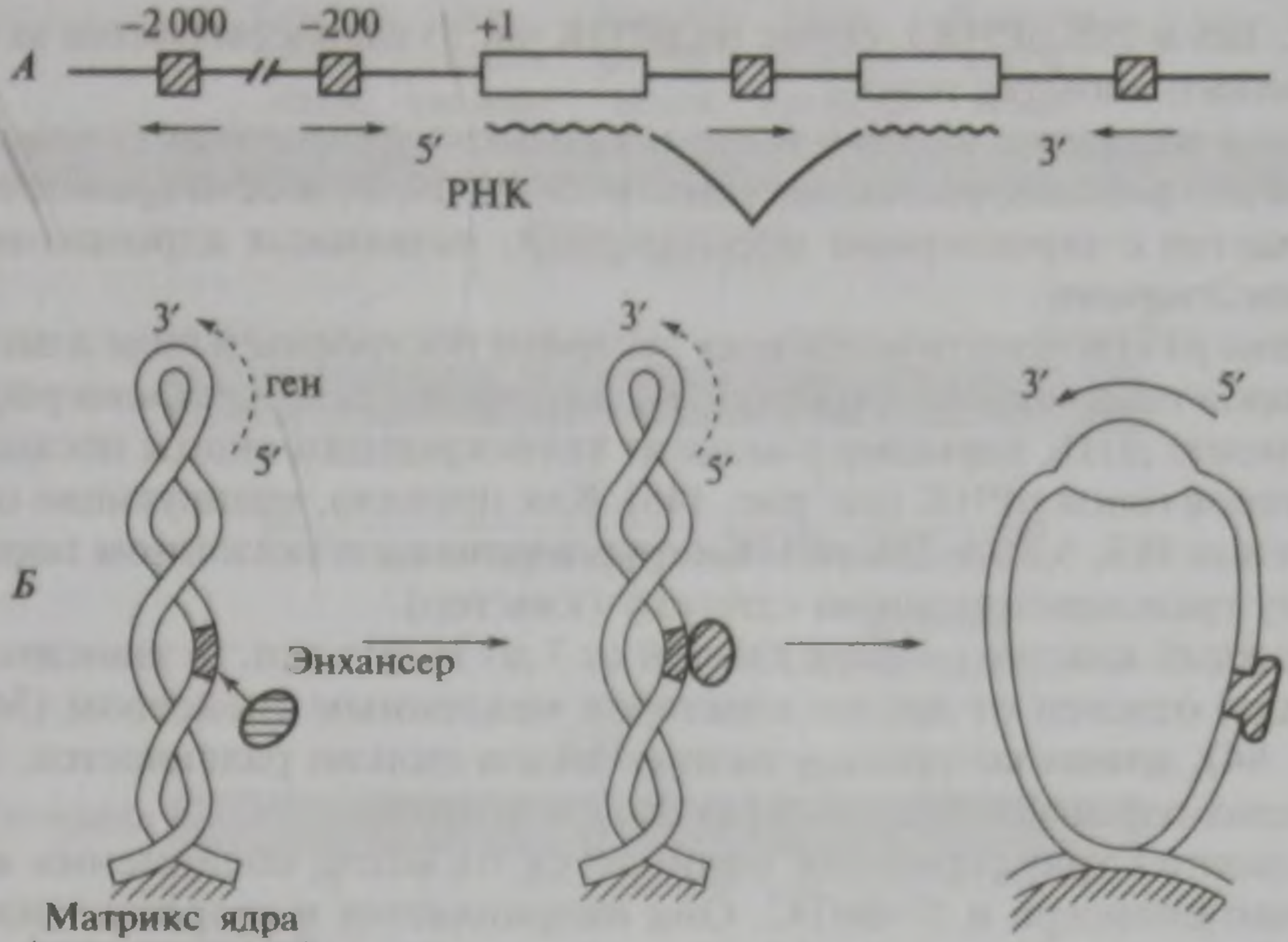


Рис. 83. Энхансеры эукариотических генов:

А — возможные районы локализации энхансера ( ) с разной ориентацией (→) относительно направления транскрипции; *Б* — взаимодействие белка с энхансером; стрелка — транскрипция гена

гена. Негативное действие сайленсеров, как и в случае энхансеров, не зависит от положения и ориентации относительно сайта инициации транскрипции.

Таким образом, регуляторные элементы генов обладают рядом общих функциональных характеристик. На физической карте регуляторной области гена они могут располагаться в одном районе, создавая сложную мозаику регуляторных сигналов. Отсутствие требований к строгой локализации и ориентации регуляторных последовательностей относительно генов расширяет возможности создания новых регуляторных связей путем перетасовки регуляторных элементов в процессе эволюции. В таком случае эволюция регуляторных элементов эукариот может быть основана на изменчивости мозаики ограниченного числа регуляторных нуклеотидных мотивов.

Рибосомные гены

Число генов, кодирующих рРНК (р-гены, рДНК), колеблется от сотен до нескольких тысяч в зависимости от вида эукариот. Так как транскрипт р-гена дает только одну копию конечного продукта

(5,8S, 18S и 28S рРНК), спрос на рРНК часто удовлетворяется за счет огромного избытка генов.

Гены находятся в одной или в нескольких хромосомах (у человека около 200 р-генов, расположенных в 13, 14, 15, 21 и 22-й хромосомах) в областях с характерной морфологией, названных ядрышковыми организаторами.

Гены рРНК практически всех эукариот построены в виде длинных повторов типа «голова к хвосту». На электронных микрофотографиях фрагмента ДНК ядрышек в момент транскрипции видны несколько кластеров генов рРНК (см. рис. 148). Как правило, кодирующие области генов 18S, 5,8S и 28S рРНК сгруппированы в указанном порядке в одну транскрипционную единицу (кластер).

Каждый кластер р-генов длиной от 7 до 14 тыс. н.п. (в зависимости от вида) отделен от других кластеров межгенным спейсером (МГС) (рис. 84), длина которого у разных видов сильно различается, но в пределах одного вида относительно постоянна.

Единица транскрипции начинается от места соединения межгенного спейсера и 5′-ВнТС. Она направляется в сторону гена 18S рРНК и заканчивается в разных сайтах межгенного спейсера за последовательностью, кодирующей 28S рРНК.

Кодирующие области р-генов у разных видов являются консервативными. Длина и нуклеотидные последовательности межгенных спейсеров весьма сильно различаются от вида к виду. Общей особенностью последних является то, что они обычно содержат прямые тандемные повторы, сходные с теми, которые окружают сайты начала транскрипции, т. е. сигнальные последовательности для РНК-полимеразы I. Разное число актов кроссинговера, происходящих между тандемно повторяющимися единицами в межгенном спейсере, приводит к различиям в числе самих этих единиц, а также в числе полноразмерных р-генов. Таким образом, межгенные спейсеры эволюционируют как согласованно с р-геном, в который они входят, так и независимо от него. В результате кроссинговера общая длина повторяющейся единицы р-генов варьирует от одного близкородственного вида к другому и даже в пределах одного вида.

Единица транскрипции р-генов в ДНК человека имеет размер около 14 тыс. н.п., приближаясь к самой длинной из известных. Примерно половина областей р-генов *Drosophila*, кодирующих 26S рРНК, имеют вставки различных размеров; они не являются интронами, и кластеры с такими вставками не транскрибируются.

После завершения транскрипции про-рРНК расщепляется под действием эндонуклеазы на три зрелые молекулы рРНК: 5,8S, 18S и 28S.

Гены 5S рРНК представлены сотнями и даже тысячами тандемно повторяющихся копий, разделенных нетранскрибируемыми спейсерами. Они собраны в одном или нескольких районах генома и, как правило, сцеплены с генами 18S, 5,8S и 28S рРНК. Ген 5S рРНК,

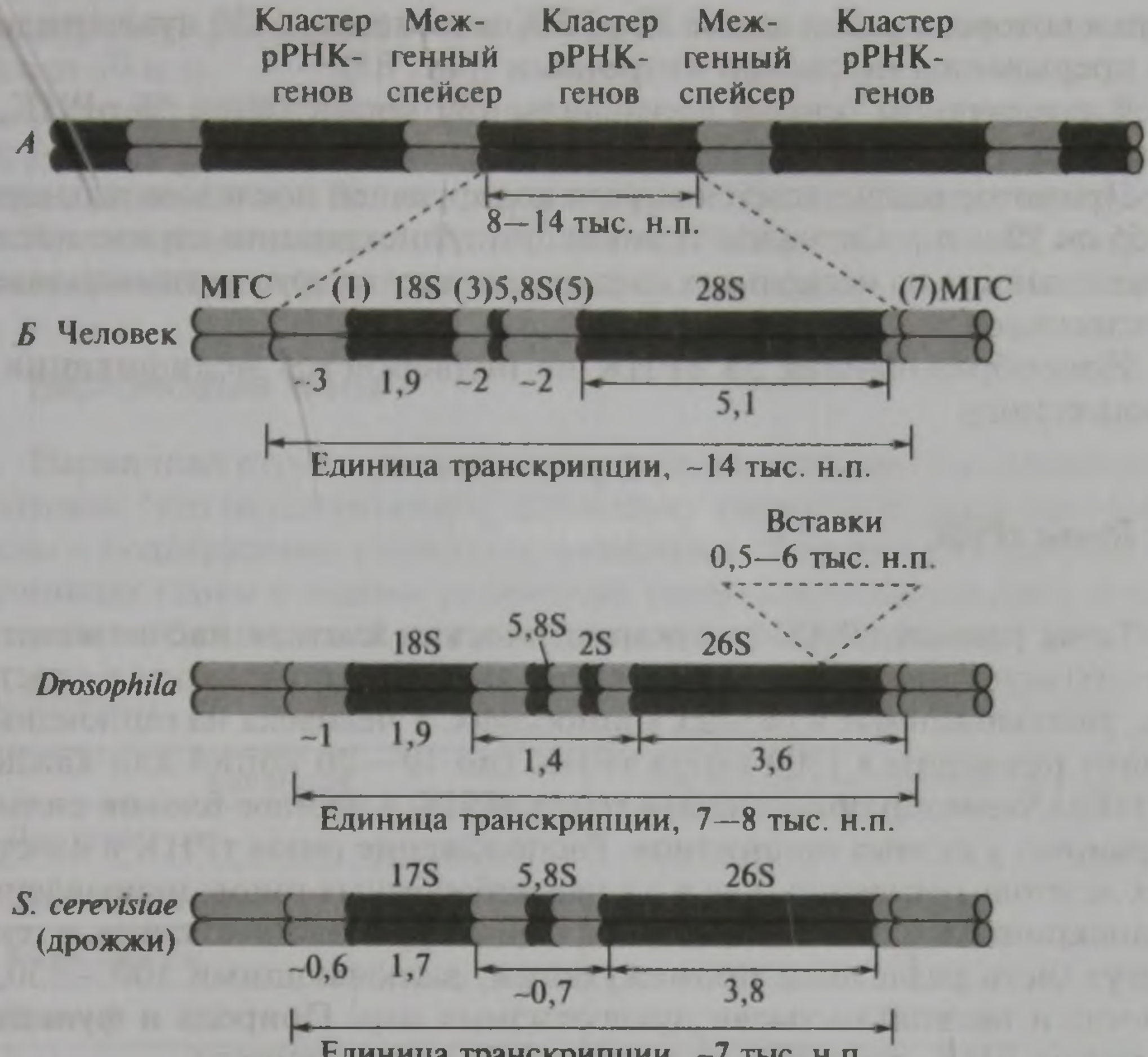


Рис. 84. Организация кластеров р-генов эукариот:

А — каждый кластер р-генов длиной от 8 до 14 тыс. н.п., отделенный от других кластеров межгенными спейсерами (МГС); *Б* — кластер р-генов (единица транскрипции) человека, дрозофилы, дрожжей, состоящий из последовательностей, кодирующих 18S, 5,8S и 28S рРНК, а также транскрибируемых спейсеров: 5′-внешнего (1); 5′-внутреннего (3); 3′-внутреннего (5); 3′-внешнего (7); ген 26S рРНК дрозофилы содержит вставку длиной 0,5 — 6 тыс. н. п.

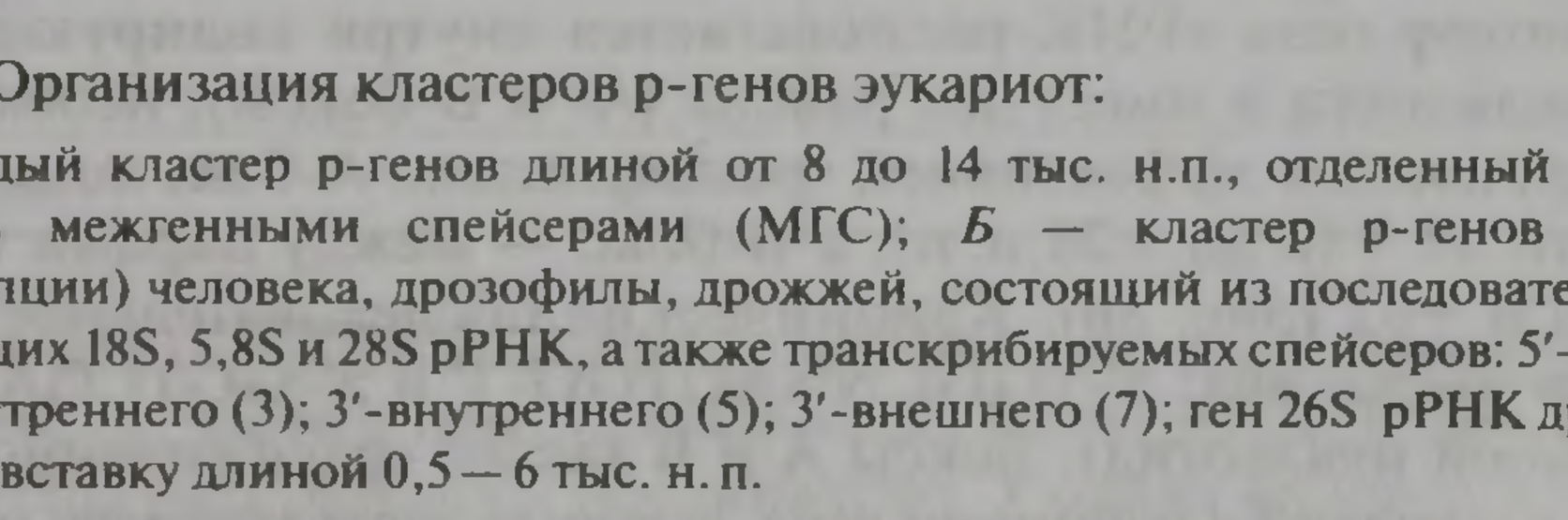


Рис. 85. Гены 5S рРНК

Промотор (■) располагается внутри гена (25 н.п.)

длина которого равна длине 5S рРНК и составляет 120 нуклеотидов, не прерывается никакими интронами (рис. 85).

В гаплоидном геноме дрозофилы 160 копий генов 5S рРНК, у человека — 2 000 (локализация на 1-й хромосоме).

Промотор располагается внутри кодирующей последовательности (с 55 по 80 н. п.). Сигналом терминации транскрипции служит последовательность из нескольких следующих друг за другом тимидиловых нуклеотидов.

Новообразованная 5S рРНК не подвергается модификации и процессингу.

Гены тРНК

Гены разных тРНК (в эукариотических клетках насчитывается 40 — 60 основных типов тРНК) также часто сгруппированы в кластеры, расположенные в разных хромосомах. У человека на гаплоидный геном приходится 1300 генов тРНК (по 10 — 20 копий для каждой тРНК). Характер организации генов тРНК в составе блоков сильно варьирует у разных организмов. Расположение генов тРНК в кластерах не столь регулярно, как в случае рибосомных генов: направление транскрипции близлежащих генов бывает противоположным, а гены могут быть разделены промежутками, включающими 100 — 150, а иногда и несколько тысяч нуклеотидных пар. Природа и функции участков ДНК, разделяющих гены тРНК, не выяснены.

Промотор гена тРНК располагается внутри кодирующей последовательности и имеет два района (А- и В-боксы), необходимые для правильной и эффективной транскрипции. А-бокс локализован в области от +10 до +20 н. п., а В-бокс — между парами оснований +50 и +65 (рис. 86). Канонические последовательности А- и В-боксов имеют вид: 5′-TGGCNNAGTGG-3′ и 5′-GGTCGANNC-3′ (N — любой нуклеотид). Боксы А и В такого «расщепленного промотора» совпадают с районами гена, которые соответствуют наиболее консервативным районам молекулы тРНК в дигидроуридиловой и псевдоуридиловой петлях, содержащих наибольшее количество инвариантных нуклеотидов. Для активной транскрипции требуется

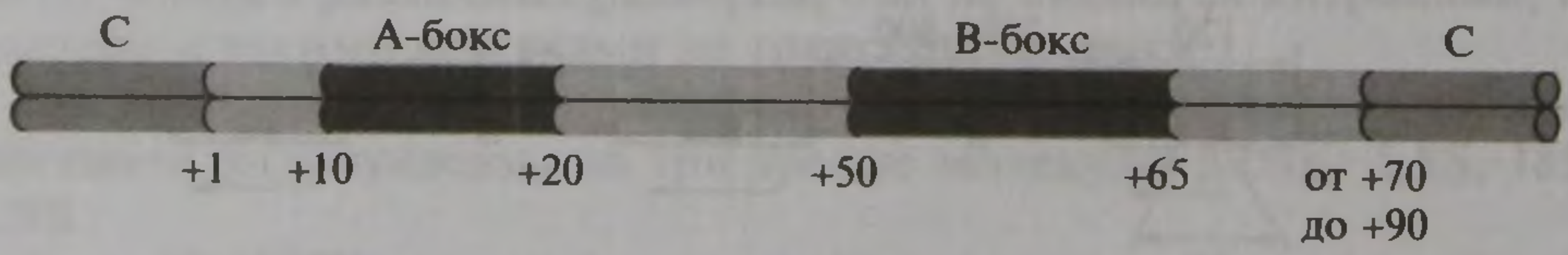


Рис. 86. Структура гена тРНК.

Промотор располагается внутри гена (А- и В-боксы). (+1) — сайт инициации транскрипции; С — нетранскрибируемые спейсеры

сохранность расстояния между А- и В-боксами, которое составляет около 50 н. п.

В генах тРНК многих организмов имеется один интрон размером от 14 до 60 н. п. (в зависимости от вида тРНК). Он всегда локализуется в одном и том же месте у разных тРНК: через один нуклеотид от 3′-конца антикодона. Район генов тРНК, кодирующий интрон, выщепляется в результате процессинга.

Гистоновые гены

Первичная структура гистонов у разных эукариот высококонсервативна. Это неудивительно, поскольку гистоны играют ключевую роль в поддержании структуры хроматина. Частоты повторения гистоновых генов в геноме различных организмов изменялись в ходе эволюции без видимой связи с усложнением организмов (табл. 14).

Таблица 14. **Частоты повторения гистоновых генов у эукариот**

Организм	Число генов на гаплоидный геном
Дрожжи	2
Дрозофила	110
Морской еж	470
Тритон	700
Курица	10
Мышь	10 — 20
Человек	40

Гены разных гистонов сцеплены. Все они транскрибируются по отдельности, часто с противоположных цепей ДНК. В геноме гистоновые гены распределены неодинаково. У одних видов они образуют тандемные повторы, у птиц и млекопитающих (рис. 87) рассеяны по всему геному. Большинство гистоновых генов лишено интронов, а у мРНК отсутствуют 3′-поли(А)-концы.

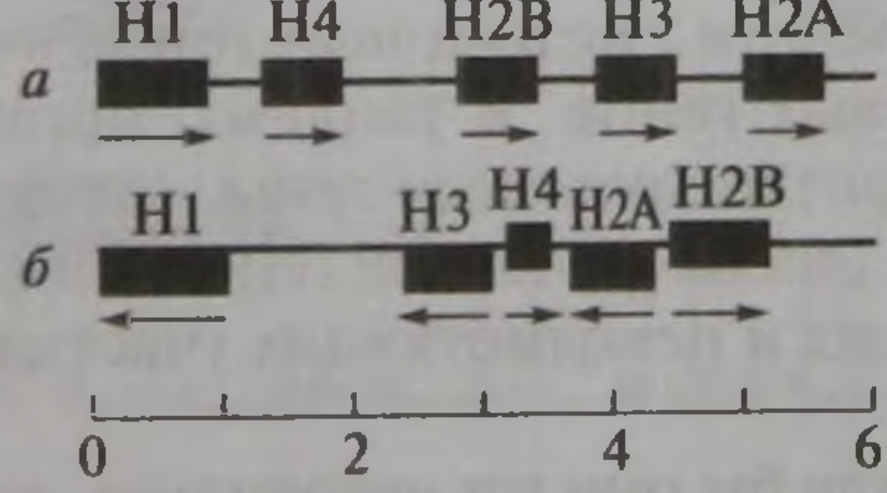


Рис. 87. Расположение гистоновых генов в повторяющейся единице морского ежа (*а*) и дрозофилы (*б*) (стрелками показано направление транскрипции)

Таким образом, большая часть изученных генов у эукариот представлены или множеством сходных копий, или ограниченным набором дуплицированных и дивергировавших в разной степени последовательностей. Те и другие носят название «мультигенные семейства». Повторяемость генов в мультигенных семействах необходима либо для усиления *дозы* генов, либо для обеспечения синтеза различных *вариантов* родственных полипептидов. Дозовая повторяемость характерна для генов, кодирующих РНК рибосом, тРНК, а также гистоновых генов. Именно множественность генов, обслуживающих аппарат трансляции, обеспечивает физиологическую реактивность клетки, так как темп белкового синтеза часто лимитируется количеством рибосом и тРНК. Очевидна целесообразность множественности и гистоновых генов, так как в определенные моменты жизни клетка должна за короткий интервал времени одновременно с репликацией синтезировать полный комплект гистонов.

Вариантная повторяемость свойственна таким генам, как глобиновые, актиновые, иммуноглобулиновые, хорионные и т. д. Число генов в этих мультигенных семействах, как правило, небольшое, а степень гомологии отдельных его членов варьирует в широких пределах. Члены мультигенных семейств происходят, по-видимому, от одной общей предковой последовательности путем дупликации и последующей дивергенции.

Мультигенные семейства отличаются друг от друга по степени сложности организации. Наиболее простые мультигенные семейства образуют гены, кодирующие рРНК. Более сложно устроены гены для гистонов и тРНК. И наиболее сложное строение характерно для регулируемых мультигенных семейств, контролирующих ключевые стадии развития организмов.

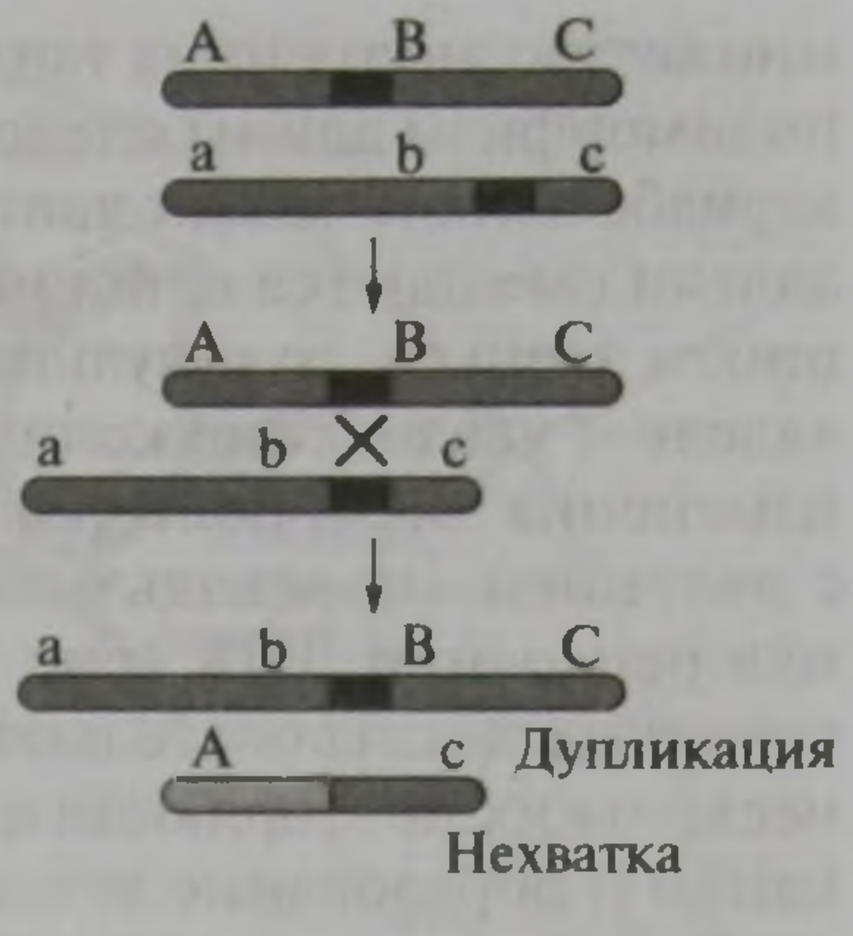
Тандемные повторы

Тандемно организованные повторяющиеся последовательности ДНК различных типов широко представлены в эукариотических геномах. Они варьируют по длине кластера и размеру повторяющегося звена.

В случае генов рРНК или гистоновых генов тандемно повторяются группы одинаковых генов. В тандемы организованы такие родственные, но все же различающиеся гены, как члены мультигенных семей глобиновых и овальбулиновых генов. Тандемные повторы встречаются в кодирующих и некодирующих участках, между генами и в интронах.

Тандемные повторы, где бы они ни находились, как правило, связаны с межвидовым полиморфизмом. Спейсеры между генами рРНК и между транскрипционными единицами рРНК обычно содержат

Рис. 88. Неравный кроссинговер



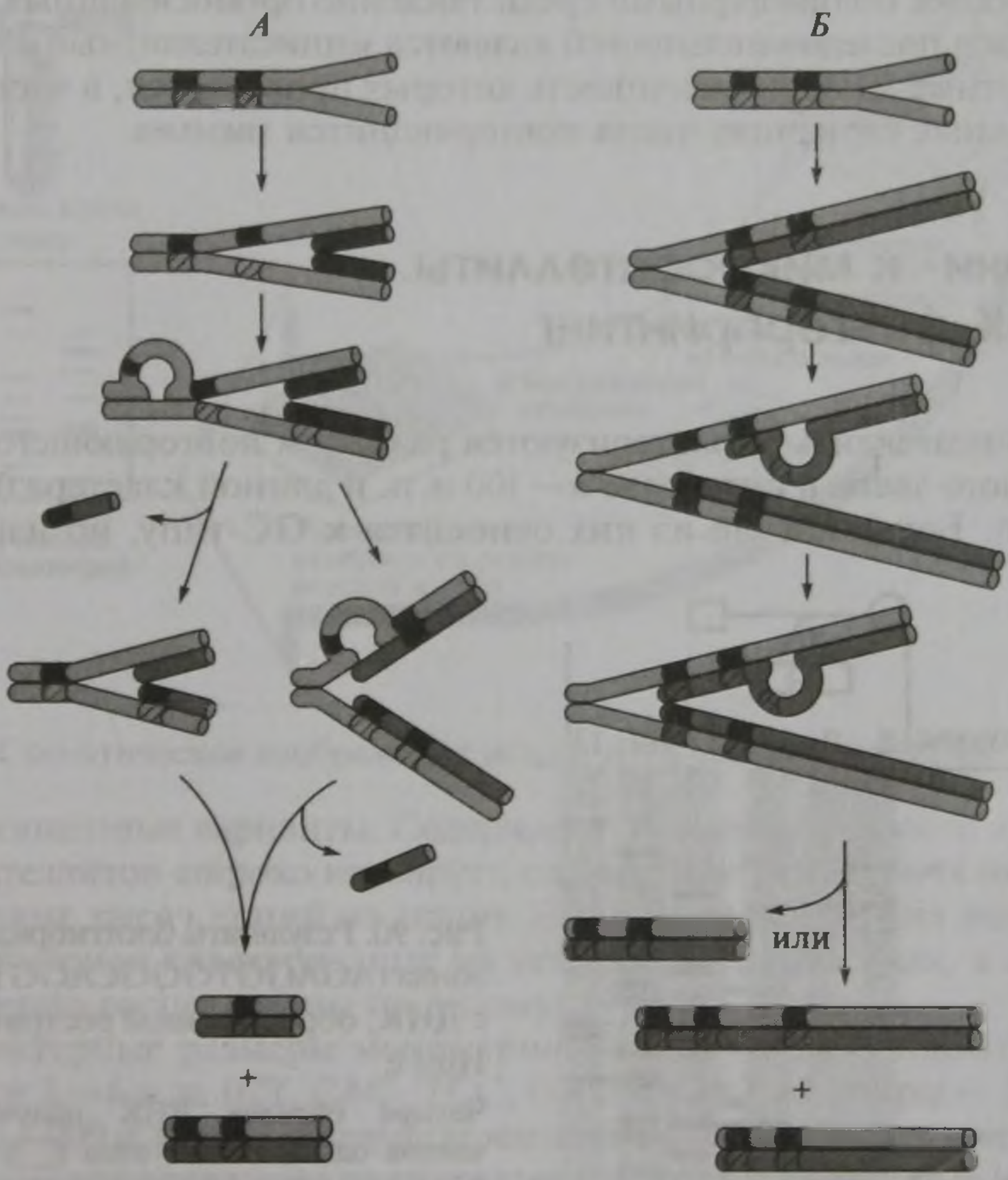


Рис. 89. Делеции и амплификации, возникающие при репликации или репарации:

А — «проскальзывание» матричной цепи, вырезание петли или репликация (минуя петлю) с последующей репарацией, приводящие к делеции; *Б* — «проскальзывание» синтезируемой цепи и последующая репарация, приводящие к амплификации

множество внутренних тандемных повторов, которые ответственны за полиморфизм длины спейсеров. Причиной происхождения подобной вариабельности может явиться неравный кроссинговер (рис. 88). Если аллели смещаются относительно друг друга на несколько повторяющихся звеньев, то результатом кроссинговера будет делеция в одном аллеле и увеличение количества повторов в другом. Наиболее часто изменения числа повторов коротких последовательностей связывают с явлением «проскальзывания» нити ДНК в процессе репликации или репарации ДНК (рис. 89). При наличии тандемно повторенных коротких последовательностей могут наблюдаться «выпетливание» нескольких повторяющихся единиц отстающей цепи в вилке репликации и образование дуплекса с матричной нитью со смещением на соответствующее количество тандемных повторов.

Наиболее полиморфными среди тандемно организованных повторяющихся последовательностей являются минисателлитные и микросателлитные ДНК, изменчивость которых проявляется, в частности, в аллельных вариациях числа повторяющихся звеньев.

Мини- и микросателлиты. ДНК-фингерпринтинг

Минисателлиты характеризуются размером повторяющегося мономерного звена в интервале 6 — 100 н. п. и длиной кластера 0,2 — 20 тыс. н.п. Большинство из них относится к GC-типу, но известны



Рис. 90. Результаты блотгибридизации зонда $(AGAGGTGGGCAGGTGG)_{29}$ с ДНК, обработанной рестриктазой Hinf I.

Четыре образца ДНК получены от членов одной семьи: отца (□), матери (○) и их детей — однояйцовых близнецов. Дорожки 12 и 13 — неродственные индивидуумы мужского пола. Стрелками отмечены «родительские» сателлитные полосы на электрофореграммах ДНК близнецов (*Jeffreys A.* и др. // Nature. — 1985. — V. 316. — P. 76)

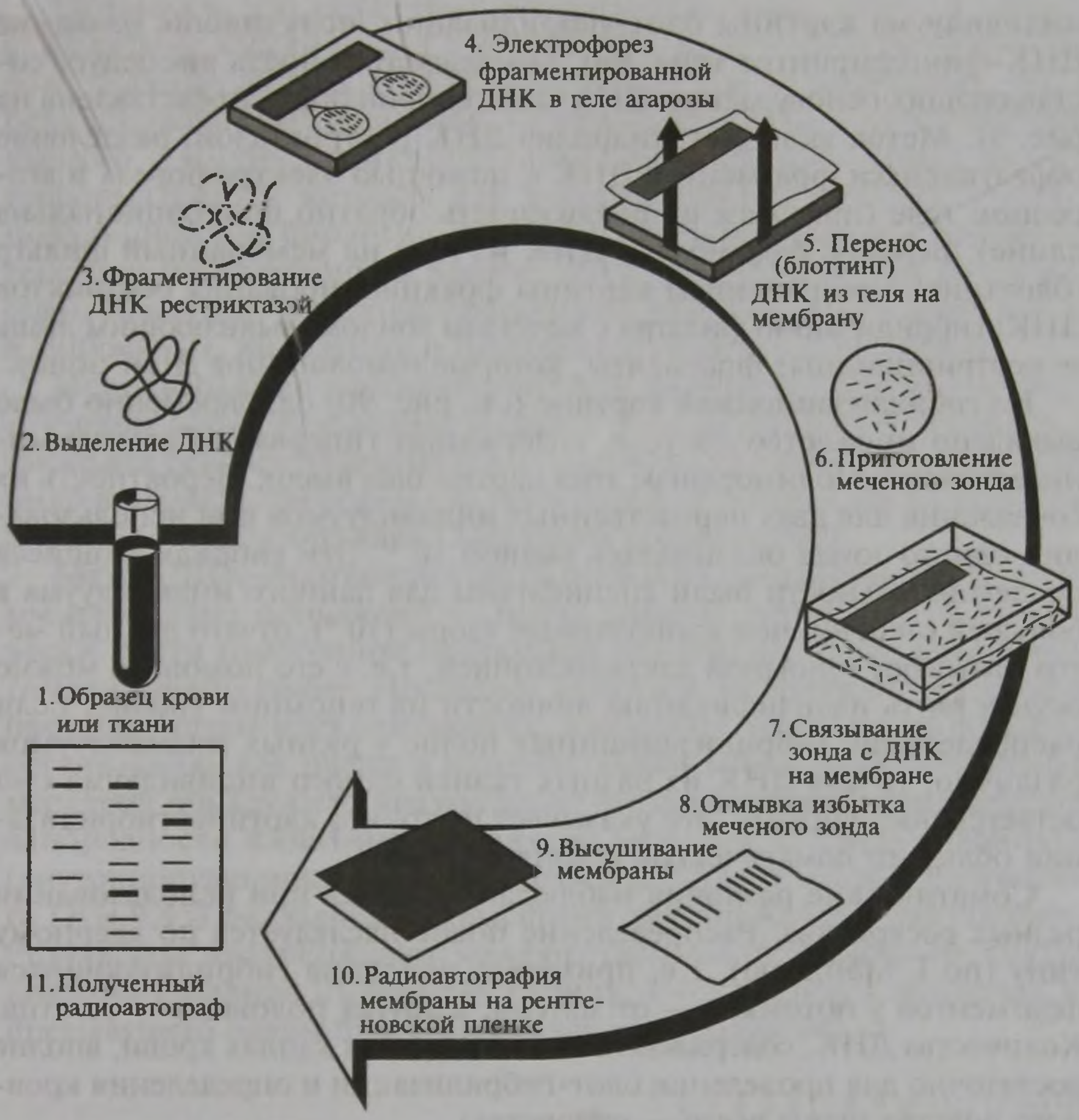


Рис. 91. Схематическое изображение метода ДНК-фингерпринтинга

АТ-обогащенные варианты. Содержание различных классов и типов минисателлитов широко варьирует, составляя от нескольких сотен до нескольких тысяч копий на геном. Для некоторых из них показана прителомерная кластеризация на некоторых хромосомах, а другие равномерно распределены по геному.

Характерные размеры мономерного звена микросателлитов составляют 2—6 н.п. (CT, CAC, TCC, GACA и др.), кластеры — 20—60 н.п. (до 1 000 н.п. при некоторых наследственных заболеваниях, когда происходит экспансия триплетных повторов). Некоторые микросателлиты насчитывают 10^4—10^5 копий, которые распределяются по всему геному равномерно.

А. Джеффрис и сотр. (1985), используя минисателлитные пробы в качестве зонда для детектирования множества гомологичных локусов в ДНК человека, обнаружили специфические для каждого

индивидуума картины блот-гибридизации, получившие название ДНК-фингерпринтов (рис. 90). Последовательность процедур, составляющих основу метода ДНК-фингерпринтинга, представлена на рис. 91. Метод включает: гидролиз ДНК-рестриктазой; разделение образующихся фрагментов ДНК с помощью электрофореза в агарозном геле (при этом их подвижность обратно пропорциональна длине); перенос фрагментов ДНК из геля на мембранный фильтр (блоттинг) с сохранением картины фракционирования рестриктов ДНК; гибридизацию фильтра с меченым зондом, выявляющим лишь те рестрикционные фрагменты, которые гомологичны ДНК-зонду.

На гибридизационной картине (см. рис. 90) одновременно было выявлено множество локусов, содержащих гипервариабельные минисателлиты. Полиморфизм этих картин был высок. Вероятность их совпадения для двух неродственных индивидуумов при использовании одного зонда оказывалась равной 10^{-11}. Эти гибридизующиеся последовательности были специфичны для данного индивидуума в большей степени, чем капиллярные узоры (10^{-6}), отчего данный метод называют геномной дактилоскопией, т.е. с его помощью можно осуществить идентификацию личности на геномном уровне. Если распределение гибридизационных полос у разных индивидуумов различно, то для ДНК из разных тканей одного индивидуума оно остается постоянным. Это указывает на то, что картины гибридизации обладают соматической стабильностью.

Соматические различия наблюдаются лишь при использовании разных рестриктаз. Распределение полос наследуется по ядерному типу (по Г. Менделю), т.е. примерно половина гибридизующихся фрагментов у потомков — от матери, а другая половина — от отца. Количества ДНК, содержащегося в нескольких каплях крови, вполне достаточно для проведения блот-гибридизации и определения кровного родства (чаще всего — отцовства).

В 1987 г. бельгийский исследователь Дж. Вассарт, а также российский ученый А. П. Рысков с сотр. обнаружили, что минисателлиты фага M13 (15-членное звено — GAGGGTGGXGGGXTCT (X — любой нуклеотид) образует в геноме фага два кластера, повторяясь в одном участке 10, а в другом 4 раза, и находится в гене белка III) широко представлены у разных организмов, включая человека, беспозвоночных животных, растения, низших эукариот и бактерий. Использование этой пробы для блот-гибридизационного анализа ДНК неродственных индивидуумов выявило чрезвычайно высокий уровень полиморфизма (20—25 полос гибридизации) (рис. 92). Тот факт, что гомологичные минисателлиты обнаруживаются в столь разных организмах (человек и фаг M13), предполагает широкое распространение этих последовательностей в живой природе и возможность использования ДНК-фингерпринтинга с зондом ДНК фага M13 для исследования индивидуальности, генетического разнообразия и родства любых живых существ.

Оказалось, что некоторые гипервариабельные мини- и микросателлитные аллели имеют изменчивость в 100 — 1000 раз большую, чем обычные локусы генома, вследствие как мутационных, так и рекомбинационных процессов. В связи с этим ДНК-фингерпринтинг именно с мини- и микросателлитами в качестве зонда явился информативным методом для фундаментальных исследований в популяционной биологии в плане изучения биоразнообразия на генетико-популяционном уровне. ДНК-фингерпринтинг позволяет оценить внутри- и межпопуляционную генетическую изменчивость в аутбредных, инбредных и партеногенетических популяциях.

Использование минисателлитных маркеров эффективно при изучении дифференциации малых и изолированных популяций, а также для реконструкции их филогенетических связей, что прекрасно было продемонстрировано на популяции островных лисиц (*Urocyon littoralis*) и горбатого кита (*Megaptera novaeangliae*).

Метод ДНК-фингерпринтинга позволил охарактеризовать геномы инбредных популяций — линии мышей, породы кур и генетически почти гомогенной природной популяции голых землекопов *Heterocephalus glaber*. Последнее животное является единственным представителем млекопитающих, имеющим такую же социальную организацию, как пчелы и термиты. Применение минисателлитных маркеров обнаружило чрезвычайно высокий уровень генетической однородности животных как внутри вида, так и между колониями (внутрипопуляционные коэффициенты родства ($r \approx 0{,}8$) и инбридинга ($F \approx 0{,}45$) оказались наивысшими для диких млекопитающих). Результаты подтверждают межродственное скрещивание внутри колонии (колония имеет одну царицу; 1 — 3 самцов и молодняк из предыдущего потомства, обслуживающий колонию).

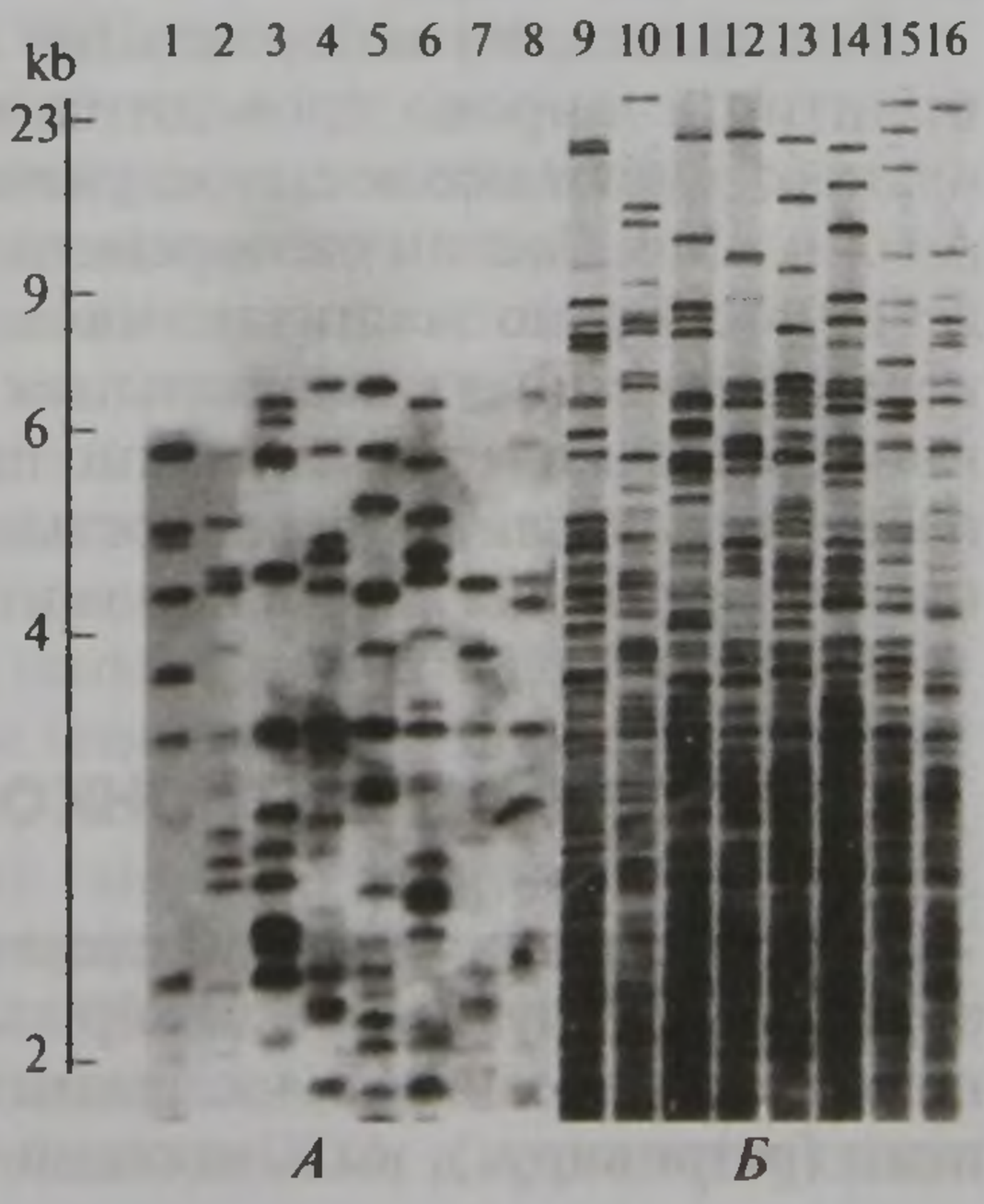


Рис. 92. Гипервариабельные участки ДНК человека, детектируемые с помощью ДНК фага M13 (*Просняк М. И.* и др. // Генетика. — М., 1990. — Т. 26. — С. 134):

А — ДНК фага M13, меченная ^{32}P; *Б* — ДНК фага M13, безызотопный вариант. Использована ДНК восьми индивидуумов

ДНК-фингерпринтинг оказался очень информативным методом в популяционной биологии для решения многочисленных задач в программах, связанных с оценкой, рациональным использованием и консервацией генетических ресурсов, особенно для выработки научно обоснованных рекомендаций по сохранению и поддержанию редких и исчезающих видов.

Выбор популяционных образцов для консервации (при создании заповедника или криогенных банков клеток и ДНК) определяется уникальностью их генотипа. При снижении популяционной численности значительное влияние на генетическую структуру оказывают дрейф и инбридинг. Все эти параметры определяются полиморфными маркерами и в первую очередь гипервариабельными мини- и микросателлитами. Здесь следует отметить работы по исследованию исчезающей популяции волка (*Canis lupus*) в международном биосферном заповеднике на о. Айл-Ройал (Великие озера, Северная Америка); по восстановлению исчезающего пуэрториканского попугая (*Amazona vittata*); сибирского журавля (*Grus leucogeranus*) или стерха; асконийской популяции лошади Пржевальского (*Equus przewalskii*) и др.

Исследование методом ДНК-фингерпринтинга геномов клонально размножающихся однополых видов позвоночных на примере партеногенетических видов кавказских ящериц рода *Lacerta* позволило вплотную подойти к выяснению природы и механизмов генетической изменчивости, а также факторов, ее определяющих.

Большой интерес метод представляет для медицины в плане ДНК-диагностики. Обнаружена взаимосвязь между так называемыми динамическими мутациями и патогенезом ряда неврологических наследственных заболеваний (хорея Гентингтона, синдром ломкой Х-хромосомы и др.).

Генетическая паспортизация населения методом ДНК-фингерпринтинга широко проводится в развитых странах: генетический паспорт имеют все военнослужащие США и все мужское население Англии. Для России своевременной явилась бы паспортизация всех детей, особенно воспитанников детских домов и специнтернатов, нередко склонных к асоциальным поступкам, в том числе к дромомании (бродяжничеству). Именно наличие у таких детей генетического паспорта позволит с уверенностью проводить их идентификацию, что несомненно будет иметь положительные социальные последствия.

Онкогены и антионкогены

Некоторые вирусы в определенных условиях способны вызывать образование опухолей. Эти вирусы были названы *онкогенными*. Первым среди описанных онкогенных вирусов был вирус саркомы Рауса птиц (ретровирус), вызывающий опухоли мышечной ткани цыплят.

Развитие опухоли обусловливал один ген ретровируса — *онкоген*. Оказалось, что в каждом случае ретровирусный онкоген имеет своего «двойника» в нормальных клетках. Вирусный ген обозначают приставкой *v* (от англ. *virus* — вирус), а соответствующий *клеточный ген* приставкой *c* (от англ. *cellular* — клеточный). Клеточный ген называется *протоонкогеном*. Протоонкогены кодируют белки, участвующие в важных регуляторных клеточных механизмах, нарушение которых может привести к неконтролируемому делению клеток. Некоторые из них кодируют факторы роста и транскрипции, мембранные рецепторы для факторов роста, белки контроля клеточного цикла (циклины), белки, участвующие в апоптозе, «белок бессмертия» теломеразу и т. д. В эволюционном плане ретровирусные онкогены могли произойти от ретровирусов в результате захвата генов хозяина. Один из путей активации онкогена — структурная мутация, например замена нуклеотида в ДНК, ведущая к замене аминокислоты в белковой цепи и к изменению функции белка. Другой путь — это усиление активности онкогена, ведущее к усилению синтеза белкового продукта онкогена и его накоплению в клетке. Активация онкогена могла быть результатом увеличения числа копий данного гена в геноме (амплификация) или следствием такой перестройки, когда рядом с потенциальным онкогеном появляется новый регуляторный элемент, обеспечивающий его более активную транскрипцию. Число известных онкогенов непрерывно растет и, по данным секвенирования ДНК человека, на февраль 2001 г. выявлено 130 онкогенов.

Вскоре после открытия первых онкогенов появились сообщения о существовании генов, утрата или подавление активности которых тоже приводит к развитию опухолей. Иными словами, белковые продукты этих генов препятствуют клетке превратиться в раковую. Эти гены были названы *антионкогенами*, или *генами-супрессорами опухолей* (*ГСО*). Число известных ГСО тоже быстро растет, хотя и уступает числу открытых онкогенов.

Многие гены, кодирующие белки передачи сигнала для роста и размножения клеток, являются потенциальными онкогенами. В отлаженной системе нормальной клетки они обеспечивают строго регулируемые рост и размножение, но при повреждении или нерегулируемой активации этих генов клетки превращаются в раковые.

Поскольку онкогенов и ГСО чрезвычайно много, набор мутаций в различных опухолях человека весьма разнообразен. В одних опухолях активированы или подавлены одни гены, в других — другие. Каждая опухоль, таким образом, имеет свой генетический портрет, который и характеризует ее свойства.

Известно, что апоптоз (клеточная смерть) определяется большим числом белков. Центральное место среди них занимает ген p53, кодирующий белок с мол. массой ~53 000 Да. Ген p53 играет важную роль в направлении клеток с поврежденной ДНК по пути апоптоза. Гены, способствующие апоптозу, относят к ГСО, а препятствующие

ему гены — к онкогенам. Когда ген p53 выведен из строя, клетки с поврежденной ДНК (мутантные клетки) перестают выбраковываться и в них происходит накопление новых мутаций, которые могут затрагивать ГСО и онкогены. Поэтому повреждение p53 резко усиливает аккумуляцию других изменений и дефект p53 наблюдается в опухолях очень часто. Так как опухоль — болезнь генома, она незаразна. С другой стороны, имеются вирусы — носители онкогенов. ДНК таких вирусов может встраиваться в геном клетки, и в такой клетке появляется активный онкоген. Таким образом, некоторые вирусы привносят в клетки новые онкогены, тем самым повышают вероятность возникновения рака у индивидуума. Конкретные механизмы развития каждой отдельной опухоли различны и знание истинной причины клеточной трансформации крайне важно для разработки методов эффективной терапии.

Подвижные генетические элементы эукариот

Существенную часть генома эукариот (10 — 30 %) составляют повторяющиеся последовательности, имеющие определенную структурную организацию и способные перемещаться в геноме как в пределах одной хромосомы, так и между хромосомами. Они получили название *подвижных генетических элементов*. Элементы включают в свой состав от тысячи до десятков тысяч нуклеотидных пар. Перемещение подвижных элементов осуществляется либо путем вырезания элемента из одного места и встраивания его в другое, либо путем образования копии подвижного элемента, внедряющейся в новое место, тогда как родительская копия остается на прежнем месте. В последнем случае увеличивается их число в геноме. Открытие подвижных элементов показало, что последовательность нуклеотидов ДНК по длине хромосомы изменяется благодаря перемещению этих элементов. Встраиваясь в гены или в соседние с генами области, они либо вызывают мутации, либо повышают уровень активности близлежащих генов. Все это позволило взглянуть по-новому на природу мутационных процессов и молекулярных механизмов эволюции генома.

Подвижные элементы эукариот представлены отдельными семействами, сходными по своей структуре и поведению. Внутри семейства различаются подсемейства идентичных или очень сходных подвижных элементов, число которых колеблется от нескольких копий до нескольких тысяч копий на геном. Наибольшее количество подвижных элементов обнаружено в геноме растений (до 50 %). Подвижные генетические элементы обычно рассеяны по геному, но могут концентрироваться в отдельных участках хромосом.

Различают два основных класса подвижных генетических элементов: *транспозоны* и *ретротранспозоны*. Такая классификация основана на молекулярных механизмах, с помощью которых перемещаются эти элементы.

Транспозоны. Эти элементы ограничены инвертированными повторами (последовательностями, направленными навстречу друг другу), как и некоторые транспозоны прокариот (см. с. 148). Примерами их могут служить Р-элемент дрозофилы и Ас-элемент кукурузы. В геномах этих организмов насчитывается по 30—50 копий таких элементов. Полные копии (часть копий дефектна, так как имеет внутренние делеции) содержат открытые рамки считывания, кодирующие транспозазу (рис. 93).

Схема перемещения транспозона показана на рис. 94. Инвертированные повторы необходимы для перемещения элемента, которое осуществляется благодаря их сближению друг с другом и узнаванию транспозазами. Инвертированные повторы сближаются и точно отрезаются от соседних участков ДНК хозяина. Вырезанный транспозон внедряется в район вносимого транспозазой разрыва в молекуле-мишени и сшивается с ДНК хозяина в новом месте. Разрыв и зашивание осуществляются транспозазой и вспомогательными белками. Транспозаза может кодироваться как самим подвижным элементом, который будет перемещаться, так и другой копией элемента, локализованной в том же геноме в отдалении. Р-элементы обычно встраиваются в определенные сайты, консенсусная последовательность которых такова: 5′-GGCCAGAC. При встраивании Р-элемента этот сайт-мишень дуплицируется.

Р-элемент содержится только в некоторых линиях *D. melanogaster*. Скрещивание самок без Р-элемента с самцами, несущими Р-элементы, приводит у гибридов к транспозициям Р-элемента, которые наблюдаются только в клетках зародышевого пути. Совокупность наблюдаемых эффектов называют *гибридным дисгенезом*. Дисгенные мухи часто оказываются стерильными, а если они фертильны, то у их потомков обнаруживается множество мутаций и эти признаки

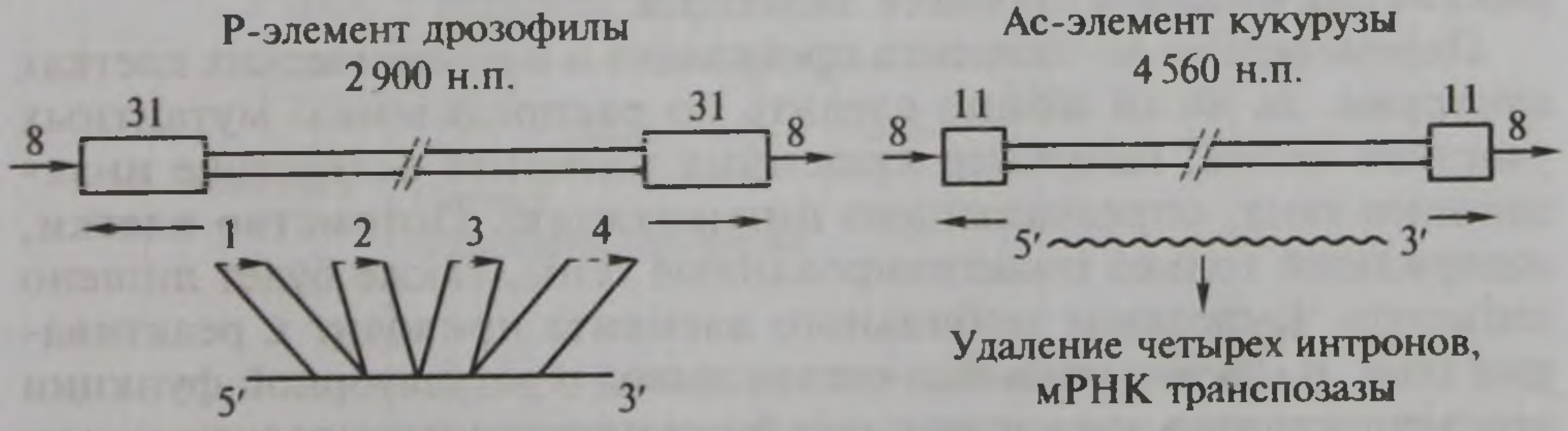


Рис. 93. Транспозоны эукариот

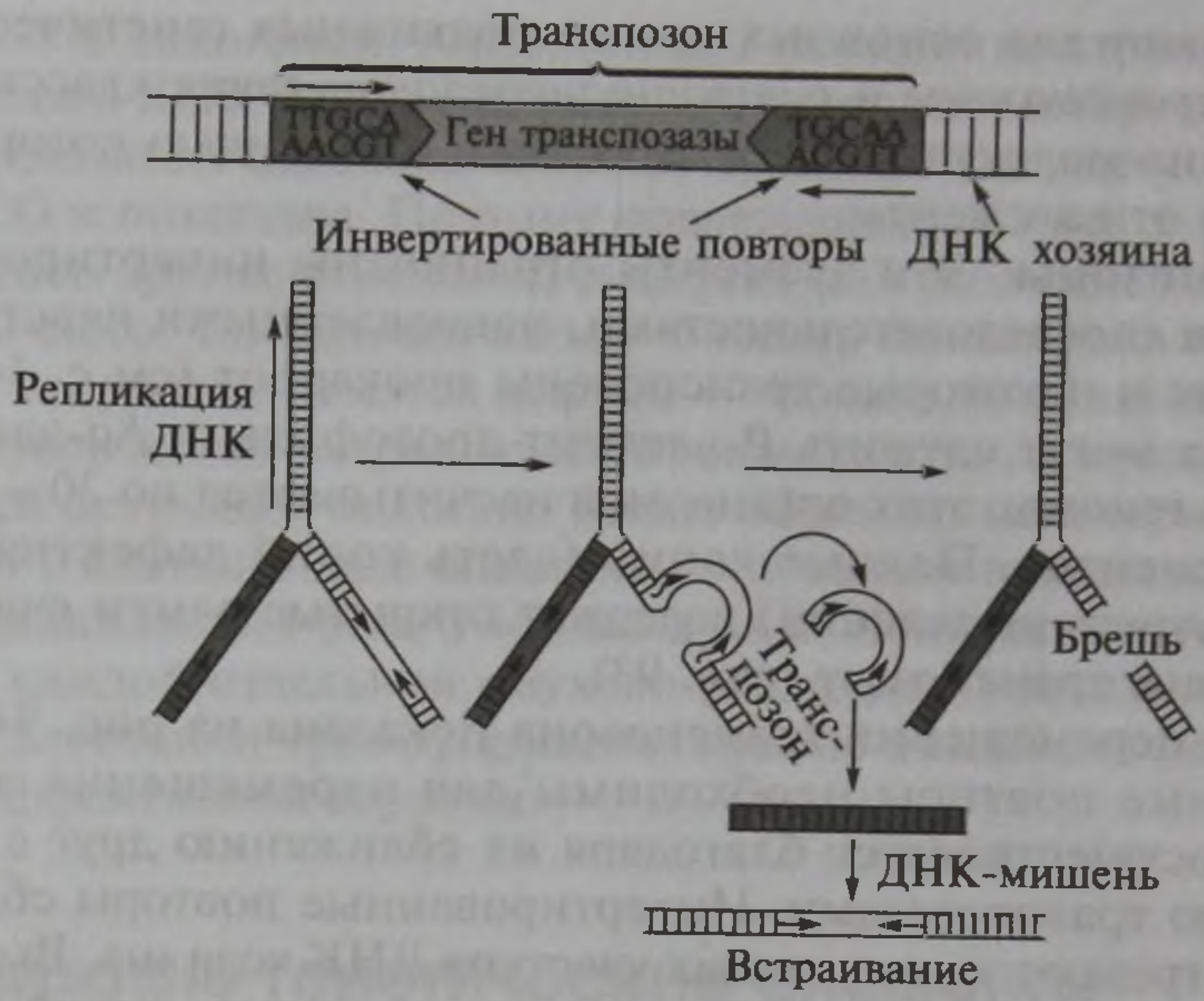


Рис. 94. Перемещение транспозона

передаются следующим поколениям. Поэтому линии с Р-элементом и без него выглядят как репродуктивно изолированные, по крайней мере частично. Биологическая изоляция играет огромную роль в процессе эволюции. Изоляция линий вызвана активацией транспозиций Р-элемента, присутствующего в одной из них. Причина, почему транспозиции Р-элемента ограничены зародышевыми клетками, объясняется тем, что только в клетках-предшественниках гамет осуществляется такой тип сплайсинга, который приводит к образованию непрерывной открытой рамки трансляции, кодирующей транспозазу (см. рис. 93). Ограничение транспозиции зародышевыми клетками имеет определенный смысл, поскольку обеспечивает выживание особей, несущих гаметы, в которых произошли геномные перестройки вследствие транспозиции Р-элемента. Подобный «геномный шок», сопровождающийся высокой частотой мутагенеза, может обеспечить большую степень геномной изменчивости, которая послужит материалом для отбора в процессе эволюции.

Перемещения Ас-элемента происходят и в соматических клетках кукурузы. За ними можно следить по распределению мутантных участков ткани, например лишенных пигмента вследствие инактивации гена, определяющего пигментацию. Потомство клетки, содержащей только инактивированные гены, также будет лишено пигмента. Вырезание мобильного элемента приводит к реактивации гена. В связи с этим был сделан вывод о регуляторной функции перемещающихся элементов и они были названы *контролирующими* (Б. Мак-Клинток).

Брешь в ДНК после вырезания транспозона может заделываться либо с помощью гомологичного отрезка ДНК, либо за счет сшивания

разорванных концов (это особенно характерно для транспозонов растений).

В целом подвижные элементы эукариот представляют собой чрезвычайно разнородную популяцию. Существует мнение о том, что они не влияют на фенотип организма и размножаются в геноме лишь благодаря особенностям своей структуры, в результате чего постепенно заселяют геном. Предполагается, что они составляют часть так называемой эгоистичной ДНК, размножение которой ограничивается естественным отбором. Отбор устраняет те случаи, когда внедрение эгоистичной ДНК сопровождается вредными последствиями. На определенных стадиях эволюции эти элементы могут использоваться для целей, которые в дальнейшем обеспечат селективные преимущества организму.

Большой класс подвижных элементов представляют *ретротранспозоны*, сходные по своей структуре с проретровирусами, которые внедряются в геном, используя механизмы обратной транскрипции (см. с. 236). Эти элементы содержат «тело» размером 5—8 тыс. н.п., ограниченное прямыми длинными концевыми повторами (ДКП, или LTR — от англ. *long terminal repeats*), обычно содержащими по 300—400 н. п. (рис. 95). Число копий этих элементов, принадлежащих к одному семейству, достаточно постоянно для вида, но варьирует от нескольких копий до сотен тысяч копий в зависимости от типа ретротранспозона. Ретротранспозоны, относящиеся к разным семействам, не имеют протяженных гомологий нуклеотидных последовательностей ни в ДКП, ни в теле. В составе «тела» элемента обнаруживаются открытые рамки считывания для обратной транскриптазы и нуклеазы (интегразы). Разные копии одного семейства могут иметь вставки и делеции как в ДКП, так и в самом элементе. По флангам ДКП находятся короткие прямые повторы. Ретротранспозоны широко распространены у эукариот.

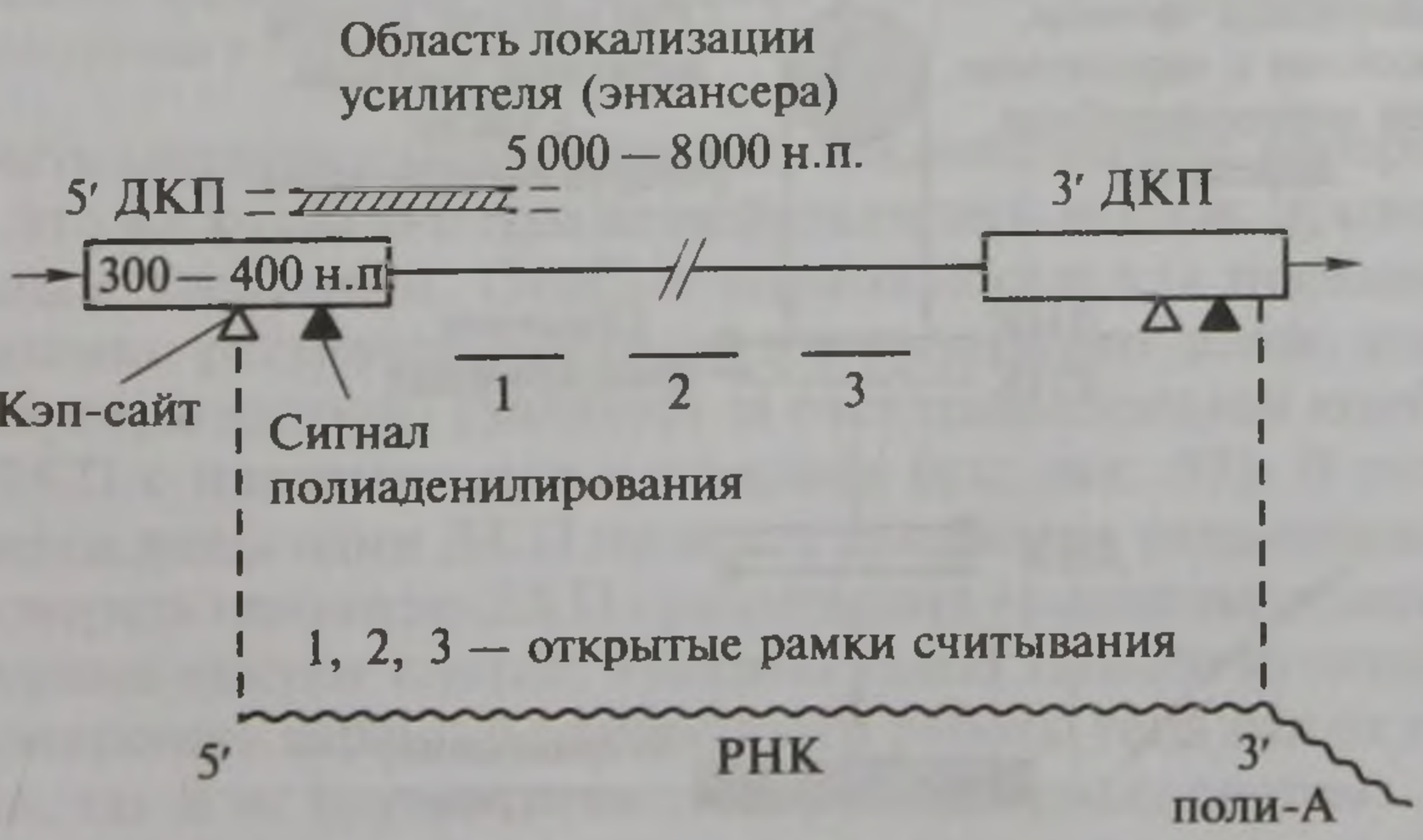


Рис. 95. Структура ретротранспозона с длинным концевым повтором

Один из способов перемещения ретротранспозонов предполагает его транскрипцию с помощью РНК-полимеразы II. В составе ДКП имеются сайты инициации транскрипции и сигналы полиаденилирования. ДКП могут служить активными промоторами. Транскрипция начинается в одном ДКП (условно левом, 5′-ДКП), а заканчивается в другом (правом, 3′-ДКП). Синтезируемая молекула РНК транслируется с образованием ферментов, необходимых для синтеза ДНК и внедрения ее в геном. Эта схема полностью повторяет схему образования и интеграции провируса (рис. 96). Остается открытым вопрос: произошли ли ретровирусы из ретротранспозонов или, напротив, ретротранспозоны возникли из вирусов в результате потери способности к заражению? Подавляющая часть ретротранспозонов (при сравнении их с ретровирусами) либо потеряла ген оболочки вируса, либо еще не приобрела его. Таким образом, ретротранспозоны представляют собой внутригеномные, неинфекционные элементы, способные лишь к самовоспроизведению и «подзаражению» того же генома.

Внедрение ретротранспозона внутрь гена или около гена вызывает разные эффекты. Во многих случаях происходит инактивация гена (рис. 97). При интеграции в район промотора на 5′-фланге гена ретротранспозон может резко активировать экспрессию гена, обеспечивая транскрипцию с собственного промотора. Если в этом случае подвижный элемент оказался около протоонкогена, то результатом

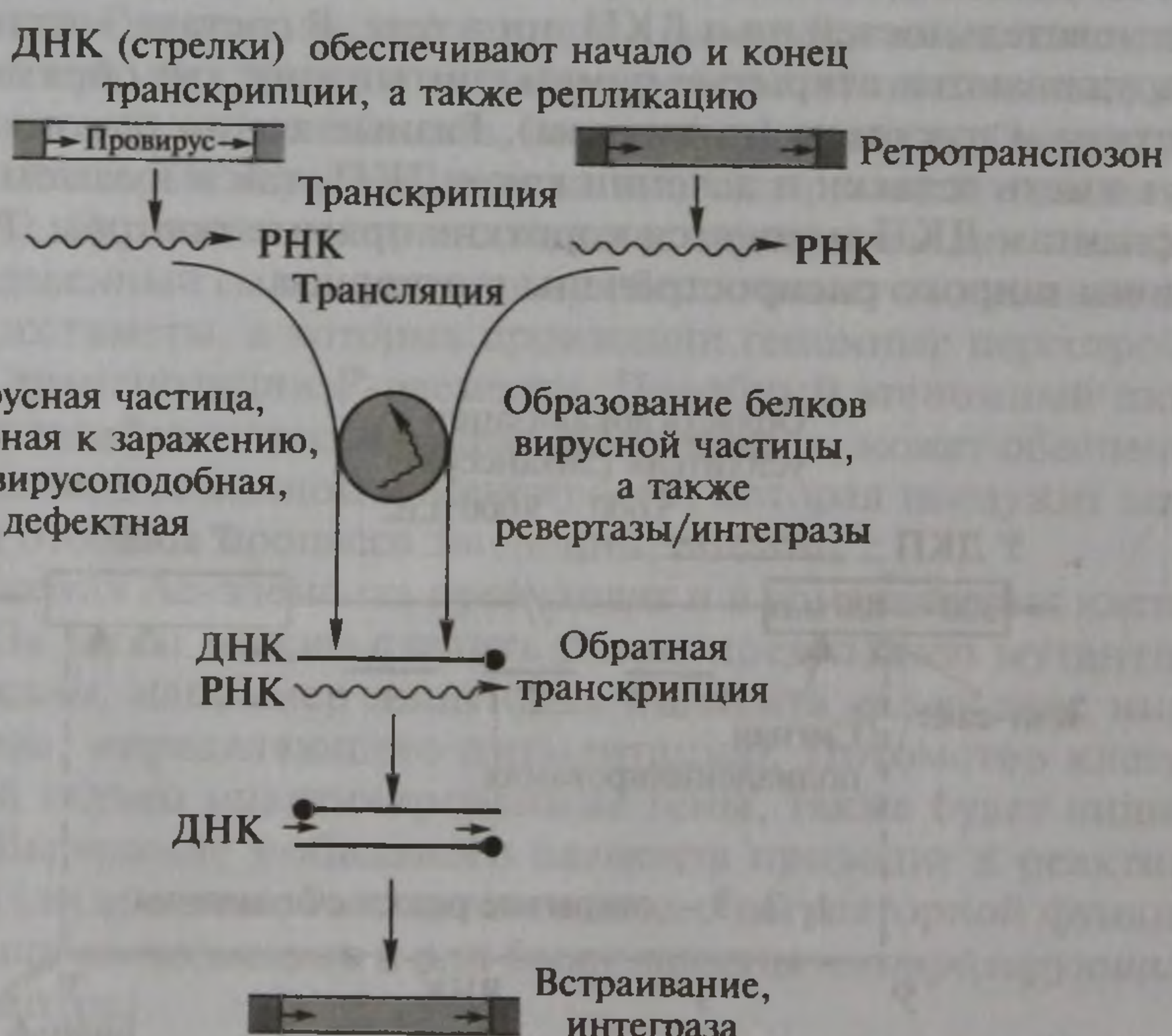


Рис. 96. Перемещение ретротранспозона с длинным концевым повтором

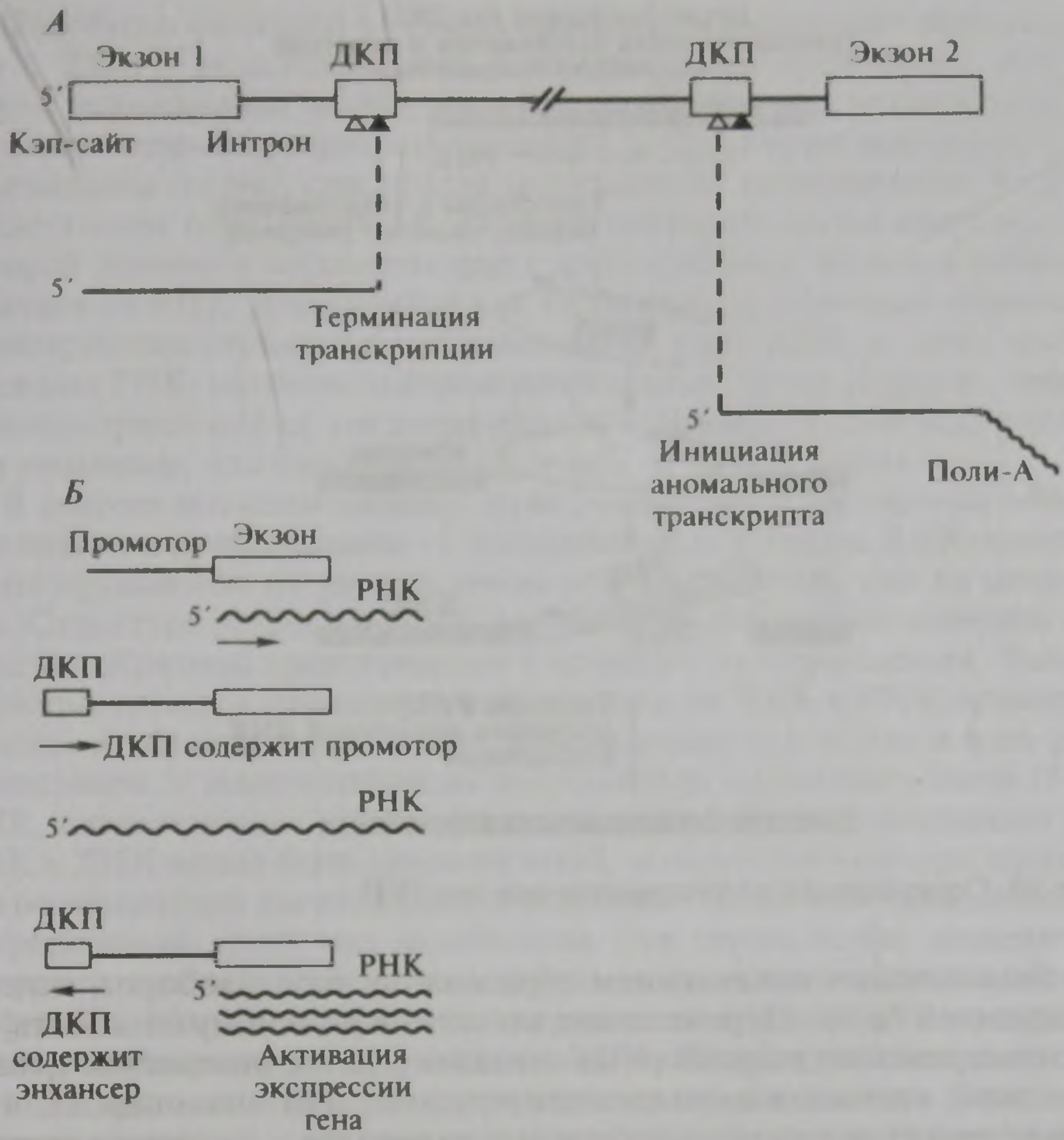


Рис. 97. Эффекты внедрения ретротранспозона внутрь гена или около гена:

А — нарушение транскрипции при внедрении в интрон; *Б* — усиление экспрессии гена при внедрении в 5′-район гена

может быть сверхпродукция белка и злокачественное перерождение клетки. Это же возможно при воздействии энхансеров, привносимых мобильным элементом. Особые возможности для перенесения и приобретения регуляторных сигналов возникают тогда, когда элемент (ретротранспозон) удаляется за счет гомологичной комбинации между ДКП с идентичными повторами (см. рис. 97). В результате сохраняется лишь один ДКП на месте внедрения ретротранспозона. В свою очередь одинокие ДКП существенно влияют на изменчивость регуляторных систем клетки. Именно такой промотор может перепрограммировать характер зависимости работы гена как от внешних сигналов, так и от внутриклеточных регуляторных систем.

Представленные случаи внедрения элемента сопровождаются мутациями с самыми разными фенотипическими проявлениями,

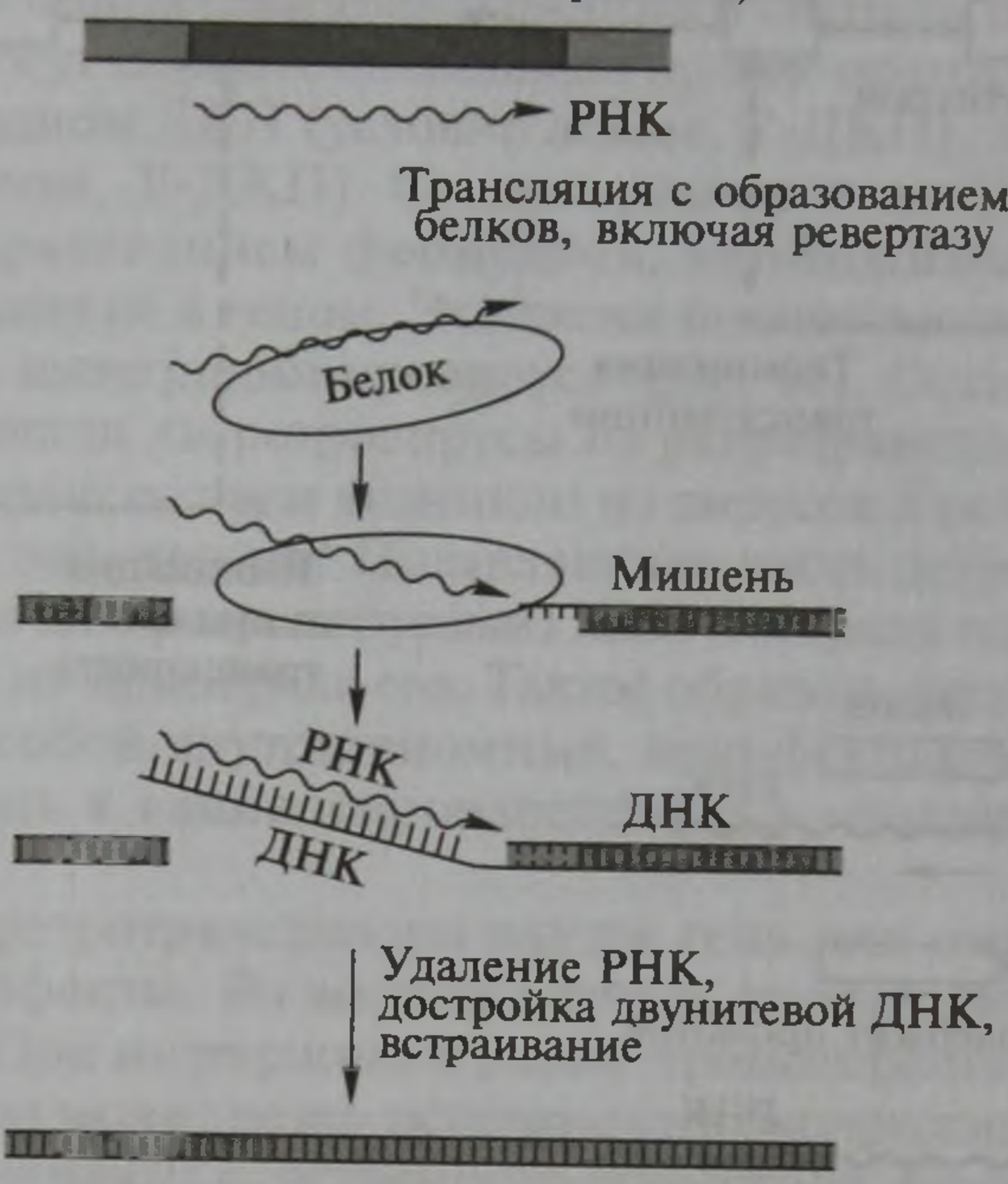


Рис. 98. Перемещение ретротранспозона без ДКП

обусловленными подавлением образования или, наоборот, гиперпродукцией белка. Перемещение элементов по геному способствует распространению регуляторных сигналов (сайтов инициации транскрипции, сигналов полиаденилирования, или энхансеров), что делает значительной роль мобильных элементов в эволюции систем регуляции.

Существует другой большой класс ретротранспозонов, которые не несут длинных концевых повторов. Механизм внедрения их несколько иной, хотя он также осуществляется с помощью обратной транскрипции. К ним относятся представители семейства L1, населяющие геном человека. Репликация элемента без ДКП сопряжена непосредственно с районом будущего внедрения ретротранспозона (рис. 98). РНК, образовавшаяся при транскрипции элемента, перемещается к достаточно случайному месту разрыва ДНК-мишени, часто сшивается с одной из нитей ДНК. Сюда устремляются и необходимые для интеграции белки — обратная транскриптаза и интеграза. Другая, комплементарная нить ДНК служит затравкой для копирования РНК-копии элемента с участием обратной транскриптазы. Фермент копирует сначала небольшой участок ДНК-мишени, а затем меняет матрицу и копирует РНК, после чего РНК удаляется и образуется вторая комплементарная нить ДНК. Ретротранспозоны без ДКП участвуют в сохранении концов хромосом в ряду поколений.

Подобную функцию у многих организмов выполняет теломераза (см. с. 232). У ряда насекомых, в том числе и у дрозофилы, отсутствует теломеразная активность. Концы ДНК у них удлиняются за счет перемещений ретротранспозонов и в этом случае выступают как компоненты генома, спасающие хромосому от укорачивания. Ретротранспозоны перемещаются, образуя повторяющуюся структуру, в которой элементы соединены друг с другом по типу «голова к хвосту». Сначала на РНК-транскрипте как на матрице с помощью обратной транскрипции строится комплементарная нить ДНК, а затем после удаления РНК-матрицы достраивается другая. Таким образом, геном хозяина приспособил эти элементы для сохранения концевых участков хромосом, для спасения хромосомы от потери генов.

В геноме млекопитающих, птиц, амфибий и насекомых обнаруживаются *ретропозоны* — внедрившиеся в геном ДНК-копии, синтезированные на разных типах клеточных РНК, как на матрицах. Структура ретропозонов позволяет с уверенностью говорить об участии обратной транскрипции в процессе их образования. Таким образом, наряду с переносом информации от ДНК к РНК осуществляется и обратный процесс — возвращение ее в геном в виде ретропозонов. У млекопитающих ретропозоны составляют более 10 % ДНК, следовательно, мощность встречного потока информации от РНК к ДНК может быть существенной, по крайней мере при оценке его во временном эволюционном масштабе. Различают разные типы ретропозонов, среди них *псевдогены*. Это копии генов, лишенные интронов и обладающие последовательностью поли-дА-поли-дТ на 3′-фланге. Матрицей при их копировании послужила «процессированная» полиаденилированная мРНК. Рамки трансляции таких «генов» часто «испорчены» стоп-кодонами и делециями. Они, как и проретровирусы, ограничены короткими прямыми повторами, представляющими собой дупликацию геномной последовательности в области сайта внедрения ДНК-копии, образованной на мРНК. Если ретропозон правильно процессирован, ДНК-копия содержит открытую рамку считывания и возможна его экспрессия, то такие ретропозоны называют *ретрогенами*.

Большую массу повторяющихся элементов позвоночных представляют ДНК-копии клеточных РНК, кодирующих белки неизвестной природы, а также аномально процессированных клеточных транскриптов тРНК, 7S РНК и UРНК. Они представляют собой длинные и короткие повторы. Семейства повторяющихся последовательностей часто получают свои названия по расщепляющим их рестриктазам (например, короткие Alu-повторы или длинные Kpn-повторы в геноме человека).

Каждый вид млекопитающих обладает собственным главным семейством длинных повторов, отличающих его от других видов. Предполагают, что разные семейства длинных повторов возникли сравнительно недавно, уже после эволюционной дивергенции видов млекопитающих.

Короткие повторы, например повторы Alu-семейства у человека, составляют $3 \cdot 10^5$, что соответствует 5—6 % массы ДНК клетки. Их нуклеотидная последовательность гомологична последовательности отдельных участков 7S РНК. Такие повторы рассеяны по геному, их обнаруживают в интронах, на 5′-флангах генов, в составе 3′-нетранслируемых участков гена.

Между особями одного вида иногда наблюдаются различия (полиморфизм) в структуре генов и их ближайшего окружения, обусловленные внедрением коротких или длинных повторов. В геноме существует множество неактивных дефектных копий этих элементов, они не способны кодировать транспозазу или обратную транскриптазу. С другой стороны, они сохраняют способность к перемещениям, если в случае транспозонов имеются инвертированные повторы, узнаваемые транспозазой, а в случае ретротранспозонов сохранены промотор и возможность транскрипции элемента. Множество таких дефектных копий начнут перемещаться, если ферменты, ответственные за перемещение, будут кодироваться другими полноценными элементами. В геноме человека источником активной обратной транскриптазы является ретротранспозон L1, число копий которого достигает 100 тыс. Из них активно перемещаются 30—60 тыс., остальные, поврежденные, перемещаться не могут.

Мобильные генетические элементы могут принимать участие в перестройках хромосом. Наличие в хромосомах нескольких оди-

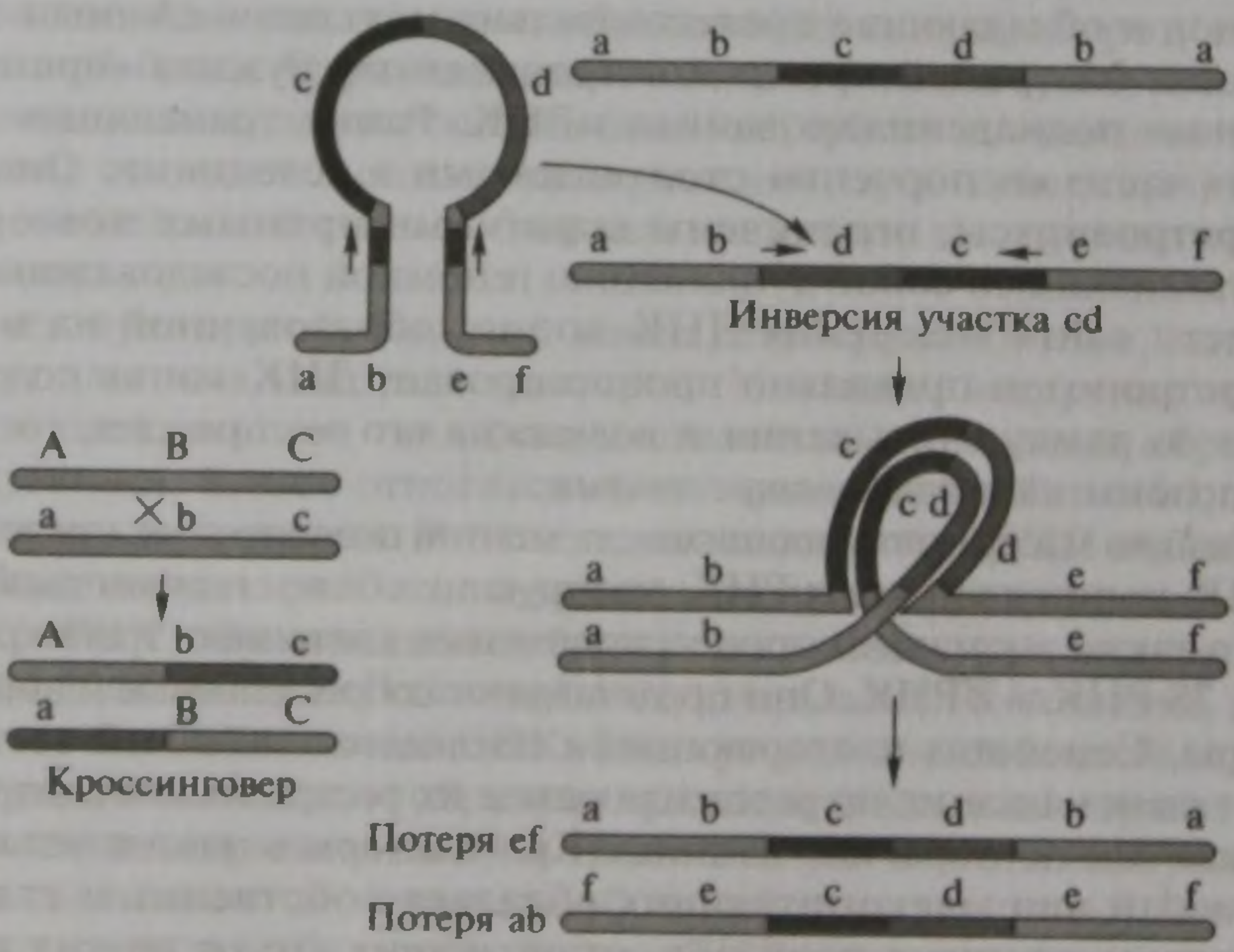


Рис. 99. Перестройки хромосом, обусловленные присутствием в них одинаковых повторяющихся последовательностей (обозначены буквами), представленных подвижными элементами



Рис. 100. Устранение двунитевого разрыва ДНК с помощью ретротранспозона

наковых по нуклеотидной последовательности копий подвижного элемента позволяет в редких случаях осуществить рекомбинацию по районам их локализации. В результате неравного кроссинговера возникают делеции отдельных участков или, наоборот, дупликации (рис. 99). Неравный кроссинговер по районам локализации транспозона mariner (моряк) у человека приводит к делеции участка в коротком плече 17-й хромосомы. Если это событие происходит в зародышевой клетке при созревании гамет, то хромосома с делецией будет передана потомкам. Это приводит к наследственным заболеваниям нервной системы — невропатиям и параличам. Внутрихромосомная рекомбинация между двумя элементами приводит к инверсии — повороту участка хромосомы на 180°. Инверсия может быть вредна для организма, а может способствовать эволюции генома, поскольку помогает передать потомству случайно сложившееся благоприятное сочетание генов, препятствуя кроссинговеру.

Ретротранспозоны с ДКП участвуют в ликвидации двунитевых разрывов ДНК. Обычно такой разрыв залечивается с помощью гомологичной молекулы ДНК, например сестринской, только что реплицированной нити. Участие реплицирующейся ДНК ретротранспозонов с ДКП в процессе заживления двунитевых разрывов обнаружено у дрожжей (рис. 100). Брешь в двунитевой спирали, заделанная ретротранспозоном, сохраняет целостность хромосомы, но изменяет ее нуклеотидную последовательность. Если район хромосомы, где была брешь, не содержит существенного гена, то клетка остается жизнеспособной.

Подвижные элементы не стоит рассматривать только как «эгоистическую» ДНК, паразитирующую на ДНК клетки-хозяина. Они приспособлены и для нужд генома клетки: способствуют сохранению

структурной целостности хромосомной ДНК; внедряясь в гены, не только инактивируют, вызывая мутации, но и меняют характер их экспрессии. В связи с этим подвижные генетические элементы, будучи важными факторами изменчивости генов и участвуя в перестройках структуры хромосом, имеют огромное значение в процессах эволюции геномов.

Программа «Геном человека»

Для всестороннего исследования генома человека в 1990 г. была разработана и в настоящее время воплощается в жизнь специальная международная программа «Геном человека» («Human Genome Project») — это один из самых дерзновенных, дорогостоящих и потенциально важных проектов в истории цивилизации.

В ядре каждой соматической клетки человека содержится 23 пары хромосом: на каждую хромосому приходится по одной молекуле ДНК. Длина всех 46 молекул ДНК в одной клетке человека равна почти 2 м, количество нуклеотидных пар составляет 6,4 млрд.

Основной задачей программы является построение исчерпывающих генетических карт большого разрешения каждой из 24 хромосом человека, которое должно завершиться определением полной первичной структуры ДНК всех хромосом. Для решения этой задачи возник Международный консорциум по изучению генома человека, объединивший 20 лабораторий и сотни ученых во всем мире.

В случае успешного завершения программы (по плану это должно было произойти в 2005 г.) у человечества появятся перспективы понять механизмы функционирования каждого из его генов, принципы развития организма, генетические причины многих наследственных болезней и механизмы старения.

Картирование генома человека

Решение основной задачи программы «Геном человека» включает следующие этапы.

• На первом этапе необходимо завершить составление детальной *генетической карты* и отметить гены, отстоящие друг от друга на расстоянии, не превышающем в среднем 2 млн оснований (1 млн оснований равен 1 мегабазе — 1 Мб, от англ. *base* — основание).

• Второй этап предполагает составление физических карт низкого разрешения каждой хромосомы (разрешение 0,1 Мб).

• На третьем этапе следует получить физическую карту высокого разрешения всего генома в виде охарактеризованных по отдельности клонов (клон содержит 5 Кб).

• Четвертый этап посвящен определению полной первичной структуры (секвенированию) всей ДНК генома человека (разрешение — 1 основание). По плану это должно было произойти к 2004 г.

• На пятом, заключительном, этапе необходимо в найденных последовательностях нуклеотидов локализовать все гены организма и определить их функциональное значение (к 2005 г.).

Прохождение вышеперечисленных этапов требуется для получения исчерпывающих характеристик не только генома человека, но и любого другого генома большого размера.

Генетическое картирование

Генетические карты сцепления. Генетические карты сцепления определяют хромосомную принадлежность и взаимное расположение генетических маркеров относительно друг друга. Картирование в узком смысле — определение положения гена или мутации в хромосоме. Позднее этот термин получил более широкое толкование. Он относится не только к гену, но к любому *маркеру*, под которым подразумевают ген, мутацию, участок ДНК с неопределенной функцией, точку расщепления ДНК рестриктазами. Таким образом, маркер — это любой наследуемый признак, доступный идентификации тем или иным способом. Установление локализации какого-либо маркера позволяет использовать его для определения положения другого маркера.

Маркеры, выбранные для картирования, должны быть *полиморфными*, т. е. должны существовать альтернативные формы признака, которые можно было бы легко отличить и описать у членов семьи. В последовательности ДНК варьирует (т. е. полиморфен) в среднем один из 300 — 500 нуклеотидов. Мутации в экзонах могут приводить к видимым изменениям — различию в цвете волос и глаз, группе крови, особенностях строения тела, деятельности физиологических систем, подверженности болезням и т. п. Мутации в интронах имеют минимальный эффект или не имеют его вообще. В то же время эти изменения легко определяются на уровне ДНК и могут использоваться как маркеры. Среди маркеров такого типа — ПДРФ (полиморфизм длин рестрикционных фрагментов) — изменение последовательности ДНК в сайтах рестрикции; различия в числе тандемных повторов, что приводит к различиям по длине ДНК, которую можно легко определить. Возможность построения генетических карт сцепления обусловлена двумя фундаментальными характеристиками генома: линейным характером локализации генов в хромосоме (это определяется линейностью молекулы ДНК) и относительной стабильностью расположения облигатных (обязательных) элементов генома в пределах вида. Термин «сцепление» употребляется в генетике для

обозначения вероятности совместной передачи двух признаков от одного из родителей потомству.

Генетические признаки, локализованные в разных хромосомах, не сцеплены друг с другом, т.е. передаются от родителей детям независимо. Гены, локализованные в одной хромосоме, рекомбинируют за счет кроссинговера, за счет обмена участками гомологичных хромосом в процессе их спаривания при мейозе. Вероятность кроссинговера между двумя генами зависит от расстояния между ними. Чем ближе гены расположены друг к другу, т.е. чем больше они сцеплены, тем эта вероятность меньше. На генетических картах сцепления расстояние между маркерами определяется в сантиморганидах (сМ). Когда частота рекомбинации между двумя маркерами равна 1 %, это значит, что они находятся на расстоянии 1 сМ. Генетическое расстояние в 1 сМ примерно равно физической протяженности в 1 млн н.п. — 1 Мб. Общая длина генома человека в этих единицах составляет около 3 300 сМ. Генетические карты сцепления правильно отражают порядок расположения генетических маркеров на хромосомах, но значения расстояний между ними приблизительны и не соответствуют реальным физическим расстояниям. Этот факт связывают с тем, что эффективность рекомбинации на отдельных участках хромосом может сильно различаться. Существуют так называемые «горячие» точки рекомбинации и районы генома, где рекомбинация подавлена (центромерные, теломерные и гетерохроматиновые участки хромосом). Кроме того, частота рекомбинации у мужчин меньше, чем у женщин, так что общая длина мужского генома, измеренная в единицах рекомбинации, составляет лишь 3 000 сМ. Таким образом, генетические карты сцепления являются наименее точными и дают весьма ориентировочную информацию о реальном расстоянии, выражаемом в парах нуклеотидов. Тем не менее на практике именно они и только они позволяют локализовать сложные генетические маркеры (например, ассоциированные с симптомами заболевания) на первых этапах исследования и дают возможность их дальнейшего изучения.

До начала 70-х годов XX в. построение генетических карт человека продвигалось очень медленными темпами. Первый ген человека (ген цветной слепоты) был картирован на Х-хромосоме в 1911 г., а первый аутосомный ген — только в 1968 г. К 1973 г. на хромосомах человека было картировано 64 гена, а к 1994 г. — 5 000 структурных генов и свыше 60 000 маркерных ДНК-последовательностей. Столь стремительный прогресс в картировании генов человека связан с появлением новых технологий в цитогенетике, в клеточных культурах и особенно в молекулярной генетике.

Гибридизация соматических клеток. Одним из наиболее популярных методов отнесения генетического маркера (функционально активного гена) к конкретной группе сцепления является гибридизация (слияние друг с другом) соматических клеток разных

биологических видов организмов, один из которых — исследуемый. Гибридные клоны получают путем искусственного слияния клеток человека и различных грызунов: китайского хомячка, мыши, крысы. Культивирование таких соматических гибридов, как оказалось, сопровождается утратой хромосом человека. Потеря хромосом носит случайный характер, и образующиеся клоны клеток содержат оставшиеся хромосомы в разных сочетаниях. Так получают панели гибридных клеточных клонов, содержащих всего одну или несколько хромосом человека и полный набор хромосом другого вида. Обнаружение человеческих белков, специфических мРНК или последовательностей ДНК в таких клонах позволяет однозначно определить хромосомную принадлежность соответствующих генов.

Гибридизация in situ. Этот метод дает возможность локализовать определенные последовательности нуклеотидов на хромосомах. Они выступают в качестве зондов. Препараты фиксированных хромосом гибридизуют с исследуемыми последовательностями, меченными радиоактивной или флуоресцентной меткой. Меченые молекулы оказываются ассоциированными с участками хромосом, содержащими последовательности, комплементарные меченому зонду. Полученные гибриды анализируют с помощью микроскопа либо непосредственно, либо после радиоавтографии. Этот метод по частоте использования в последнее время прочно выходит на первое место. Наиболее популярной оказалась группа методов, получивших название *флуоресцентной гибридизации in situ* — метод FISH (от англ. *fluorescence in situ hybridization*).

Полимеразная цепная реакция (ПЦР) позволила быстро и эффективно амплифицировать почти любой участок генома человека, а полученные продукты ПЦР использовать в качестве зондов для картирования соответствующих участков на хромосомах путем гибридизации in situ. В этом плане успешно разработана концепция сайтов, привязанных к последовательностям, — STS (от англ. *sequence-tagged sites*). Все фрагменты ДНК, которые используются для построения генетических и физических карт, можно однозначно идентифицировать с помощью последовательности нуклеотидов длиной в 200 — 500 н. п., которая является уникальной для данного фрагмента. Эти сайты амплифицируют с помощью ПЦР и применяют в качестве зондов. STS позволили создать основу для разработки единого языка, дающего возможность разным лабораториям описать свои клоны. Конечным результатом разработки концепции STS является создание исчерпывающей карты STS генома человека. Для получения маркеров в настоящее время часто применяют праймеры, соответствующие диспергированным повторяющимся последовательностям, среди которых первыми стали использовать Alu-последовательности, так как они характерны именно для генома человека. Поскольку в геноме человека больше 90 % умеренно повторяющихся последовательностей представлены семействами Alu и Kpn I (последние повторяются реже

и обладают характерной локализацией в хромосомах), они и используются для получения соответствующих зондов в ПЦР-реакции.

Физические карты низкого разрешения. Физические карты генома отражают реальное расстояние между маркерами, выражаемое в парах нуклеотидов. Физическую карту низкого разрешения часто называют *хромосомной* (*цитогенетической*) *картой генома*.

В начале 70-х годов XX в. появилась реальная возможность точной идентификации не только всех хромосом в кариотипе человека, но и их отдельных сегментов. Это связано с появлением метода дифференциального окрашивания препаратов метафазных хромосом. Хромосомные препараты окрашивают некоторыми флуорохромами после соответствующей протеолитической обработки или нагревания. При этом на хромосомах выявляется характерная поперечная исчерченность — так называемые диски (бэнды), расположение которых специфично для каждой хромосомы. Величина небольших дисков на прометафазных хромосомах соответствует примерно 1 млн н.п. на физических картах. Каждая хромосома после дифференциальной окраски может быть разделена на сегменты, нумерация которых начинается от центромерного района вверх (короткое плечо *p*) либо вниз (длинное плечо — *q*) (рис. 101). Полосы в каждом сегменте также пронумерованы в аналогичном порядке. Запись положения гена на карте включает номер хромосомы, плечо, номер сегмента, бэнда и его субъединицы. Запись 7q21.1 означает, что ген локализован в субъединице 1-го бэнда 2-го сегмента длинного плеча хромосомы 7. Подобная запись удобна для цитогенетического картирования метода гибридизации in situ, позволяющего локализовать ген с точностью до одного бэнда и даже его субъединицы.

Хромосомные карты генома человека получают также локализацией генетических маркеров, чаще всего методом FISH: для метафазных хромосом разрешающая способность хромосомных карт находится в пределах 2—5 млн н.п.; для интерфазных хромосом (генетический материал находится в менее компактной форме) — приближается к 100 тыс. н.п. Для этого уровня картирования характерны *карты кДНК* (с. 361). Они отражают положение экспрессирующихся участков ДНК (экзонов) относительно известных цитогенетических маркеров (бэндов) на метафазных хромосомах. Поскольку такие карты дают представление о локализации транскрибирующихся участков генома, в том числе и генов с неизвестными функциями, они могут быть использованы для поиска новых генов. Этот подход полезен при поиске генов, повреждение которых вызывает заболевания человека, в том случае, если приблизительная локализация таких участков хромосом уже проведена на генетических картах сцепления (см. рис. 101).

Физические карты высокого разрешения. Для построения физических карт высокого разрешения экспериментально реализуется два альтернативных подхода: картирование сверху вниз и картирование снизу вверх (рис. 102). Для картирования сверху вниз препарат

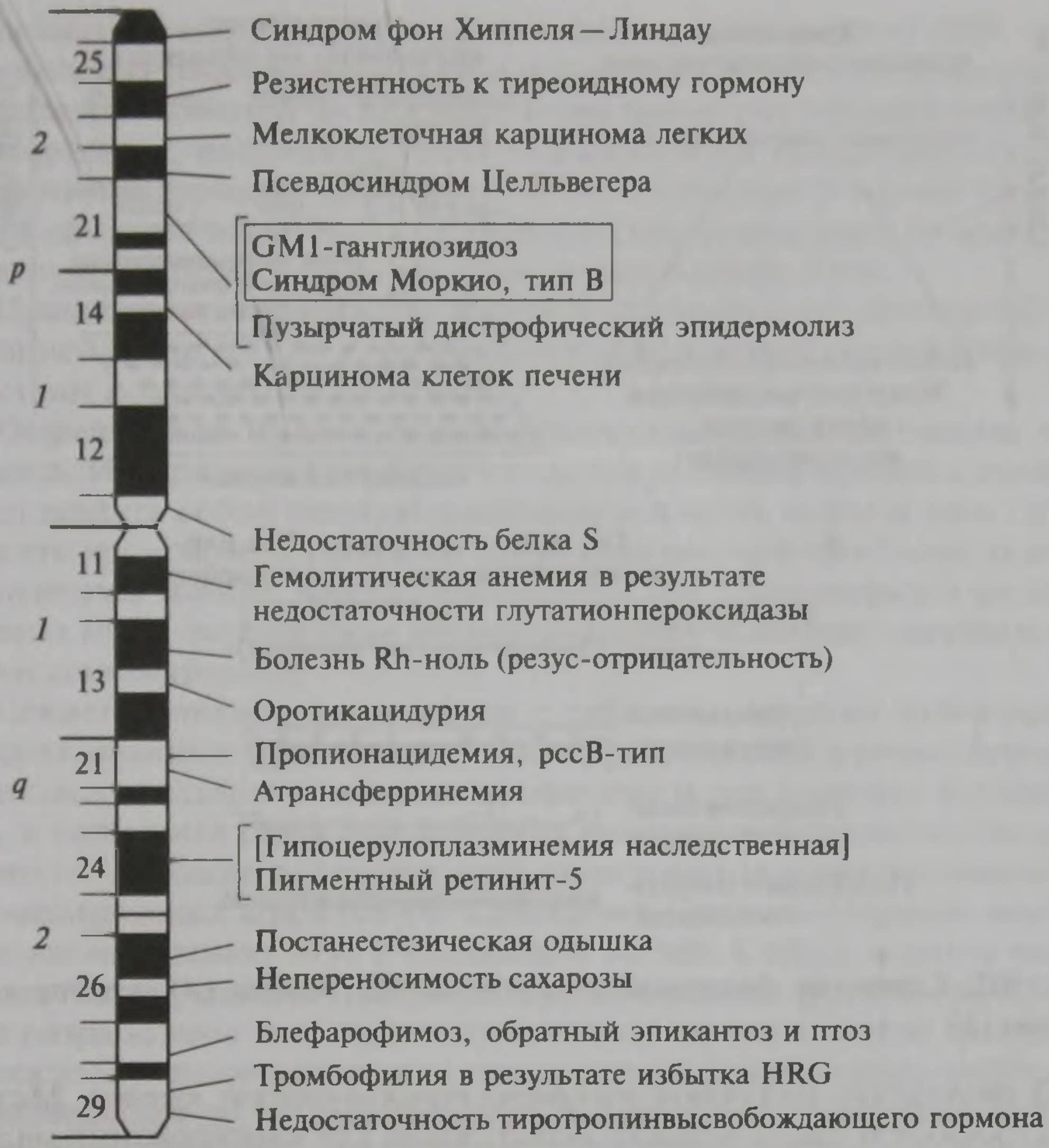


Рис. 101. Генетическая карта 3-й хромосомы человека

ДНК индивидуальной хромосомы человека разрезают крупнощепящими рестриктазами (например, Not I) на длинные фрагменты, которые после разделения методом электрофореза в пульсирующем поле подвергаются дальнейшей обработке другими рестриктазами.

Методом электрофореза под действием однонаправленного постоянного поля в агарозном или полиакриламидном гелях удается разделить фрагменты ДНК размером не более 30—50 тыс. н.п. Продвижение больших фрагментов ДНК в геле при пульсирующем изменении направления электрического поля происходит за счет конформационных изменений, обусловленных скручиванием и раскручиванием молекул ДНК в момент переключения направления поля. В этом случае удается разделить молекулы ДНК размером от 50 тыс. н.п. до 10 млн н. п.).



Рис. 102. Стратегия физического картирования генома (*А*) и типы карт генома (*Б*)

В результате получают *макрорестрикционную карту*. Метод электрофореза был с успехом использован для картирования малых геномов.

Для картирования генома человека снизу вверх на основе препарата суммарной ДНК генома или индивидуальной хромосомы получают серию случайных клонов протяженных последовательностей ДНК (10 — 1 000 тыс. н.п.), часть из которых перекрывается друг с другом. В качестве вектора для клонирования в этом случае используют искусственные минихромосомы дрожжей (YAC). Последовательный набор клонов, содержащих частично перекрывающиеся и дополняющие друг друга фрагменты ДНК из определенного района генома, получил название *скользящего зондирования*, или «прогулки по хромосоме». Каждый раз отобранный фрагмент используется в качестве ДНК-зонда для последующего поиска. В результате получают набор клонированных фрагментов ДНК, полностью перекрывающих исследуемый участок генома, получивший название «контиг». Эта стратегия впервые была успешно применена для изучения 3-й хромосомы дрозофилы. С ее помощью редко удается пройти более 200 — 300 тыс. н. п. в одном направлении из-за наличия в геноме по-

вторяющихся и трудно клонируемых последовательностей ДНК. Для преодоления таких ограничений и ускорения процесса поиска генных последовательностей Ф. Коллинз, ныне президент Международного консорциума, предложил метод «прыжков» по хромосоме, позволяющий изолировать фрагменты ДНК, отстоящие в геноме друг от друга на сотни тысяч пар нуклеотидов (длина прыжка), не выделяя при этом все промежуточные последовательности ДНК.

Правильность полученных контигов подтверждают обычно гибридизацией in situ (FISH) с одновременной привязкой к определенным участкам исследуемых хромосом.

Определение нуклеотидной последовательности генома человека. Исчерпывающая физическая карта генома человека должна представлять собой полную последовательность нуклеотидов ДНК всех его хромосом. К решению такой грандиозной по объему задачи привлечены многие хорошо финансируемые лаборатории в разных странах мира, оснащенные автоматическими высокопроизводительными секвенаторами.

Секвенирование начинается с субклонирования небольших рестрикционных фрагментов ДНК, образующихся в результате рестрикции предварительно клонированных более крупных фрагментов, и первичная структура коротких фрагментов определяется, как правило, методом полимеразного копирования с использованием терминирующих аналогов (Ф. Сангером). Средняя скорость секвенирования к концу 1998 г. составляла 90 Мб. Сейчас анализу чаще подвергают продукты ПЦР, а не клонированные фрагменты ДНК. Уже разработаны и внедрены гигантские автоматические фабрики по секвенированию генома — геномотроны (секвенаторы-роботы), производящие до 10^5 полимеразных цепных реакций в час и секвенирующие несколько миллионов нуклеотидов в неделю. Ожидается, что внедрение технологий третьего поколения, среди которых флуоресцентное детектирование отдельных меченых оснований в проточных цитофотометрах, прямое чтение оснований в последовательностях ДНК с использованием сканирующих туннельных микроскопов, прямой масс-спектрометрический анализ последовательностей ДНК, гибридизация изучаемых высокомолекулярных фрагментов ДНК с панелями (ДНК-чипами) коротких фрагментов ДНК, позволит секвенировать по 500 Мб в год. Необходимо иметь в виду, что длина ДНК самой маленькой Y-хромосомы человека составляет 50 Мб, самой большой (хромосома 1) — 300 Мб. Эти цифры наглядно демонстрируют масштабы задачи, которую предстоит решить.

После определения первичной структуры генома человека нужны будут долгие годы для ее обработки. Одна только запись всей последовательности нуклеотидов генома человека обычным шрифтом потребует около 200 томов книг по 1000 страниц каждая. Хранение и анализ такого объема информации влечет за собой необходимость

создания уникальных алгоритмов обработки информации с помощью вычислительной техники.

Крупнейшей базой данных по структуре генома человека в настоящее время является GDB (Genome Data Base), созданная и поддерживаемая в университете Дж. Гопкинса (Балтимор, США). Кроме известных последовательностей нуклеотидов генома человека в ней хранится вся полученная информация о генетических маркерах, зондах и контигах, ассоциированных с генетическими заболеваниями.

Четыре других базы данных хранят все известные последовательности нуклеотидов, включая последовательности нуклеотидов генома человека: *GenBank* и Genome Sequence Data Base (*GSDB*) в США; European Molecular Biology Laboratory (*EMBL*) Nucleotide Sequence Database в Великобритании; DNA Data Bank of Japan (*DDBJ*) в Японии. В России Институтом молекулярной биологии им. В.А. Энгельгардта поддерживается аналогичная база данных по геному человека (Hugene).

На ноябрь 2011 г. в базе данных содержится 2 696 полностью расшифрованных геномов прокариот и 168 генома эукариот (http://www.ebi.ac.uk/GenomeReviews/stats/cellular/; http://www.ebi.ac.uk/genomes/; http://www.genomesonline.org/). В состоянии чернового варианта находится еще множество геномов про- и эукариот.

Описание и интерпретация информации, которая зашифрована в геноме, называется *аннотацией генома*. Этой проблемой занимается биоинформатика. Аннотация — это процесс связывания последовательностей с биологической информацией, которая в них заложена. Методы аннотации быстро развиваются, но они еще далеко недостаточны, чтобы извлечь полную информацию из генома. Аннотация генома включает в себя как экспериментальные, так и компьютерные анализы и представляет собой многоступенчатый процесс, который призван получить три группы информации:

— информацию о нуклеотидных последовательностях, имеющих то или иное функциональное значение;

— информацию о белках или РНК, которые кодируются геномом;

— информацию о тех процессах, в которые вовлечены различные участки генома и кодируемые ими продукты.

Структура генома человека (по данным секвенирования на 2001—2003 гг.)

Чтобы установить последовательность генома человека, понадобилось объединить усилия множества групп из разных стран.

Начиная с 1998 г. работы по секвенированию генома человека проводят параллельно Международный консорциум и частная коммерческая компания Целера Геномикс (Celera Genomics, президент Крейг

Таблица 15. **Количество и плотность генов в пяти отсеквенированных эукариотических геномах**

Организм	Количество отсеквенированных оснований, тыс. н. п.	Степень законченности, %	Количество предсказанных генов	Количество генов на 10^6 оснований
Дрожжи (*S. cerevisiae*)	12 068	93	5 885	483
Круглый червь (*C. elegans*)	97 000	99	19 099	197
Плодовая мушка (*D. melanogaster*)	116 000	64	13 601	117
Растение (*A. thaliana*)	115 000	92	25 498	221
Человек (*H. sapiens*) (консорциум)	2 693 000	84	31 780	12
Человек (*H. sapiens*) (Celera)	2 654 000	83	39 114	15

Вентор). В феврале 2001 г. были опубликованы две предварительные версии структуры генома человека[1]. Описание генома человека пока что не отличается высокой точностью, но уже сейчас можно сделать ряд важных заключений. Большое значение приобретает сравнение генома человека с отсеквенированными геномами четырех других эукариотических организмов — дрожжей, круглого червя, плодовой мушки и растения (табл. 15).

На основе компьютерных алгоритмов, построенных на современных представлениях об общей структуре гена и о белковых доменах, было рассчитано количество генов, кодирующих белки в геноме человека. Международный консорциум определил 31 780 белок-кодирующих генов, а фирма Целера Геномикс обнаружила 39 114 таких генов.

Эти величины являются предварительными, и они много меньше ранее постулированного числа белок-кодируемых генов, которое

---

[1] *Ventor J. C., Adams M. D., Myers E. W. et al.* The sequence of the human genome // Science, 2001. — V. 291. — P. 5507. — P. 1304—1351.

The International Human Genome Sequencing Consortium. The Human Genome: Initial Sequencing and Analysis // Nature, 2001. — V. 409. — P. 806—921.

оценивалось в 100 000 — 140 000. Отмечается более низкая плотность белок-кодирующих генов в геноме человека.

Самые общие характеристики генома человека приведены в табл. 16.

Таблица 16. **Общие структурные характеристики генома и протеома человека** (Ventor et al., 2001)

Геном		Протеом, %	
Длина	2,91 Мб	Ферменты	10,2
G+C	38 %	Трансдукция сигнала	12,2
Повторы	35—50 %	Факторы связывания нуклеиновых кислот	13,5
Аннотированные гены (несколько методов, идентификации совпадают)	26 383	Клеточная адгезия	1,9
Доля аннотированных генов с полностью неизвестной функцией	42 %	Цитоскелетные структурные белки	2,8
Число генов (аннотированных и гипотетических)	39 114	Внеклеточный матрикс	1,4
Доля аннотированных и гипотетических генов с неизвестной функцией	59 %	Иммуноглобулины	0,9
Средний размер гена	27 000	Ионные каналы	1,3
Наиболее богатая генами хромосома	19 (23 гена/Мб)	Протоонкогены	2,9
Наименее богатые генами хромосомы	13, Y (5 генов/Мб)	Внутриклеточный транспорт	1,1
Доля генома, занятая генами	25,5—37,8 %	Факторы транскрипции	6
Доля генома, занятая экзонами	1,1—1,4 %	Регуляторные молекулы	3,2

Окончание табл. 16

Геном		Протеом, %	
Доля генома, занятая интронами	24,4—36,4 %	Транспортеры	1,7
Доля генома в межгенном пространстве	74,5—63,6 %	Рецепторы	5
Плотность мононуклеотидного (SNP) полиморфизма	1/1 250 н. о.	Функция неизвестна	41,7

По сравнению с геномами других эукариотических организмов у человека большее распространение получили гены, участвующие в обеспечении иммунной защиты; в развитии нервной системы (нейротрофические факторы, факторы роста нервов), сигнальных молекул, миелиновых белков, потенциал-управляемых ионных каналов и синаптических рецепторных белков; в построении цитоскелета и движении везикул, обеспечении внутри- и межклеточной сигнализации, поддержании гомеостаза. У человека значительно большее количество генов участвует в транскрипции и трансляции. Из 2 000 таких генов 900 относятся к семейству белков, содержащих «цинковые пальцы».

В геноме человека число генов, кодирующих белки, всего в 2 раза больше числа генов в геноме червя, мушки, растения. Это обусловлено возникновением у человека большого числа новых белков. Существует предположение, что *H. sapiens* не «изобрел» новые гены, а использовал существующие структурные домены, собирая из них новые белки с новыми функциями. Геном человека кодирует больше паралогов и многодоменных белков с большим разнообразием функций и архитектуры доменов по сравнению с геномами других эукариот.

В целом на долю генов, кодирующих белки, приходится 2 % генома; на области, кодирующие РНК, — около 20 % генома, повторяющиеся последовательности занимают более 50 % генома, причем значительная часть этой ДНК возникла за счет обратной транскрипции РНК.

До 10 % генома человека составляют повторяющиеся элементы, по своему строению напоминающие интегрированную форму инфекционных ретровирусов птиц и млекопитающих. Число обнаруженных семейств эндогенных ретровирусов в геноме человека достигло 30 и продолжает расти. Пока их биологическая роль в геноме человека неизвестна. Крайне интересно, что у обезьян эндогенных вирусов намного меньше или нет вообще. Получается, что по чужеродным элементам генома человек и обезьяна различаются гораздо сильнее,

чем по самим геномам. Это дало основание одному из известных исследователей генома академику Е. Д. Свердлову высказать мысль о том, что вирусы могли сыграть важную роль в «очеловечивании» обезьяны. В связи с этим такая древняя и постоянно привлекающая внимание проблема происхождения человека получила новое звучание.

Идентифицированы гены практически всех наиболее часто встречающихся (примерно 320) и сравнительно редких (около 170) наследственных заболеваний, 30 рецессивных и более 100 доминантных онкогенов. Тем не менее сегодня еще нельзя исключить вероятность того, что дополнительные гены располагаются в неотсеквенированных до сих пор областях генома. Международный консорциум 14 апреля 2003 г. сообщил, что расшифровано 99,99 % генома человека. Получена завершенная последовательность генома, имеющая минимальное число ошибок (менее 1 на 10 000 звеньев ДНК), она максимально полная — пробелы остаются только там, где последовательность не может быть установлена имеющимися технологиями

Сделаны общие заключения по структуре генома человека.

1. «Ландшафт» генома поражает разнообразием распределения различных элементов: генов, псевдогенов, мобильных элементов, содержания G+C, островков CpG и скоростей рекомбинации. Это распределение отличается от мыши и даже от обезьяны. Что в нем — случайные ли причуды эволюции, или что-то функционально важное, нам еще предстоит узнать.

2. 20 000—30 000 кодирующих белки генов никак не объясняют различий в сложности человека и червяка или дрозофилы.

3. Сотни человеческих генов, возможно, произошли в результате горизонтального переноса от бактерий. Множество других являются продуктами деятельности транспозонов.

4. Геном, особенно в перицентромерных и прителомерных областях, наполнен дупликациями длинных сегментов, которые, по-видимому, играли существенную роль в эволюции и являются полиморфными в человеческой популяции. Сегментные дупликации у человека гораздо больше распространены, чем у дрожжей, мухи или червя.

5. Скорости рекомбинаций, как правило, выше в концевых областях хромосом, на более коротких плечах и на более коротких хромосомах, обеспечивая, по крайней мере, один кроссинговер на мейоз на каждое плечо хромосом.

6. Геном чрезвычайно вариабелен, так что разные индивидуумы отличаюся громадным количеством нуклеотидных замен, положениями транспозонов и даже длинными сегментными повторами.

7. Геном насыщен разного рода повторяющимися элементами, некоторые из них способны перемещаться в геноме, создавая новые гены и новые регуляторные элементы.

Исследование структуры генома ряда прокариот и эукариот, и человека в частности, способствовало созданию науки о геномах — геномики. В нее включают изучение геномов на молекулярном, хромосомном, биохимическом и фенотипическом уровнях. На рис. 103 представлена схема, поясняющая взаимоотношения между геномикой человека и другими научными направлениями в современной биологии. Структурная и сравнительная геномика через биоинформатику переходит в новый раздел — функциональную геномику, главной задачей которой является выяснение биологических функций генных продуктов и в первую очередь белков.

У многих современных исследователей, работающих в области геномики, нет сомнений, что первое десятилетие XXI в. будет эрой функциональной геномики и биоинформатики.

В сети Интернет можно найти большое число адресов, содержащих разнообразную информацию, касающуюся генома человека:

http://www.nhgri.nih.gov
http://genlink.wustl.edu
http://www.geneletter.org
http://www.hhmi.org
http://www.mpimg-berlin-dahlem.mpg.de/~cytogen
http://ftp.tigr.org
http://cancernet.nci.nih.gov/ord/index
http://www.wiley.co.uk/genetherapy
http://www.genom.gov.

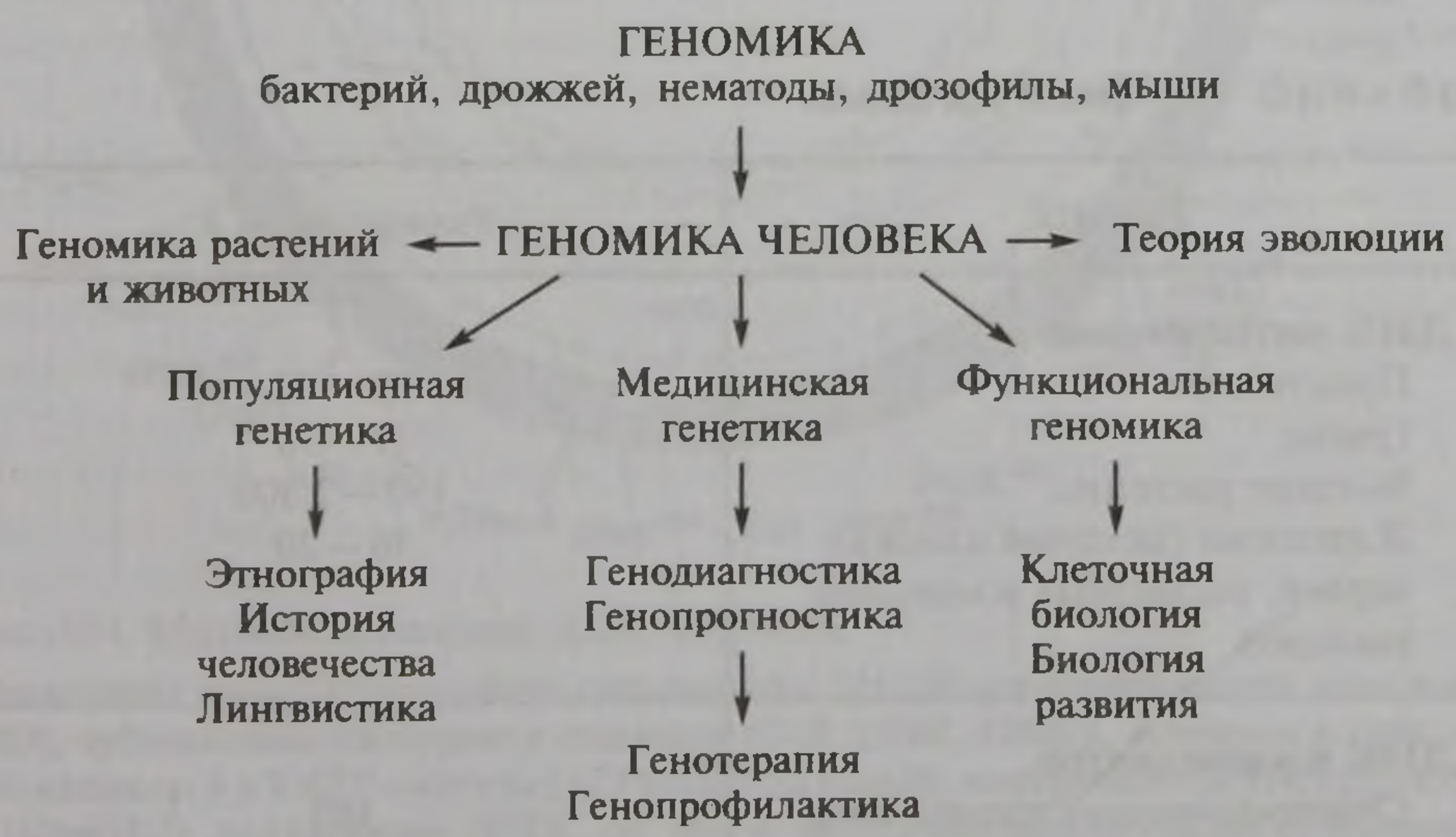


Рис. 103. Схема, поясняющая взаимоотношения между геномикой человека и другими научными направлениями

Геномы органелл эукариот: ДНК митохондрий и хлоропластов

Существуют два типа цитоплазматических ДНК: одни находятся в митохондриях эукариот, другие — в хлоропластах зеленых растений и водорослей.

Как и все цитоплазматические элементы, они наследуются по материнской линии, а не по законам Менделя. Большая часть белков этих органелл, закодированная в ядерной ДНК, синтезируется в цитоплазме и затем переходит в органеллу. Однако некоторые белки митохондрий и хлоропластов и все их РНК кодируются в ДНК самих органелл и в них же синтезируются. Таким образом, органеллы — это результат объединенных усилий двух геномов и двух трансляционных аппаратов. РНК-компоненты рибосом органелл, а также тРНК, использующиеся при трансляции, кодируются геномами митохондрий и хлоропластов.

Размеры генома хлоропластов у всех исследованных организмов сходны, тогда как митохондриальные геномы у растений намного больше, чем у животных (табл. 17).

Все митохондрии и хлоропласты содержат по несколько копий собственной геномной ДНК. Эти молекулы ДНК обычно распределены в виде отдельных групп в матриксе митохондрий и в строме хлоропластов, где они прикреплены к внутренней мембране. Способ упаковки ДНК неизвестен. По структуре геном более сходен с бактериальным геномом: например, как и у бактерий, у них нет гистонов.

Таблица 17. **Геном органелл**

Тип ДНК	Размер, тыс. н. п.
ДНК митохондрий	
Простейшие	22 — 40
Грибы	17 — 78
Высшие растения	150 — 2 500
Животные (включая плоских червей, насекомых и млекопитающих)	16 — 20
ДНК хлоропластов	
Chlamydomonas (зеленая водоросль)	180
Высшие растения	120 — 200

Геномы митохондрий

В клетках млекопитающих митохондриальная ДНК составляет меньше 1 % всей клеточной ДНК. В клетках других организмов — в листьях высших растений или в очень крупных яйцах амфибий и других — доля ДНК энергетических органелл, особенно в последнем случае, может достигать 99 % всей ДНК клетки.

Многие митохондриальные геномы представляют собой замкнутые кольцевые сверхспиральные дуплексные ДНК. Большинство из них кодируют сходные функции. Сделать еще какие-то обобщения относительно митохондриальных ДНК пока довольно трудно.

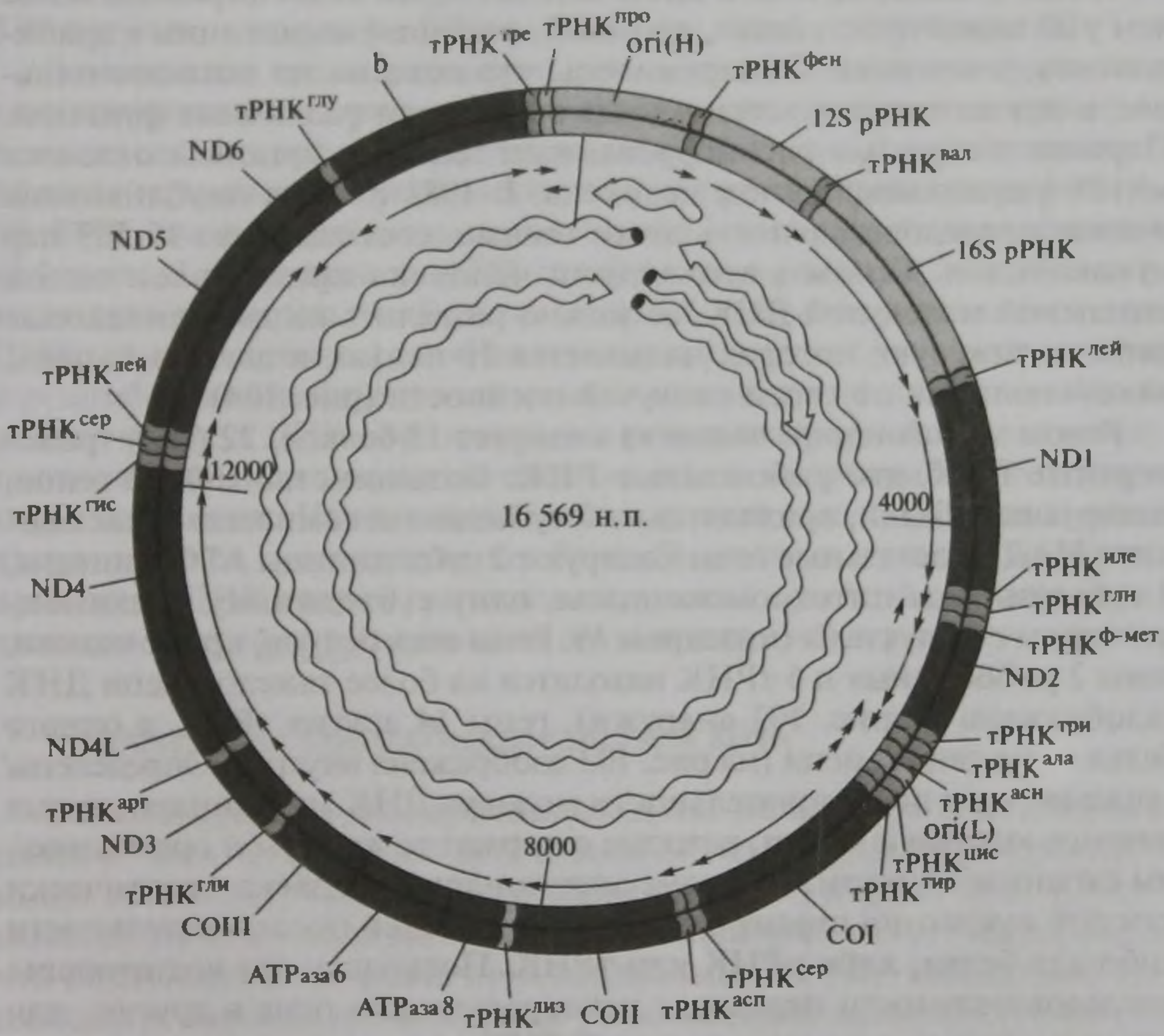


Рис. 104. Митохондриальная ДНК человека.

Обозначены гены, кодирующие субъединицы НАДН-дегидрогеназного комплекса (ND), субъединицы цитохром-*c*-оксидазы (COI, COII, COIII), цитохром *b* (цит. *b*), субъединицы 6 и 8 АТР-синтазы (АТРаза 6 и АТРаза 8), кодирующие рРНК (12SрРНК и 16SрРНК); кодирующие тРНК; ori (H) и ori (L)-сайты инициации репликации тяжелой и легкой цепей ДНК. Точки внутри круга обозначают сайты инициации транскрипции. Маленькие стрелки указывают 5′→3′-направление смысловой цепи в различных генах

Иногда даже близкородственные виды имеют совершенно разные рестриктные карты и лишь частично совпадающие нуклеотидные последовательности. Расположение гомологичных генов в кольцевой ДНК также различается. Иногда у отдельного растения присутствуют митохондриальные геномы разных размеров, и представлены они как линейными, так и кольцевыми молекулами. Митохондрии некоторых грибов и простейших имеют линейные геномы.

Все митохондриальные ДНК не метилированы. В очень крупных митохондриальных геномах подавляющую долю избыточной ДНК составляют некодирующие последовательности с неизвестной функцией. Напротив, митохондриальные ДНК млекопитающих организованы рационально, и между генами почти нет промежутков.

К настоящему времени ДНК митохондрий секвенирована более чем у 20 видов простейших, дрожжей, растений маршанции и арабидопсиса, у человека. Обнаружилось, что сходные по последовательности нуклеотидов участки иногда выполняют различные функции. Первым полностью расшифрованным геномом органелл оказался митохондриальный геном человека. В 1981 г. была опубликована полная последовательность этого генома, состоящая из 16 569 пар нуклеотидов. Геном митохондрий человека представлен одной кольцевой молекулой ДНК, ее можно разделить на две одинаковые цепи — тяжелую, которая называется H-цепью, и легкую L-цепь, различающиеся по своей плавучей плотности (рис. 104).

Геном митохондрий человека кодирует 13 белков, 22 (все) транспортные РНК, две рибосомные РНК. Большинство (60 %) генов, кодирующих белки, приходятся на 7 субъединиц комплекса, окисляющего НАДН, остальные гены кодируют 2 субъединицы АТФ-синтазы, 3 субъединицы цитохромоксидазы, одну субъединицу убихинон-цитохром *c*-редуктазы (цитохром *b*). Гены всех белков, кроме одного, гены 2 рибосомных и 6 тРНК находятся на более тяжелой цепи ДНК (изображена на рис. 103 снаружи), гены 14 других тРНК и одного белка — на легкой цепи (на рис. 103 изображена внутри). Определены также полные последовательности молекул ДНК митохондриальных геномов коровы и мыши, которые содержат те же гены и организованы сходным образом. В геноме митохондрий человека практически каждый нуклеотид входит в состав кодирующей последовательности либо для белка, либо рРНК или тРНК. Поскольку эти кодирующие последовательности переходят непосредственно одна в другую, для регуляторных последовательностей ДНК остается очень мало места. Промежутки между генами, как правило, не превышают 25 н.п. и в большинстве случаев составляют менее 3 н.п. и иногда перекрываются (например, ген $тРНК^{иле}$, $тРНК^{глу}$ и гены АТФ-синтазных субъединиц 6 и 8).

Митохондриальная ДНК дрожжей и человека содержит примерно равное число генов, но у дрожжей она больше почти в 5 раз — около 80 тыс.н.п. Основная ее часть приходится на долю АТ-богатых

участков с неизвестной функцией, а также интронов, которые прерывают кодирующие участки и отсутствуют в митохондриальных ДНК млекопитающих. Различия в структуре митохондриальной ДНК у разных штаммов дрожжей определяются в основном изменчивостью некодирующих участков.

Количество митохондрий в растительной клетке составляет от 50 до 2 000, и каждая митохондрия содержит от 1 до 100 копий генома. Геном митохондрий растений значительно больше генома митохондрий животных. У арабидопсиса (сем. крестоцветных) митохондриальный геном содержит 370 тыс. н.п., т. е. в 20 раз больше, чем геном митохондрий человека. Размер митохондриального генома растений сильно варьирует даже внутри одного семейства иногда в 5 — 10 раз. Так, в семействе тыквенных у арбуза в ДНК митохондрий содержится 330 тыс. н.п., у тыквы — 850 тыс. н.п., а у дыни — 2 400 тыс. н.п.

В настоящее время полностью секвенирована последовательность митохондриального генома печеночника маршанции, завершена работа по секвенированию митохондриального генома арабидопсиса. Митохондриальный геном растений кодирует три рибосомные РНК, 16 тРНК, около 10 рибосомальных белков, некоторые белки дыхательной цепи, 3 субъединицы АТФ-синтазы, четыре белка, участвующих в синтезе цитохрома *c*. Большая вариабельность размера молекул ДНК, их множественность в митохондриях высших растений связана с рекомбиногенной природой самого генома у этих объектов. Множество молекул ДНК образовалось в результате рекомбинаций между повторяющимися последовательностями и их рекомбиногенными перестановками. Часть последовательностей имеет ядерное и даже хлоропластное происхождение. Рекомбинации и мутации митохондриальной ДНК ведут к цитоплазматической мужской стерильности, вызывая нарушения созревания пыльцевых трубок.

Репликация митохондриальной ДНК

Репликация митохондриальной ДНК не связана с клеточным циклом. Каждая из цепей дуплексной молекулы ДНК митохондрий имеет свою точку начала репликации (точка ori на рис. 104). Репликация ДНК в митохондриях человека однонаправленна и асинхронна, что обусловливается различной локализацией точек инициации репликации. Сначала инициируется репликация H-цепи ДНК, а синтез L-цепи — только после того, как синтезируется около 67 % H-цепи.

В дрожжевой митохондриальной ДНК имеется семь и более мест начала репликации, рассеянных по всему геному; каждое из таких мест содержит три коротких GC-богатых сегмента, разделенных двумя длинными участками, почти целиком состоящими из AT-пар.

Синтез митохондриальной ДНК осуществляется при участии ДНК-полимеразы γ. Сведения о других ферментах и белковых факторах, участвующих в репликации митохондриальной ДНК, отсутствуют.

Полиморфизм митохондриальной ДНК и эволюция человека

Уникальным свойством митохондриальной ДНК человека, как и других позвоночных, является наследование ее по материнской линии, что было продемонстрировано с помощью семейного анализа.

Возникновение мутаций в одной из молекул митохондриальной ДНК создает внутриклеточную смесь мутантных и нормальных молекул — состояние, известное как гетероплазмия. Последовательные деления гетероплазмической клетки сдвигают соотношение мутантных и нормальных митохондриальных ДНК либо к чисто мутантным, либо к нормальным (гомоплазмия) молекулам ДНК. Несмотря на то что гетероплазмия в настоящее время — единственный механизм формирования разнообразия митохондриальных геномов, молекулярные основы поддержания состояния гетероплазмии остаются невыясненными. Высокая скорость эволюции митохондриальной ДНК в 10—20 раз больше, чем у ядерных генов, и это повышает частоту ее мутаций. Из-за отсутствия систем репарации и измененного генетического кода (в большинстве случаев замены оснований в третьем положении кодонов не отражаются на аминокислотных последовательностях белков) мутации фиксируются с большей вероятностью. Высокая скорость накопления мутаций в митохондриальной ДНК наряду с отсутствием рекомбинационных процессов является крайне важным свойством, которое позволяет использовать митохондриальную ДНК в эволюционных и популяционно-генетических исследованиях.

Вариабельность нуклеотидной последовательности митохондриальной ДНК человека наибольшая в районе главной некодирующей области (контрольного региона), где находятся все структуры, ответственные за инициацию и регуляцию процессов транскрипции и репликации, точка ori для тяжелой цепи, промоторы транскрипции обеих цепей митохондриальной ДНК, сайты связывания транскрипционных факторов. Вариабельность этого района обусловлена преимущественно нуклеотидными заменами и короткими делециями или инсерциями. У человека, как и других млекопитающих, он состоит из трех доменов, различающихся по степени вариабельности и насыщенности функциональными элементами. Наиболее вариабельной частью контрольного региона мтДНК является домен, расположенный между позициями 16 024 — 16 400 н. п., получивший название гипервариабельного сегмента I митохондриальной ДНК. Из-за высокого уровня изменчивости именно этого сегмента контрольного региона митохондриальная ДНК наиболее часто используется в эволюционных исследованиях.

К настоящему времени исследования полиморфизма мтДНК выполнены во многих этнических группах человека. Полученные данные о генетической дифференциации современного человечества

широко используются для установления родства между этносами и реконструкции расогенетических процессов.

Исследования вариабельности мтДНК продемонстрировали высокий уровень митохондриального генофонда и наличие корреляций между типами мтДНК и этногеографическим происхождением индивидуумов. Согласно полученным данным, максимальное разнообразие типов митохондриальной ДНК было зафиксировано в популяциях негроидов. С помощью филогенетического анализа было установлено, что все типы этой ДНК человека являются ветвями одного дерева, причем корневые типы его характерны для современных негроидов, что, по сути, является подтверждением гипотезы об африканском происхождении современного человека.

ДНК хлоропластов

Хлоропластная ДНК (хлДНК) представляет собой замкнутую кольцевую двуспиральную молекулу. Ее размеры варьируют у разных видов растений преимущественно в интервале 130 — 160 тыс. н.п. В настоящее время полностью расшифрована нуклеотидная последовательность хлДНК ряда видов, в том числе табака и риса. При этом обнаружены общие принципы организации хлоропластной ДНК и ее консервативность (неизменность первичной структуры) в ходе эволюции. ХлДНК содержит около 130 генов. В ней представлены по два гена четырех типов (4,5S, 5S, 16S и 23S) рибосомальных РНК (рРНК), гены всех транспортных РНК (около 30 видов), гены рибосомальных белков (около 20). Гены субъединиц РНК-полимеразы — фермента, осуществляющего синтез РНК на хлДНК. Хлоропластный геном кодирует около 40 белков тилакоидной мембраны, участвующих в формировании комплексов электронтранспортной цепи. Это составляет около половины входящих в них белков. Остальные белки тилакоидной мембраны кодируются в ядре. ХлДНК содержит ген большой субъединицы ключевого фермента фотосинтеза рибулозодифосфаткарбоксилазы (РДФК).

Организация генетического аппарата хлоропластов и бактерий имеет много общего. По прокариотическому типу организованы промоторы, регулирующие начало транскрипции и локализованные в области 35 — 10 н.п. до точки начала транскрипции, и терминаторы, определяющие ее окончание. Вместе с тем в отличие от прокариот в хлДНК обнаружены интроны, характерные для генов эукариот. Хлоропластные рибосомы относятся к 70S-типу, характерному для прокариот.

Для каждого вида растений характерно определенное число хлоропластов в клетке, варьирующее у разных видов от нескольких единиц до величин, превышающих сотню. Число хлоропластов в клетке, а следовательно, их деление контролируется ядром. Так, ДНК-полимераза, осуществляющая репликацию хлДНК, кодируется

в ядре, синтезируется на 80S рибосомах цитоплазмы и затем проникает в хлоропласт, где и обеспечивает синтез ДНК. Формирование ферментативного аппарата темновых реакций фотосинтеза, в ходе которых происходит усвоение CO_2, также находится под контролем двух геномов, ядра и хлоропласта. Большая субъединица (54 кДа) ключевого фермента цикла Кальвина — Бенсона рибулозодифосфаткарбоксилазы кодируется и синтезируется в хлоропласте, а малая (14 кДа) кодируется в ядре и синтезируется в цитоплазме в виде предшественника, а затем проникает в хлоропласт, где оба полипептида образуют структуру функционально активного фермента рибулозодифосфаткарбоксилазы, состоящего из 8 больших и 8 малых субъединиц. Таким образом, формирование всех важнейших структур хлоропласта зависит и от ядра, и цитоплазмы. Это объясняет невыполнимость идеи создания культуры изолированных хлоропластов, где бы они самостоятельно размножались на питательной среде.

Происхождение ДНК органелл

Широко распространено мнение, что митохондрии и хлоропласты произошли от прокариотических эндосимбионтов, которые обитали в цитоплазме предшественников эукариот. Как полагают, митохондриям дали начало пурпурные бактерии, а хлоропластам (позднее) — цианобактерии (синезеленые водоросли) или близкие к ним организмы.

Симбионты проникли в эукариотические клетки и в ходе эволюции потеряли свою автономность, передав большое число важнейших генов в ядерный геном. В результате независимая бактериальная клетка превратилась в полуавтономную органеллу, сохранившую главную исходную функцию — способность к фотосинтезу (у хлоропластов) и систему окислительного фосфорилирования (у митохондрий).

Хотя многие гены этих древних бактерий все еще используются для синтеза белков органеллы, большая их часть по неясным причинам включилась в ядерный геном, где они кодируют ферменты, которые сходны с бактериальными и синтезируются на рибосомах в цитоплазме, а затем переходят в органеллу.

Существуют экспериментальные данные, свидетельствующие о том, что последовательности ДНК «путешествуют» между органеллами и ядром, а также между геномами самих органелл. Подтверждением первого является тот факт, что две субъединицы рибулозодифосфаткарбоксилазы у простейшего *Cyanophora paradoxa* кодируются неядерным геномом, входящим в состав хлоропластоподобной структуры — цианеллы, типичной для цианобактерий. Однако механизмы перемещения и интеграции ДНК органелл в новые геномы пока не установлены.

Глава 7
РЕПЛИКАЦИЯ ДНК

Генетическая программа клеточных организмов записана в нуклеотидной последовательности ДНК. Для сохранения уникальных свойств организма необходимо точное воспроизведение этой последовательности в каждом последующем поколении. Во время деления клетки содержание ДНК должно удвоиться, чтобы каждая дочерняя клетка могла получить полный спектр ДНК. Так, кишечная палочка при образовании каждого последующего поколения должна дуплицировать практически без ошибок полный геном размером $4 \cdot 10^6$ н. п., точно так же должны быть скопированы почти 6 млрд н. п. в любой соматической клетке человека.

Процесс удвоения родительских молекул геномной ДНК во время воспроизводства клеток живого организма получил название репликации, или репликативного синтеза ДНК.

Репликация ДНК является примером матричного синтеза биологических макромолекул. Основу хромосомы составляет одна непрерывная двуцепочечная молекула ДНК. Во время репликации каждая из цепей родительской ДНК служит матрицей для синтеза комплементарной дочерней цепи. Положение каждого нуклеотида в строящейся цепи ДНК по правилам комплементарности (G—C и A—T) однозначно определяется положением соответствующего нуклеотида матрицы.

Механизм происходящей при репликации ДНК химической реакции заключается в переносе остатка дезоксирибонуклеозидмонофосфата от дезоксирибонуклеозидтрифосфата на концевой нуклеотидный остаток растущей в процессе синтеза нуклеотидной цепи (рис. 105). Перенос идет на место атома H гидроксильной группы, стоящей при 3-м углеродном атоме дезоксирибозы концевого нуклеотида, и сопровождается выделением пирофосфата, который затем расщепляется неорганической пирофосфатазой, что делает реакцию полимеризации практически необратимой.

Так как формирование нового полинуклеотида идет на полинуклеотидной матрице при непрерывном замыкании водородных связей между комплементарными пуриновыми и пиримидиновыми основаниями матрицы и дезоксирибонуклеозидтрифосфатов, то условием функционирования этого механизма является одноцепочечная

структура матрицы. В связи с этим существенным моментом в биосинтезе является расхождение биспирального полидезоксирибонуклеотида на одиночные полидезоксирибонуклеотидные цепи, на которых осуществляется сборка комплементарных им полинуклеотидов. В итоге из одной биспиральной молекулы ДНК образуются две биспиральные молекулы ДНК, совершенно идентичные друг другу и исходной молекуле ДНК. После одного раунда репликации одна цепь в каждой из двух дочерних молекул является родительской, т.е. консервативной, а другая — синтезированной заново. Такой способ удвоения молекул ДНК называется *полуконсервативным*. В 1958 г. М. Мезельсон и Ф. Сталь получили экспериментальное доказательство полуконсервативного способа репликации ДНК.

Белки и ферменты, участвующие в репликации ДНК

Процесс репликации ДНК осуществляется с участием множества белков, которые образуют сложный и эффективно работающий репликативный комплекс.

ДНК-полимеразы. Комплементарное копирование одноцепочечной матрицы катализируют ДНК-зависимые *ДНК-полимеразы* или просто ДНК-полимеразы. Некоторые про- и эукариотические ДНК-полимеразы получены в чистом виде, охарактеризованы их физические и ферментативные свойства. И хотя эти свойства не совсем идентичны, механизм их катализа для всех указанных ферментов в общих чертах одинаков.

Полимеризация нуклеотидов в процессе репликации происходит только в одном направлении: от 5′-конца к 3′-концу (5′ → 3′) строящейся цепи, и синтезированная полинуклеотидная цепь ДНК антипараллельна по отношению к ДНК-матрице (рис. 105).

ДНК-полимеразы безразличны к последовательности нуклеотидов матрицы. ДНК синтезируется чрезвычайно быстро: скорость ее полимеризации колеблется в пределах от 50 нуклеотидов/с у эукариот до 500 нуклеотидов/с у бактерий.

Нормальное размножение клеток требует высокой точности копирования ДНК-матрицы, для чего существуют специальные механизмы, один из которых — *механизм коррекции*. ДНК-полимеразы проверяют комплементарность каждого нуклеотида матрицы дважды: один раз — перед включением его в состав растущей цепи и второй раз — перед тем, как включить следующий нуклеотид.

Очередная фосфодиэфирная связь в процессе репликации образуется лишь в том случае, если последний (3′-концевой) нуклеотид синтезируемой цепи комплементарен матрице. Если же на предыдущей стадии полимеризации произошла ошибка, то репликация

останавливается до тех пор, пока неправильный нуклеотид не будет удален. Для этого фермент перемещается в обратном направлении и вырезает последнее добавленное звено. Оказалось, что некоторые ДНК-полимеразы обладают наряду с полимеразной также 3′→5′-экзонуклеазной активностью (см. с. 218), благодаря которой и происходит отщепление не спаренного с матрицей нуклеотида синтезируемой цепи, после чего полимеризация восстанавливается. В результате такого устранения ошибок считывания их уровень сводится к минимуму. Так, частота ошибочных включений нуклеотидов в ДНК у кишечной палочки не превышает $10^{-9} — 10^{-10}$ на нуклеотид за один раунд репликации.

ДНК-праймаза. ДНК-полимеразы не способны инициировать синтез новых цепей ДНК, они могут лишь добавлять дезоксирибонуклеотидные звенья к 3′-концу уже имеющейся полинуклеотидной цепи. Чтобы молекулы ДНК-полимераз могли начать синтез ДНК, им необходима *затравка*, или *праймер* (от англ. *primer* — затравка), короткий олигодезоксирибонуклеотид или олигорибонуклеотид, комплементарный соответствующему участку ДНК-матрицы, у которого на конце имеется свободная 3′-ОН-группа. Неспособность молекул ДНК-полимераз самостоятельно, без затравки начинать синтез ДНК принципиально отличает эти ферменты от других ферментов матричного синтеза — РНК-полимераз.

На стадии инициации репликации короткую РНК-затравку из рибонуклеозидтрифосфатов синтезирует фермент, называемый ДНК-праймазой. ДНК-праймаза может быть отдельным ферментом (как у бактерий) или входить в качестве субъединицы в ДНК-полимеразу (как у ДНК-полимеразы α эукариот). В любом случае праймаза — это фермент, отличный от РНК-полимераз, которые синтезируют разнообразные клеточные РНК. После того как будет синтезирован РНК-праймер, подключается ДНК-полимераза и продолжает наращивать цепь. Новосинтезированные цепи ДНК всегда содержат на 5′-конце несколько рибонуклеотидов: у прокариот — от двух до пяти нуклеотидов, у эукариот их в два раза больше. В дальнейшем короткие праймеры замещаются сегментами ДНК.

Поскольку две цепи в молекуле ДНК антипараллельны, а ДНК-полимеразы способны наращивать полинуклеотидную цепь только в направлении 5′→3′, один и тот же фермент, ДНК-полимераза, не может обеспечить сборку дочерних цепей одновременно в направлениях 5′→3′ и 3′→5′ в соответствии с перемещением репликативной вилки при расплетании биспиральной молекулы ДНК. Направление расплетания двойной спирали при репликации совпадает с направлением синтеза ДНК лишь для одной матричной цепи, в то же время оно противоположно направлению синтеза ДНК на комплементарной матрице. Это значит, что лишь на одной матричной цепи, ориентированной в направлении 3′→5′, синтез ДНК может проходить непрерывно. На второй матричной цепи, ориентированной

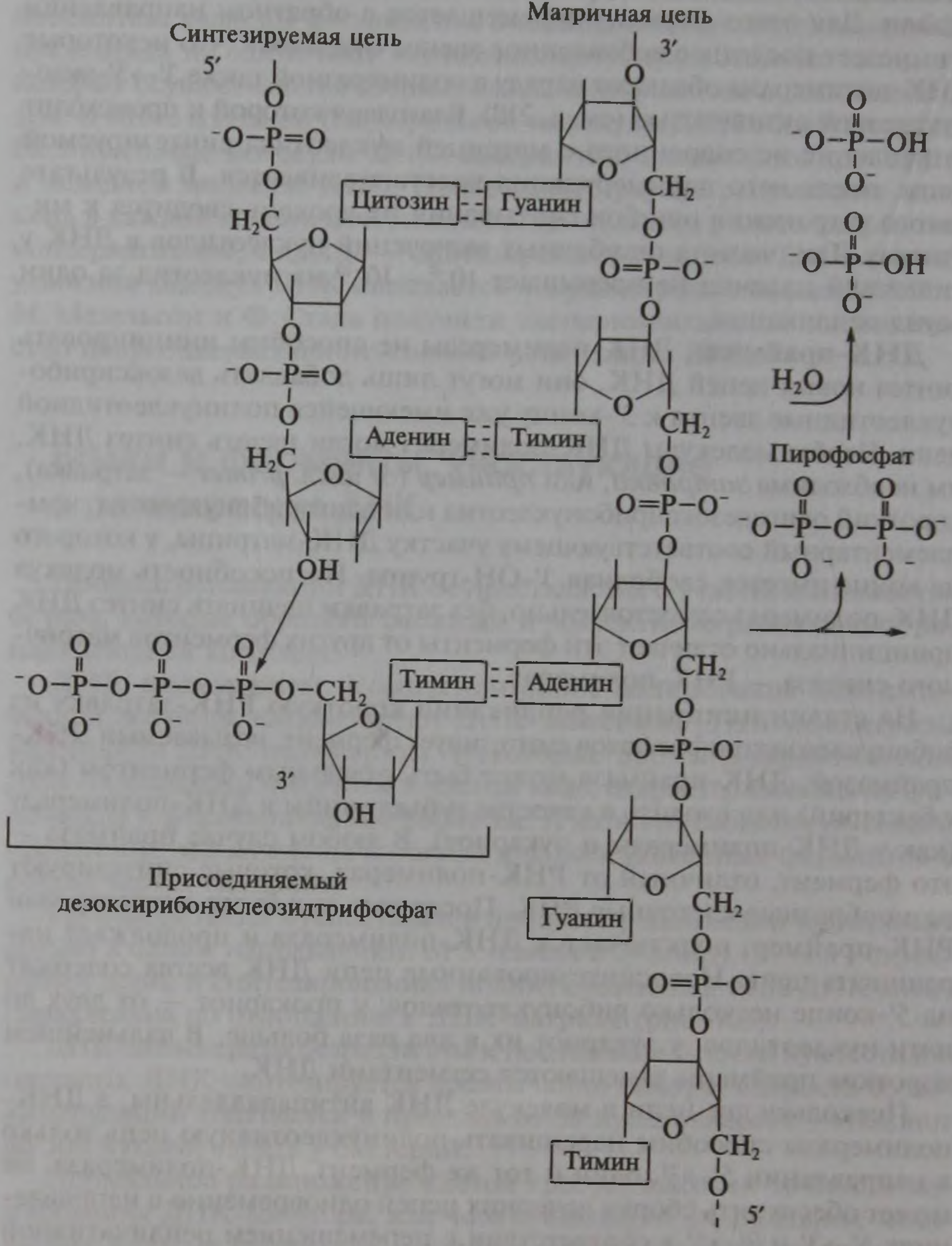
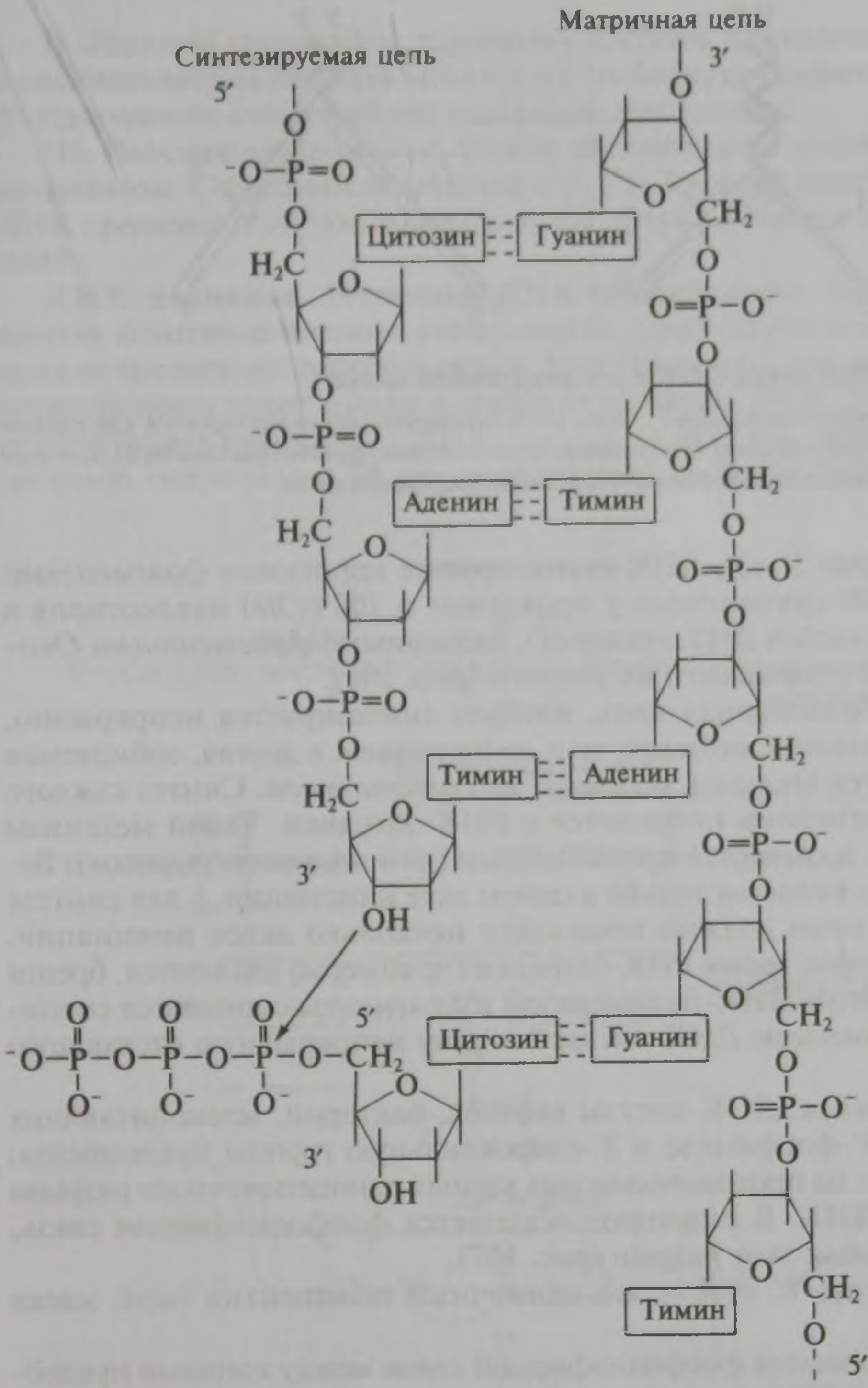


Рис. 105. Схема удвоения молекулы ДНК



при репликации

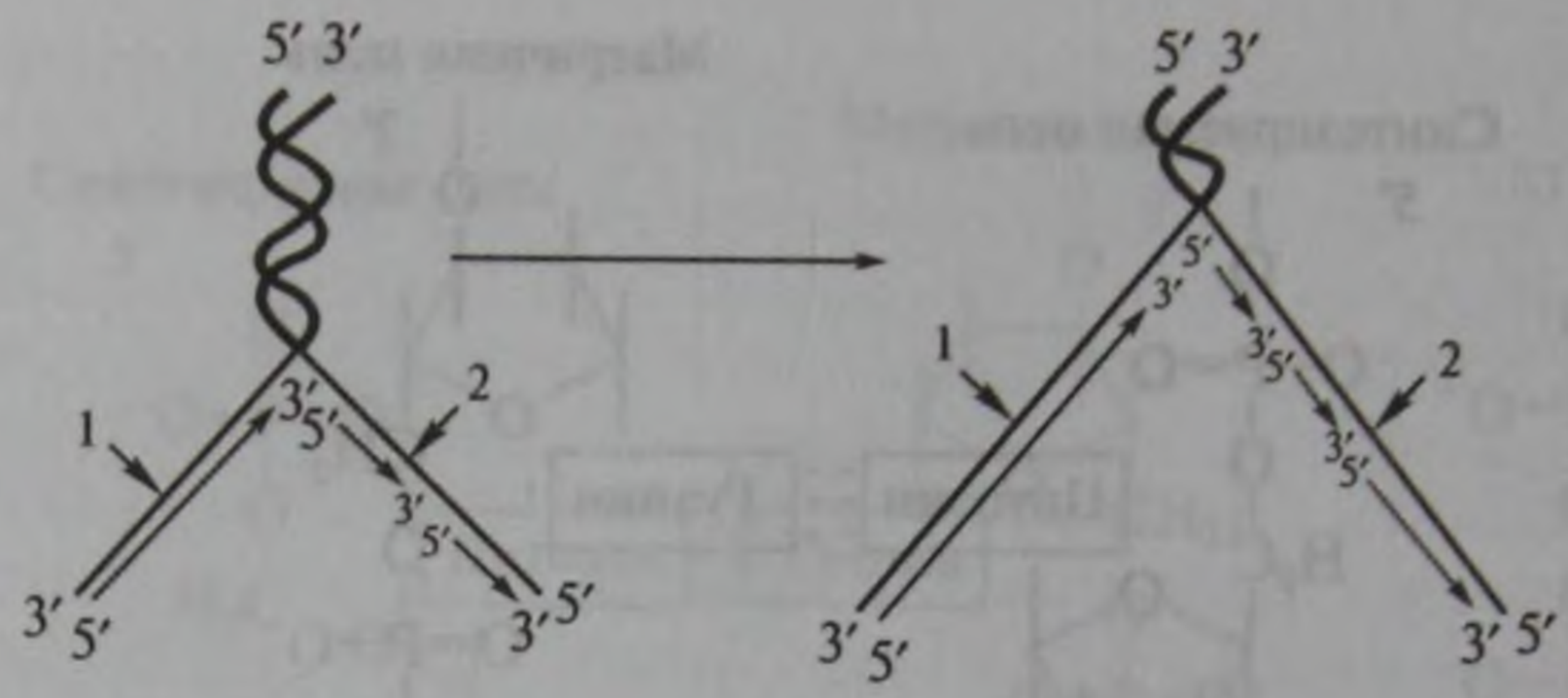


Рис. 106. Синтез цепей ДНК в репликативной вилке.

Направление репликации 5′→3′. Одна цепь (лидирующая) синтезируется как единое целое, а другая (отстающая) — короткими фрагментами (фрагменты Оказаки), которые впоследствии соединяются с образованием непрерывной цепи

в направлении 5′→3′, ДНК синтезируется короткими фрагментами (1 000 — 2 000 нуклеотидов у прокариот и 100 — 200 нуклеотидов в реплицирующейся ДНК эукариот), названными *фрагментами Оказаки* в честь открывшего их ученого (рис. 106).

Вновь образованная цепь, которая синтезируется непрерывно, получила название *ведущей*, или *лидирующей*, а другая, собираемая из фрагментов Оказаки, *ведомой*, или *отстающей*. Синтез каждого из этих фрагментов начинается с РНК-затравки. Такой механизм репликации называется *прерывистым* (или *полунепрерывным*). Ведущая цепь нуждается только в одном акте инициации, а для синтеза отстающей цепи должно произойти несколько актов инициации. Через некоторое время РНК-затравки (праймеры) удаляются, бреши «застраиваются» ДНК-полимеразой и фрагменты сшиваются специальным ферментом ДНК-лигазой в одну непрерывную отстающую цепь.

ДНК-лигаза. ДНК-лигазы вирусов, бактерий, млекопитающих соединяют 5′-фосфатную и 3′-гидроксильную группы нуклеотидов, находящихся на противоположных концах одноцепочечного разрыва в дуплексе ДНК. В результате образуется фосфодиэфирная связь, ликвидирующая этот разрыв (рис. 107).

ДНК-лигаза *E. coli* — это одиночный полипептид (мол. масса 75 000 Да).

Для образования фосфодиэфирной связи между концами нуклеотидных цепей ДНК-лигазы используют энергию гидролиза ATP либо NAD. Реакция протекает в три стадии:

I. *Реакция аденилирования ДНК-лигазы.* Фермент взаимодействует с ATP (ДНК-лигаза фагов и млекопитающих) или NAD (ДНК-лигаза *E. coli*) с образованием фермента — аденилатного производного по ε-NH_2-группе остатка лизина полипептидной цепи с одновременным высвобождением пирофосфата (в случае ATP) или никотинамидмононуклеотида (в случае NAD).

II. *Реакция трансаденилирования.* Остаток адениловой кислоты с аденилатлигазы перебрасывается на свободную фосфатную группу 5′-углеродного атома рибозы концевого нуклеотида.

III. *Реакция лигирования.* Между сближенными активированной аденилатом 5′-фосфатной группой и 3′-ОН-группой фрагмента цепи ДНК происходит образование фосфодиэфирной связи с выделением AMP.

ДНК-хеликаза. Нативные ДНК биспиральны. Для осуществления комплементарного копирования цепей двуцепочечная ДНК должна постепенно раскручиваться. Раскручивание, или расплетание, спирали происходит только в локальном участке ДНК. Эту реакцию осуществляет хеликаза — ДНК-зависимая АТРаза, использующая энергию гидролиза АТР для расплетания двойной спирали ДНК.

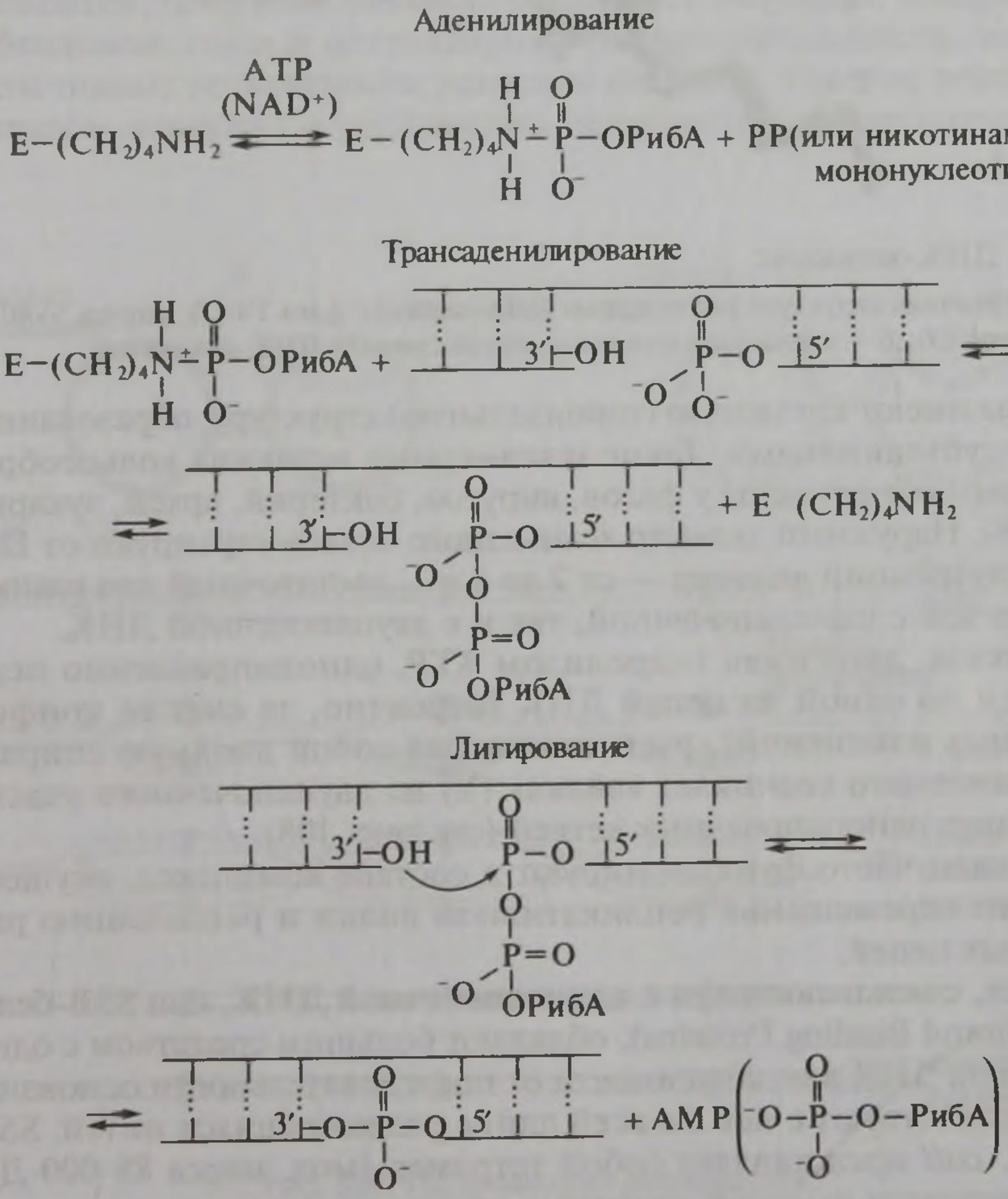


Рис. 107. Механизм действия ДНК-лигазы:

Риб — рибоза; А — аденин; PP — пирофосфат

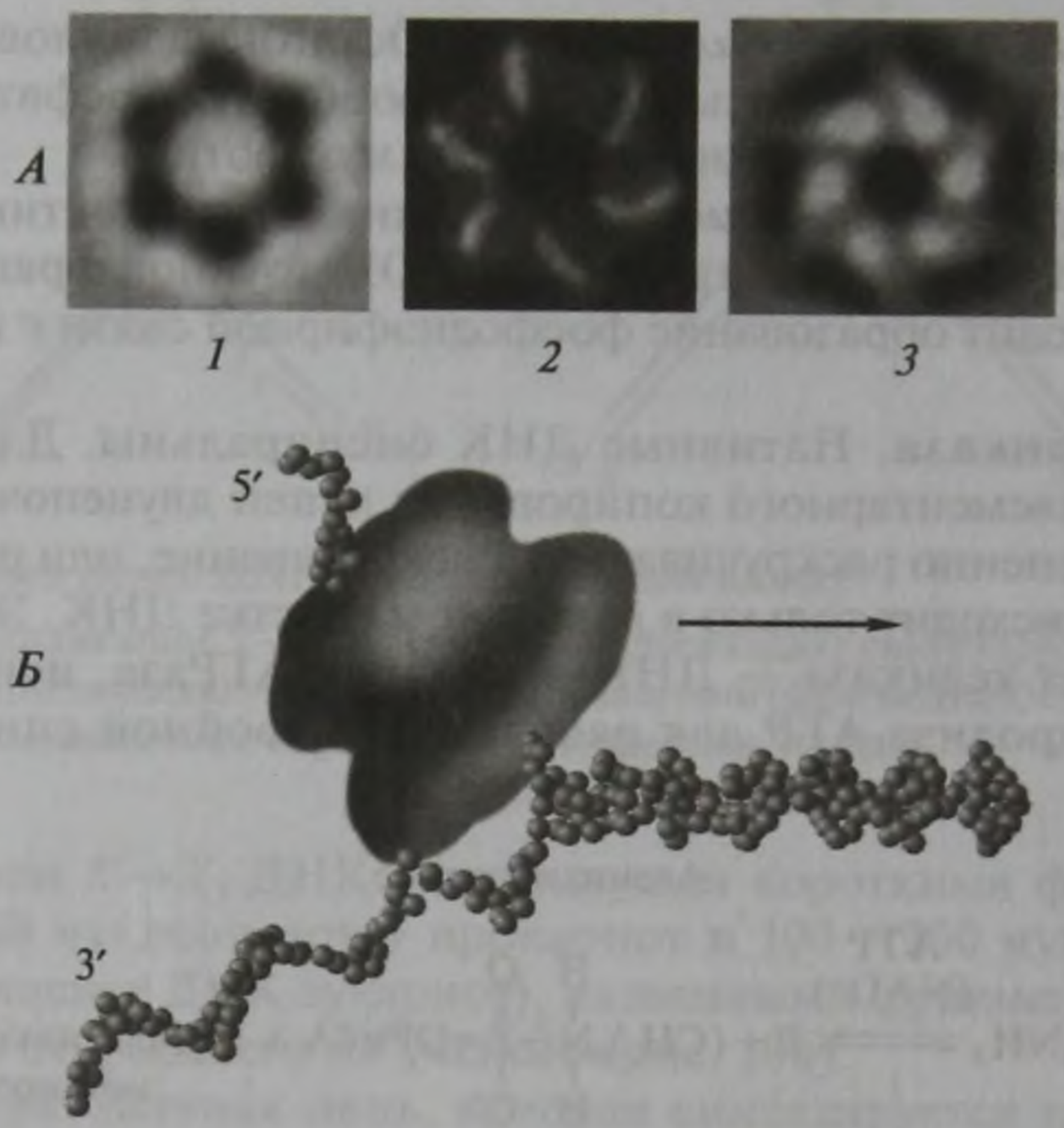


Рис. 108. ДНК-хеликазы:

А — четвертичная структура гексамерных ДНК-хеликаз: фага Т4 (*1*), вируса SV40 (*2*), Dna B *E. coli* (*3*); *Б* — схема расплетания двойной спирали ДНК-хеликазой

Хеликазы имеют кольцевую (тороидальную) структуру, образованную шестью субъединицами. Такие гексамерные хеликазы кольцеобразной формы обнаружены у фагов, вирусов, бактерий, архей, эукариот (рис. 108). Наружный диаметр хеликазного кольца варьируют от 12 до 14 нм, внутренний диаметр — от 2 до 4 нм, достаточный для взаимодействия как с одноцепочечной, так и с двуцепочечной ДНК.

Хеликаза, движимая гидролизом АТР, однонаправленно перемещается по одной из цепей ДНК (вероятно, за счет ее конформационных изменений), расплетая перед собой двойную спираль, в результате чего возникает «вилка» (Y) из двуцепочечного участка ДНК и двух одноцепочечных ветвей (см. рис. 108).

Хеликазы часто функционируют в составе комплекса, осуществляющего перемещение репликативной вилки и репликацию расплетенных цепей.

Белки, связывающиеся с одноцепочечной ДНК, или SSB-белки (Single Strand Binding Proteins), обладают большим сродством с одноцепочечной ДНК вне зависимости от последовательности оснований и взаимодействуют с ней по всей длине разделившихся нитей. SSB-белок *E. coli* представляет собой тетрамер (мол. масса 88 000 Да), характеризуется высокой степенью асимметрии молекулы, преобладанием дикарбоновых аминокислот над диаминокислотами. Белок кооперативно связывается с одноцепочечной ДНК, сохраняя ее в рас-

тянутом состоянии. Суть кооперативного способа связывания состоит в том, что связывание одной молекулы белка облегчает связывание другой, пока вся одноцепочечная ДНК не будет покрыта белком SSB. Если хеликазы вызывают локальное раскручивание двойной спирали, то SSB-белки стабилизируют одноцепочечную ДНК, обеспечивая условия для комплементарного спаривания. SSB-белки удаляют возможные элементы вторичной структуры ДНК; связывание одноцепочечной ДНК с SSB-белками стимулирует ДНК-полимеразу и повышает точность ее работы. У эукариот таковым белком является ядерный репликативный белок А (RPA), представляющий гетеротример с субъединицами 70, 32 — 34 и 11 — 14 кДа.

Топологические проблемы репликации ДНК рассмотрены в гл. 3 (с. 96).

Начало репликации. Молекула ДНК, способная к автономной репликации, получила название *репликона*. Репликон содержит все необходимые гены и регуляторные последовательности, которые обеспечивают регулируемое удвоение его ДНК. Участок репликона, в котором начинается репликация, называется репликатором, или

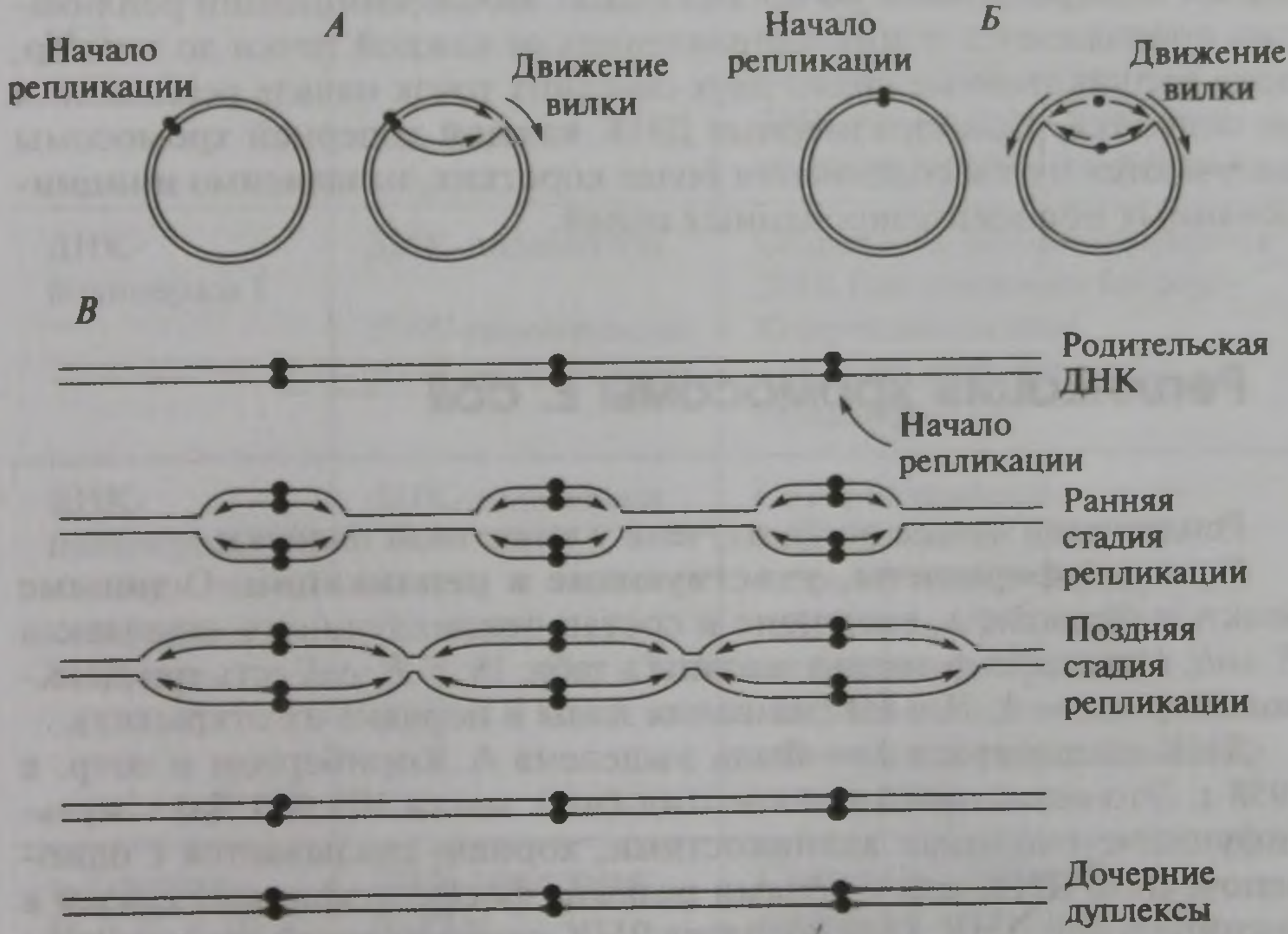


Рис. 109. Направление репликации:

А — однонаправленная репликация в кольцевой молекуле ДНК; *Б* — двунаправленная репликация в кольцевой молекуле ДНК; *В* — множественная двунаправленная репликация в эукариотической хромосоме (полирепликон)

областью начала репликации (англ. *replication origin*), у *E. coli* — ori C.

При инициации репликации *инициатор*, кодируемый репликоном (у *E. coli* — белок Dna A), взаимодействует с репликатором. Целая хромосома E. coli является репликоном; у эукариот репликоном называют отрезок ДНК, репликация которого протекает под контролем единственной точки начала репликации. Точка начала репликации имеет специфическую последовательность оснований, богатую парами А—Т, что, вероятно, облегчает разделение цепей.

В результате инициации раунда репликации в точке ori образуется одна или две репликативные вилки. При однонаправленной репликации возникает одна репликативная вилка (плазмида ColF1), и синтез заканчивается в точке ori (рис. 109).

При двунаправленной репликации (хромосома *E. coli*) инициируются две репликативные вилки, перемещающиеся в противоположных направлениях до тех пор, пока они не встретятся на противоположной стороне кольца.

Для геномов эукариотических клеток обычно характерно наличие множественных точек начала репликации, разбросанных по хромосоме на расстоянии 20 тыс. н.п., т. е. эукариотические хромосомы имеют полирепликонную организацию. После инициации репликация продолжается в двух направлениях от каждой точки до тех пор, пока репликативные вилки двух соседних точек начала репликации не сольются. Полноразмерные ДНК каждой дочерней хромосомы получаются путем соединения более коротких, независимо инициированных новосинтезированных цепей.

Репликация хромосомы *E. coli*

Репликация лучше всего изучена у кишечной палочки.

Белки и ферменты, участвующие в репликации. Основные белки и ферменты, входящие в состав репликативного комплекса *E. coli*, а также их функции указаны в табл. 18. У *E. coli* есть три ДНК-полимеразы — I, II и III (названия даны в порядке их открытия).

ДНК-полимераза I — была выделена А. Корнбергом и сотр. в 1958 г. Это одиночный полипептид (мол. масса 109 000 Да) с мультифункциональными активностями, хорошо связывается с одноцепочечной ДНК или с зонами разрыва фосфодиэфирных связей в биспиральной ДНК. Она обладает ДНК-полимеразной активностью (см. рис. 105), а также осуществляет гидролиз фосфодиэфирных связей как в одной цепи ДНК, так и на неспаренном конце дуплексной ДНК, причем за один акт удаляется один нуклеотид, начиная с 3′-конца цепи (3′→5′-экзонуклеаза) (рис. 110).

Таблица 18. **Белки, входящие в состав репликативного комплекса *E. coli***

Белки	Энзиматическая активность	Биологическая функция
Dna A		Узнавание области начала репликации и привлечение к месту сборки остальных белковых компонентов репликосомы
Dna B	ДНК-хеликаза	Раскручивание двойной спирали ДНК в репликативной вилке
Dna C		Обеспечение взаимодействия хеликазы и праймазы с ДНК, находящейся в комплексе с SSB-белком
Dna G SSB	Праймаза	Синтез РНК-затравок Взаимодействие с одноцепочечной ДНК; активация ДНК-полимеразы и стабилизация расплетенного дуплекса
ДНК-полимераза I	ДНК-полимераза 3′—5′-экзонуклеаза 5′—3′-экзонуклеаза	Созревание реплицирующейся ДНК (застраивание брешей) Коррекция ошибок Удаление РНК-затравки (праймера)
ДНК-полимераза III ($\alpha\theta\varepsilon$) — кофермент β-белок	ДНК-полимераза	Синтез лидирующей и отстающей цепей ДНК Активация ДНК-полимеразы и ДНК-зависимой АТРазы; осуществление функции «скользящего зажима», обеспечивающей процессивность репликации
γ-Комплекс (γ, δ, δ', χ, ψ)	ДНК-зависимая АТРаза	Связывание затравки с матрицей; активация ДНК-полимеразы
τ-Белок		Сборка и димеризация холофермента ДНК-полимеразы

Окончание табл. 18

Белки	Энзиматическая активность	Биологическая функция
Лигаза	Лигаза	Лигирование фрагментов ДНК
Топоизомераза I	Топоизомераза	Релаксирование отрицательной суперспирализации
Топоизомераза II (ДНК-гираза)	Топоизомераза	Индуцирование образования отрицательных сверхвитков, разделение катенанов

Вторая реакция также состоит в отщеплении нуклеотидов, но гидролиз начинается с 5′-конца дуплексной ДНК в направлении к 3′-концу (5′→3′-экзонуклеаза). Эти различные активности присущи разным сайтам полипептидной цепи ДНК-полимеразы I. При обработке фермента трипсином полипептидная цепь расщепляется на большой и малый фрагменты. Большой, С-концевой фрагмент («фрагмент Кленова») сохраняет ДНК-полимеразную и 3′ → 5′-экзонуклеазную активности; малый, N-концевой фрагмент обладает только 5′ → 3′-экзонуклеазной активностью.

ДНК-полимераза I и присущие ей экзонуклеазные активности играют большую роль в репликации и репарации хромосомной ДНК *E. coli*. Экзонуклеазная активность 3′ → 5′ обеспечивает контроль за присоединением каждого нуклеотида и удаление ошибочных нуклеотидов с растущего конца цепи. Если эта активность подавлена в результате каких-то мутаций в гене, кодирующем ДНК-полимеразу I,

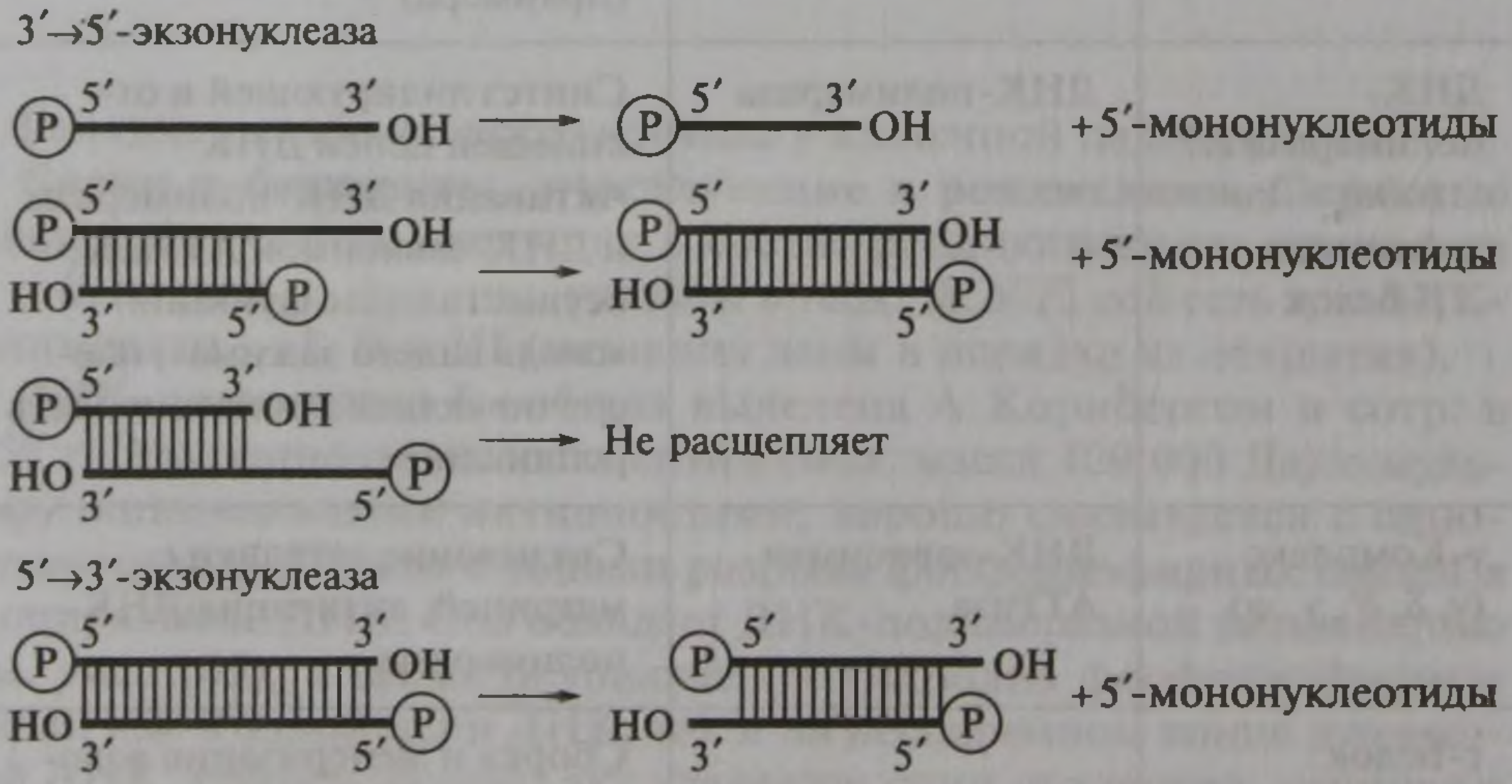


Рис. 110. Экзонуклеазные реакции ДНК-полимеразы I

то при репликации генома часто происходят мутации — замены оснований. Экзонуклеазная активность 5′ → 3′ осуществляет разрушение праймера, необходимого для синтеза фрагментов Оказаки (см. с. 212).

В клетке *E. coli* имеется несколько сотен молекул ДНК-полимеразы I. В целом ДНК-полимераза I имеет большее отношение к созреванию реплицирующейся ДНК, чем непосредственно к полимеразным процессам в репликативной вилке.

ДНК-полимераза II (мол. масса 90 кДа) — представлена одной полипептидной цепью, обладающей полимеразной и 3′ → 5′-экзонуклеазной активностями. Она плохо соединяется с одноцепочечными ДНК, но лучше работает с биспиральной ДНК, имеющей одноцепочечные бреши длиной в несколько десятков нуклеотидов, обладает лишь 10 %-й ДНК-полимеразной активностью по сравнению с ДНК-полимеразой I. Предполагают, что основной функцией ДНК-полимеразы II является достраивание поврежденных участков в молекуле ДНК, т. е. репарация ДНК.

ДНК-полимераза III (мол. масса 10^3 кДа) — играет главную роль в репликации ДНК у *E. coli*. В каждой клетке содержится только 10—20 копий фермента, приблизительно столько же, сколько репликативных вилок. ДНК-полимераза III является основным компонентом мультиферментного комплекса, инициирующего формирование репликативных вилок в точках начала репликации, участвующего в элонгации лидирующей цепи в вилке и удлиняющего РНК-праймеры с образованием фрагментов Оказаки.

ДНК-полимераза III состоит из десяти типов субъединиц (α, β, γ, δ, δ′, ε, θ, τ, χ, ψ). Репликацию проводит полная форма фермента — холофермент, содержащий все субъединицы. Холофермент не обладает 5′→3′-экзонуклеазной активностью, в связи с чем для репликации отстающей цепи необходимо участие ДНК-полимеразы I. Полимеразную реакцию осуществляет каталитический кор из α-, ε-, θ-субъединиц, в котором α-субъединица обладает полимеразной активностью, ε-субъединица — 3′→5′-экзонуклеазной активностью, функция θ-субъединицы пока неясна. Помимо субъединиц, составляющих полимеразный кор, ДНК-полимераза III-холофермент содержит еще семь субъединиц (τ, γ, β, δ, δ′, χ, ψ). Эти полипептиды существуют во множестве копий, являются регуляторными и усиливают действие каталитического ядра (кора) ДНК-полимеразы III.

Отличительная черта холофермента ДНК-полимеразы III — исключительно высокая процессивность. Мерой процессивности является длина фрагмента вновь синтезированной макромолекулы, которую комплекс (или индивидуальные ферменты) способен образовывать в одном цикле, не диссоциируя от матрицы. Установлено, что холофермент ДНК-полимеразы III синтезирует ведущую цепь ДНК длиной в 50 000 нуклеотидов со скоростью более 500 нуклеотидов в секунду в одном цикле, ни разу не диссоциируя от ДНК-матрицы.

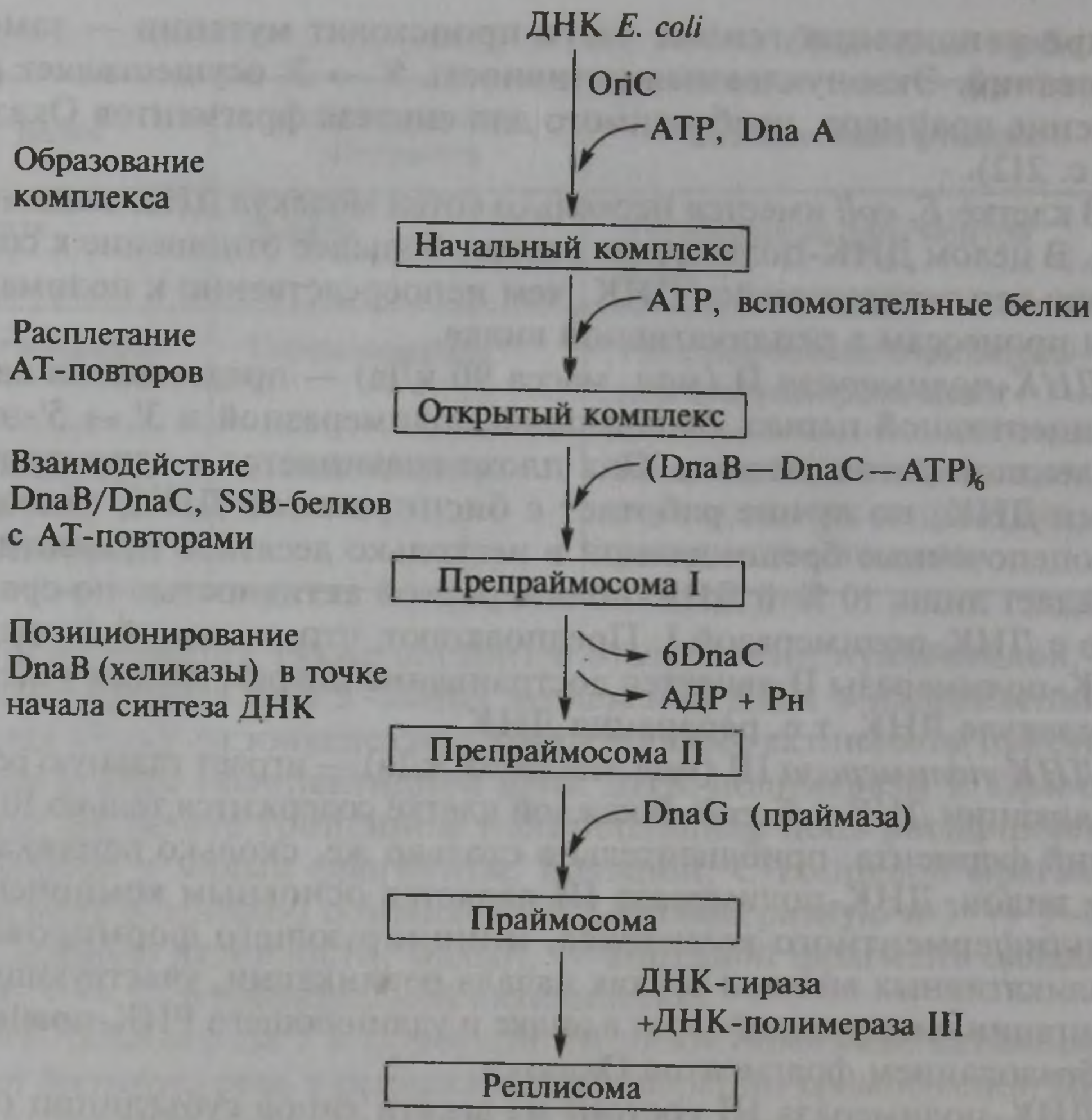


Рис. 111. Схема инициации репликации хромосомы *E. coli*

Инициация репликации (рис. 111). Хромосома *E. coli* содержит единственную область начала репликации (oriC), размер которой составляет 258 н.п. В oriC имеется пять консенсусных девятинуклеотидных сайтов связывания инициаторного белка Dna A, названных Dna A-боксами. В левой части oriC наряду с Dna A-боксами находятся АТ-богатые 13-нуклеотидные повторы. Белок Dna A распознает область начала репликации и привлекает к oriC остальные белковые компоненты, участвующие в инициации репликации.

Сначала белок Dna A в комплексе с АТР взаимодействует с Dna A-боксами. С помощью электронной микроскопии исходный комплекс обнаруживается в виде компактной эллипсоидной структуры, содержащей ~20 мономеров Dna A, которая закрывает oriC. В этом комплексе частично расплетаются АТ-богатые повторы и формируется открытый комплекс. В этом процессе участвуют некоторые вспомогательные белки, которые помогают инициатору Dna A раскручивать и изгибать ДНК (ранее они обозначались n, n′, n″, i, a. В последнее время их обозначают IME, FIS-факторы и т.д.).

Белок Dna B (хеликаза) в виде гексамеров (см. рис. 108) в комплексе с шестью мономерами белка Dna C, каждый из которых связывает одну молекулу ATP, а именно (Dna B—Dna C—ATP)$_6$, взаимодействует с одноцепочечными участками частично расплетенной ДНК. В этом комплексе хеликазная активность Dna B блокирована. Транслокация Dna B (хеликазы) от места ее первоначального вхождения в комплекс к месту старта репликативной вилки, которая ассоциируется с ATP-зависимым освобождением из комплекса белка Dna C, вызывает активацию хеликазы. Далее хеликаза взаимодействует с белком Dna G (праймазой), и этот комплекс играет ключевую роль в инициации репликации на oriC. Оба фермента обеспечивают сопряженное функционирование двух репликативных вилок, движущихся в противоположные стороны: хеликаза начинает расплетать дуплекс ДНК, праймаза синтезирует первые затравки. В сформированном праймирующем комплексе присутствие белка Dna A не требуется, и он после освобождения из комплекса может быть повторно использован для репликации на другом oriC.

Сложный комплекс белков, осуществляющий инициацию репликации, получил название *праймосомы*. У *E. coli* мол. масса праймосомы равна 10^6 Да. Праймосома в свою очередь является компонентом еще более сложного комплекса — реплисомы, осуществляющей процесс полной репликации. Схема инициации репликации хромосомы *E. coli* и образование реплисомы представлена на рис. 111. При образовании реплисомы происходит ATP-зависимое формирование димерного комплекса холофермента ДНК-полимеразы III, связанного с 3′-концами праймеров.

Элонгация репликации. Модель полуконсервативной прерывистой репликации ДНК в репликативной вилке представлена на рис. 112. Дуплекс родительской молекулы ДНК расплетает ATP-зависимая хеликаза. Образующиеся одноцепочечные участки кооперативно покрывает SSB-белок. ДНК-полимераза III движется по одной из матричных цепей в направлении раскрывания вилки и синтезирует ведущую цепь ДНК. По другой, матричной цепи в том же направлении движется праймосома. Время от времени, входящая в состав праймосомы праймаза (белок Dna G), синтезирует РНК-затравки для отстающей цепи. Вторая молекула холофермента ДНК-полимеразы III удлиняет эти затравки до тех пор, пока не достигнет предыдущей затравки, т. е. она синтезирует фрагменты Оказаки длиной от 1 000 до 2 000 нуклеотидов. Удаление сегментов РНК с 5′-конца каждого фрагмента Оказаки осуществляет ДНК-полимераза I, используя свою 5′→3′-экзонуклеазную активность. Заполнение пробелов (брешей) между фрагментами Оказаки катализируется той же ДНК-полимеразой I. Она низкопроцессивна, непрочно удерживается на матричной цепи и довольно быстро отделяется от нее после удаления РНК. После полимеразного действия указанного фермента между двумя соседними фрагментами Оказаки



Рис. 112. Схема синтеза ДНК в репликативной вилке у прокариот

остается только одноцепочечный разрыв, который зашивает ДНК-лигаза. Таким образом, из серии фрагментов Оказаки образуется непрерывная новая цепь.

Холофермент ДНК-полимеразы III может образовывать димер при участии τ-белка, в связи с чем ведущая и отстающая цепи реплицируются одновременно. Преодоление антипараллельности цепей при репликации в этом случае осуществляется за счет возникновения «петли» (модель А. Корнберга), где чередование фосфодиэфирных связей на ее восходящем отрезке изменяется на обратное и не препятствует ДНК-полимеразе III вести синтез фрагментов Оказаки на отстающей цепи родительской ДНК в том же направлении, что и на ведущей цепи (рис. 113).

Необычность ситуации во время синтеза ДНК в репликативной вилке заключается в том, что один и тот же белковый комплекс

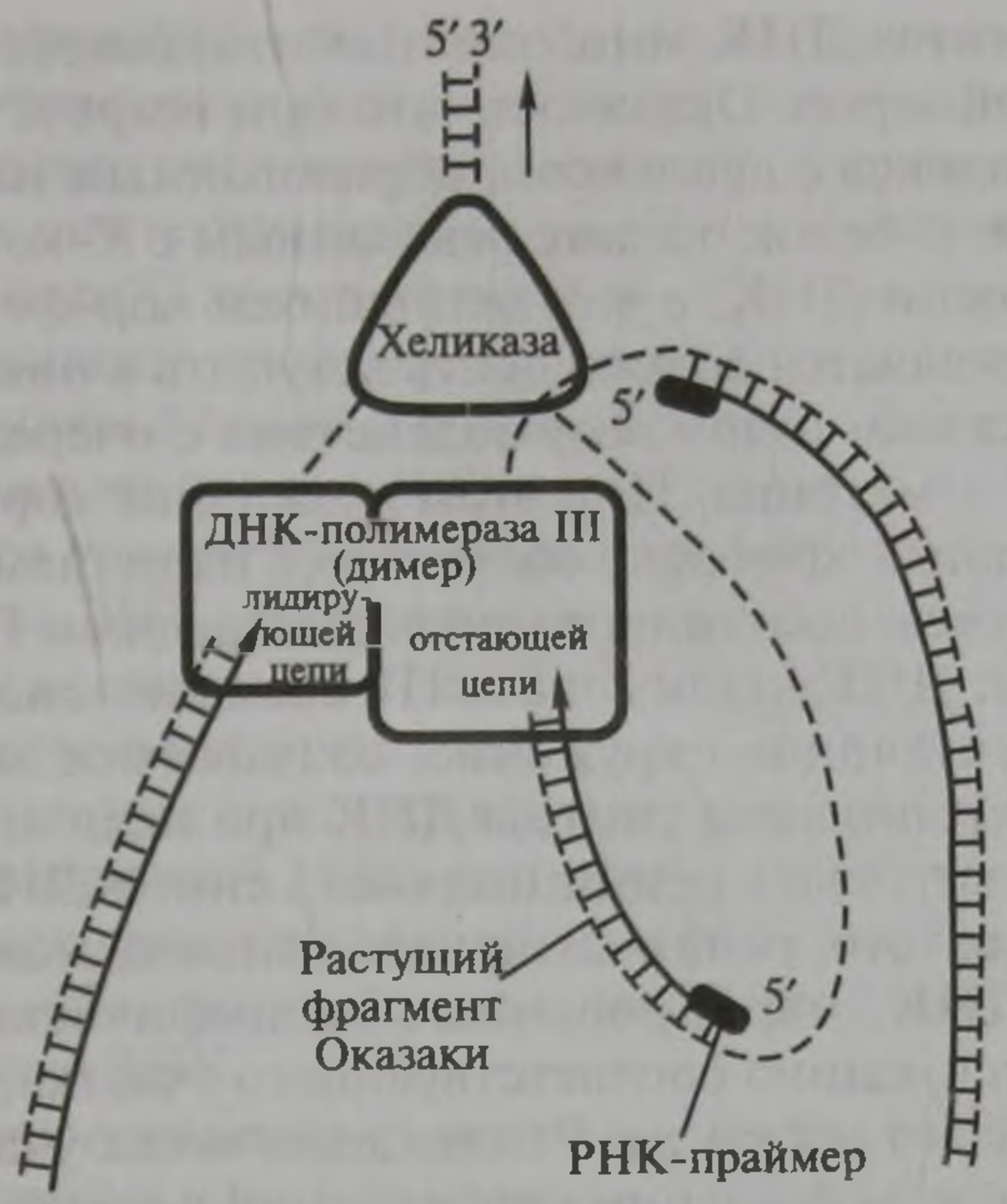


Рис. 113. Преодоление антипараллельности цепей при репликации за счет возникновения петли (модель А. Корнберга)

осуществляет как высокопроцессивный непрерывный синтез ведущей цепи ДНК, так и прерывистый синтез фрагментов Оказаки отстающей цепи, претерпевая во втором случае периодическую диссоциацию от матрицы для инициации синтеза ДНК с каждого нового праймера. Известно, что холофермент ДНК-полимеразы III включает кор-фермент, подвижно связывающий β-белок и белки γ-комплекса (см. табл. 18).

Роль γ-комплекса заключается в распознавании РНК-затравок на матричной ДНК. Он связывается с единственным праймером ведущей цепи ДНК или с каждым из праймеров фрагментов Оказаки отстающей цепи, что делает возможным присоединение к промаркированным таким образом праймерам кор-фермента и β-белка.

Две молекулы β-белка входят в состав репликативного комплекса вслед за белками γ-комплекса, связываясь с ДНК позади белков γ-комплекса и оставляя 3′-конец праймера доступным для ДНК-полимеразы. Димер β-белка образует кольцо вокруг молекулы ДНК и стимулирует АТРазную активность белков γ-комплекса. Оба вида белка, будучи связанными с дуплексом праймер-матрица, обеспечивают присоединение к этому комплексу кор-фермента ДНК-полимеразы III. Затем ДНК-полимераза III при наличии четырех дезоксирибонуклеозидтрифосфатов, используя праймер, с высокой эффективностью синтезирует ведущую цепь ДНК. Те же самые белки участвуют и в синтезе отстающей цепи ДНК. Но только в

этом случае синтез ДНК многократно инициируется на большом количестве праймеров. Оказалось, что при встрече репликативного белкового комплекса с дуплексом, образованным матричной ДНК и РНК-затравкой, β-белок остается связанным с 3′-концом вновь синтезированной цепи ДНК, а у отделившихся кор-фермента и белков γ-комплекса появляется возможность вступить в новый цикл синтеза фрагментов Оказаки путем взаимодействия с очередным дуплексом РНК-затравка — матрица. При этом вхождение кор-фермента в новый репликативный комплекс облегчается наличием в нем β-белка и белков γ-комплекса, ассоциированных с очередным РНК-праймером. Таким образом, ДНК-полимераза III обладает способностью распознавать молекулярное окружение, создаваемое матричной ДНК, осуществлять терминацию синтеза ДНК при наличии сигнала в виде дуплекса ДНК-затравка и реинициировать синтез ДНК на следующем праймере. Кроме того, репликативный комплекс может распознавать повреждения ДНК, маркированные специфическими белками, и прекращать репликацию соответствующего участка, останавливаясь или диссоциируя от матрицы. Репликация таких участков ДНК возобновляется после ликвидации повреждений ферментами репарации. Для диссоциировавшей ДНК-полимеразы это становится возможным благодаря тому, что 3′-конец вновь синтезированной цепи ДНК в месте остановки репликации остается связанным с β-белком, облегчающим повторное вхождение ДНК-полимеразы в репликативный комплекс.

Терминация репликации происходит тогда, когда встречаются две репликативные вилки при удвоении кольцевых молекул ДНК. Непрерывный рост лидирующей и отстающей цепей вдоль кольцевой матрицы неизбежно приводит к совмещению 3′-гидрокси- и 5′-фосфорильного концов одной цепи, либо в точке начала репликации (однонаправленная репликация), либо — при двунаправленной репликации — в середине кольца (см. рис. 109). Кольца в этих местах встречи соединяются ДНК-лигазой, при этом обычно они оказываются попарно сцепленными, т. е. образуют *катенан*. ДНК-гираза может расцепить зацепленные кольца, используя свою возможность вносить временный двуцепочечный разрыв (см. рис. 51).

Регуляция репликации. При каждом клеточном делении каждая молекула ДНК должна удваиваться, т. е. на каждом oriC должен происходить в точности один акт инициации репликации. В противном случае постепенно происходила бы утеря репликона или его бесконтрольное накопление. Таким образом, к регуляции репликации предъявляются довольно жесткие требования: регуляторная система должна чувствовать отклонения в обе стороны от среднего числа копий данного репликона и соответствующим образом менять частоту инициации на oriC. Очевидно, что частота инициации должна быть согласована также со скоростью роста клеток. Существуют тонкие механизмы регуляции репликации на уровне инициации новых раун-

дов. Позитивным регулятором этого процесса является белок Dna A. Среди моделей, объясняющих механизм регуляторного действия белка Dna A, наибольшее распространение получила модель титрования Dna A: весь вновь синтезируемый белок Dna A связывается (титруется) Dna A-боксами oriC хромосомы. Как только количество молекул инициатора превышает число внутриклеточных Dna A-боксов (все Dna A-боксы заняты белком), инициируется синтез ДНК. После запуска инициации на одном oriC освобождаются молекулы Dna A, резко повышаются его внутриклеточная концентрация и синхронная инициация синтеза ДНК на других доступных областях начала репликации. При этом ассоциация с мембранами первой oriC защищает ее от использования в реинициации.

Отмечена активация репликации транскрипцией, при которой перед транскрибирующей РНК-полимеразой имеет место локальная положительная сверхспирализация ДНК, а вслед за ней — отрицательная. Это облегчает образование открытых комплексов при инициации синтеза ДНК.

Сегрегация бактериальных репликонов по дочерним клеткам при делении. Стабильное поддержание любого репликона требует не только согласования его репликации с клеточным делением, но и упорядоченного распределения молекул ДНК по дочерним клеткам. Считается, что правильная сегрегация достигается у бактерий за счет прикрепления ДНК к мембране, причем пространственная организация участков прикрепления, зоны роста мембраны и клеточной стенки обеспечивает автоматическое «растаскивание» двух копий реплицировавшейся молекулы ДНК по разным дочерним клеткам.

Репликация хромосом у эукариот

Механизмы репликации ДНК у эукариот менее изучены из-за их большей сложности. Основные результаты получены на модельной системе с ДНК вируса SV40, в которой процесс репликации исследовали в зараженных клетках человека, культивируемых in vitro. В этой системе вирусный белок, называемый Т-антигеном, выполняет многие функции, необходимые для репликации вирусной ДНК. Он является белком-инициатором, обладает ДНК-хеликазной активностью и необходим для правильного взаимодействия с ДНК ферментного комплекса, синтезирующего праймеры. В то же время вирус SV40 использует для репликации своей небольшой хромосомы и многие белки клетки-хозяина, что позволяет исследовать функционирование репликативного комплекса клеток человека в такой относительно простой системе.

ДНК-полимеразы эукариот. В клетках эукариот имеются по меньшей мере шесть различных ДНК-зависимых ДНК-полимераз: α, β, δ, ε, γ, ζ. Четыре из них — α, β, δ, ε — непосредственно участвуют в репликации хромосомной ДНК (табл. 19).

Таблица 19. **Белки, входящие в состав репликативного комплекса эукариотических организмов**

Белок	Молекулярная масса субъединиц, кДа	Энзиматическая активность	Биологическая функция
ДНК-полимераза: α (альфа)	160 — 185* 67 — 70 54 — 58 46 — 50*	ДНК-полимераза Праймаза	Синтез праймеров на обеих цепях ДНК Транспорт ДНК-полимеразы в клеточное ядро Связывание пуринового нуклеотида, присоединение праймазы к ДНК-полимеразе
β (бета)	39 — 40*	ДНК-полимераза	Заполнение брешей при эксцизионной репарации ДНК, участие в рекомбинации, процессинг фрагментов Оказаки
γ (гамма)	125 — 140* 35 — 40	ДНК-полимераза 3′→5′-экзонуклеаза	Репликация и репарация митохондриальной ДНК
δ (дельта)	125 — 130* 48 — 55	ДНК-полимераза 3′→5′-экзонуклеаза	Синтез лидирующей цепи геномной ДНК в репликативной вилке Связь с фактором процессивности PCNA
ε (эпсилон)	230 — 261* 55	ДНК-полимераза 3′→5′-экзонуклеаза	Синтез отстающей цепи геномной ДНК в репликативной вилке
ξ (дзета)	173* 29	ДНК-полимераза 3′→5′-экзонуклеаза	Синтез ДНК на поврежденной матрице при SOS-ответе

Окончание табл. 19

Белок	Молекулярная масса субъединиц, кДа	Энзиматическая активность	Биологическая функция
Т-антиген		ДНК-хеликаза	Раскручивание двойной спирали ДНК в репликативной вилке
RPA (репликативный белок А)			Белок, связывающийся с одноцепочной ДНК, обеспечивает оптимальную для синтеза ДНК конформацию матриц
RFC (репликативный фактор С)		ДНК-зависимая АТРаза	Обеспечение связывания затравки с матрицей; стимулирование ДНК-полимеразы
PCNA (ядерный антиген пролиферирующих клеток)			Фактор процессивности репликации; выполняет функции «скользящего зажима»
Лигаза I		Лигаза	Лигирование фрагментов ДНК
FEN1 или MF1 (фактор созревания)		Экзонуклеаза	Удаление РНК-затравки
РНКаза Н1		Эндонуклеаза	Удаление РНК-затравки
Топоизомераза I		Топоизомераза	Релаксация как положительных, так и отрицательных сверхвитков ДНК
Топоизомераза II		Топоизомераза	Релаксация положительных сверхвитков

* Каталитическая субъединица.

ДНК-полимераза α — первая ДНК-полимераза, обнаруженная в клетках эукариот. Она представлена в клетке в виде прочного комплекса с ДНК-праймазой — ферментом, осуществляющим синтез РНК-затравок. Комплекс ДНК-полимераза α-праймаза является единственным у эукариот ферментативным ансамблем, способным инициировать синтез ДНК de novo. На богатых пиримидином участках матрицы праймаза комплекса синтезирует РНК-затравки длиной примерно 10 нуклеотидов, которые затем элонгируются ДНК-полимеразой без диссоциации комплекса от матрицы. В ходе репликации в клеточных ядрах ДНК-полимераза α-праймаза синтезирует затравку лидирующей нити в участке ori и затравки фрагментов Оказаки запаздывающей нити.

Как правило, ДНК-полимераза α не обладает корректорской 3′→5′-экзонуклеазной активностью. По-видимому, в ходе эволюции экзонуклеазный центр в данном ферменте редуцировался.

ДНК-полимераза β является наименьшей по размеру и самой простой по строению ДНК-полимеразой в клетках эукариот. Основная функция ДНК-полимеразы β в клетке связана с эксцизионной репарацией ядерной ДНК. Существуют экспериментальные указания на возможность участия этого фермента в рекомбинационных событиях и в репликации ДНК, в частности в процессинге фрагментов Оказаки. Так, ДНК-полимераза β способна заменить ДНК-полимеразу I в клетках *E. coli*, мутантных по гену ДНК-полимеразы I, инициируя репликацию с 3′-конца РНК/ДНК гибрида в районе ori.

ДНК-полимераза δ — гетеродимер, состоящий из каталитической субъединицы (125 — 130 кДа) и субъединицы 48 — 55 кДа, необходимой для преодоления ферментом структурных барьеров в природных однонитевых матрицах и для связи с фактором процессивности PCNA. Обе ферментативные активности — ДНК-полимеразная и 3′→5′-экзонуклеазная — связаны с высокомолекулярным полипептидом. Кроме фактора, в присутствии которого процессивность этого фермента увеличивается на два порядка, холофермент включает в себя факторы RFC и RPA. Три молекулы PCNA образуют кольцевой тример с отверстием для двунитевой ДНК в центральной части, который представляет собой перемещающуюся по ДНК подвижную платформу или «скользящую скрепку» в форме тора (бублика), удерживающую ДНК-полимеразу δ в ходе полимеризации на матрице и обеспечивающую высокопроцессивный синтез ДНК. Хотя PCNA и прокариотический фактор процессивности субъединица β ДНК-полимеразы III *E. coli* имеют низкую гомологию на уровне первичной структуры, оба белка формируют близкие по пространственной геометрии структуры «скользящей скрепки».

ДНК-полимераза ε, выделенная из клеток HeLa, содержит два полипептида — каталитический 261 кДа и полипептид 55 кДа. Каталитический полипептид обладает ДНК-полимеразной и 3′→5′-экзонуклеазной активностями. Особенностью холофермента ДНК-

полимеразы ε по сравнению с ДНК-полимеразой δ является его меньшая зависимость от вспомогательных факторов (PCNA, RFC и RPA), а также низкая (почти на порядок) скорость синтеза ДНК. Это различие, возможно, связано с разной функцией ДНК-полимераз в репликативной вилке. Один холофермент, ДНК-полимераза δ осуществляет быстрый и процессивный синтез лидирующей нити, используя для элонгации единственную затравку, синтезируемую ДНК-полимеразой α-праймазой в районе ori, и диссоциирует только по достижении конца репликона, тогда как несколько холоферментов ДНК-полимеразы ε могут одновременно синтезировать фрагмены Оказаки в «зоне Оказаки», удлиняя затравки, синтезируемые ДНК-полимеразой α-праймазой в начале каждого фрагмента.

Следует отметить, что у высших эукариот участие ДНК-полимеразы ε в синтезе запаздывающей цепи не является строго доказанным. Так, при репликации ДНК вируса SV40 в бесклеточной системе обе нити на этапе элонгации синтезирует ДНК-полимераза δ .

ДНК-полимераза γ локализована в митохондриях, ее функция связана с репликацией и репарацией митохондриальной ДНК, она кодируется ядерным геномом. ДНК-полимераза γ способна направлять высокопроцессивную полимеризацию на однонитевых ДНК-матрицах в отсутствие вспомогательных факторов.

Охарактеризованы также другие ДНК-полимеразы эукариот: η, θ, ι, REV1. Все эти ферменты участвуют в репарации ДНК.

В последние годы наряду с углубленным изучением строения и свойств отдельных ДНК-полимераз эукариот большое внимание уделяется взаимодействию этих ферментов со вспомогательными факторами и механизму функционирования их в составе многокомпонентных репликативных и репаративных комплексов. Список белков, взаимодействующих с ДНК-полимеразами, постоянно растет и включает не только известные факторы PCNA, RFC и RPA, но и ключевые факторы регуляции клеточного метаболизма, такие, как белки группы MCM (minichromosome maintenance factors), факторы узнавания участков ori репликации ORC (origin recognition complex) и др.

Принципиальное значение имеет установление того факта, что функционирование репликативных ДНК-полимераз α, δ, ε в ходе клеточного цикла регулируется с помощью специфических посттрансляционных модификаций, в частности фосфорилирования. Так, негативная регуляция ДНК-полимеразы α-праймазы в ходе клеточного цикла осуществляется с помощью фосфорилирования субъединиц с молекулярной массой 180 кДа и 68—70 кДа по остаткам серина. Одной из характеристик некоторых опухолевых клеток является нарушение фосфорилирования субъединицы 180 кДа в клеточном цикле.

В клетках эукариот синтез ДНК происходит в основном в специфических плотных структурах («репликативных фабриках»),

присоединенных к диффузному ядерному матриксу. В составе «репликативных фабрик», связанных с ядерным матриксом, обнаружен мультибелковый 21S комплекс («синтесома»), состоящий не менее чем из 30 белков (см. табл. 19).

Инициация репликации у эукариот. Скорость синтеза ДНК у эукариот существенно ниже, чем у прокариот (~50 нуклеотидов в 1 с у человека). Видимо, по этой причине инициация синтеза ДНК эукариот происходит во многих точках хромосомы, т. е. эукариотические хромосомы имеют *полирепликонную организацию*.

Инициация репликации у эукариот происходит на специфических множественных последовательностях нуклеотидов — репликаторах, или ориджинах репликации. Наиболее изучены репликаторы дрожжей, впервые идентифицированные как автономно реплицирующие последовательности (ARS — autonomously replicating sequences). У дрожжей на участке ДНК длиной 5 тыс. н.п. находится три отдельных репликатора (ARS). У млекопитающих области начала репликации располагаются на расстоянии ~100 тыс. н.п. друг от друга. Установлено, что синтез ДНК в отдельных репликонах происходит по двум направлениям, причем перемещение репликативной вилки осуществляется предпочтительно в одном направлении, которое может изменяться в зависимости от стадии развития организма. Частота использования отдельных репликаторов изменяется в онтогенезе, уменьшаясь в клетках взрослого организма. Репликаторы эукариот различны между собой по последовательности, но имеют несколько участков гомологии: ДНК-расплетающие элементы, участки прикрепления к ядерному матриксу, пиримидиновые тракты, канонические последовательности, взаимодействующие с белками репликативного комплекса, АТ-богатые последовательности, участки изгибов ДНК и другие, в большинстве своем служащие, вероятно, местами узнавания ответственных за инициацию репликации ядерных белков. Инициация репликации строго регулируется. Полирепликонная организация требует, чтобы в каждом цикле клеточного деления каждый репликатор «сработал» только один раз, в противном случае на хромосоме образуются разветвленные структуры. Для дрожжевых ARS-последовательностей инициация происходит один раз в каждой S-фазе клеточного цикла. У высших эукариот достаточно протяженные участки хромосом, содержащие несколько соседних репликонов, начинают синтез ДНК в S-фазе приблизительно одновременно, но каждый такой участок имеет характерное для него время инициации репликации: одни активируются в начале S-фазы, другие — позже, а третьи — в конце. Таким образом, существует запрет на повторную инициацию репликации в том же цикле клеточного деления, хотя одновременно может происходить инициация на других, не сработавших ранее репликаторах. Некоторые данные указывают на то, что подобный запрет снимается в ходе митоза, когда ядерная оболочка распадается. Нарушение запрета на повторную инициацию

приводит к амплификации (появлению множественных копий) того или иного участка хромосомы, где произошла «лишняя» инициация репликации.

Элонгация репликации. Четырехсубъединичная ДНК-полимераза α образует функциональные комплексы с рядом ферментов и вспомогательных белков. В комплексе, осуществляющем репликацию лидирующей цепи, ДНК-полимераза α связана с ДНК-полимеразой δ, а в комплексе, синтезирующем отстающую цепь, — с ДНК-полимеразой ε. У эукариот оба комплекса связаны, вероятно, друг с другом в репликативной вилке подобно тому, как связаны холоферменты синтеза лидирующей и отстающей цепей у прокариот.

Эукариотическая ДНК-праймаза в отличие от аналогичного белка прокариот образует постоянный комплекс с ДНК-полимеразой α, роль которого ограничивается синтезом праймеров при репликации обеих цепей ДНК.

Белок PCNA и фактор репликации С (RFC) также образуют стабильный комплекс с ДНК-полимеразой δ, а в определенных условиях стимулируют и активность ДНК-полимеразы ε. Во многих отношениях PCNA и RFC являются функциональными аналогами соответственно β-белка и белков γ-комплекса *E. coli*. Механизмы репликации ДНК про- и эукариот существенно различаются в том отношении, что во втором случае синтез лидирующей и отстающей цепей ДНК осуществляют разные ДНК-полимеразы (δ и ε соответственно), тогда как у *E. coli* обе цепи ДНК синтезируются димером ДНК-полимеразы III. Созревание фрагментов Оказаки у эукариот требует удаления РНК-затравок с помощью 5′→3′-экзонуклеазы (белковые факторы FEN1 или MF1) и РНКазы H1. ДНК-полимераза β заполняет бреши, равные вырезанным праймерам, а ДНК-лигаза I соединяет фрагменты отстающей цепи.

Терминация репликации. Продвижение репликативной вилки прекращается только при столкновении с другой вилкой, движущейся в противоположном направлении, или по достижении конца хромосомы. После сборки на молекуле ДНК хромосомных белков каждая пара хромосом в процессе митоза упорядоченно разделяется по дочерним клеткам. Репликация происходит в S-фазу клеточного цикла. Известно, что в регуляции клеточного цикла участвуют белки циклины (A, B, D, E). Они активируют циклинзависимые протеинкиназы, и те могут фосфорилировать специфические белки, участвующие в подготовке клетки к делению. Так, циклин D регулирует переход клетки из G_1-фазы в S-фазу; циклины A и E — активируют синтез ДНК на начальной стадии S-фазы; циклин B регулирует переход клетки из G_2-фазы в M-фазу.

Репликация теломерных участков эукариотических хромосом. На концах хромосом эукариот находятся специализированные повторяющиеся последовательности ДНК, получившие название *теломерной ДНК*, а содержащие ее концы хромосом — *теломера-*

ми. В клетках животных количество хромосом, а следовательно, и теломерных участков невелико — они составляют лишь небольшую часть от всех остальных последовательностей.

Использование в качестве объекта исследования теломерной ДНК ресничной инфузории *Tetrahymena thermophila*, в клетках которой находятся десятки тысяч мелких хромосом, а следовательно, и множество теломер, показало, что теломеры построены из коротких (содержат по 6—8 нуклеотидных остатков) многократно повторяющихся последовательностей (блоков). При этом одна цепь ДНК обогащена остатками гуаниловой кислоты (G-богатая цепь, у тетрахимены — это блок TTGGGG), а комплементарная ее цепь обогащена остатками цитидиловой кислоты (C-богатая цепь). Теломерная ДНК человека построена из TTAGGG-блоков, т.е. отличается от простейших всего лишь одним нуклеотидом в повторе. Из TTAGGG-блоков построены теломерные ДНК (их богатые G-цепи) всех млекопитающих, рептилий, амфибий, птиц и рыб. Универсален и теломерный повтор (TTTAGGG) у всех растений.

Характеристикой теломерных ДНК является их длина. В клетках жгутиковых размер теломер составляет ~50 н.п. У человека она колеблется от 2 до 20 тыс. н.п., а у некоторых видов мышей достигает сотни тысяч пар нуклеотидов. Отмечена значительная гетерогенность в разных клетках и тканях даже одного организма: длина ДНК в теломерах хромосом человека клеток зародышевой линии составляет 10—15 тыс. н.п., а в лейкоцитах периферической крови 5—12 тыс. н.п.

Теломеры играют важную роль в создании специфической архитектуры и внутренней упорядоченности клеточного ядра. Они предотвращают деградацию и слияние хромосом, а также ответственны за их прикрепление к специальной внутриклеточной структуре (своеобразному скелету клеточного ядра).

Механизмы репликации теломерных участков эукариотических хромосом и центральных областей ДНК принципиально различаются. Все известные ДНК-полимеразы, являющиеся ферментами сложного репликативного комплекса эукариот, неспособны полностью реплицировать концы линейных молекул ДНК. Известно, что ДНК-полимеразы, синтезируя дочернюю нить ДНК, прочитывают родительскую нить в направлении от ее 3′-конца к 5′-концу. Соответственно дочерняя цепь синтезируется в направлении 5′→3′. Кроме того, ДНК-полимераза начинает синтез только со специального РНК-праймера, комплементарного ДНК. После окончания синтеза ДНК РНК-праймеры удаляются, а пропуски в одной из дочерних цепей ДНК (отстающей) заполняются ДНК-полимеразой β. Однако на 3′-концевых участках ДНК такой пропуск заполнен быть не может, и поэтому они остаются однотяжевыми, а их 5′-концевые участки — недореплицированными (рис. 114). Отсюда ясно, что при каждом раунде репликации хромосомы будут укорачиваться на 10—20 нуклеотидов

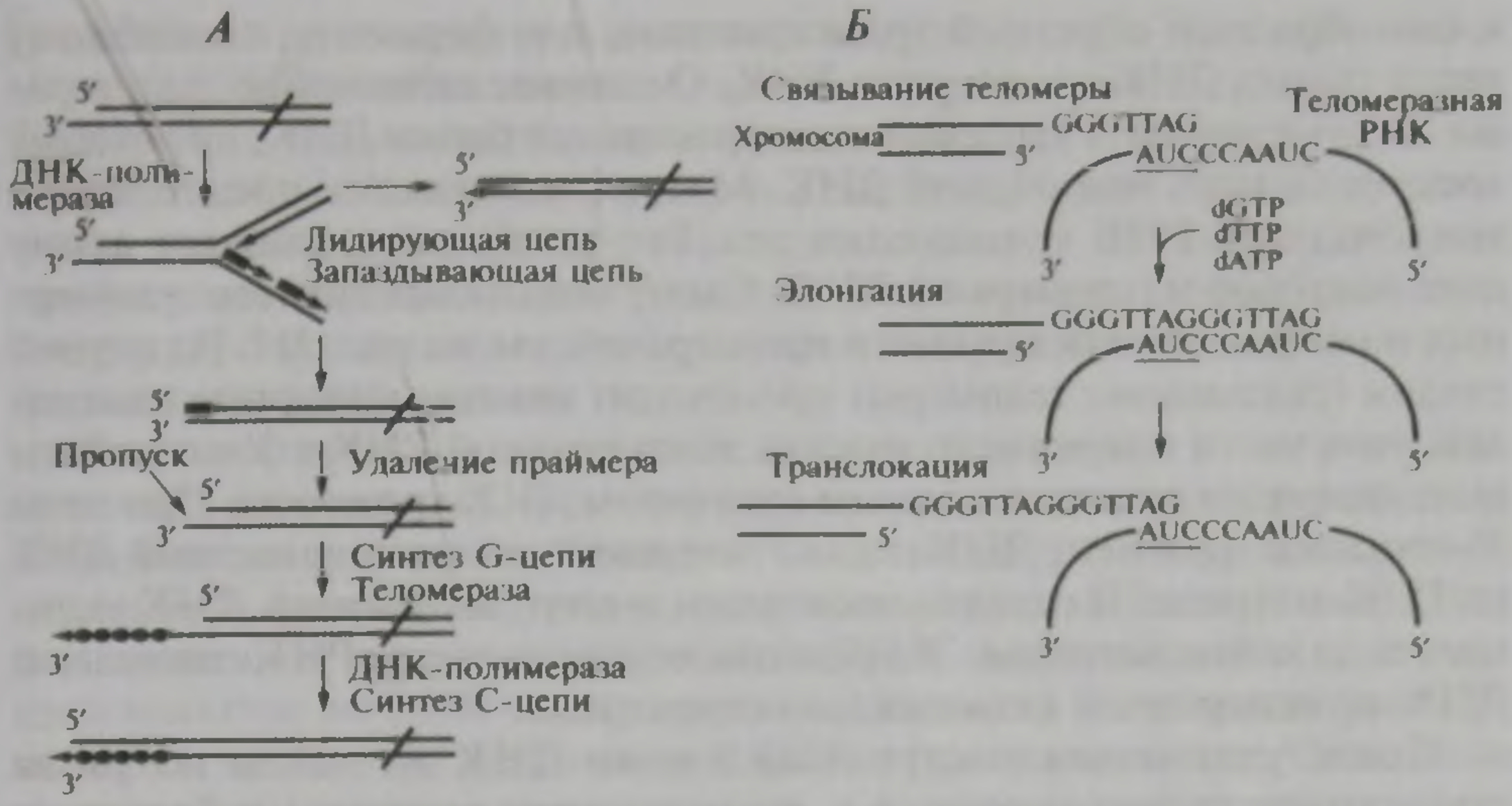


Рис. 114. Репликация теломерных участков эукариотических хромосом:

А — возникновение недореплицирования 5′-конца линейной хромосомы и синтез на этом концевом участке теломерной ДНК с помощью теломеразы; *Б* — основные этапы синтеза теломерного повтора теломеразой

(у разных видов размер РНК-затравок различен), и в первую очередь сокращать длину теломерной ДНК. Возникла проблема «концевой недорепликации ДНК». В случае репликации кольцевых бактериальных ДНК этой проблемы не существует, так как первые по времени образования РНК-праймеры удаляются ферментом, который одновременно заполняет образующуюся брешь путем наращивания 3′-ОН-конца растущей цепи ДНК, направленной в «хвост» удаляемому праймеру. Проблема недорепликации 3′-концов линейных молекул решается эукариотическими клетками с помощью специального фермента — *теломеразы*. Этот фермент был обнаружен впервые в 1985 г. у инфузории *Tetrahymena thermophila*, а впоследствии — в дрожжах, растениях и у животных, в том числе в яичниках человека и иммортализованных (бессмертных) линиях раковых клеток HeLa.

Теломераза является ДНК-полимеразой, достраивающей 3′-концы линейных молекул ДНК хромосом короткими (6—8 нуклеотидов) повторяющимися последовательностями (у позвоночных TTAGGG). Согласно номенклатуре, этот фермент называют ДНК-нуклеотидилэкзотрансферазой, или теломерной терминальной трансферазой (мол. масса 103—133 кДа). Помимо белковой части теломераза содержит РНК, выполняющую роль матрицы для наращивания ДНК повторами.

Длина теломерной РНК колеблется от 150 нуклеотидов — у простейших до 1400 нуклеотидов — у дрожжей, у человека — 450 нуклеотидов. Наличие в молекуле теломеразы РНК-последовательности, по которой идет матричный синтез фрагмента ДНК, позволяет отнести теломеразу

к своеобразной обратной транскриптазе, т.е. ферменту, способному вести синтез ДНК по матрице РНК. Основное назначение теломеразы — синтезировать тандемно повторяющиеся блоки ДНК, из которых состоит G-цепь теломерной ДНК. Матричный участок представлен в теломеразной РНК только один раз. Его длина не превышает длину двух повторов в теломерной ДНК. Схему механизма синтеза теломерных повторов, катализируемого теломеразой, см. на рис. 114. На первой стадии (связывание теломеры) происходит комплементарное взаимодействие части матричного участка теломеразной РНК с 3′-концевым выступающим одноцепочечным сегментом ДНК хромосом. При этом 3′-концевой фрагмент ДНК служит затравкой для удлинения этой ДНК на РНК-матрице. На стадии элонгации выступающая цепь ДНК удлиняется до конца матрицы. Эта реакция осуществляется РНК-зависимой ДНК-полимеразной активностью теломеразы.

После удлинения выступающей цепи ДНК до конца матрицы происходит транслокация, т.е. перемещение матрицы и белковых субъединиц фермента на заново синтезированный конец теломеразной ДНК, и весь цикл повторяется вновь. После завершения удлинения одноцепочечной 3′-концевой теломерной последовательности вторая цепь ДНК (С-цепь) достраивается с помощью обычной ДНК-полимеразы. Таким образом происходит решение проблемы концевой репликации ДНК у эукариот.

Предполагают, что в составе теломеразы кроме двух отмеченных компонентов — обратной транскриптазы и теломеразной РНК — должны быть субъединицы, отвечающие за поиск и связывание 3′-конца хромосомы, за транслокацию, за связывание продукта реакции (однотяжевой ДНК), а также субъединица с нуклеазной активностью, которая отщепляет от 3′-конца теломерной ДНК один за другим нуклеотиды до тех пор, пока на этом конце не окажется последовательность, комплементарная нужному участку матричного сегмента теломеразной РНК.

Наличие у животных тканеспецифичности в распределении теломер по размерам, а также изменение размеров этих последовательностей в онтогенезе предполагают существование механизмов, регулирующих данный процесс. Недавние исследования обнаружили резкое повышение активности теломераз, характерное для опухолевых клеток, что служит в настоящее время чувствительным физиологическим маркером их злокачественного перерождения. В связи с этим сегодня в качестве одного из подходов к терапии опухолей рассматривают подавление активности теломераз, функционирование которых необходимо для иммортализации клеток и роста опухолей.

Существование теломер у эукариот имеет глубокий смысл. Теломераза обнаружена в эмбриональных клетках, которые постоянно и быстро делятся, и непрерывное удлинение хромосом необходимо для компенсации их постоянного укорочения. В соматических клетках, которые делятся не часто, теломераза не обнаружена. Можно пред-

положить, что соматические клетки получают свою «порцию» теломерной ДНК, необходимую на время жизни клетки и ее потомства. Теломеры в большинстве клеток укорачиваются с возрастом, и это может являться важным фактором, определяющим продолжительность жизни особи. Таким образом, изучение механизма работы теломеразы и регуляции ее экспрессии в клетке помогут понять молекулярные основы процесса старения и злокачественной трансформации живой клетки.

Репликативный комплекс белков — реплисома — изучен достаточно хорошо у *E. coli*. Постоянно накапливающиеся данные о компонентах реплисомы эукариот показывают, что многие черты ее строения и функционирования у про- и эукариот общие. Инициация репликации и прерывистый синтез ДНК на отстающей цепи осуществляются по РНК-праймерному механизму и являются универсальным свойством репликации ДНК у про- и эукариот. Основные компоненты системы репликации в филогенезе функционально консервативны, и любой белковый компонент системы прокариот имеет свой прототип в системе репликации ДНК млекопитающих. Общей чертой является функционирование в репликативной вилке ДНК-топоизомераз, хеликаз, белков, связывающих и стабилизирующих односпиральную ДНК, высокополимерных ДНК-полимераз, праймаз, 3′→5′-экзонуклеаз, лигаз. Более удивительным представляется то, что у белков разных организмов, выполняющих одинаковые функции в репликации, в большинстве случаев отсутствует гомология в аминокислотных последовательностях. Так, не обнаружено сходства у белка SSB *E. coli* и белка RPA репликативной системы человека. То же самое характерно и для β-субъединицы ДНК-полимеразы III *E. coli* (β-белок) и белка PCNA человека. Это указывает на возможность выполнения одних и тех же функций полипептидными цепями с разными аминокислотными последовательностями, а также на вероятное конвергентное эволюционное происхождение таких белков и их функций из разных неродственных белков-предшественников. Репликация ДНК у прокариот и эукариот существенно различается. Для высших организмов характерны: более сложная организация инициации репликации (множественность репликонов у эукариот и однорепликонная организация у бактерий), строгий временной порядок включения репликонов, временна́я согласованность последующих актов лигирования ДНК по фланкирующим репликоны концам, сложная связь реплисом с матриксом ядра и, наконец, своеобразный механизм репликации теломерных участков эукариотических хромосом.

Именно репликация ДНК обеспечивает воспроизведение наследственной информации при образовании новых клеток. Примерно в то же время происходит удвоение количества и хромосомных белков. Следовательно, репликация ДНК — составная часть более сложного процесса — репликации хромосом.

Биосинтез ДНК на РНК-матрице (обратная транскрипция)

Биосинтез ДНК на РНК-матрице получил название обратной транскрипции, а фермент, катализирующий комплементарное копирование РНК с образованием ДНК, — обратной транскриптазой, ревертазой или РНК-зависимой ДНК-полимеразой. В про- и эукариотических клетках могут встречаться различные по происхождению РНК-зависимые ДНК-полимеразы, выполняющие разные биологические функции. Эти обратные транскриптазы объединяют в две группы: 1) гомогенную, представленную экзогенными ретровирусными энзимами, появляющимися в инфицированных эукариотических клетках; 2) гетерогенную, в состав которой входят энзимы, кодируемые ядерным (клеточные гены и мобильные генетические элементы) или митохондриальным (интроны) геномами клеток, митохондриальными плазмидами, эндогенными провирусами, интегрированными в ядерный геном.

РНК-зависимые ДНК-полимеразы были впервые идентифицированы в 1970 г. независимо друг от друга Д. Балтимором — в составе вируса Раушера лейкемии мышей и Г. Теминым и С. Мизутани — в составе вируса саркомы Рауса. Эти ретровирусные энзимы, появляющиеся в эукариотических клетках при репродукции ретровирусных частиц, широко исследованы в фундаментальном и прикладном аспектах. Очищенный фермент вируса саркомы Рауса (RSV), кодируемый геном pol, является мультифункциональным энзимом, с которым ассоциировано несколько энзиматических активностей: РНК-зависимая ДНК-полимераза, ДНК-зависимая ДНК-полимераза, РНКаза Н, эндонуклеаза (интеграза), полинуклеотид-расплетающая активность. Функция энзиматических активностей, кодируемых геном pol, связана с репликацией вирусов и интеграцией их ДНК-копий в геном клетки-хозяина.

Каждый вирион ретровирусов содержит две идентичные цепи РНК размером от 8 000 до 10 000 нуклеотидов. Области 5′- и 3′-концов обеих цепей модифицированы, как у всех эукариотических мРНК (5′-кэпы, 3′-полиаденилловые хвосты). Вирусные РНК имеют 5 структурных элементов: 1) прямые повторы на 5′- и 3′-концах РНК (R); 2) последовательность из 80 — 120 нуклеотидов, находящуюся около 5′-концевого повтора (U5); 3) последовательность из 170 — 1 200 нуклеотидов — около 3′-концевого повтора (U3); 4) последовательность из 15 — 20 нуклеотидов (P), в пределах которой клеточная тРНК комплементарно взаимодействует с ретровирусной РНК, что создает праймер для синтеза первой цепи ДНК; 5) сегмент Pu, находящийся непосредственно перед повтором U3 и являющийся сайтом для праймирования второй цепи ДНК, — такой сегмент одинаков у РНК всех ретровирусов определенного типа. Этапы обратной транскрипции представлены на рис. 115.

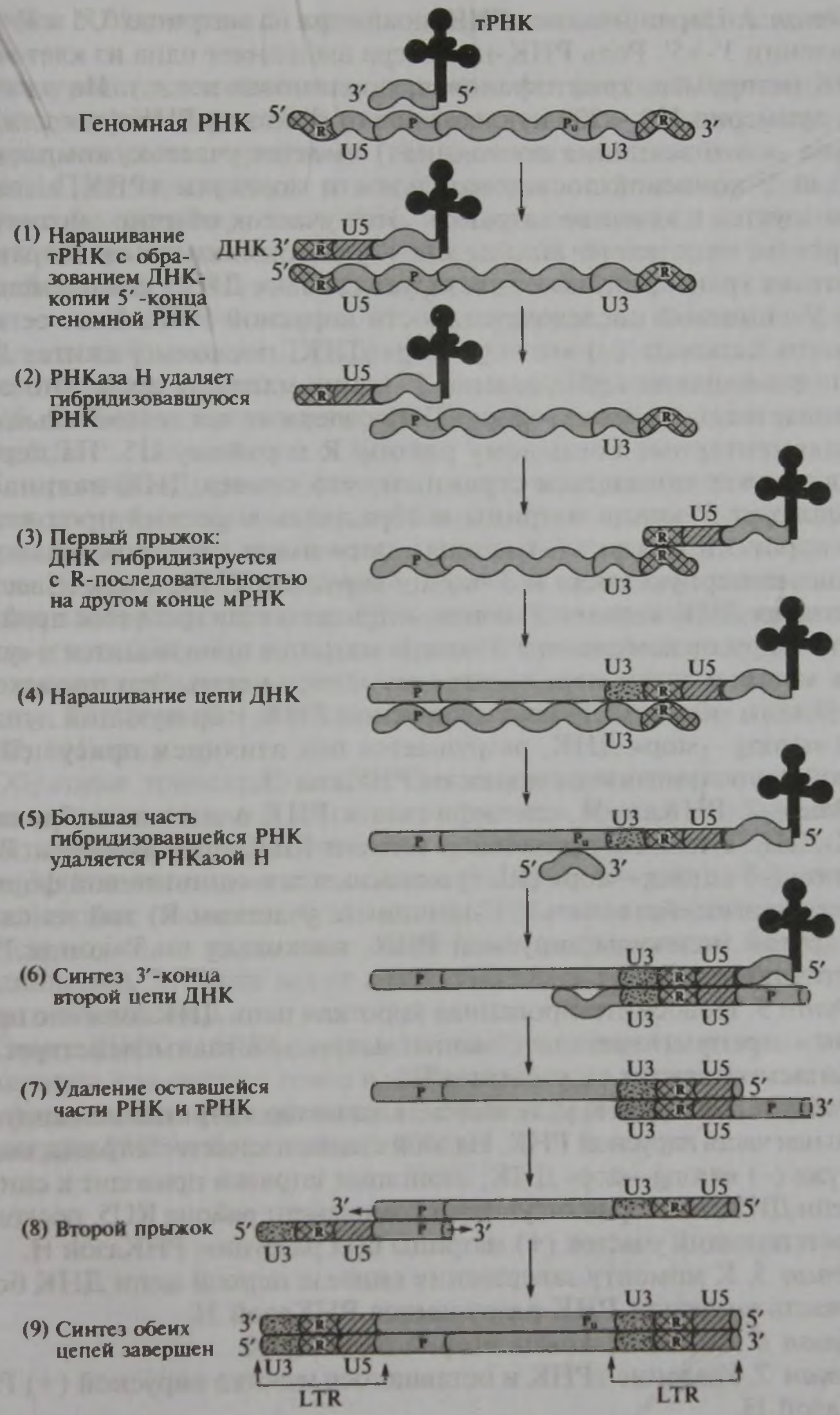


Рис. 115. Схема обратной транскрипции ретровирусной РНК с образованием линейной двуцепочечной ДНК (*Darnell J., et. al.* // Molecular Cell Biology. — N.Y.: Scientific Amer. Books, 1986. — P. 1052)

Этап 1. Наращивание тРНК-праймера на матрицах U5 и R в направлении 3′→5′. Роль РНК-праймера выполняет одна из клеточных тРНК (например, триптофановая, пролиновая и т.д.). На расстоянии примерно 100—200 нуклеотидов от 5′-конца РНК (для каждого вируса — это величина постоянная) имеется участок, комплементарный 3′-концевой последовательности молекулы тРНК, который используется в качестве затравки. Этот участок обычно обозначают как pbs (от англ. *primer binding site* — участок связывания затравки). Обратная транскриптаза синтезирует сегмент ДНК, комплементарный 5′-концевой последовательности вирусной РНК. Этот сегмент принято называть (–) «strong—stop» ДНК, поскольку синтез ДНК после завершения копирования 5′-конца матрицы временно останавливается. (–) «strong—stop» ДНК содержит последовательности, комплементарные концевому району R и району U5. На первый взгляд может показаться странным, что синтез ДНК начинается недалеко от 5′-конца матрицы и образуется короткий продукт. Но этот короткий продукт (–) «strong—stop» имеет последовательность, комплементарную также и 3′-концу вирусной РНК, а как известно, для снятия ДНК-копии с 3′-конца матрицы всегда требуется праймер. У ретровирусов комплемент 3′-конца матрицы производится в «удобном» месте, а затем переносится на «свое» место. Это происходит следующим образом: 5′-конец вирусной РНК, образующий дуплекс с (–) «strong—stop» ДНК, разрушается под влиянием присущей обратной транскриптазе активности РНКазы Н.

Этап 2. РНКаза Н, специфичная к РНК в составе гибридного РНК-ДНК-дуплекса, расщепляет сегмент РНК этого дуплекса. В результате (–) «strong—stop» (RU5) оказывается в однонитевой форме и может взаимодействовать с 3′-концом (с участком R) той же самой или другой молекулы вирусной РНК, поскольку на 3′-конце РНК имеется повтор R.

Этап 3. Новосинтезированная короткая цепь ДНК вместе с праймером «перепрыгивает» на 3′-конец матрицы и взаимодействует там с комплементарным ей участком R.

Этап 4. Цепь ДНК удлиняется, в качестве матрицы используется остальная часть вирусной РНК. На этой стадии в качестве затравки выступает уже (–) «strong—stop» ДНК; элонгация затравки приводит к синтезу (–) цепи ДНК, в которой отсутствует комплемент района RU5, поскольку соответствующий участок (+) матрицы был разрушен РНКазой Н.

Этап 5. К моменту завершения синтеза первой цепи ДНК большая часть вирусной РНК разрушается РНКазой Н.

Этап 6. Синтез 3′-конца второй цепи ДНК.

Этап 7. Удаление тРНК и оставшегося участка вирусной (+) РНК РНКазой Н.

Этап 8. Второй прыжок, в результате которого новосинтезированная вторая цепь ДНК комплементарно взаимодействует с тРНК-связывающей последовательностью первой цепи.

Этап 9. Удлинение 3′-концов каждой цепи, образование дуплекса ДНК.

Вся представленная последовательность реакций протекает без явного участия ферментов репликации клетки-хозяина (топоизомеразы, хеликазы, праймазы, ДНК-связывающего белка, лигазы и т.д.). При этом следует отметить, что молекулы вирусных ДНК длиннее молекул вирусных РНК, которые послужили матрицей для обратной транскрипции. Действительно, к 5′-концу (+) цепи вирусной ДНК добавилась последовательность U3, а к 3′-концу этой цепи — последовательность U5. В результате на концах молекулы вирус-специфической ДНК появился длинный (несколько сотен нуклеотидов) концевой повтор (ДКП или LTR), имеющий структуру U3U5. События, происходящие с двуцепочечной формой ретровирусной ДНК, описаны на с. 129.

Гетерогенная группа обратных транскриптаз малоизучена и представляет огромный интерес в области исследования структурной организации генома, эволюции организмов и регуляции экспрессии генов. Наиболее исследованной обратной транскриптазой из этой группы является теломераза. РНК включена в структуру самого фермента (см. с. 234). С помощью обратных транскриптаз гетерогенной группы синтезируются неактивные копии клеточных генов — псевдогены, LINE- и Alu-последовательности, ретропозоны с ДКП (см. с. 183, 184) и др.

Обратные транскриптазы in vivo индуцируются в определенные периоды жизни клетки и всего организма и их функционирование может быть резко ограничено во времени и/или некоторыми тканями/клетками многоклеточного организма. Разная субклеточная локализация последовательностей ДНК, кодирующих обратные транскриптазы и различные свойства этих ферментов, позволяют предполагать, что они могут контролировать различные локальные функции в клетке.

Синтез ДНК на РНК-матрице in vitro используется в генетической инженерии для синтеза генов и их фрагментов, а также синтеза кДНК для физического картирования геномов и расшифровки первичной структуры РНК и белков (см. с. 361).

Глава 8
ГЕНЕТИЧЕСКАЯ РЕКОМБИНАЦИЯ

Под *генетической рекомбинацией* понимают процессы, приводящие к перераспределению нуклеотидных последовательностей в геномах. Эти процессы лежат в основе генетической изменчивости, позволяющей организмам приспосабливаться к окружающей среде и обеспечивать длительное существование видов.

На возможность генетической рекомбинации как обмена участками гомологичных хромосом впервые указал в 1903 г. Т. Бовери. Доказательства обмена генами между хромосомами были впервые получены Т. Х. Морганом в 1919 г. В настоящее время генетическую рекомбинацию подразделяют на два больших класса: *общую* рекомбинацию и *сайт-специфическую* («незаконную») рекомбинацию.

Общая рекомбинация. В ходе общей рекомбинации обмен генетическим материалом происходит между гомологичными нуклеотидными последовательностями ДНК, чаще всего между двумя копиями одной и той же хромосомы. Самый известный пример такого рода рекомбинации — это обмен участками гомологичных хромосом в процессе *мейоза*. В профазе 1 мейоза происходит кроссинговер (перекрест), когда хромосомы обмениваются участками. Кроссинговер способен таким образом «перетасовывать» гены любой хромосомы. У прокариот, не обладающих мейозом, тем не менее также наблюдается общая рекомбинация, приуроченная к митотическому делению.

В процессе рекомбинации у *E. coli* принимают участие несколько белков, обеспечивающих поиск и рекомбинацию гомологичных последовательностей ДНК. Так, белок RecBCD присоединяется к двойной спирали ДНК и движется вдоль молекулы, осуществляя ее расплетание в поиске сайта рекомбинации (рис. 116). Этот белок действует как *хеликаза*, расплетая двойную спираль ДНК, и перемещается вдоль нее со скоростью 300 нуклеотидов в секунду пока не обнаружит консенсусную восьминуклеотидную последовательность в сайте рекомбинации. Затем белок RecBCD, действуя как нуклеаза, осуществляет разрыв в одной из цепей ДНК, в результате чего высвобождается небольшой одноцепочечный участок ДНК, который может провзаимодействовать с комплементарным участком другой молекулы ДНК. Для обеспечения комплементарного взаимодействия

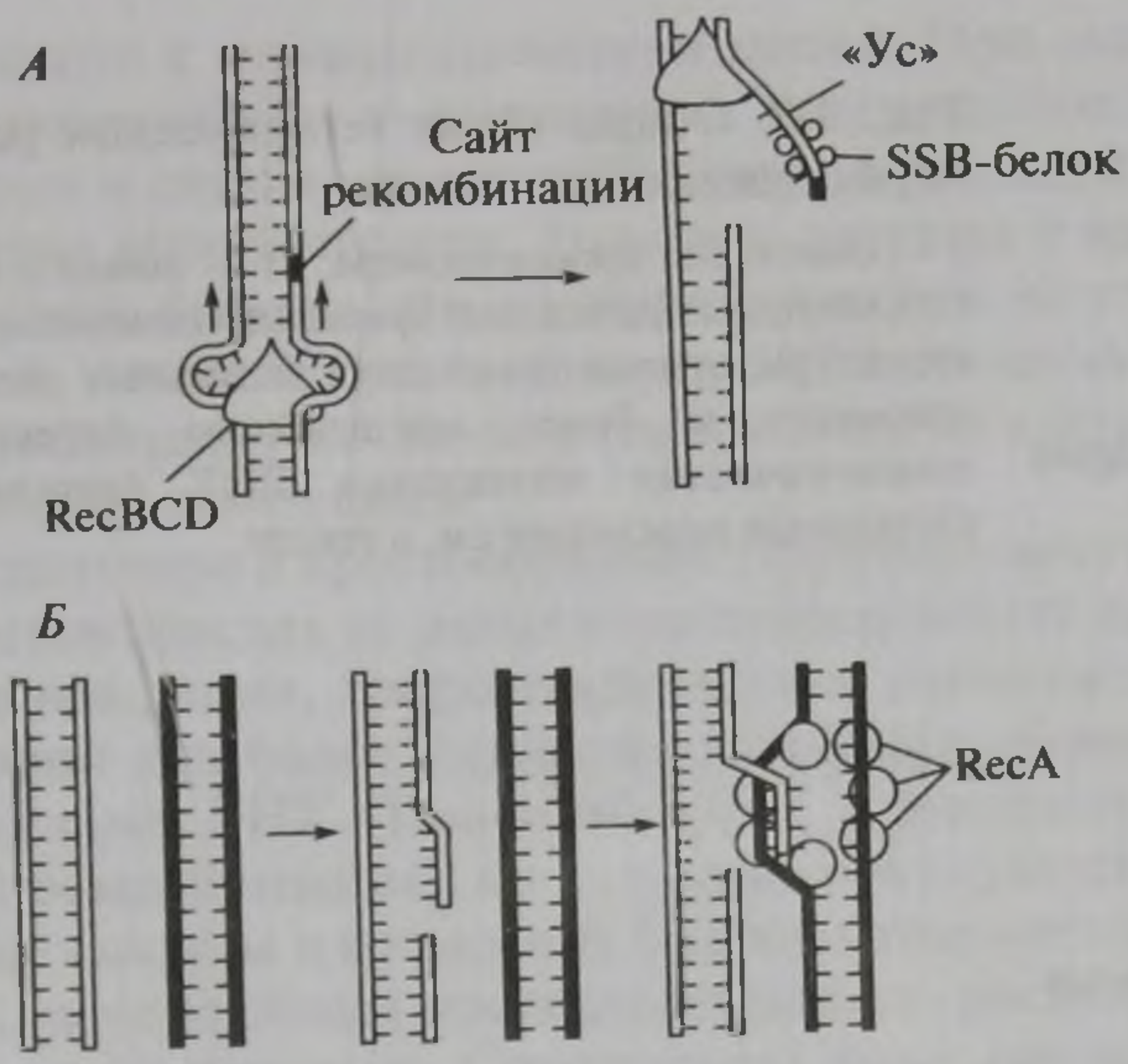


Рис. 116. Начальные этапы общей рекомбинации между двумя гомологичными двойными спиралями ДНК:

А — белок RecBCD движется вдоль молекулы ДНК, осуществляя ее расплетание, а затем разрыв одной из цепей ДНК; *Б* — после разрыва одной из цепей ДНК образовавшийся «ус» внедряется во вторую спираль, образуя короткий спаренный участок. Белок RecA удерживает вместе одиночную цепь и двойную спираль ДНК. Белок RecA и SSB-белок присоединяются к ДНК в виде групп молекул

высвободившийся участок одной из цепей ДНК, называемый «ус», поддерживается в вытянутом (линейном) состоянии с помощью SSB-белка. Этот белок участвует как в репликации бактериальных ДНК (см. гл. 7), так и в их рекомбинации, кооперативно взаимодействуя с сахарофосфатным остовом всех одноцепочечных участков ДНК и создавая таким образом условия для комплементарного взаимодействия азотистых оснований. В ходе рекомбинации SSB-белок действует в комплексе с белком RecA, который удерживает вместе одиночную цепь и двойную спираль ДНК (см. рис. 116). Небольшой по размеру белок RecA (мол. масса 38 кДа) имеет два различных участка для связывания взаимодействующих (спаривающихся) в ходе рекомбинации цепей ДНК. Белки RecBCD и RecA также обладают ДНК-зависимой АТРазной активностью, и интенсивный гидролиз ATP до ADP наблюдается в процессе присоединения этих белков к ДНК. Так же как SSB-белок, белок RecA присоединяется к взаимодействующим цепям ДНК в виде групп молекул (кластеров), и совокупное действие этих двух белков создает условие для эффективной рекомбинации. В точке рекомбинации возникает гетеродуплексное (ступенчатое) соединение молекул ДНК (рис. 117), которое может включать взаимодействие нескольких тысяч пар оснований. Затем происходит разрыв в другой цепи ДНК, создающий предпосылки

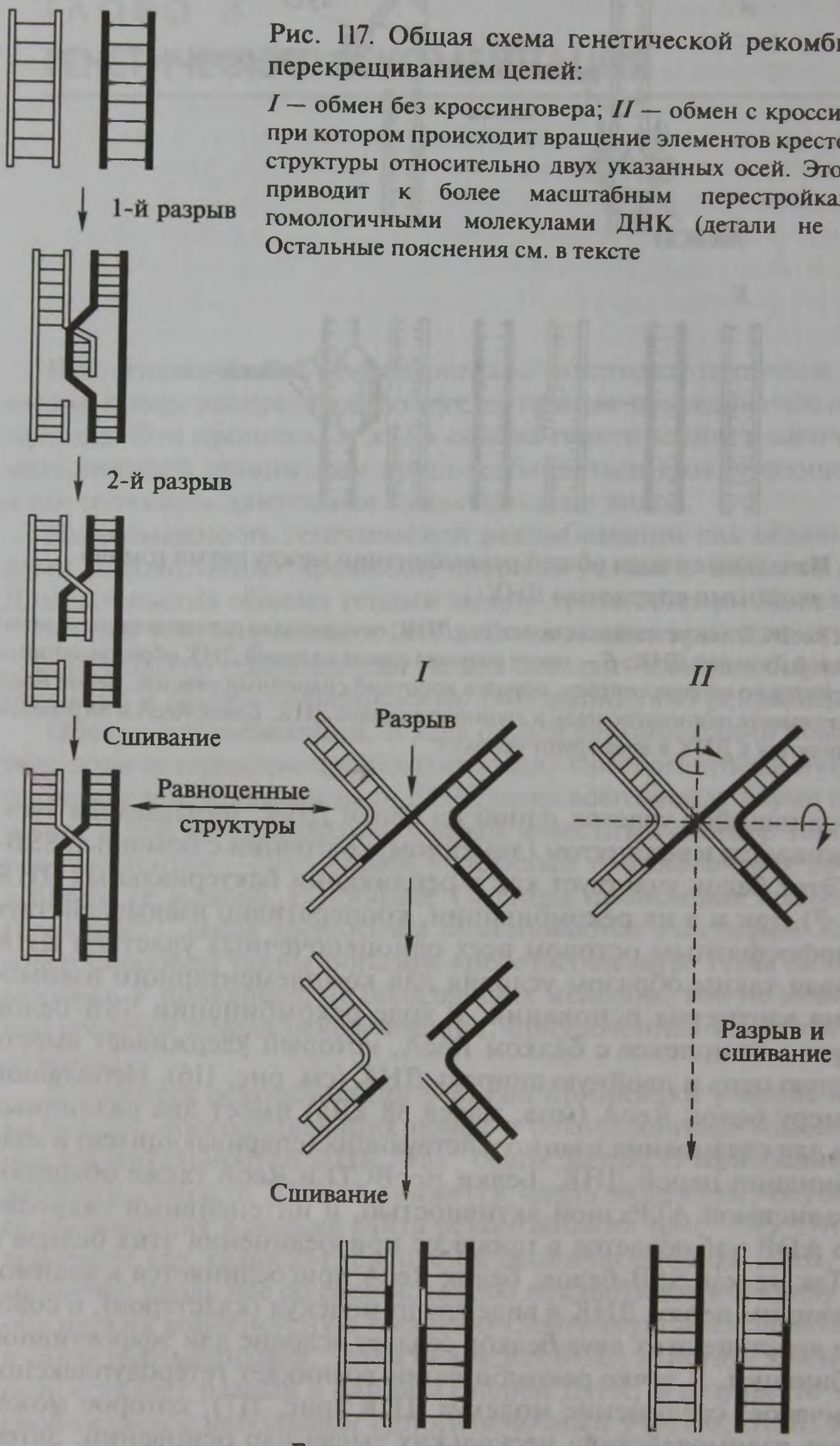


Рис. 117. Общая схема генетической рекомбинации с перекрещиванием цепей:

I — обмен без кроссинговера; *II* — обмен с кроссинговером, при котором происходит вращение элементов крестообразной структуры относительно двух указанных осей. Этот вариант приводит к более масштабным перестройкам между гомологичными молекулами ДНК (детали не указаны). Остальные пояснения см. в тексте

для рекомбинации с *перекрещиванием цепей*. При этом возникает характерная *крестообразная структура*, предложенная Р. Холлидеем в 1964 г., разрыв и сшивание которой завершает общую рекомбинацию в отсутствие кроссинговера. Точности разрыва и воссоединения оказываются столь высоки, что ни один нуклеотид не утрачивается и не добавляется, и вместе с тем молекулы ДНК обмениваются значительными по размерам участками, а содержащие их клетки обновляют свой генетический потенциал.

При кроссинговере в крестообразной структуре могут происходить вращения составляющих ее элементов относительно друг друга (см. рис. 117). Эти вращения, сопровождающиеся изомеризацией цепей, создают условия для более крупномасштабного обмена участками между молекулами ДНК. Таким образом, у эукариот в результате генетической рекомбинации с кроссинговером осуществляется перемещение из хромосомы в хромосому больших гомологичных последовательностей, однако общая последовательность расположения генов в хромосоме не нарушается. Совершенно иные последствия может вызвать сайт-специфическая рекомбинация.

Сайт-специфическая рекомбинация не требует гомологии ДНК. В этом случае в обмен вступают короткие специфические нуклеотидные последовательности ДНК. При этом изменяется распределение нуклеотидных последовательностей в геноме. Существует особый фермент рекомбинации, который «узнает» специфические последовательности в одной или двух рекомбинирующих молекулах ДНК, а комплементарного спаривания протяженных участков не происходит.

Первая исследованная система такой рекомбинации — это включение бактериофага λ в ДНК *E. coli*. Фермент, осуществляющий этот процесс, называют *лямбда-интегразой*. Данный фермент связывает между собой ДНК бактериофага и определенный участок ДНК бактерии, разрывая цепи обеих молекул, а после взаимодействия фаговой и бактериальной ДНК сшивает их (рис. 118). Таким образом ДНК фага оказывается интегрированной в геном клетки-хозяина.

Интеграза кодируется геномом бактериофага и может осуществлять обратный выход его ДНК из ДНК бактерии. Лямбда-интеграза представляет собой «*обратимую нуклеазу*»: этот фермент способен как нуклеаза разрывать фосфодиэфирные связи в молекулах ДНК, а затем снова сшивать их, действуя как лигаза. Таким образом, лямбда-интеграза действует аналогично *ДНК-топоизомеразе* (*гиразе*) бактерий, которая, как описано в гл. 3, вносит обратимые разрывы в цепи ДНК в процессе репликации.

Сайт-специфическая рекомбинация лежит в основе перемещения генов в пределах одной молекулы ДНК и объясняет, таким образом, феномен «прыгающих» генов (подвижных генетических элементов), свойственных как про-, так и эукариотам. Подвижные элементы (IS-элементы и транспозоны) прокариот и мобильные диспергиро-



Рис. 118. Сайт-специфическая рекомбинация, посредством которой кольцевая молекула ДНК фага λ внедряется в хромосому клетки бактерии (П-специфические последовательности, определяющие место внедрения фаговой ДНК)

ванные гены (ретротранспозоны) эукариот также представляют собой мощный элемент рекомбинации. Она не только создает условия для сохранения видов, но и является (по крайней мере, для эукариот) мощным фактором эволюции. Массовые перестройки в геноме могут, вероятно, приводить к «эволюционным взрывам» — скачкообразному возникновению новых видов и/или гибели ранее существующих организмов. Особенности структуры, механизмы перемещения подвижных генетических элементов и молекулярные последствия этих процессов рассмотрены в гл. 7.

В связи с тем что рекомбинация может затрагивать как кодирующие, так и регуляторные последовательности геномов, она может выступать как один из способов регуляции экспрессии генов. Данные перестройки наблюдаются, в частности, в повторяющихся последовательностях геномов многих организмов, которые могут находиться по отношению друг к другу либо в прямой, либо в обратной ориентации, и носят название *обращенных повторов* (см. гл. 5). Такого рода сайт-специфические инверсии участков хромосом наблюдаются у бактерий (*E. coli*, *Salmonella typhimurium* и др.) и бактериофагов (P1, P7). Они состоят в поворотах на 180° определенных сегментов хромосом, происходящих при участии особых ферментов — инвертаз, изменяют ориентацию генов относительно их промоторов и тем самым альтернативно влияют на их транскрипцию. Существуют определенные свидетельства того, что такие перестройки, несмотря на их случайный характер, имеют приспособительное значение для бактерий.

Перестройки в геномах могут носить запрограммированный характер на определенных стадиях дифференцировки клеток, что выявлено, в частности, при дифференцировке гетероцист у циано-

бактерий. В этом случае наблюдается удаление вставочных ДНК-элементов, которые встроены в кодирующие последовательности определенных генов. Эти элементы, несущие на концах короткие прямые повторы, удаляются на определенной стадии развития клеток специфическими ферментами — *рекомбиназами*, способными далее осуществлять объединение разделенных частей генов путем рекомбинации между их повторами.

К запрограммированным перестройкам относится и формирование генов иммуноглобулинов в клетках-предшественниках В-лимфоцитов позвоночных животных. Сборка этих генов осуществляется путем рекомбинации и совмещения разобщенных фрагментов, а затем (после объединения со специальным промотором) происходят транскрипция и биосинтез разнообразных молекул иммуноглобулинов в В-лимфоцитах. Сходные процессы имеют место и при дифференцировке Т-лимфоцитов. Во всех других соматических клетках животных фрагменты генов иммуноглобулинов остаются разобщенными и биосинтез иммуноглобулинов не происходит.

В заключение отметим, что генетическая рекомбинация включает в себя возможность рекомбинации РНК-геномов, свойственных многим вирусам. Механизмы рекомбинации РНК еще не достаточно изучены. У ряда вирусов, например вирусов гриппа, рекомбинация осуществляется как обмен целыми молекулами РНК, что обеспечивает высокую изменчивость вирусов и соответственно затрудняет борьбу с ними. РНК-содержащие бактериофаги (фаг Qβ и родственные ему) также способны к рекомбинации, приводящей к образованию новых молекул РНК, причем партнерами в такой рекомбинации могут выступать как природные, так и полученные искусственным путем (синтетические) молекулы РНК. Последнее обстоятельство используется при синтезе *антисмысловых РНК* и *рибозимов* с целью создания специфических препаратов для борьбы с вирусными инфекциями. Важные перспективы имеют исследования в области получения рекомбинантных РНК и для разработки систем внеклеточного синтеза белков (см. гл. 14).

Глава 9
ТРАНСКРИПЦИЯ

Транскрипцией называют биосинтез РНК на матрице ДНК. Это первая стадия реализации генетической информации, в процессе которой определенные участки нуклеотидной последовательности ДНК «переписываются» в комплементарные одноцепочечные молекулы РНК. В результате транскрипции образуются мРНК, кодирующие аминокислотные последовательности белков, а также тРНК, рРНК и другие виды РНК, выполняющие структурные, регуляторные и каталитические функции. В основе транскрипции лежит фундаментальный *принцип комплементарности* азотистых оснований полинуклеотидных цепей ДНК и РНК, а сам процесс осуществляется с участием соответствующих ферментов — РНК-полимераз, и большой группы белков — регуляторов транскрипции. Относительно детально изучены этапы и механизмы регуляции транскрипции прокариотических геномов, тогда как детали транскрипции у эукариот известны пока далеко не в полной мере. Вместе с тем в общей системе контроля реализации генетической информации контроль на уровне транскрипции представляет собой, несомненно, наиболее важный этап, ответственный за дифференциальную активность генов в онтогенезе любого организма, и сложность организации этого контроля, прогрессивно увеличивающаяся в эволюции, лишний раз подтверждает фундаментальное значение транскрипции в реализации программы жизни.

Поскольку ДНК у подавляющего большинства организмов представляет собой двуцепочечные молекулы, составленные из комплементарных (но не идентичных) цепей, транскрипции подвергается обычно одна из них, носящая название *матричной цепи* (матрицы):

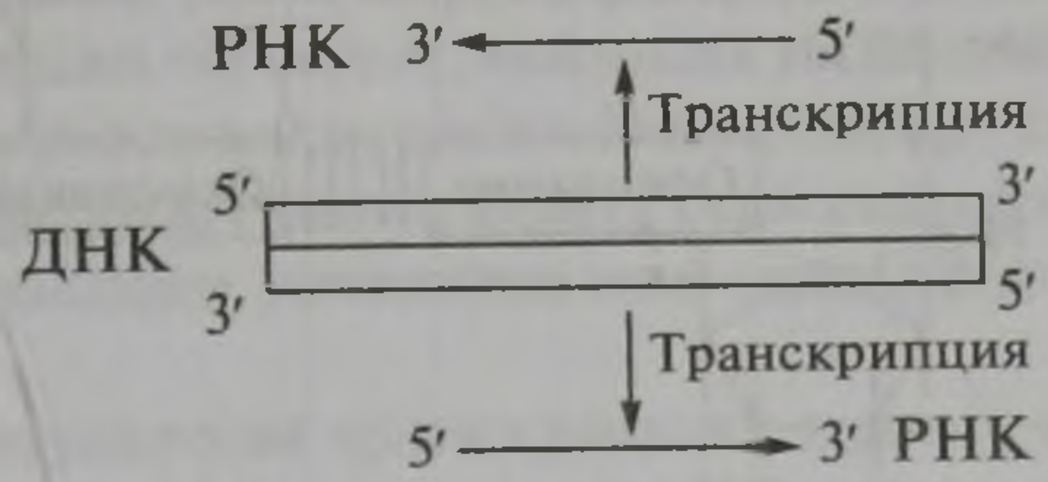
ДНК 5′-ATGATTGGGGCTCTA-3′ смысловая цепь
3′-TACTAACCCCGAGAT-5′ матричная цепь

↓ Транскрипция

РНК AUGAUUGGGGCUCUA

Различные РНК обычно транскрибируются с одной из цепей дуплексного генома, однако у некоторых геномов (как у бактериофага

T_4) матрицей могут служить разные цепи ДНК. В случае перекрывающихся генов матрицами для молекул РНК, транскрибируемых в противоположных направлениях, служат обе цепи ДНК в одной и той же области ее молекулы, как это имеет место у бактериофага λ:



Синтез РНК на матрице ДНК ведут РНК-полимеразы (ДНК-зависимые), которые, как и ДНК-полимеразы, относятся к группе нуклеотидилтрансфераз, использующих в качестве субстратов нуклеозид-5′-трифосфаты. Так же как ДНК-полимеразы, РНК-полимеразы активны только в присутствии ионов Mg^{2+}, для связывания которых в их молекуле существует консенсусная аминокислотная последовательность из 10 аминокислотных остатков. РНК-полимеразы в отличие от ДНК-полимераз не нуждаются в праймере и используют рибонуклеозид-5′-трифосфаты (ATP, GTP, CTP, UTP). Рост цепи РНК происходит путем последовательного присоединения рибонуклеозид-5′-монофосфатов к 3′-гидроксильной группе рибозы предшествующего нуклеотида (т.е. в направлении 5′ → 3′). Последовательность нуклеотидов в синтезируемом РНК-транскрипте определяется комплементарными взаимодействиями азотистых оснований, используемых для синтеза рибонуклеотидов, с азотистыми основаниями матричной цепи ДНК (рис. 119). Разные виды РНК у эукариот синтезируются различными РНК-полимеразами, тогда как у бактерий один фермент синтезирует все виды РНК.

У всех без исключения организмов процессу транскрипции подвергается единовременно не вся молекула ДНК, а только определенные ее участки — *транскриптоны*. Они ограничены двумя последовательностями, которые называются *промотором* (зона начала транскрипции) и *терминатором* (зона остановки транскрипции). Транскриптоны бактерий (их структура рассмотрена ниже) называют *оперонами*. Опероны обычно включают в себя нуклеотидные последовательности, кодирующие структуру нескольких белков, называемые *цистронами*, или *структурными генами*. Поэтому синтезируемая на оперонах бактерий мРНК является *полицистроновой* и может быть использована для синтеза нескольких белков, в отличие от *моноцистроновых* мРНК высших организмов, служащих матрицами для синтеза одного определенного белка.

На интенсивность транскрипции существенное влияние оказывает пространственная структура ДНК, в которой нередко наблюдаются изгибы, петли (домены), а также сверхспирализованные участки. Все эти мотивы структуры ДНК нередко способствуют усилению

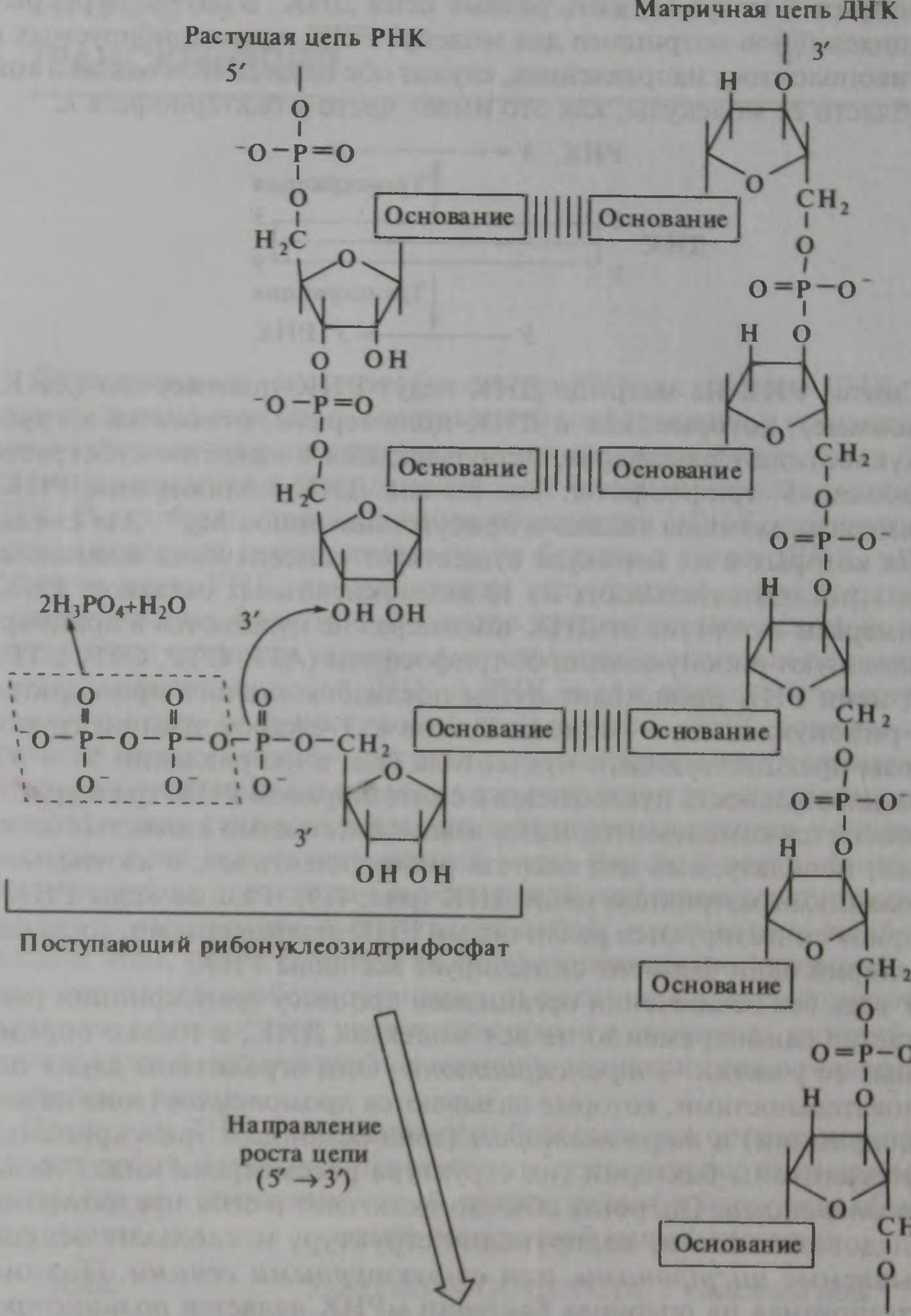


Рис. 119. Синтез РНК на матричной цепи ДНК.

Последовательность нуклеотидов в синтезируемой молекуле РНК определяется нуклеотидной последовательностью матричной цепи ДНК. В основе процесса лежит принцип комплементарности. Субстратами для синтеза служат рибонуклеозидтрифосфаты. К 3′-концу растущей цепи РНК присоединяется нуклеозидмонофосфат, а высвободившийся пирофосфат быстро гидролизуется до двух молекул неорганического фосфата, что делает реакцию энергетически выгодной и необратимой

транскрипции, поскольку создают дополнительные возможности для присоединения к ДНК разнообразных белков-регуляторов транскрипции. Эти белки способны связываться с определенными нуклеотидными последовательностями ДНК, расположенными как в самих транскриптонах, так и в других, нередко достаточно удаленных от зоны транскрипции областях, носящих название *энхансеров* (усилителей) и *адапторных элементов*. В последнем случае воздействие регуляторных элементов генома может опосредоваться особыми белками, способствующими усилению или ослаблению интенсивности транскрипции.

Сами белки-регуляторы транскрипции приобретают или утрачивают функциональную активность в результате модификаций, вызываемых либо каталитически активными белками-ферментами, либо возникающих при соединении с веществами небелковой природы, модулирующими их регуляторные свойства. У примитивных форм жизни в регуляции транскрипции непосредственно участвуют различные клеточные метаболиты (углеводы, аминокислоты, нуклеотиды и др.). У эукариот транскрипция непосредственно связана с изменением структуры *хроматина* — белково-нуклеинового комплекса, в котором постоянно пребывают их ДНК. Переход этого комплекса в активную (транскрибируемую) форму представляет собой дополнительный элемент регуляции транскрипции у высших организмов.

Транскрипция у прокариот

Бактериальные РНК-полимеразы — сложные белки, состоящие из нескольких субъединиц. Наиболее изученный из них фермент *E. coli* содержит пять постоянно присутствующих протомеров: две субъединицы α и по одной β, β′ и ω (конституитивные субъединицы, составляющие основу фермента, называемую *кором*). После присоединения к кору еще одной субъединицы — σ-фактора — образуется *холофермент РНК-полимераза*, способная узнавать промоторную область в оперонах бактерий и инициировать процесс транскрипции. Молекулярные массы α-, β-, β′- и ω-субъединиц составляют соответственно 36,5; 150; 155 и 70—80 кДа, а холофермента — 487 кДа. По нуклеотидной последовательности генов реконструирована полная аминокислотная последовательность каждой субъединицы РНК-полимеразы *E. coli*. Методом футпринтинга установлено, что при взаимодействии с ДНК фермент «закрывает» 60 пар нуклеотидов. β-Субъединица участвует в связывании рибонуклеозидтрифосфатов и построении полинуклеотидной цепи РНК, β′-субъединица осуществляет связывание ДНК, а комплекс α- и β′-субъединиц ($\alpha_2\beta'$) служит для специфического взаимодействия с *промоторами* — нуклеотидными последовательностями ДНК, детерминирующими инициацию

транскрипции. Функции ω-субъединицы не установлены. Узнавание промоторов происходит только после присоединения к кор-ферменту σ-фактора. После инициации синтеза РНК σ-субъединица отделяется от холофермента и дальнейшая транскрипция (элонгация цепи РНК) осуществляется кор-ферментом. Здесь следует заметить, что столь сложное структурно-функциональное устройство РНК-полимераз бактерий (и еще более сложная организация РНК-полимераз эукариот, в составе которых, по разным оценкам, содержится от 12 до 18 различных субъединиц) пока не имеет отчетливого объяснения, поскольку известно, что РНК-полимеразы некоторых бактериофагов вполне компетентны в синтезе РНК, несмотря на то что состоят всего-навсего из одной полипептидной цепи с молекулярной массой не более 100 кДа.

Синтез РНК требует локального расплетания двуцепочечной молекулы ДНК с тем, чтобы РНК-полимераза получила доступ к азотистым основаниям матрицы. Имеющиеся данные говорят о том, что РНК-полимераза расплетает (раскручивает) двуцепочечную молекулу ДНК с помощью особого ДНК-раскручивающего сайта, вступая во взаимодействие с матрицей благодаря наличию в ее структуре «цинковых пальцев» (см. рис. 130). Перемещаясь вдоль двойной спирали ДНК, РНК-полимераза непрерывно раскручивает спираль впереди того участка, где происходит синтез РНК и где на короткое время образуется так называемый *открытый комплекс*, внутри которого возникает РНК-ДНК-спираль длиной около 20 нуклеотидов. Затем фермент (вероятно, с помощью специального сайта) вновь закручивает ДНК позади участка полимеризации РНК, и готовая часть РНК-транскрипта выводится из комплекса через особый канал, свойственный РНК-полимеразе (рис. 120).

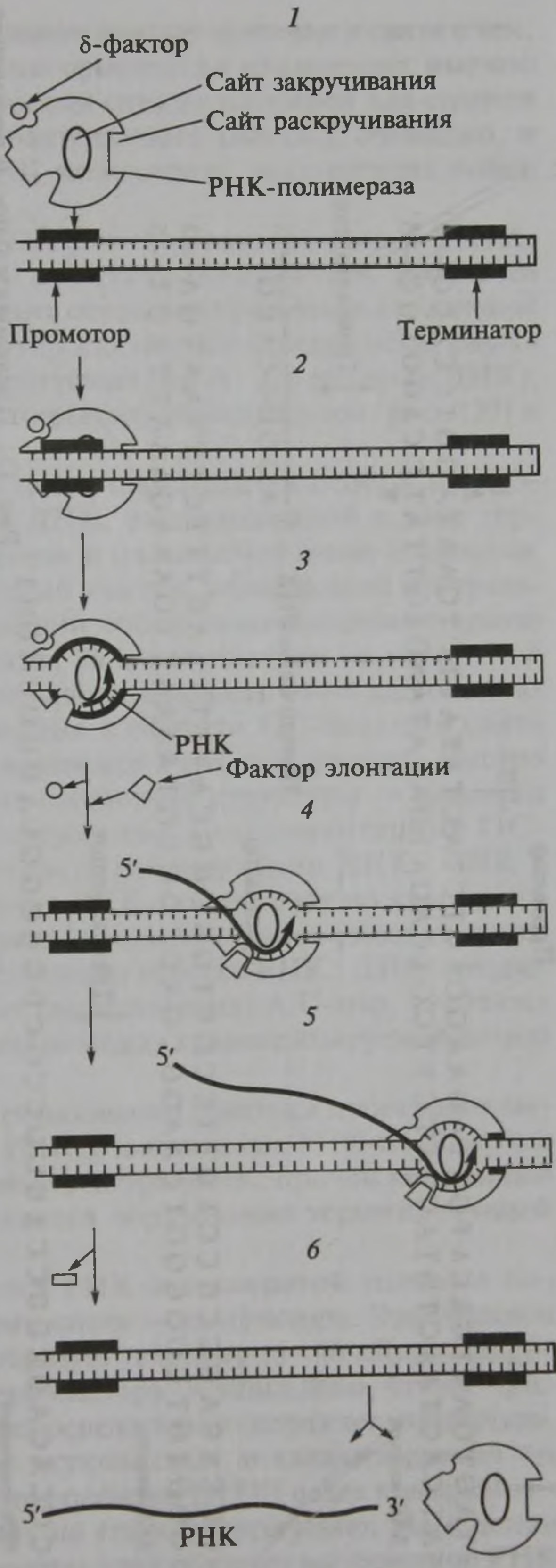
Скорость синтеза РНК у бактерий составляет ~30 нуклеотидов в 1 с, однако она не постоянна и может несколько снижаться. Такие периоды носят название *пауз транскрипции*, их причины не вполне ясны.

Для эффективного продолжения (элонгации) транскрипции изначально присоединившийся к РНК-полимеразе σ-фактор должен быть замещен фактором элонгации (одним из этих факторов является белок NusA *E. coli*).

Инициация транскрипции происходит в зоне промотора — особой нуклеотидной последовательности ДНК, которая содержит сигнал начала транскрипции (старт-сигнал). В области старт-сигнала располагаются две *консенсусные последовательности*, выявленные у многих сильных промоторов *E. coli*: последовательность 5′-ТАТААТ-3′ (расположена на расстоянии 10 нуклеотидов от точки начала транскрипции), последовательность 5′-TTGACA-3′ (удалена от точки начала транскрипции на 35 нуклеотидов). Обе эти последовательности выведены на основе анализа нескольких десятков различных промоторов и их принято обозначать на *нематричной* (*смысловой цепи*)

Рис. 120. Схема транскрипции, катализируемой РНК-полимеразой *E. coli*:

1 — осуществление РНК-полимеразой после присоединения σ-фактора (фактор инициации) поиска промотора; *2* — образование закрытого промоторного комплекса: РНК-полимераза связывается с зоной промотора ДНК; *3* — образование открытого комплекса: РНК-полимераза раскручивает двойную спираль ДНК, получает доступ к азотистым основаниям матричной цепи и начинает синтез РНК (стадия инициации); *4* — отсоединение σ-фактора и присоединение фактора элонгации создают условия для элонгации транскрипции: РНК-полимераза быстро продвигается по ДНК, раскручивая ее двойную спираль, синтезирует РНК и вновь закручивает спираль ДНК позади места синтеза РНК; *5* — завершение (терминация) синтеза РНК: РНК-полимераза достигает зоны терминатора, где синтез РНК прекращается (см. рис. 121); *6* — высвобождение готовой РНК-транскрипта в виде копии матричной цепи ДНК



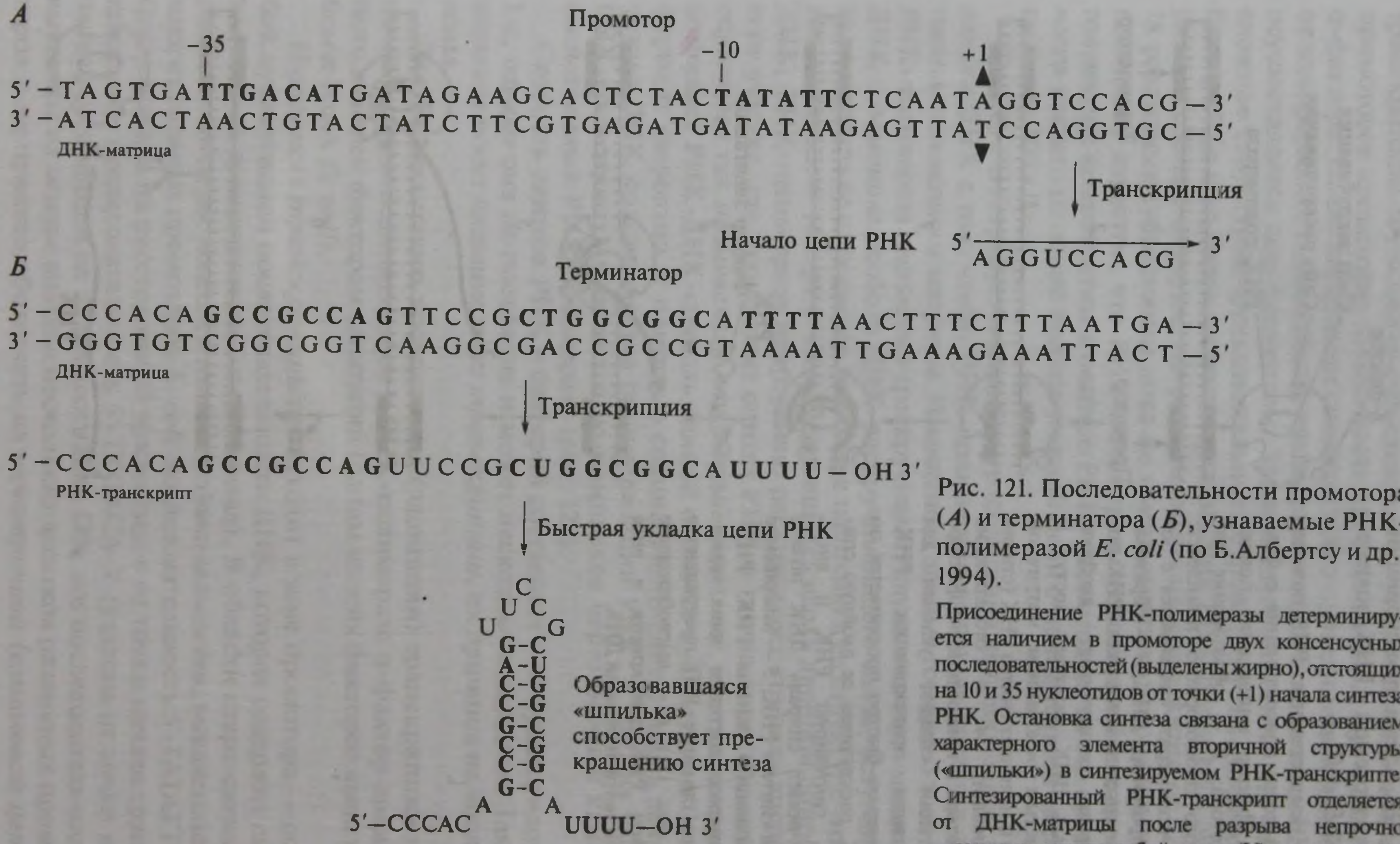


Рис. 121. Последовательности промотора (*А*) и терминатора (*Б*), узнаваемые РНК-полимеразой *E. coli* (по Б.Албертсу и др., 1994).

Присоединение РНК-полимеразы детерминируется наличием в промоторе двух консенсусных последовательностей (выделены жирно), отстоящих на 10 и 35 нуклеотидов от точки (+1) начала синтеза РНК. Остановка синтеза связана с образованием характерного элемента вторичной структуры («шпильки») в синтезируемом РНК-транскрипте. Синтезированный РНК-транскрипт отделяется от ДНК-матрицы после разрыва непрочно связанных между собой олиго(U)- и олиго(A)-последовательностей в гибриде РНК : ДНК

ДНК. Последнее обстоятельство имеет важное значение в связи с тем, что для транскрипции существенна ориентация *промотора*: именно она определяет ту цепь ДНК, которая служит матрицей для синтеза РНК. На консенсусные последовательности (боксы), очевидно, и ориентируется холофермент РНК-полимераза, осуществляя поиск точки начала транскрипции.

Последовательность 5′-ТАТАТТ-3′ называется также *ТАТА-боксом*, или *Прибнов-боксом* — в честь ее открывателя. В области ТАТА-бокса, очевидно, происходит открытие промотора (локальное плавление относительно непрочно связанных между собой *двумя водородными* связями комплементарных пар А : Т в дуплексе ДНК), что и приводит к образованию открытого комплекса (см. рис. 120) и инициации транскрипции.

Терминация транскрипции также определяется особой нуклеотидной последовательностью в ДНК, расположенной в зоне терминатора бактериальных оперонов и называемой *стоп-сигналом*. В стоп-сигнале имеется GC-богатый участок, обладающий центральной симметрией и представляющий собой самокомплементарную нуклеотидную последовательность. За этим участком на матричной цепи расположена олиго(А)-последовательность (4—8 адениловых нуклеотида подряд). Транскрипция в области GC-богатого сайта приводит к тому, что в образовавшемся РНК-транскрипте быстро образуется устойчивый элемент вторичной структуры — *шпилька* (спирализованная область, содержащая комплементарные GC-пары), которая нарушает прочность гибридизации ДНК : РНК в открытом комплексе и вытесняет РНК-полимеразу из комплекса (рис. 121). Кроме этого транскрипция олиго(А)-последовательности в матричной цепи ведет к образованию участка РНК : ДНК дуплекса, составленного из непрочных (легкоплавких) A:U-пар, что также способствует разрушению контакта между транскрибируемой цепью ДНК и молекулой РНК.

Таким образом, сразу после образования шпильки происходит высвобождение синтезированной РНК, отсоединение РНК-полимеразы и восстановление двуспиральной структуры ДНК, причем решающим моментом в этом процессе является образование терминирующей шпильки в составе РНК.

Часть терминаторов узнается РНК-полимеразой только с помощью белкового фактора терминации — ро-фактора. Этот белок с молекулярной массой 46 кДа образует гексамер (6 · 46 кДа), способный связываться с областью терминатора. Установлено, что он присоединяется к РНК-продукту непосредственно перед терминатором. Ро-фактор обладает NTPазной активностью, и катализируемая им реакция гидролиза нуклеозидтрифосфатов (NTP) ведет к высвобождению энергии, которая влияет на его конформацию. Изменение конформации ро-фактора, вероятно, способствует вытеснению РНК или РНК-полимеразы из транскрипционного комплекса.

Регуляция транскрипции

Вопрос о регуляции транскрипции имеет важное значение для понимания механизмов *дифференциальной активности* (*экспрессии*) генов в онтогенезе любого про- или эукариотического организма. Путем регуляции генной экспрессии клетки приспосабливают свой фенотип к изменяющимся условиям внешней и внутренней среды. Фенотип клеток в свою очередь определяется количеством и свойствами продуцируемых ими структурных, каталитических, регуляторных или других классов белков. Иными словами, должна существовать эффективная *система контроля реализации генетической информации* в клетке, позволяющая в каждый момент ее существования синтезировать и поддерживать в функционально активном состоянии необходимое и достаточное количество разных белков. У бактерий набор белков регулируется в основном наличием мРНК, пригодных для биосинтеза белка (трансляции), регуляцией самого процесса трансляции в рибосомальном аппарате клетки, а также регуляцией процессов деградации (протеолиза) белков. У эукариот белковый пул в клетке может дополнительно контролироваться с помощью процессов созревания (процессинга) первичных транскриптов и транспорта зрелых РНК из ядра в цитоплазму. Дополнительные возможности контроля клетки эукариот приобрели в результате разобщения в пространстве и времени процессов транскрипции (происходит в ядрах) и трансляции (происходит в цитоплазме). У прокариот транскрипция и трансляция происходят практически единовременно в едином неструктурированном клеточном содержимом: синтез РНК на ДНК-матрице еще продолжается, а уже начинается трансляция содержащейся в ней генетической информации в структуру белков с участием рибосом, тРНК и белковых факторов трансляции (рис. 122).

В любом случае первым и главным элементом контроля реализации потока генетической информации на пути от ДНК к белку у всех организмов является контроль на уровне транскрипции.

Регуляция транскрипции у прокариот

Транскрипция у прокариот регулируется преимущественно на стадии инициации и связана с деятельностью регуляторных белков — *активаторов* и *репрессоров транскрипции*. В ходе элонгации транскрипции важное значение приобретает вторичная структура синтезированной РНК: от нее во многом зависит, будут ли синтезированы полноразмерные молекулы, либо процесс остановится на синтезе коротких функционально непригодных транскриптов. Регуляция транскрипции у бактерий обычно охватывает группу (кластер) генов, кодирующих функционально родственные белки, участвующие

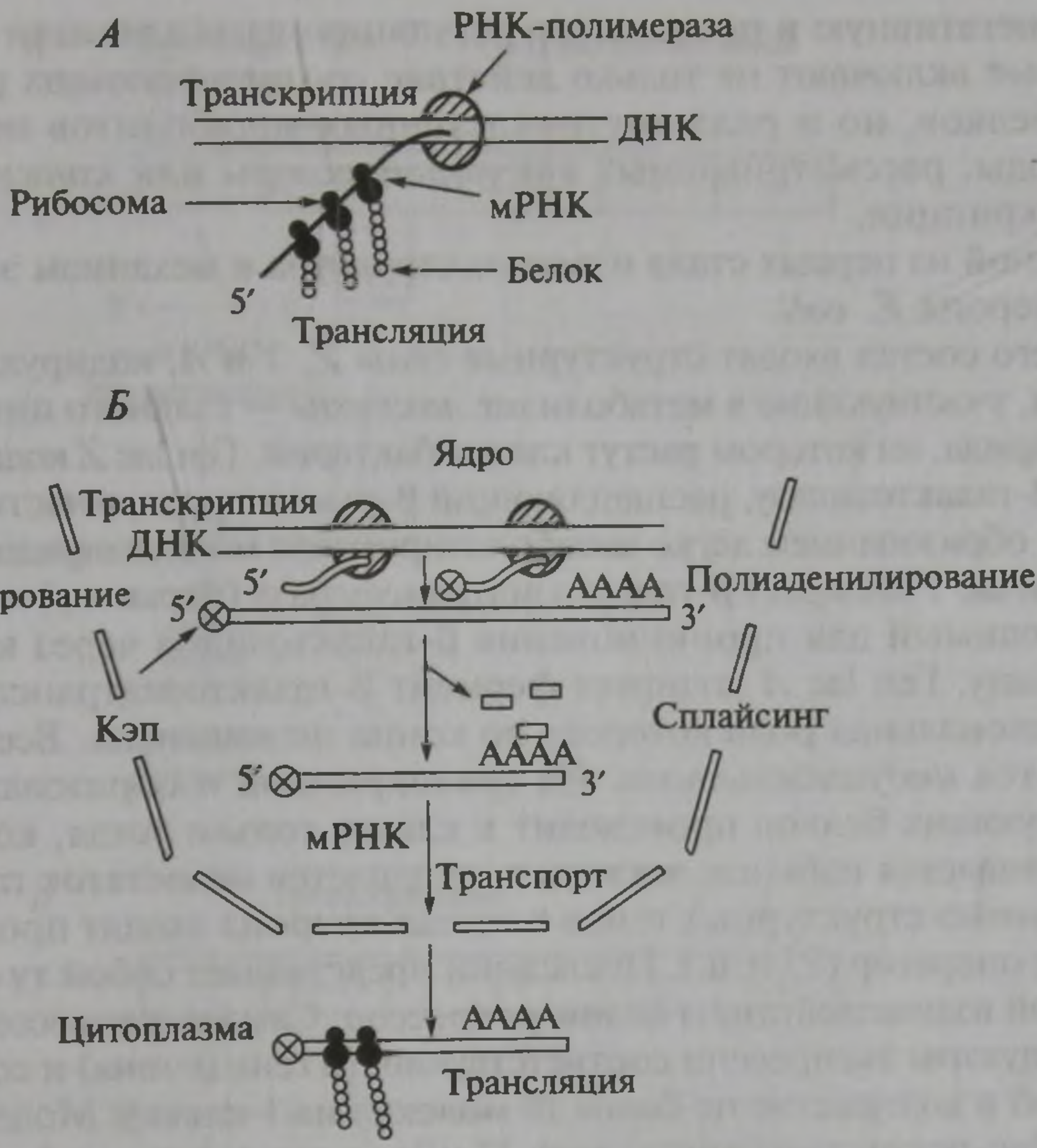


Рис. 122. Различия в транскрипции и трансляции у прокариот (*А*) и эукариот (*Б*).

Транскрипция и трансляция у эукариот разделены во времени и пространстве: транскрипция происходит в ядре, а трансляция (биосинтез белка) — в цитоплазме. Между транскрипцией ДНК и трансляцией РНК происходит процессинг и транспорт РНК. Транскрипция и трансляция у прокариот совершаются одновременно

в осуществлении тесно связанных между собой химических превращений в клетке. Такими белками обычно являются ферменты, занятые осуществлением серии реакций (метаболическим путем), ведущих к распаду или синтезу тех или иных соединений. Группа согласованно регулируемых генов, кодирующих эти ферменты, называется *опероном*. К наиболее хорошо изученным оперонам бактерий относятся лактозный (lac)-оперон, галактозный (gal)-оперон и триптофановый (trp)-оперон *E. coli*. В состав каждого из этих оперонов входят несколько структурных генов, кодирующих ферменты, а также регуляторные области — промотор (место присоединения РНК-полимеразы) и *оператор* — участок оперона, к которому присоединяются белки-репрессоры и/или активаторы транскрипции. Зоны промотора и оператора часто перекрываются, и это объясняет, как белки-регуляторы влияют на инициацию транскрипции. Разли-

чают негативную и позитивную регуляцию транскрипции оперонов, которые включают не только действие соответствующих регуляторных белков, но и ряда внутриклеточных метаболитов небелковой природы, рассматриваемых как *корепрессоры* или *коактиваторы* транскрипции.

Одной из первых стала известна структура и механизм экспрессии lac-оперона *E. coli*.

В его состав входят структурные гены *Z*, *Y* и *A*, кодирующие ферменты, участвующие в метаболизме *лактозы* — главного питательного дисахарида, на котором растут клетки бактерий. Ген lac *Z* кодирует фермент β-галактозидазу, расщепляющий β-галактозиды, в частности лактозу, с образованием легко метаболизируемого моносахарида — глюкозы. Ген lac *Y* кодирует β-галактозидтрансферазу (пермеазу) — фермент, необходимый для проникновения β-галактозидов через клеточную мембрану. Ген lac *A* кодирует фермент β-галактозидтрансацетилазу, функциональная роль которого до конца не выяснена. Все эти гены являются *индуцибельными*. Их транскрипция и образование соответствующих белков происходит в клетке только тогда, когда в ней накапливается избыток лактозы и ощущается недостаток глюкозы.

Помимо структурных генов в состав оперона входят промотор (40 н. п.) и оператор (27 н. п.). Последний представляет собой ту область, с которой взаимодействует белок-репрессор. Сам lac-репрессор является продуктом экспрессии соответствующего гена (i-гена) и содержится у *E. coli* в количестве не более 10 молекул на 1 клетку. Молекулярная масса lac-репрессора составляет 37 кДа, а его активная форма представляет собой гомотетрамер, который связывается с зоной оператора и блокирует действие РНК-полимеразы (рис. 123). Такая *негативная регуляция* транскрипции наблюдается при достаточной концентрации глюкозы в клетке и в отсутствие избытка лактозы.

При избытке β-галактозида (лактозы) его молекулы связываются с субъединицами репрессора с образованием *репрессор-индукторного комплекса*, в котором индуктор (β-галактозид) выступает в роли *аллостерического регулятора*, изменяющего конформацию белка-репрессора, что ведет к инактивации последнего. У инактивированного репрессора резко снижается сродство к зоне оператора, в результате чего репрессор отсоединяется от промотора, открывая «вход» для РНК-полимеразы. Вслед за этим начинается транскрипция оперона, а затем и синтез ферментов, метаболизирующих лактозу.

Позитивная регуляция оперона состоит в индукции транскрипции, достигаемой посредством присоединения к промотору комплекса, состоящего из CAP-белка (белок-активатор катаболитических оперонов) и cAMP — универсального внутриклеточного регулятора многих метаболических процессов. Комплекс CAP-cAMP связывается с ДНК в самом начале промотора непосредственно перед сайтом связывания РНК-полимеразы и способен многократно (почти в 50 раз) усиливать транскрипцию оперона (рис. 124).

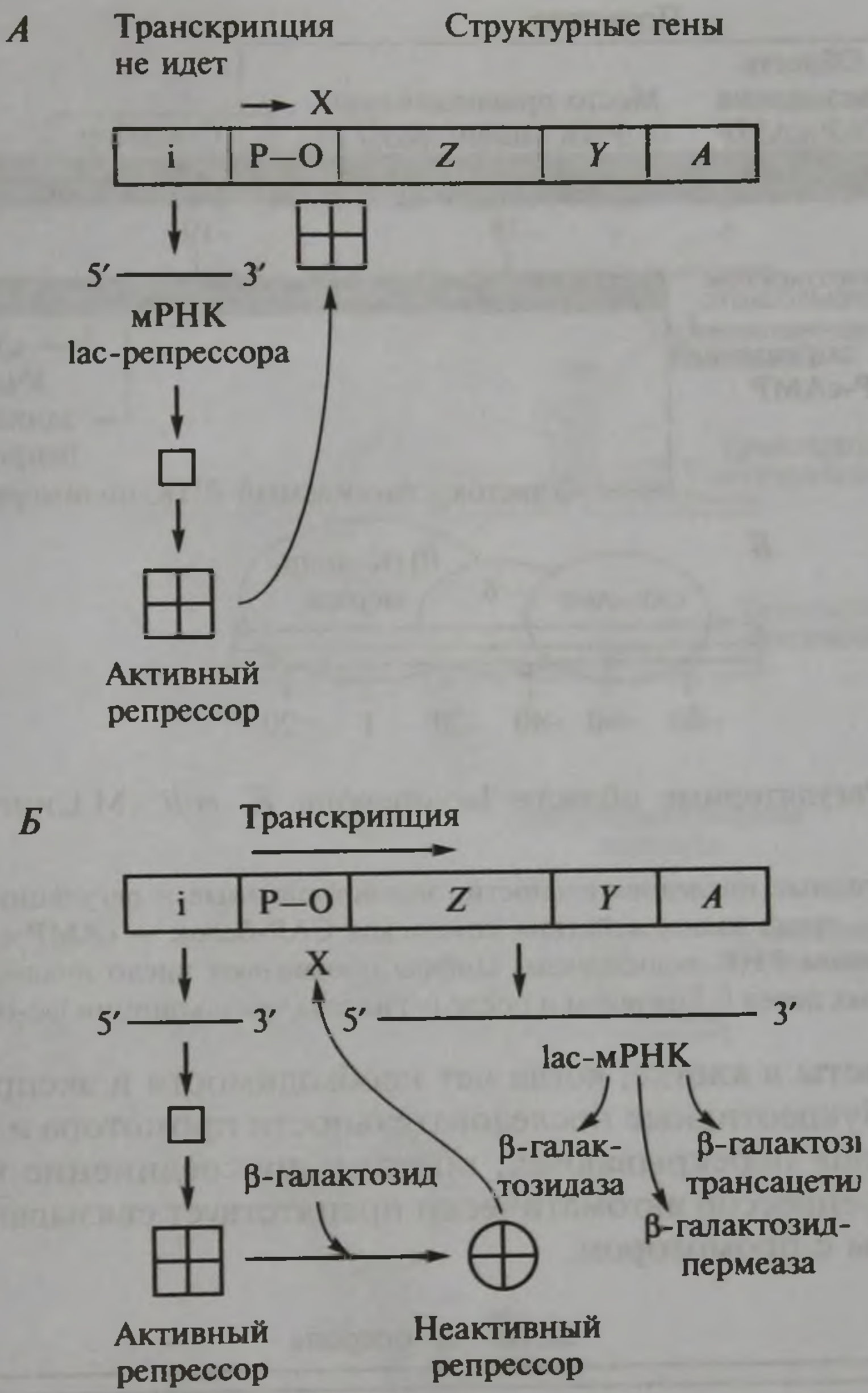


Рис. 123. Схема регуляции лактозного оперона *E. coli*:

А — лактозный оперон репрессирован присоединением белка-репрессора к зоне оператора (транскрипция не идет); i — ген белка-репрессора; P — промотор; O — оператор; *Б* — дерепрессия оперона путем инактивации белка-репрессора. Происходит транскрипция структурных генов и синтезируется полицистроновая lac-мРНК

Несколько более сложная и своеобразная система регуляции транскрипции свойствена триптофановому (trp-) оперону. Он содержит пять генов, кодирующих ферменты синтеза триптофана. Репрессия этого оперона, как и lac-оперона, осуществляется с помощью белка-репрессора с молекулярной массой 58 кДа, кодируемого геном trp R, расположенным на значительном расстоянии от trp-оперона. Белок-репрессор блокирует область оператора в комплексе с самим триптофаном только при высокой концентрации этой



Рис. 124. Регуляторные области lac-оперона *E. coli* (М.Сингер, П.Берг, 1998):

А — нуклеотидные последовательности, задействованные в регуляции экспрессии оперона; *Б* — схема взаимодействия комплекса CAP-белок — cAMP с σ-фактором и кор-ферментом РНК-полимеразы. Цифры обозначают число нуклеотидных пар, расположенных перед (–) началом и после (+) начала транскрипции lac-оперона

аминокислоты в клетке, когда нет необходимости в экспрессии trp-оперона. Нуклеотидные последовательности промотора и оператора в trp-опероне перекрываются, поэтому присоединение комплекса trp-белок-репрессор автоматически препятствует связыванию РНК-полимеразы с промотором.

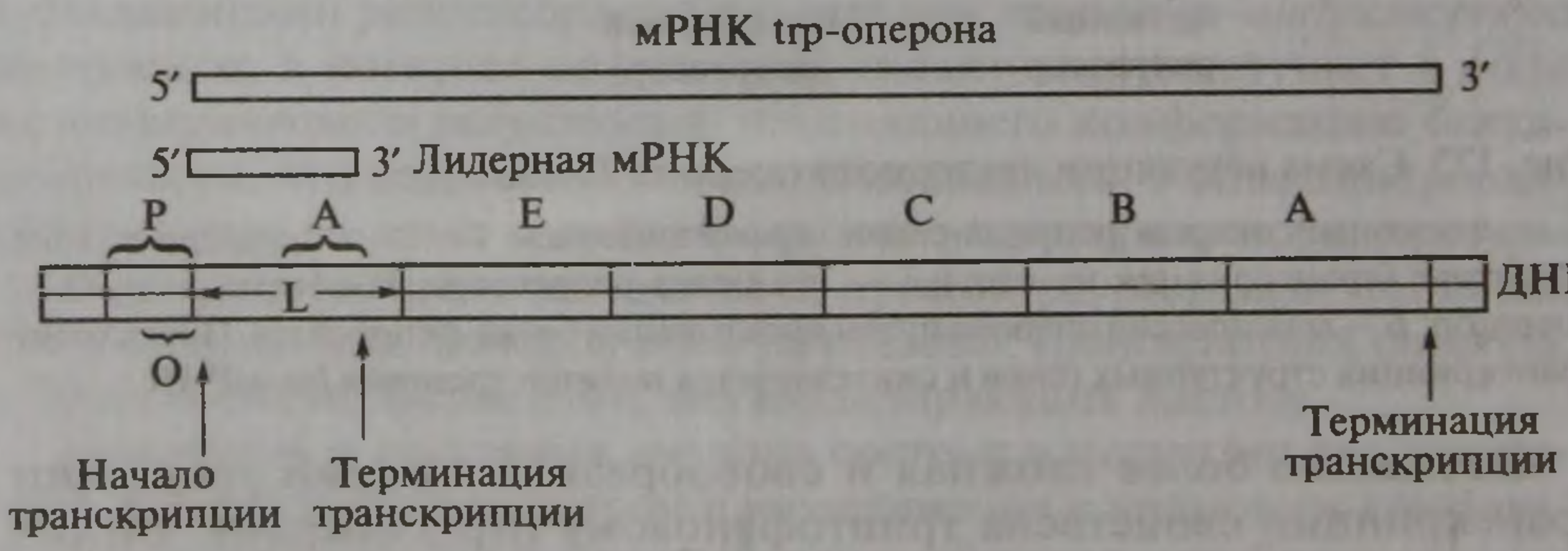


Рис. 125. Схема триптофанового оперона *E. coli*:

P — промотор; O — оператор; A — аттенюатор; L — лидерная последовательность; E, D, C, B, A — цистроны, кодирующие структуру ферментов, участвующих в биосинтезе триптофана. Показаны укороченная (лидерная) мРНК, возникающая в результате терминации транскрипции в зоне аттенюатора, а также полноразмерная полицистроновая мРНК trp-оперона

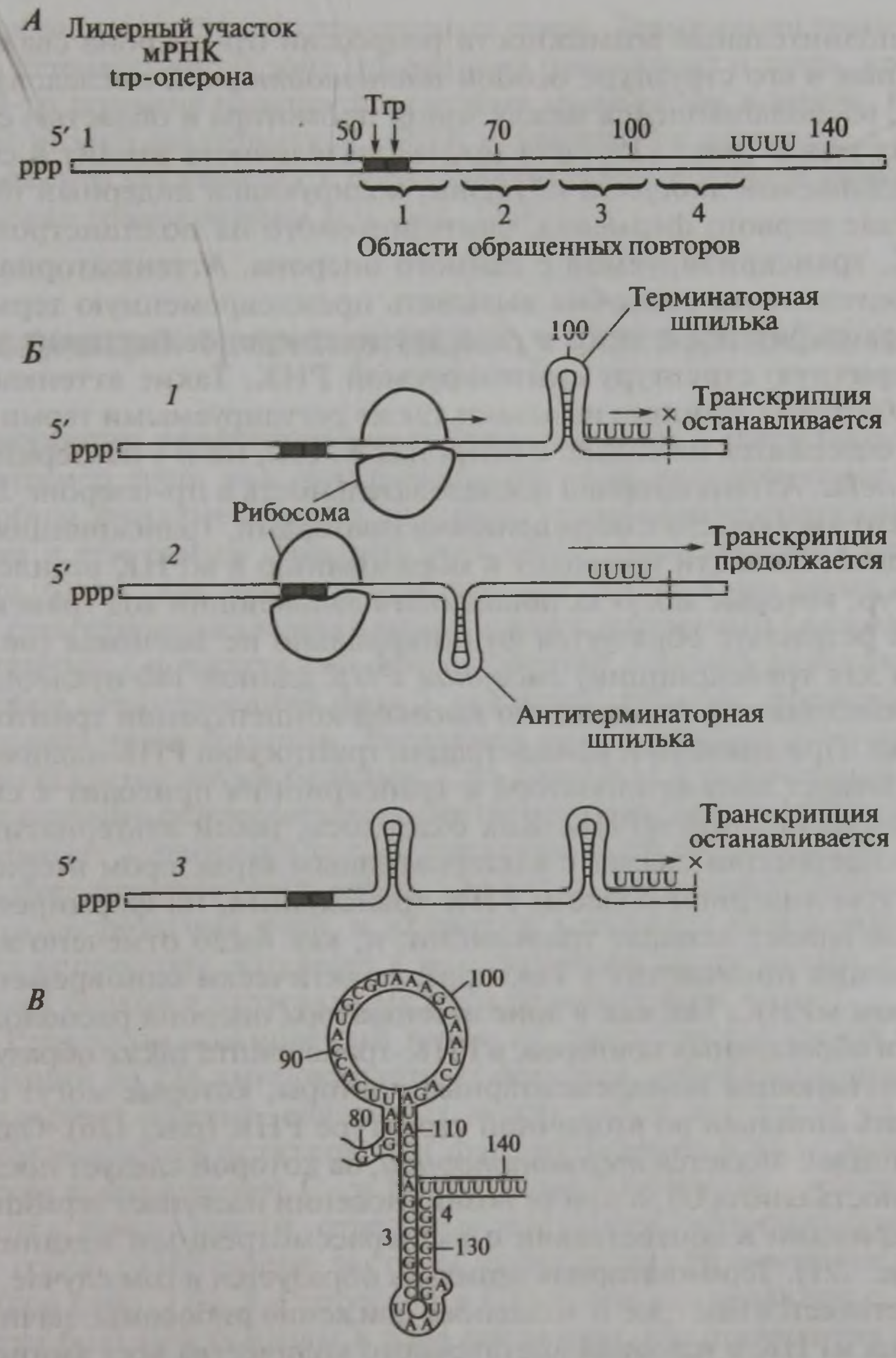


Рис. 126. Схема регуляции транскрипции мРНК trp-оперона с участием рибосомы:

А — лидерный участок мРНК trp-оперона, содержащий четыре области обращенных повторов, способных образовывать шпилечные структуры. Стрелками показано положение двух рядом расположенных кодонов, детерминирующих включение аминокислотных остатков триптофана (Trp) в синтезируемый полипептид; *Б* — упрощенная схема трех вариантов формирования структуры лидерного участка мРНК trp-оперона, один из которых приводит к продолжению транскрипции (в центре), а два других — к остановке транскрипции (*1* — достаточное количество всех аминокислот, включая Trp; *2* — недостаток Trp; *3* — нарушение системы трансляции); *В* — структура терминаторной шпильки мРНК trp-оперона

Дополнительные возможности репрессии trp-оперона связаны с наличием в его структуре особой *аттенюаторной* последовательности, располагающейся между зоной промотора и областью структурных генов (рис. 125). Эта последовательность входит в состав так называемой *лидерной области*, кодирующей лидерный пептид в составе первого фермента, синтезируемого на полицистроновой мРНК, транскрибируемой с данного оперона. Аттенюаторная последовательность способна вызывать преждевременную терминацию транскрипции с этого и ряда других оперонов бактерий, влияя на вторичную структуру синтезируемой РНК. Такие аттенюаторы («ослабители»), которые называют также регулируемыми терминаторами, содержатся не только в оперонах *E. coli*, но и у бактерий рода *Salmonella*. Аттенюаторная последовательность в trp-опероне *E. coli* содержит две области с обращенными повторами. Транскрипция этой последовательности приводит к образованию в мРНК шпилечных структур, которые могут останавливать дальнейший ход транскрипции. В результате образуется функционально не значимая (не пригодная для транскрипции) *лидерная РНК* длиной 140 нуклеотидов, что происходит при достаточно высокой концентрации триптофана в клетке. При снижении концентрации триптофана РНК-полимераза преодолевает зону аттенюатора и транскрипция приводит к синтезу полноразмерной мРНК. Как оказалось, такой альтернативный ход транскрипции связан с альтернативным характером вторичной структуры лидерной области РНК-транскрипта, на формирование которой влияет аппарат трансляции, и, как было отмечено выше, трансляция происходит у бактерий практически одновременно с синтезом мРНК. Так как в зоне аттенюатора оперона расположены области обращенных повторов, в РНК-транскрипте также образуются соответствующие комплементарные повторы, которые могут образовывать шпильки во вторичной структуре РНК (рис. 126). Одна из этих шпилек является *терминаторной*, за которой следует последовательность олиго(U), и при ее возникновении наступает терминация транскрипции в соответствии с вышерассмотренным механизмом (см. рис. 121). Терминаторная шпилька образуется в том случае, если осуществляется быстрое и «плавное» движение рибосомы, начиная с 5′-конца мРНК в условиях достаточного количества всех аминокислот, включая триптофан. В этом случае в ходе трансляции рибосома закрывает два из четырех обращенных повторов, а два другие взаимодействуют с образованием терминаторной шпильки. Напротив, при недостатке триптофана (и соответственно $тРНК^{Тр}$) рибосома останавливается в том участке лидерной последовательности мРНК, где находятся подряд два триплета (кодона), кодирующих данную аминокислоту в составе синтезируемого лидерного пептида. При этом рибосома закрывает первую из четырех областей обращенных повторов, что ведет к возникновению другой *антитерминаторной* шпильки, терминаторная шпилька не образуется, и транскрипция

продолжается в область структурных генов. Терминация транскрипции в аттенюаторной зоне trp-оперона происходит и тогда, когда по какой-то причине нарушается система трансляции в целом. В этом случае вторичная структура мРНК формируется независимо от рибосомы, что также ведет к образованию терминаторной шпильки и остановке транскрипции (см. рис. 126).

Регуляция транскрипции у бактериофага λ

Регуляция транскрипции у бактериофага λ представляет собой пример того, как отдельные *ключевые регуляторные белки* способны фундаментально изменить направление транскрипции генома и тем самым изменить весь модус существования вируса в клетке. Фаг λ, как известно (см. рис. 61), располагает двумя способами существования в зараженной клетке: пассивным (*лизогенным*) и активным (*литическим*). Молекулярной основой переключения способов существования фага λ являются события, происходящие в аппарате транскрипции. Регуляция транскрипции данного фага осуществляется двумя белками: Cro-белком и λ-репрессором. Они могут взаимно репрессировать синтез друг друга таким образом, что связывание λ-репрессора с соответствующей областью оператора блокирует транскрипцию мРНК Cro-белка и определяет лизогенный путь существования фага, и напротив, Cro-белок может выступать в роли репрессора, что ведет к транскрипции генов лизиса, синтезу фаговых белков и литическому пути развития фага.

λ-Репрессор, кодируемый геном *cI*, представляет собой белок, состоящий из 236 аминокислотных остатков, организованных в два глобулярных домена, связанных между собой мостиком из серии аминокислотных остатков. Он функционирует в виде димера, в котором контактируют C- и N-концевые домены. Белок Cro также является гомодимером, отдельные протомеры которого представляют собой глобулярные частицы, содержащие 66 аминокислотных остатков (рис. 127). λ-Репрессор и Cro-белок связываются с ДНК в области большой борозды в зоне оператора, где происходит тонкое и точное взаимодействие структурных элементов (α-спиралей) белков с азотистыми основаниями в узнаваемой белками нуклеотидной последовательности ДНК (рис. 128). У обоих белков имеется характерный структурный элемент «спираль — петля — спираль», благодаря которому они контактируют с ДНК. Одна из этих спиралей входит в зону большого желоба ДНК, а другая (запирающая спираль) фиксирует первую в этом положении.

У фага λ идентифицированы два промотора: сильный (P_R), с которого открывается транскрипция гена белка Cro и генов лизиса, и слабый (P_{RM}), присоединение к которому РНК-полимеразы детерми-

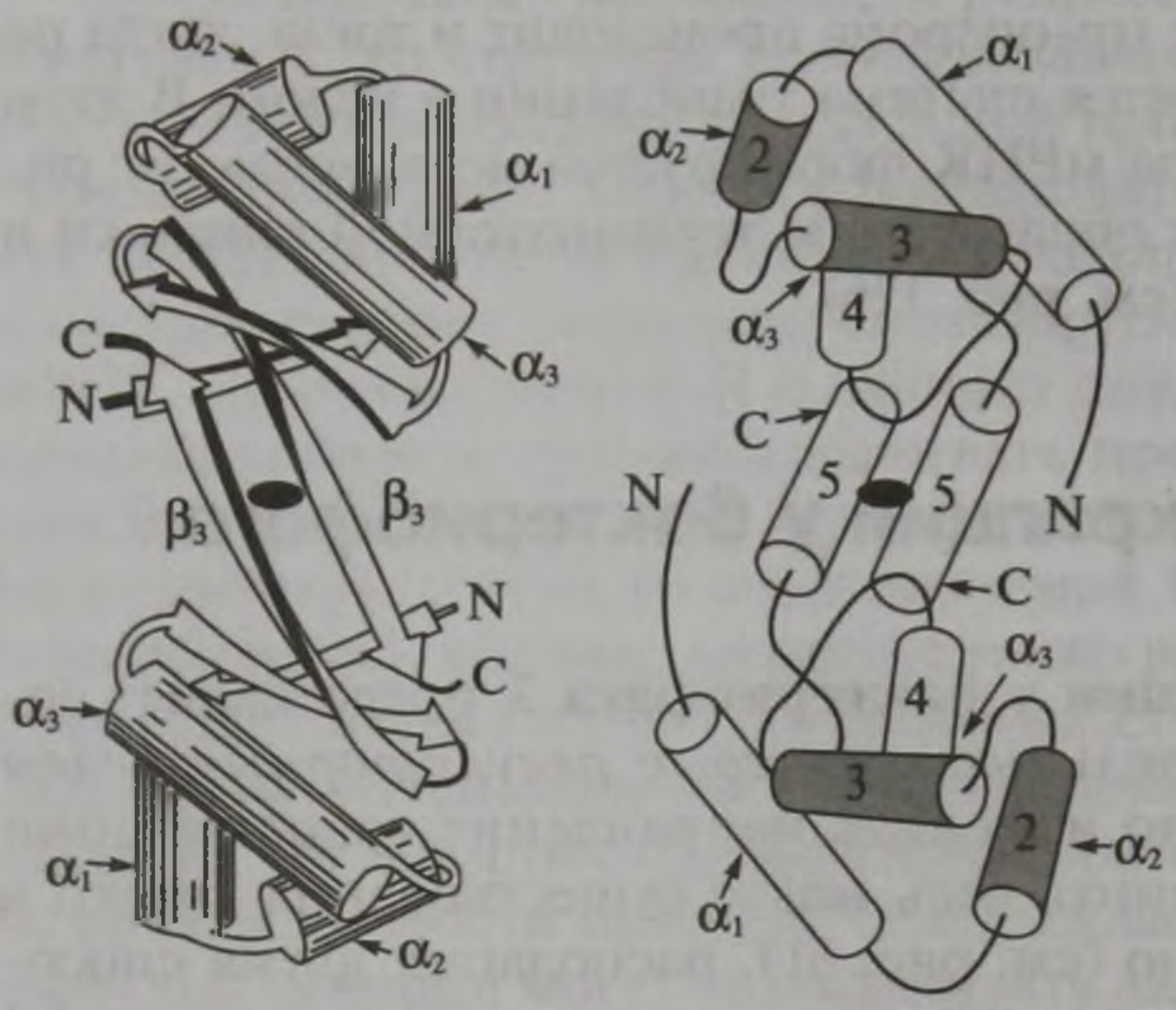


Рис. 127. Димерные молекулы белка Cro(*А*) и репрессора фага λ (*Б*).

α-Спирали показаны в виде цилиндров, β-слои — в виде стрелок. Точкой обозначена ось симметрии. Спирали α_2 и α_3 λ-репрессора участвуют во взаимодействии с ДНК (см. рис. 128)

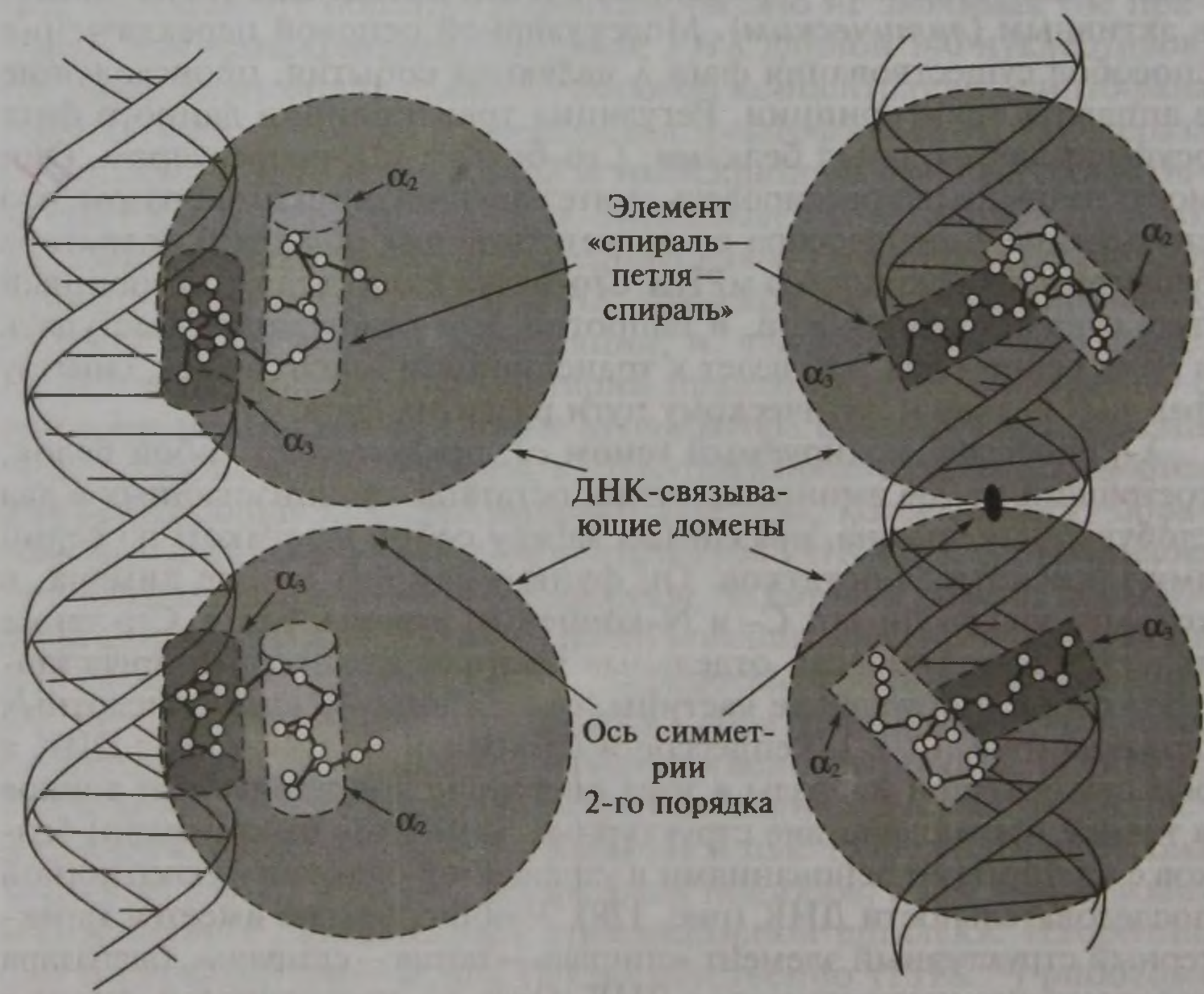


Рис. 128. Взаимодействие с ДНК белков, содержащих структурный элемент «спираль — петля — спираль» (к ним относятся, в частности, белок Cro и λ-репрессор) (М.Сингер, П.Берг, 1998).

Большие круги — домены, связывающиеся с ДНК; α_3 — спираль, входящая в большой желоб ДНК; α_2 — фиксирующая (запирающая) спираль

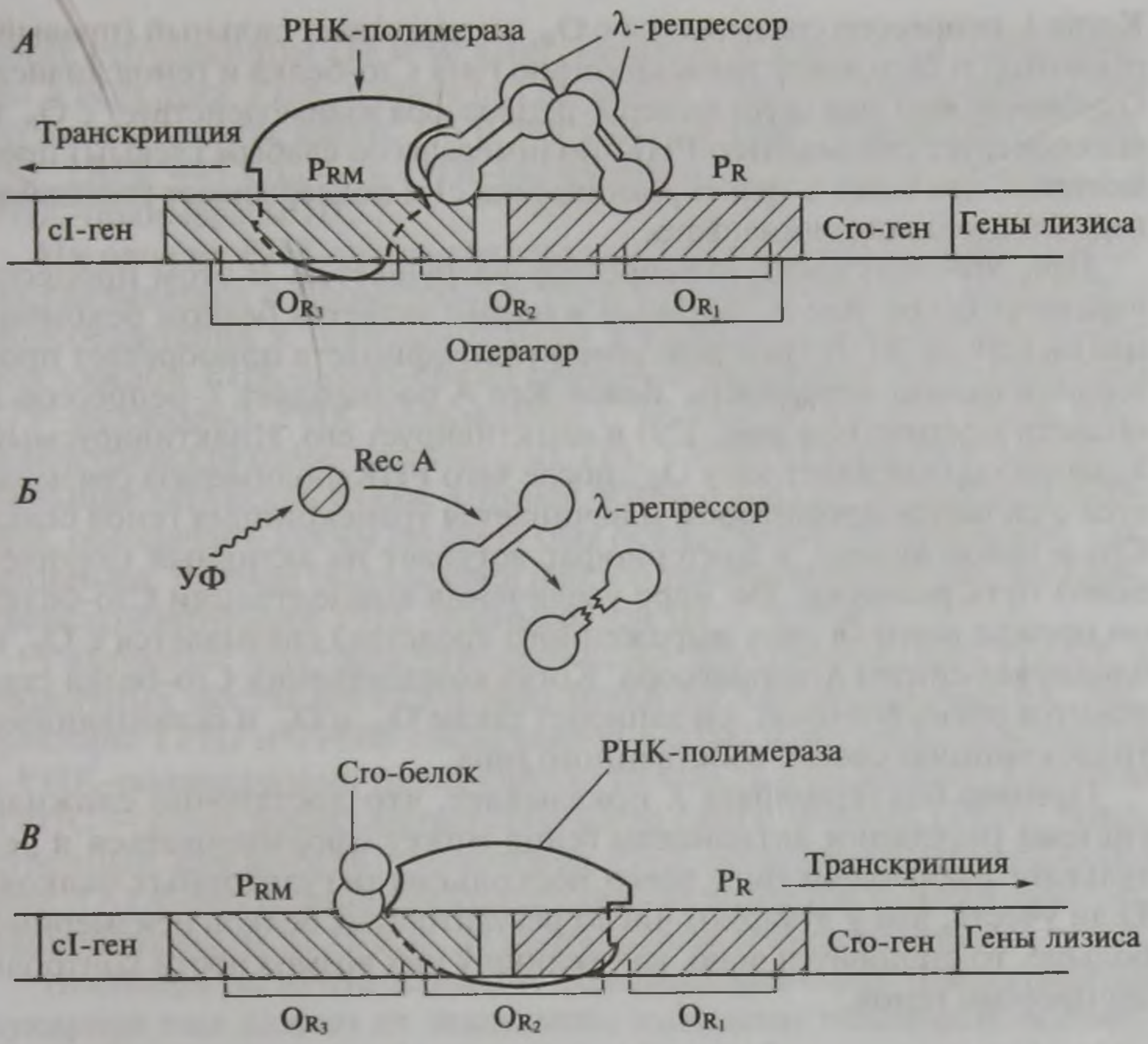


Рис. 129. Схема взаимодействия молекул λ-репрессора и Cro-белка с зоной оператора ДНК фага λ:

А — λ-репрессор определяет лизогенный путь существования фага; *Б* — белок Rec A, под действием УФ-облучения приобретающий протеолитическую активность и расщепляющий молекулу λ-репрессора; *В* — Cro-белок, определяющий активный (литический) путь существования фага; P_{RM} — левый (слабый) промотор; P_R — правый (сильный) промотор; O_{R_1}, O_{R_2}, O_{R_3} — участки оператора; cI — ген, кодирующий λ-репрессор

нирует синтез λ-репрессора. Оба промотора перекрываются с зоной оператора, в которой выделяют три участка: O_{R_1}, O_{R_2}, O_{R_3} (рис. 129). Участок O_{R_1} перекрывается с сильным (правым) промотором (P_R); участок O_{R_3} перекрывается со слабым (левым) промотором (P_{RM}); участок O_{R_2} перекрывается с обоими промоторами. Разные операторные участки имеют разное сродство к белкам-регуляторам, а эффективность связывания этих белков зависит от их концентрации в клетке. Так, λ-репрессор при невысокой концентрации связывается только с участком O_{R_1}, а при высокой может связаться с участками O_{R_2}, O_{R_3} и тем самым закрыть вход для РНК-полимеразы, что блокирует транскрипцию гена cI (т. е. блокирует транскрипцию своей мРНК).

Когда λ-репрессор связывается с O_{R_1} он закрывает сильный (правый) промотор и блокирует транскрипцию гена Cro-белка и генов лизиса. Одновременно еще один димер λ-репрессора взаимодействует с O_{R_2} и способствует связыванию РНК-полимеразы со слабым (левым) промотором, что ведет к транскрипции гена cI и поддерживает пассивное (лизогенное) состояние фага.

При УФ-облучении λ-репрессор разрушается. В этом процессе участвует белок Rec A, который в норме является белком рекомбинации (см. гл. 8), но под действием ультрафиолета приобретает протеолитическую активность. Белок Rec A расщепляет λ-репрессор в области мостика (см. рис. 129) и инактивирует его. Инактивируемый λ-репрессор покидает зону O_{R_1}, после чего РНК-полимераза связывается с сильным промотором и начинается транскрипция генов белка Cro и генов лизиса, а бактериофаг вступает на активный (литический) путь развития. По мере увеличения концентрации Cro-белка он прежде всего (в силу выраженного сродства) связывается с O_{R_3} и блокирует синтез λ-репрессора. Когда концентрация Cro-белка становится очень большой, он запирает также O_{R_1} и O_{R_2} и останавливает транскрипцию своего собственного гена.

Пример бактериофага λ показывает, что достаточно сложная система регуляции активности генов может сформироваться в результате взаимодействия всего нескольких регуляторных белков. Если учесть, что у эукариот число регуляторных белков неизмеримо больше, то становится ясно, как велики у них возможности контроля экспрессии генов.

Транскрипция у эукариот

РНК-полимеразы эукариот изучены значительно слабее, чем соответствующие ферменты бактерий. Синтез РНК у эукариот осуществляют три различные фермента: РНК-полимеразы I, II и III, структура которых изучена далеко не полностью.

Изначально эукариотические РНК-полимеразы были идентифицированы на основании их чувствительности к аманитину — высокотоксичному антибиотику, содержащемуся в бледной поганке. Позднее было установлено, что РНК-полимераза I (не чувствительная к аманитину) синтезирует большие (28S и 18S) рибосомальные РНК, РНК-полимераза II (наиболее чувствительная к аманитину) транскрибирует гены, кодирующие белки, а также гены малых ядерных РНК, а РНК-полимераза III отвечает за синтез 5S рРНК и тРНК.

РНК-полимеразы I, II и III дрожжей имеют три идентичные по молекулярной массе субъединицы (27, 23 и 14,5 кДа). Дрожжевые РНК-полимеразы I и II содержат также две другие одинаковые по

массе субъединицы (14 и 40 кДа). Остальные известные субъединицы у всех трех РНК-полимераз дрожжей существенно различаются по молекулярной массе: 135 и 190 кДа — у РНК-полимеразы I; 146, 150, 220 и 445 кДа — у РНК-полимеразы II; 82, 128 и 160 кДа — у РНК-полимеразы III.

Ни одна из РНК-полимераз эукариот не способна самостоятельно связываться с промоторами транскрибируемых ими генов. Для присоединения к транскриптонам эукариот служат специфичные для каждой РНК-полимеразы *белковые факторы транскрипции* (*TF-факторы*): РНК-полимеразы I, II и III требуют участия факторов транскрипции, называемых TFI, TFII и TFIII соответственно. Большая латинская буква, которая обычно следует за римской цифрой, обозначает, каким по счету был обнаружен данный фактор. Так, например, первым был выделен фактор транскрипции гена 5S рРНК, транскрибируемого РНК-полимеразой III — TFIIIA.

Различные факторы транскрипции присоединяются к ДНК в различных местах относительно точки начала транскрипции. Так, факторы TFID и TFIID связываются с ДНК и образуют комплексы с РНК-полимеразами перед сайтом начала транскрипции, а фактор TFIIC — сразу после соответствующего сайта. Фактор TFIID, существенный для многих промоторов РНК-полимеразы II, называют еще ТАТА-фактором, поскольку он связывается с ТАТА-боксом, расположенным за 25 н. п. до сайта начала транскрипции.

Несмотря на то что изучение белковых факторов транскрипции эукариот еще далеко не закончено, выявлены некоторые особенности структуры этих белков, способные обеспечивать специфическое взаимодействие их аминокислотных последовательностей с азотистыми основаниями ДНК. В этих белках обычно имеются два домена, один из которых отвечает за связывание ДНК, а другой — за регуляцию транскрипции. Например, у белка GAL 4, регулирующего транскрипцию в дрожжевых клетках, функцию связывания ДНК и активацию транскрипции выполняют разные домены. Этот белок-активатор включает транскрипцию любых генов, участвующих в метаболизме галактозы. Белок GAL 4 соединяется с определенной последовательностью в ДНК длиной в 17 н. п. одним доменом своей молекулы, в котором сосредоточены 73 аминокислотных остатка, тогда как другая часть молекулы (ее общая длина составляет 900 аминокислотных остатков) не участвует в присоединении к ДНК. В доменах связывания ДНК найдены характерные структурные элементы: спираль—петля—спираль, впервые детально охарактеризованные у Cro-белка и λ-репрессора бактериофага λ (см. выше).

Для связывания с ДНК белковые факторы транскрипции нередко используют «цинковые пальцы» (рис. 130). Эти элементы структуры выявлены у фактора TFIIIA, в белке GAL 4 дрожжей, факторе SP1 млекопитающих и в рецепторах стероидных гормонов (см. далее).

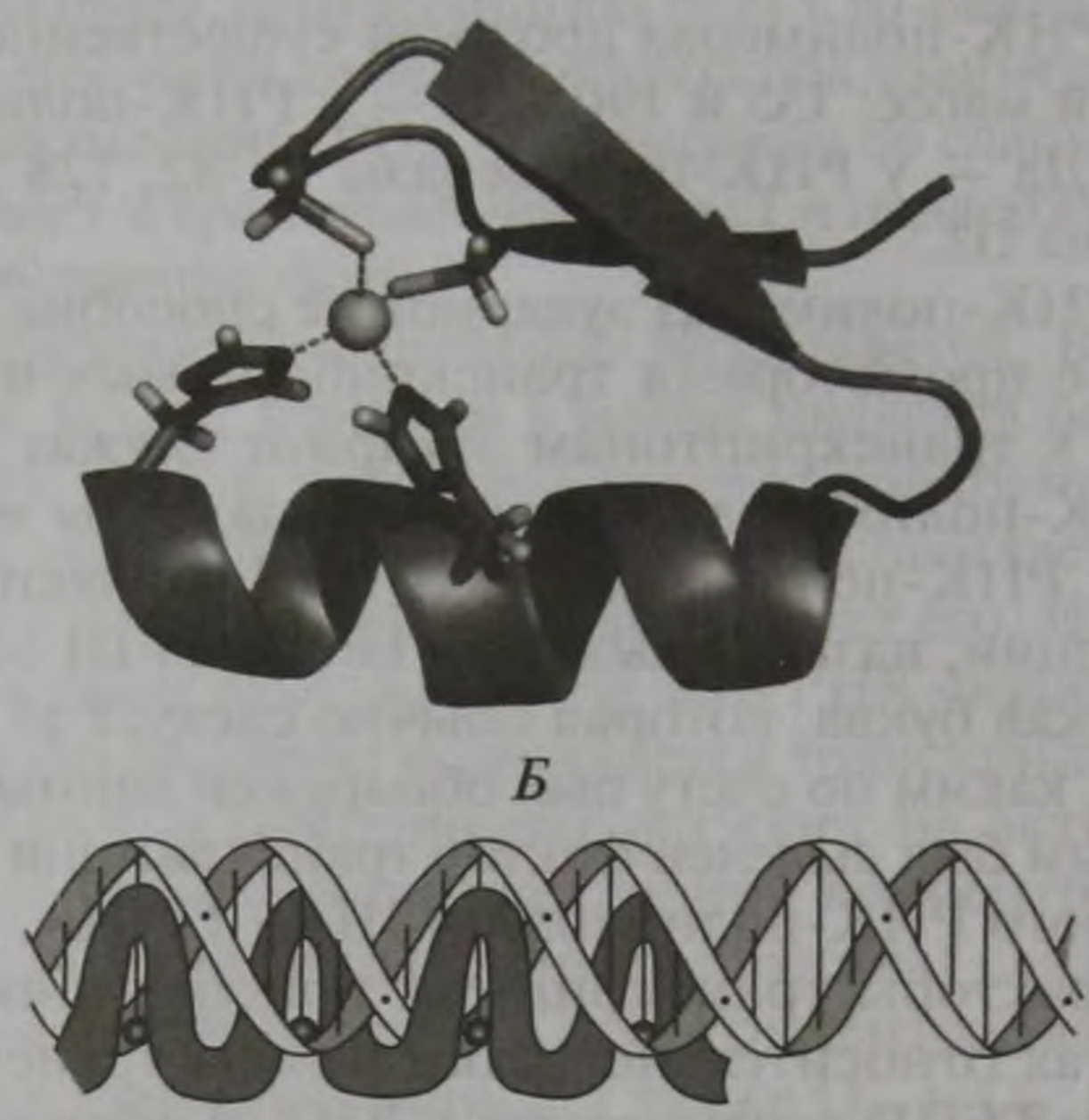


Рис. 130. «Цинковые пальцы» — элементы сверхвторичной структуры в ДНК-связывающих белках, участвующих в регуляции транскрипции:

А — цинковый палец типа Cys2His2 включает α-спираль и антипараллельную β-структуру. Ион цинка связан кооординационными связями с двумя остатками гистидина и двумя остатками цистеина; *Б* — схема взаимодействия четырех «цинковых пальцев» с большими желобами молекулы ДНК. Каждый «палец» узнает определенную последовательность из пяти нуклеотидных пар

Известно несколько типов цинковых пальцев, отличающихся числом и локализацией остатков цистеина и гистидина, координационно связанных с атомом цинка, а в целом этот элемент является отличительным признаком белков, способных связываться с ДНК.

Многие факторы транскрипции содержат особый, состоящий из двух частей структурный элемент — «*лейциновую молнию*». Этот элемент включает протяженные α-спирали, обогащенные остатками лейцина, расположенными по одну сторону спирали, а также домен, богатый положительно заряженными аминокислотными остатками. Белковые факторы транскрипции, содержащие «лейциновые молнии», взаимодействуют с ДНК в виде димеров (рис. 131). Непосредственный контакт с ДНК в этих димерах осуществляют домены, богатые положительно заряженными аминокислотными остатками, а богатые лейцином участки служат для стабилизации структуры димера за счет взаимодействия с гидрофобными остатками соответствующих α-спиралей.

Разнообразие белковых факторов транскрипции у эукариот связано с разнообразием регуляторных последовательностей, свойственных генам высших организмов. К их числу помимо ТАТА-боксов относятся *энхансеры* (усилители), *сайленсоры* (глушители), а

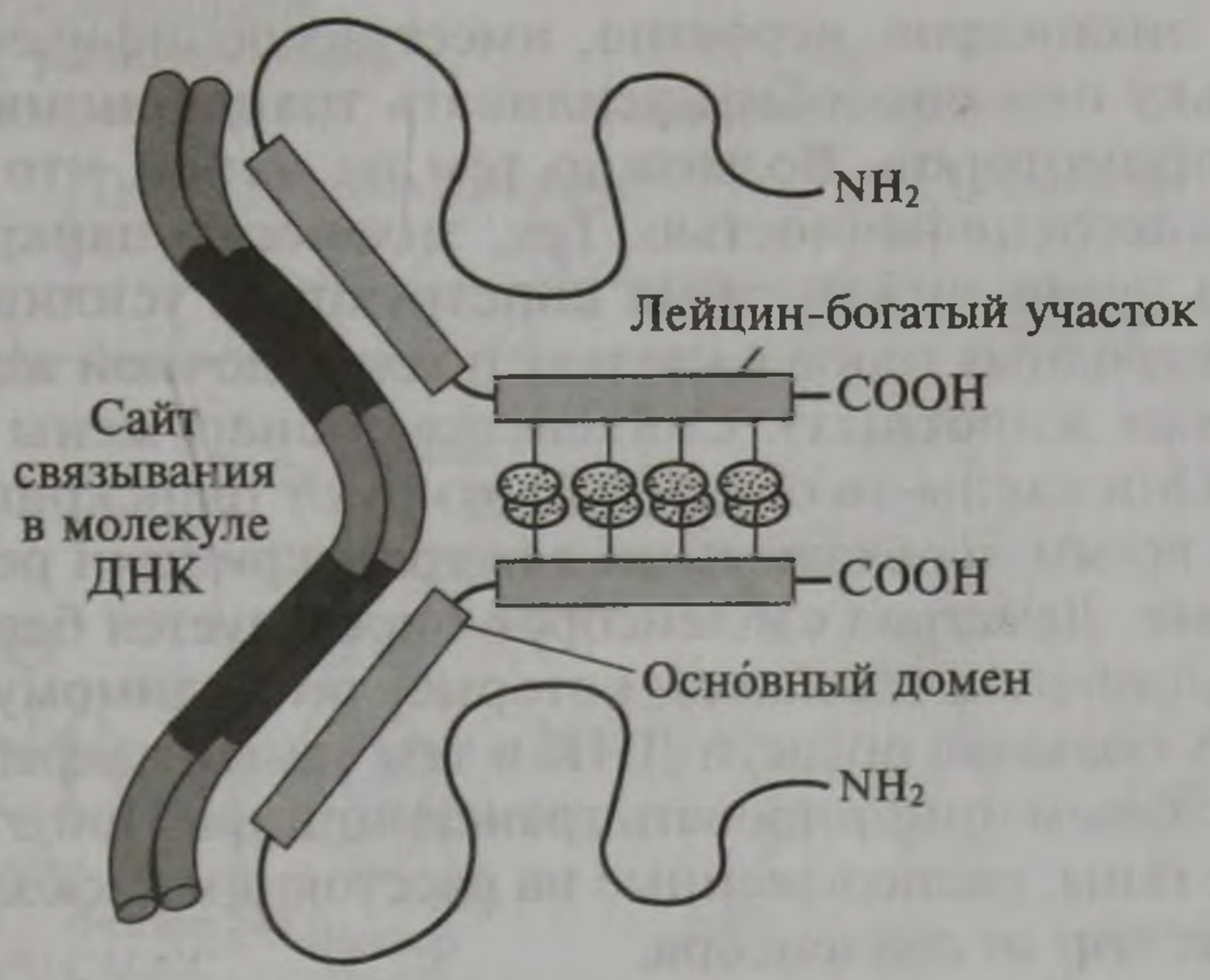


Рис. 131. Схема взаимодействия димерной молекулы фактора транскрипции, содержащей «лейциновые молнии», с молекулой ДНК (М. Сингер, П. Берг, 1998).

Димерная структура фактора транскрипции стабилизируется за счет гидрофобных взаимодействий богатых лейцином участков двух полипептидных цепей. Основные домены белкового димера фактора транскрипции связываются с соответствующим отрицательно заряженным сайтом в молекуле ДНК

также *адапторные элементы*. Эти элементы могут быть удалены на значительные расстояния от точки начала транскрипции, и их воздействие на транскрипцию осуществляется через соответствующие регуляторные белки.

Вероятно, влияние этих элементов на транскрипцию связано с изменением топологии цепей ДНК, в частности с образованием петель, что приближает регуляторные последовательности к промоторам, с которыми они взаимодействуют посредством белковых факторов (рис. 132).

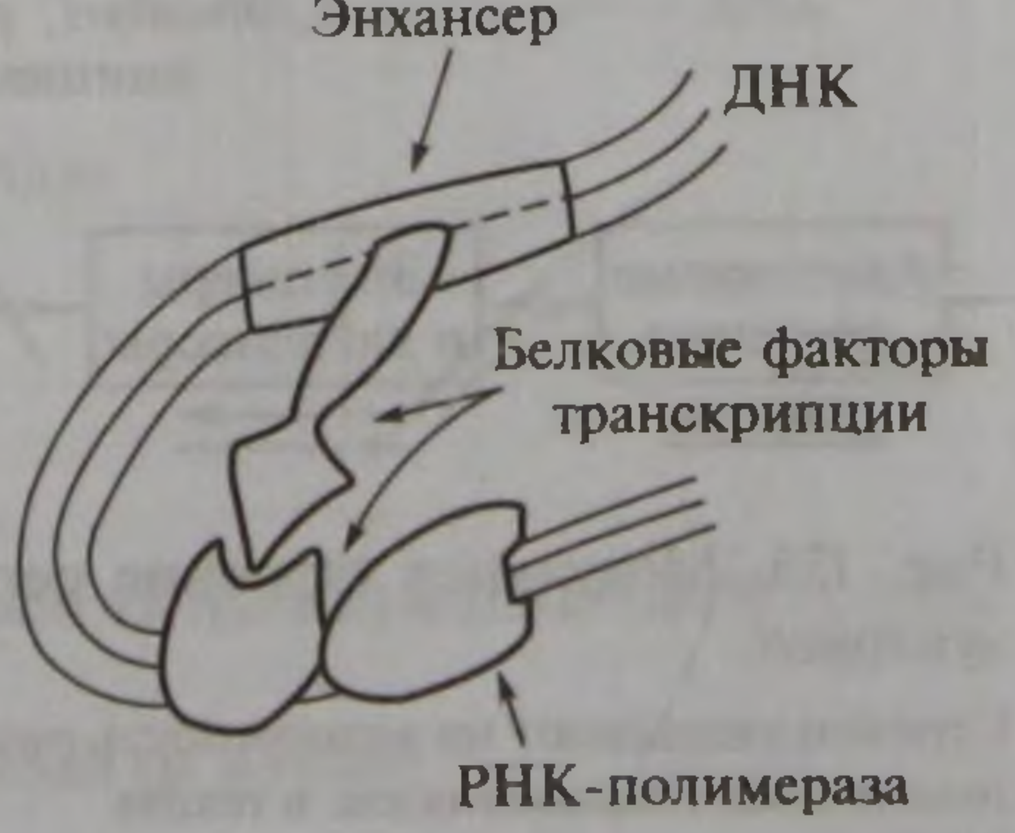


Рис. 132. Изменение топологии цепей ДНК, способствующее воздействию регуляторного элемента (энхансера) на РНК-полимеразу через белковые факторы транскрипции

Действие энхансеров, вероятно, имеет неспецифический характер, поскольку они способны усиливать транскрипцию с любых доступных промоторов. Возможно тем не менее, что энхансеры обладают тканеспецифичностью. Так, энхансеры панкреатических ферментов в генно-инженерных конструкциях усиливают транскрипцию чужеродных генов в клетках поджелудочной железы мыши (у трансгенных животных). Сайленсоры обнаружены у дрожжей *S. cerevisae*. Они каким-то образом блокируют транскрипцию генов, обладающих всеми необходимыми для транскрипции регуляторными элементами. Действие сайленсоров опосредуется белками групп SIR (silent information regulator), которые, по-видимому, способны оккупировать большие области ДНК и тем самым закрывать доступ белкам, способным инициировать транскрипцию. При этом репрессируются все гены, расположенные на расстоянии нескольких тысяч нуклеотидных пар от сайленсора.

Адапторные элементы генома эукариот проявляют избирательную чувствительность по отношению к стероидным гормонам, cAMP, глюкокортикоидам. Они специфически регулируют клеточный ответ на тепловой шок, действие металлов (Cd^{2+}, Zn^{2+}) и некоторых химических соединений (например, диоксинов).

Адапторные элементы обладают специфическими нуклеотидными последовательностями, узнаваемыми белками. Так, например, фактор транскрипции белков теплового шока узнает в ДНК последовательность CNNGAANNTTCNNG (где N-любой нуклеотид), а глюкокортикоидный белок-рецептор связывается с ДНК в области последовательности TGTTCT.

В области промоторов многих эукариотических транскриптонов выявлен *CAAT-бокс* — элемент, регулирующий частоту инициации транскрипции. Результаты изучения энхансеров, удаленных порой на значительные расстояния (до 20 000 н. п.) от промоторов, а также регуляторных участков вблизи промоторов, позволили сделать некоторые существенные обобщения. Во-первых, регуляторные последовательности в ДНК имеют модульное строение и состоят из серии определенных нуклеотидных последовательностей (рис. 133).

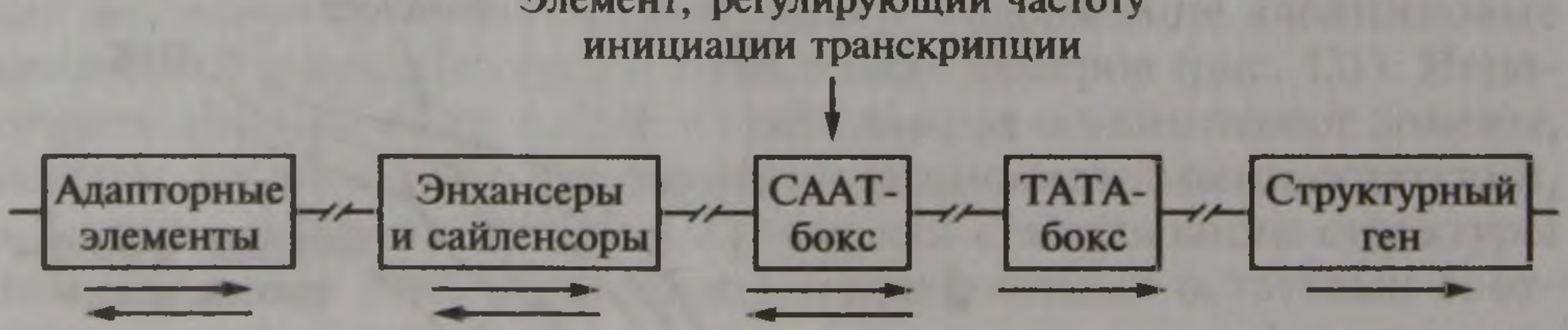


Рис. 133. Модульное строение регуляторных последовательностей в ДНК эукариот.

Стрелки указывают на возможность различной ориентации регуляторных последовательностей. Пояснения см. в тексте

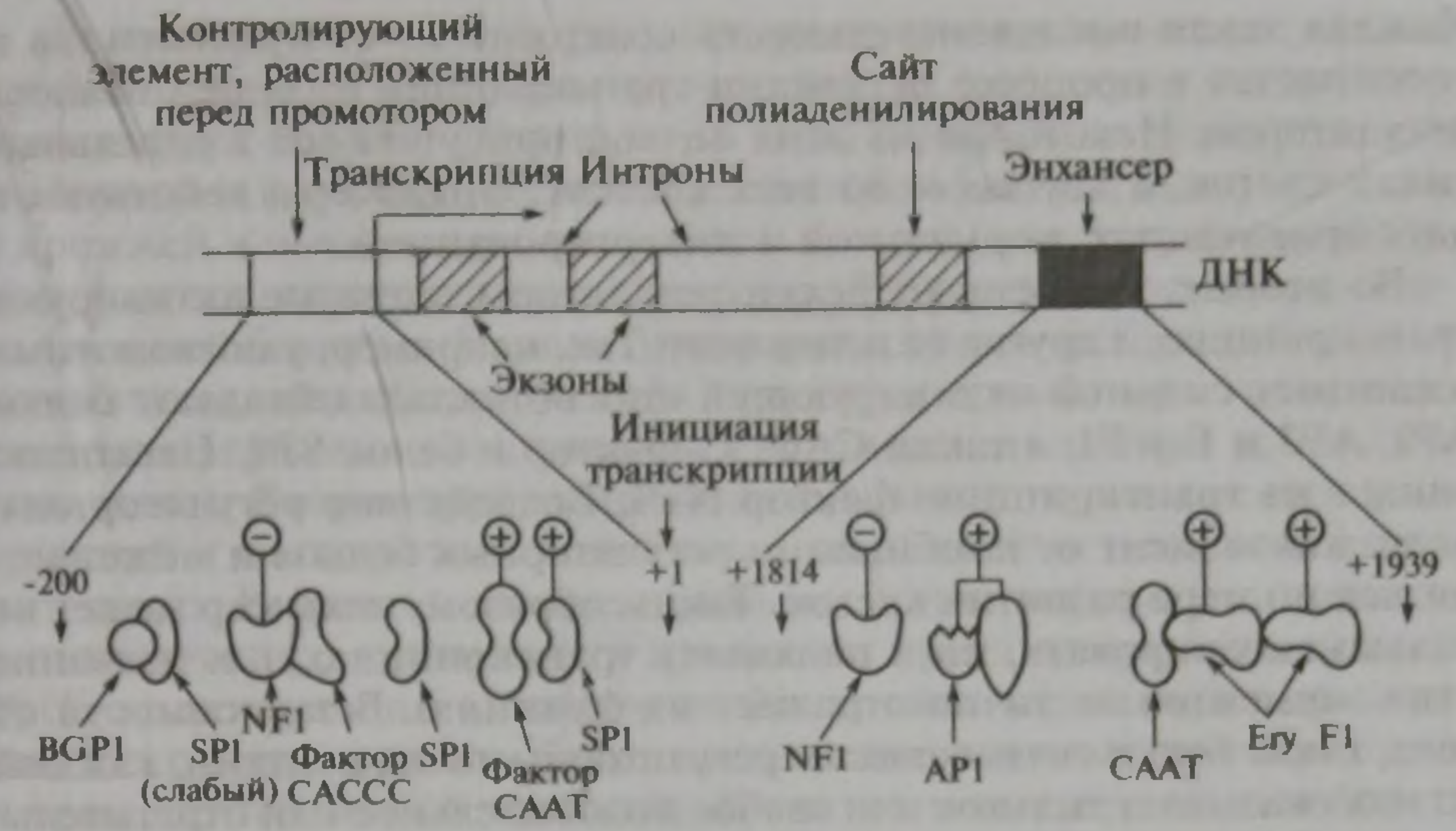


Рис. 134. Присоединение белков, усиливающих (+) и ослабляющих (–) транскрипцию, к регуляторным элементам гена β-глобина курицы (по Б. Албертсу и др., 1994). Пояснения см. в тексте

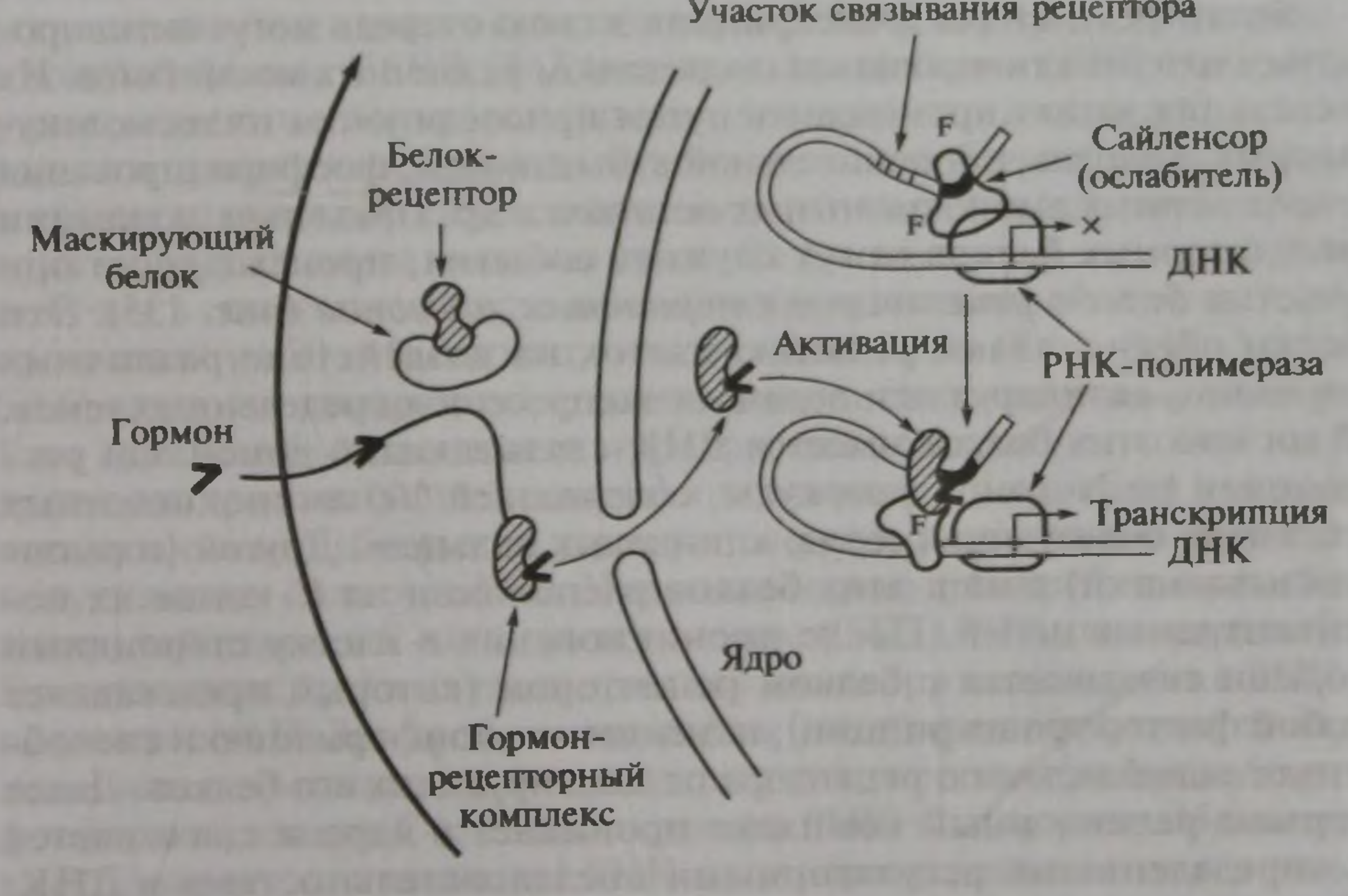


Рис. 135. Воздействие стероидных гормонов на транскрипцию (по В. А. Гвоздеву, 1996):

F — белковые факторы транскрипции. Пояснения см. в тексте

Каждая такая последовательность содержит 8—15 нуклеотидов и соединяется в процессе регуляции транскрипции с серией белков-регуляторов. Некоторые из этих белков присутствуют в отдельных типах клеток, а другие — во всех клетках. Энхансеры действуют в двух ориентациях: нормальной и инвертированной.

Во-вторых, существуют белки-регуляторы, которые активируют транскрипцию, а другие ее подавляют. Так, например, у позвоночных животных сильной активирующей способностью обладают белки AP1, AP2 и Ery F1, а также CACCC-фактор и белок SP1. Негативно влияет на транскрипцию фактор NF1. Воздействие регуляторного элемента зависит от комбинации регуляторных белков и может меняться по мере развития клеток. Таким образом, энхансер может не только активировать, но и подавлять транскрипцию, т. е. название этих элементов не точно отражает их функцию. В зависимости от того, какие белки связываются с регуляторными элементами, каждый из них оказывает сильное или слабое положительное или отрицательное действие на транскрипцию.

В-третьих, энхансеры и элементы, расположенные перед промоторами, связывают много одинаковых белков. Примером в этом отношении может служить ген β-глобина курицы. Энхансер этого гена, расположенный после кодирующей белок области, связывает группу белков, которые могут присоединяться и к регуляторному элементу, расположенному непосредственно перед промотором (рис. 134).

Белки-регуляторы транскрипции в свою очередь могут активироваться или инактивироваться посредством различных механизмов. Их активация может происходить путем присоединения низкомолекулярных лигандов, дополнительной субъединицы, фосфорилирования определенных аминокислотных остатков и др. Примером активации регуляторных белков могут служить события, происходящие при участии *белков-рецепторов стероидных гормонов* (рис. 135). Эти белки обеспечивают реакцию клеток на воздействие различных гормонов, активируя или подавляя экспрессию определенных генов. В составе этих белков имеется ДНК-связывающий домен. Он расположен на N-конце молекулы, содержащей 100 аминокислотных остатков, образующих серию «цинковых пальцев». Другой (гормон-связывающий) домен этих белков расположен на С-конце их полипептидных цепей. После проникновения в клетку стероидный гормон связывается с белком-рецептором (который представляет собой фактор транскрипции), изменяет его конформацию и способствует освобождению рецептора от маскирующих его белков. Далее гормон-рецепторный комплекс проникает в ядро и связывается с определенными регуляторными последовательностями в ДНК. Связывание гормон-рецепторного комплекса может способствовать удалению белка — сайленсора и таким образом снимать запрет на транскрипцию генов.

Долгое время считали, что у эукариот, как и у прокариот, белки-активаторы и репрессоры транскрипции действуют напрямую, контактируя с ДНК. Однако затем было обнаружено существование *медиаторов* (посредников) между этими белками и транскриптонами у дрожжей и высших эукариот. Одни из этих регуляторов действуют на хроматин, другие взаимодействуют с РНК-полимеразой и связанными с ней белками. Первая группа регуляторов включает белковые комплексы, модифицирующие структуру хроматина: гистонацетилтрансферазные комплексы и гистондеацетилазные комплексы (см. ниже). Вторая группа представлена медиаторным комплексом дрожжей и несколькими подобными комплексами в клетках млекопитающих. Вероятно, эти две группы медиаторов функционируют на разных стадиях: изменения в хроматине предшествуют стимуляции РНК-полимеразной активности. На уровне медиатора определяется в конечном итоге, будет ли инициироваться транскрипция, и если да, то с какой частотой. В состав медиатора входят белки, способные инициировать или подавлять транскрипцию. Медиатор и РНК-полимераза II эукариот проявляют свою активность в виде полимеразного холофермента. Медиатор, очевидно, интегрирует и передает положительные и отрицательные стимулы от энхансеров к промоторам, действуя непосредственно через РНК-полимеразу в промотор-зависимой транскрипции.

Впервые медиатор был обнаружен при изучении транскрипции, осуществляемой РНК-полимеразой II у дрожжей. В этих опытах было установлено, что в системе, содержащей РНК-полимеразу, транскрипционный фактор TFIIB, TATA-связывающий белок (TBP) и другие, ранее известные белки-регуляторы, транскрипция осуществлялась только в присутствии молекулы *адаптора* между белком-активатором и другими вышеперечисленными компонентами транскрипции. Этот адаптор и был назван медиатором. Затем медиатор дрожжей был получен в гомогенном состоянии. Он представляет собой комплекс, состоящий из 20 субъединиц, четыре из которых (Srb2, Srb4, Srb5 и Srb6) являются продуктами четырех доминантных генов-супрессоров. Еще девять субъединиц медиатора также представляют собой продукты экспрессии ранее идентифицированных у дрожжей генов, а остальные семь субъединиц кодируются ранее неизвестными генами, получившими название Med. Установлено, что субъединицы типа Srb взаимодействуют с С-концевым доменом (CTD) РНК-полимеразы II. Предполагают, что в клетках дрожжей медиатор связывается с РНК-полимеразой II. Затем образовавшийся холофермент входит в состав *комплекса преинициации* транскрипции, после чего происходит фосфорилирование С-концевого домена РНК-полимеразы. Медиатор высвобождается, а активная РНК-полимераза связывается с новым комплексом белков — *элонгатором*, участвующим в регуляции элонгации транскрипции. Кроме 20-субъединичного медиатора у дрожжей были найдены дополнительные белки-посредники (Srb8,



Рис. 136. Комплексы белков, участвующих в регуляции транскрипции у эукариот (упрощенно):

R — белки-регуляторы, передающие сигнал от энхансера; A, B, D, E, F, H — факторы транскрипции (TF); Srb, Gal 11, Med — белки субмедиаторного комплекса; SWI/SNF — комплекс белков, модифицирующий структуру хроматина; TBP—TATA-связывающий белок; CTD — С-концевой домен РНК-полимеразы II с присоединенными к нему фосфорными остатками — (P)

Srb9, Srb10 и Srb11), которые участвуют в отрицательной регуляции транскрипции. У млекопитающих первым очищенным до гомогенного состояния медиатором стал TRAP-*коактиваторный комплекс ядерного рецептора тироидных гормонов*. У человека и мышей выделено семь медиаторных комплексов, что позволяет считать медиаторы основными посредниками при передаче информации от энхансеров к промоторам высших эукариот (рис. 136).

Сопоставление структур медиаторов, выделенных из разных объектов, показало, что они имеют *модульное строение*, и один из модулей обычно представлен Srb-подобными белками, которые, вероятно, совершенно необходимы для инициации транскрипции всех генов, транскрибируемых РНК-полимеразой II.

Одним из важных инструментов регуляции экспрессии генетической информации в клетках эукариот является открытый Э. Файером и К. Мэлло процесс *РНК-интерференции* (RNA — interference). Это открытие считается одним из крупнейших событий в молекулярной биологии на рубеже веков и удостоено Нобелевской премии в 2006 г. (см. с. 10). РНК-интерференция представляет собой фундаментальный механизм подавления экспрессии (сайлейсинга) генов на стадиях транскрипции и трансляции. Этот механизм функционирует у млекопитающих, растений и грибов и задействован в регуляции развития, дифференцировки и защиты организма от паразитических геномов вирусов, а также перемещения транспозонов, способных вызывать негативные перестройки в геноме.

Общая схема РНК-интерференции представлена на рис. 137.

Процесс РНК-интерференции начинается с действия фермента DICER (от англ. *to dice* — нарезать в форме кубиков), который дей-

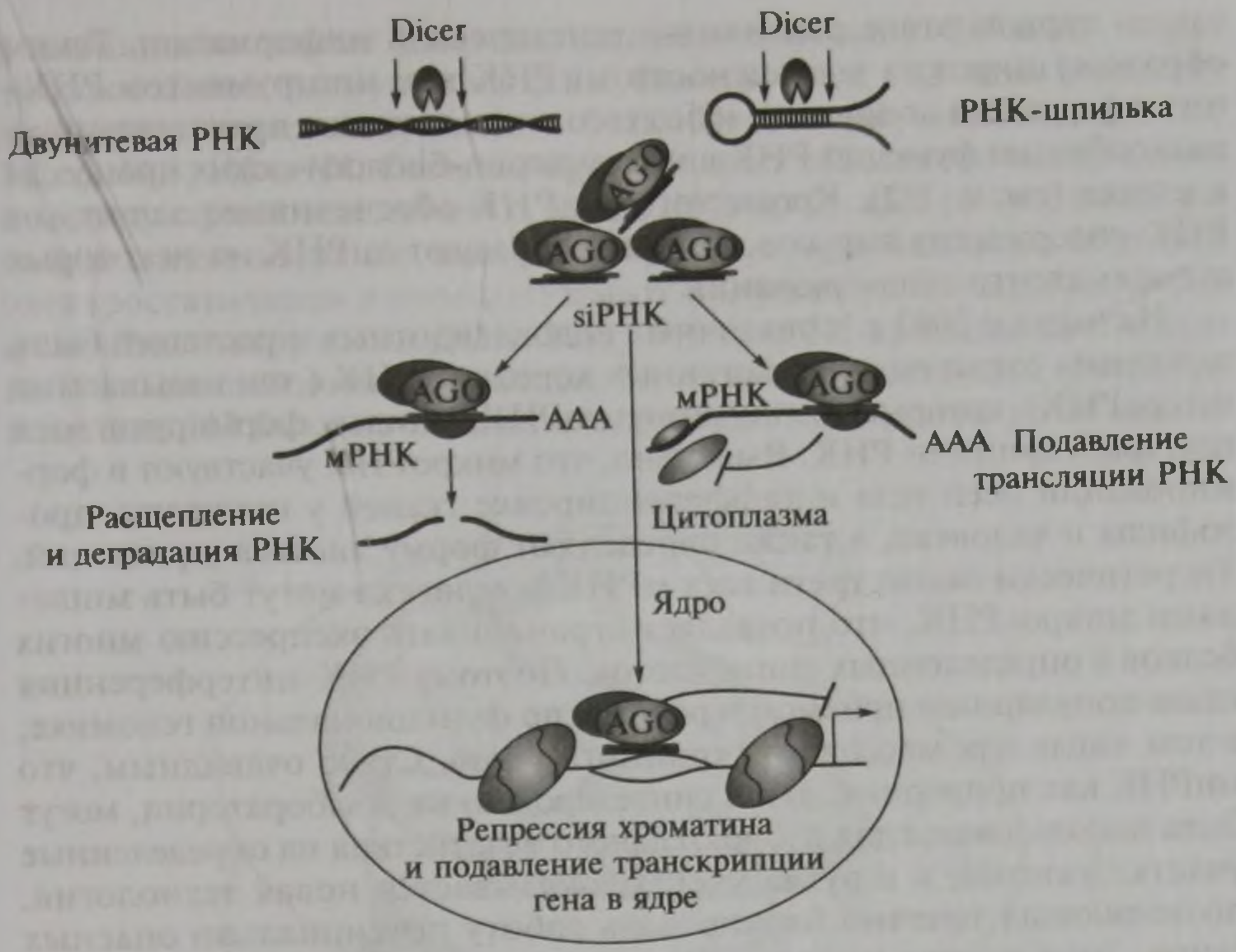


Рис. 137. РНК-интерференция. Пояснения см. в тексте

ствует как эндонуклеаза и разрезает экзогенные и эндогенные двуцепочечные РНК(либо содержащие характерную «шпильку» одноцепочечные РНК) на короткие фрагменты длиной 21 — 25 нуклеотидов. Эти фрагменты получили название *малые интерферирующие РНК* (siRNA, миРНК). Одна из цепей миРНК (ее называют ведущей цепью, англ. — *guide strand*) включается далее в состав РНК-белкового комплекса RISC (комплекс сайлейсинга, индуцированный РНК; англ. *RNA-induced silencing complex*), каталитической составляющей которого являются белки-эндонуклеазы семейства аргонавт[1] (AGO на рис. 137). Эндонуклеазы AGO вызывают расщепление и деградацию определенных мРНК. При этом ведущая цепь миРНК в комплексе с AGO находит определенные (комплементарные ей) участки нуклеотидных последовательностей молекул РНК—мишений, а AGO расщепляет эти молекулы РНК. Таким образом происходит деградация мРНК и останавливается процесс трансляции на рибосомах. Важно, что миРНК могут также подавлять транскрипцию гомологичных им по нуклеотидной последовательности участков ДНК в ядре клетки, осуществляя, таким образом, контроль за экспрессией генов на

[1] Название *аргонавт* связано с одним из первых описанных проявлений РНК-интерференции у растений, приводящее к необычной форме листьев, напоминающих щупальца головоногих моллюсков — аргонавтов.

самом первом этапе реализации генетической информации. Таким образом, широкие возможности миРНК как инструментов РНК-интерференции дополнили и без того впечатляющие представления о разнообразии функций РНК в молекулярно-биологических процессах в клетке (см. с. 112). Кроме того, миРНК обеспечивают защиту от РНК-содержащих вирусов, которые образуют дцРНК на некоторых стадиях своего существования.

Начиная с 2002 г. у различных видов животных и растений были выявлены сотни видов эндогенных коротких РНК (так называемые микроРНК), которые возникают из дцРНК-шпилек, формирующихся при транскрипции РНК. Выяснено, что микроРНК участвуют в формировании осей тела и дифференцировке тканей у нематоды, дрозофилы и человека, а также определяют форму листьев у растений. Теоретически около трети всех мРНК у человека могут быть мишенями микро-РНК, что позволяет ограничивать экспрессию многих белков в определенных типах клеток. Поэтому РНК-интерференция стала популярным приемом в работах по функциональной геномике, в том числе для массового скрининга генов. Стало очевидным, что миРНК как природные, так и синтезированные в лаборатории, могут быть использованы для избирательного воздействия на определенные участки генома, и в руках ученых оказывается новая технология, позволяющая точечно блокировать работу потенциально опасных генов. Этот подход уже применяется как лабораторный метод для выявления функций генов, в том числе и при изучении генома человека. Вероятно в ближайшем будущем могут появиться основанные на РНК-интерференции лекарства нового поколения, действующие на сугубо определенный ген и выключающие синтез кодируемого им белка. Такими мишенями могут быть гены, активные в клетках опухолей или гены, обеспечивающие размножение вирусов.

Хроматин и общая (тотальная) регуляция транскрипции у эукариот

На регуляцию транскрипции у эукариот большое влияние оказывает упаковка ДНК в составе *хроматина* — белково-нуклеинового комплекса, в составе которого постоянно пребывает ДНК высших организмов. Изменения в упаковке ДНК непосредственно влияют на экспрессию генов. Хроматин представляет собой комплекс ДНК с белками (гистонами и негистоновыми белками), который обеспечивает *компактную упаковку* гигантских, достигающих длины несколько метров, молекул ДНК эукариот в ядрах их клеток. Линейный четырехбуквенный генетический код, «словами» в котором являются тройки нуклеотидов (кодоны), позволяет хранить огромную информацию в малом объеме. Вся генетическая информация, содержащаяся

в гаплоидном геноме человека, закодированная в $3 \cdot 10^9$ н. п., теоретически может быть упакована в кубик с ребром ~1,5 мкм. В составе хроматина присутствуют пять белков-гистонов, называемых H1, H2A, H2B, H3 и H4. Гистоны представляют собой небольшие белки с молекулярной массой от 11,3 кДа (H4) до 23 кДа (H1) с высоким содержанием основных аминокислот (Lys и Arg), благодаря чему они электростатически взаимодействуют с отрицательно заряженными фосфатными группами ДНК. Так, в гистоне H1 кролика, состоящем из 213 аминокислотных остатков, 65 составляют остатки Lys и Arg, а в гистоне H3 быка из 135 аминокислотных остатков на долю основ-

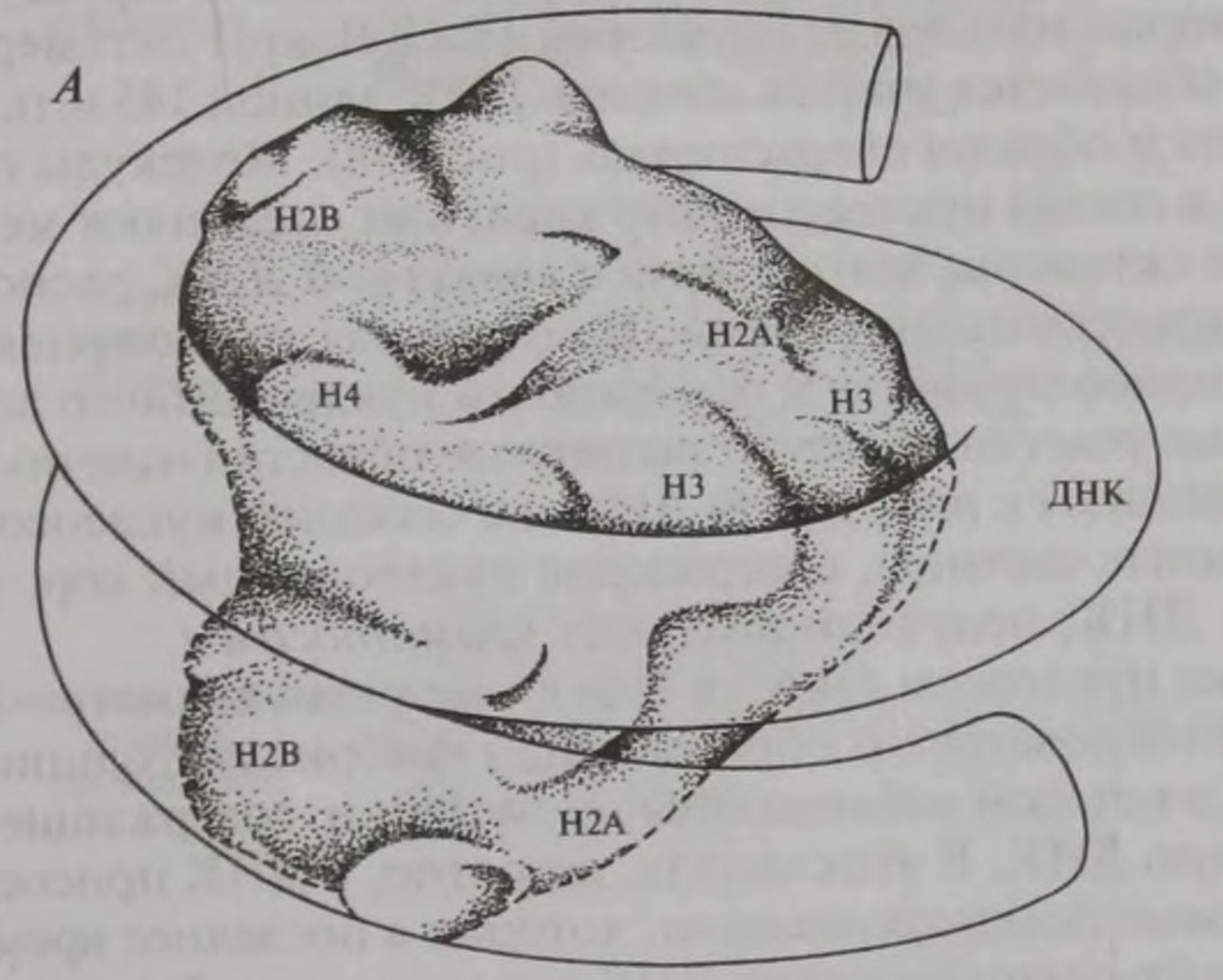


Рис. 138. Взаимодействие ДНК и гистонов в составе хроматина (*Kornberg R., Klug A.* // Sci. Amer. — 1981. — V. 244. — N 2. — P. 52):

А — модель нуклеосомного кора: участок ДНК длиной 145 н. п. обвивает октамер коровых гистонов, совершая 1¾ оборота; *Б* — схема строения 11 нм-фибриллы («нитка бус»)

ных аминокислот приходится 32 остатка. По первичной структуре гистоны весьма консервативны. Соотношение гистонов H2A, H2B, H3 и H4 в хроматине низших эукариот (дрожжи, плесневые грибы) такое же, как в хроматине млекопитающих. Исключение составляет гистон H1, который достаточно вариабелен по первичной структуре и в эритроцитах курицы замещен на другой гистон (H5). В хроматине дрожжей гистон H1 вообще отсутствует.

В изображениях, полученных методом электронной микроскопии, хроматин напоминает нитку бус, в которой каждая бусинка представляет собой октамер гистонов (включает по две молекулы гистонов H2A, H2B, H3 и H4), называемый *нуклеосомным кором*. Таким образом, гистоны помогают ДНК организоваться в структурные единицы, которые называются *нуклеосомами*. Вокруг октамера коровых гистонов обвивается участок спирали ДНК длиной 145 н.п., совершая $1^3/_4$ оборота и образуя сверхспираль (рис. 138). Молекулы гистона H1 не входят в состав нуклеосомного кора; они соединяют между собой отдельные октамеры, контактируя с *линкерной ДНК*, расположенной между нуклеосомными корами. Центральная глобулярная часть гистона H1 присоединяется к поверхности нуклеосомного кора, а N- и C-концевые участки образуют вытянутые области («плечи»), которые взаимодействуют с линкерной ДНК и с соседней нуклеосомой. Субнуклеосомные частицы, содержащие нуклеосомный кор, гистон H1 и 168 н.п. ДНК, получили название *хроматосомы*.

Цепочка нуклеосом («нитка бус»), имеющая диаметр ~11 нм, может конденсироваться с образованием фибриллы толщиной 30 нм (рис. 139), в которой наблюдаются промежутки, содержащие неплотно упакованную ДНК. В этих местах, вероятно, к ДНК присоединяются *негистоновые белки хроматина*, которые в последнее время относят к группе сайт-специфических ДНК-связывающих белков, узнающих особые последовательности в ДНК. Среди негистоновых белков выделяют так называемые *белки высокоподвижной* (при электрофорезе) *группы* (HMG-белки). Известно, что некоторые из них (HMG-14 и HMG-17) присоединяются к ДНК в тех участках, где идет активная транскрипция генов. По-видимому, присоединение этих белков предшествует и способствует связыванию гормон-рецепторных комплексов с ДНК, что приводит к активации транскрипции (рис. 140).

Вслед за 30 нм-фибриллами в хроматине наблюдаются более высокие уровни конденсации: *петельные домены*, гигантские петли суперспирализованной ДНК (*конденсированный хроматин*) и наконец *метафазные хромосомы* (см. рис. 139).

В целом компактизация ДНК у эукариот представляет собой один из способов контроля активности генов, так как характер укладки ДНК влияет на транскрипционную активность генома. Различают плотно организованный (конденсированный) хроматин — *гетерохроматин*, и менее компактный *эухроматин*, с большей активностью транскрипции. Таким образом, у эукариот наследственная передача

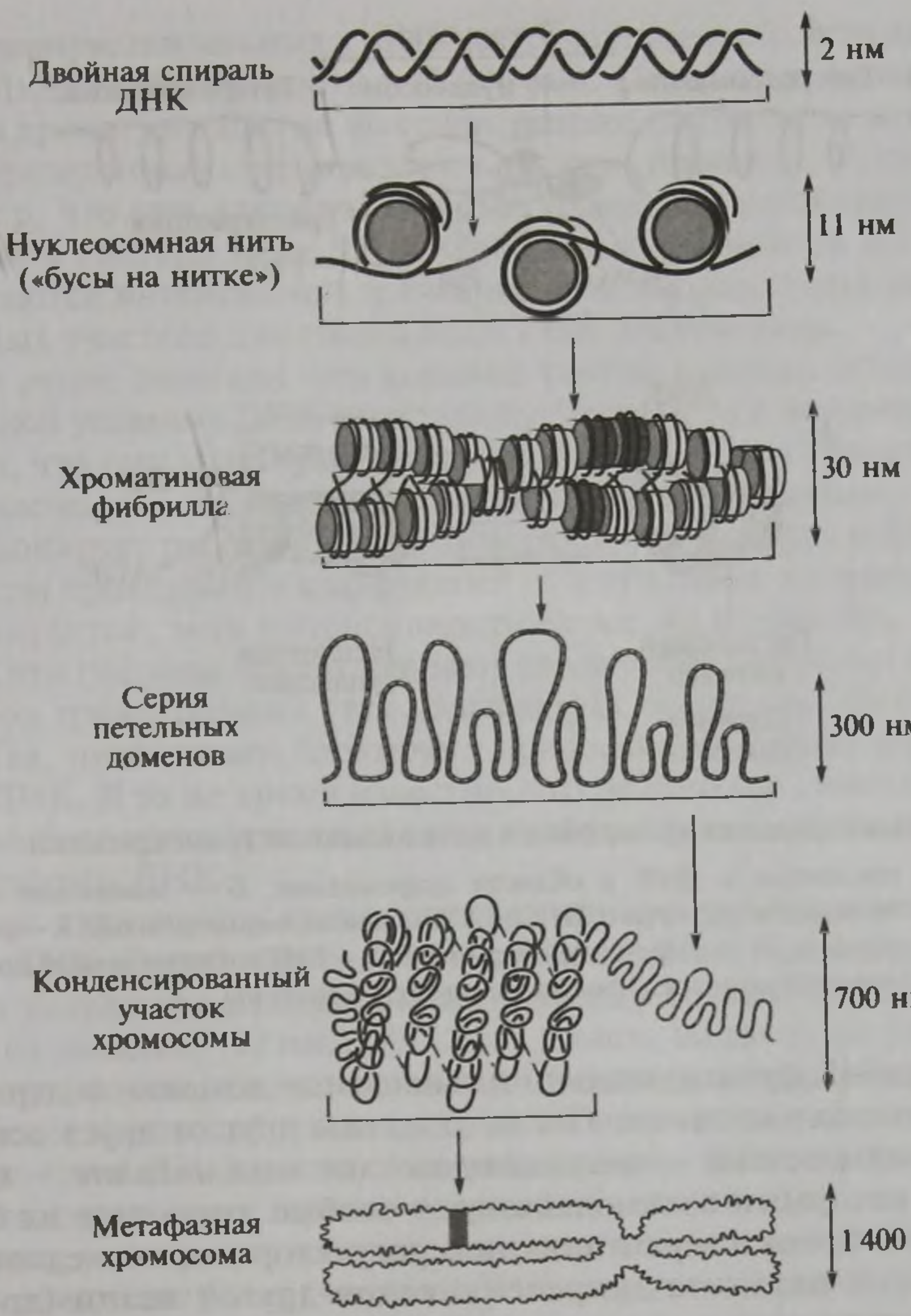


Рис. 139. Компактизация молекулы ДНК в составе хроматина (по Б. Албертсу и др., 1984)

информации о функциональном статусе генов может осуществляться в виде структурной организации хроматина. Эта особенность четко выражена в инактивации Х-хромосом у млекопитающих: известно, что женский тип у них определяется наличием двух Х-хромосом, одна из которых инактивируется (т.е. содержащаяся в ней ДНК не транскрибируется). У мышей инактивация происходит на 5—6-й день эмбрионального развития, при этом Х-хромосома является полностью конденсированной. Неактивная структура одной из Х-хромосом устойчиво наследуется.

В составе хроматина гены эукариот включены в петельные домены различного размера (длиной от 20 до 200 тыс. н.п.). Эти петли пред-

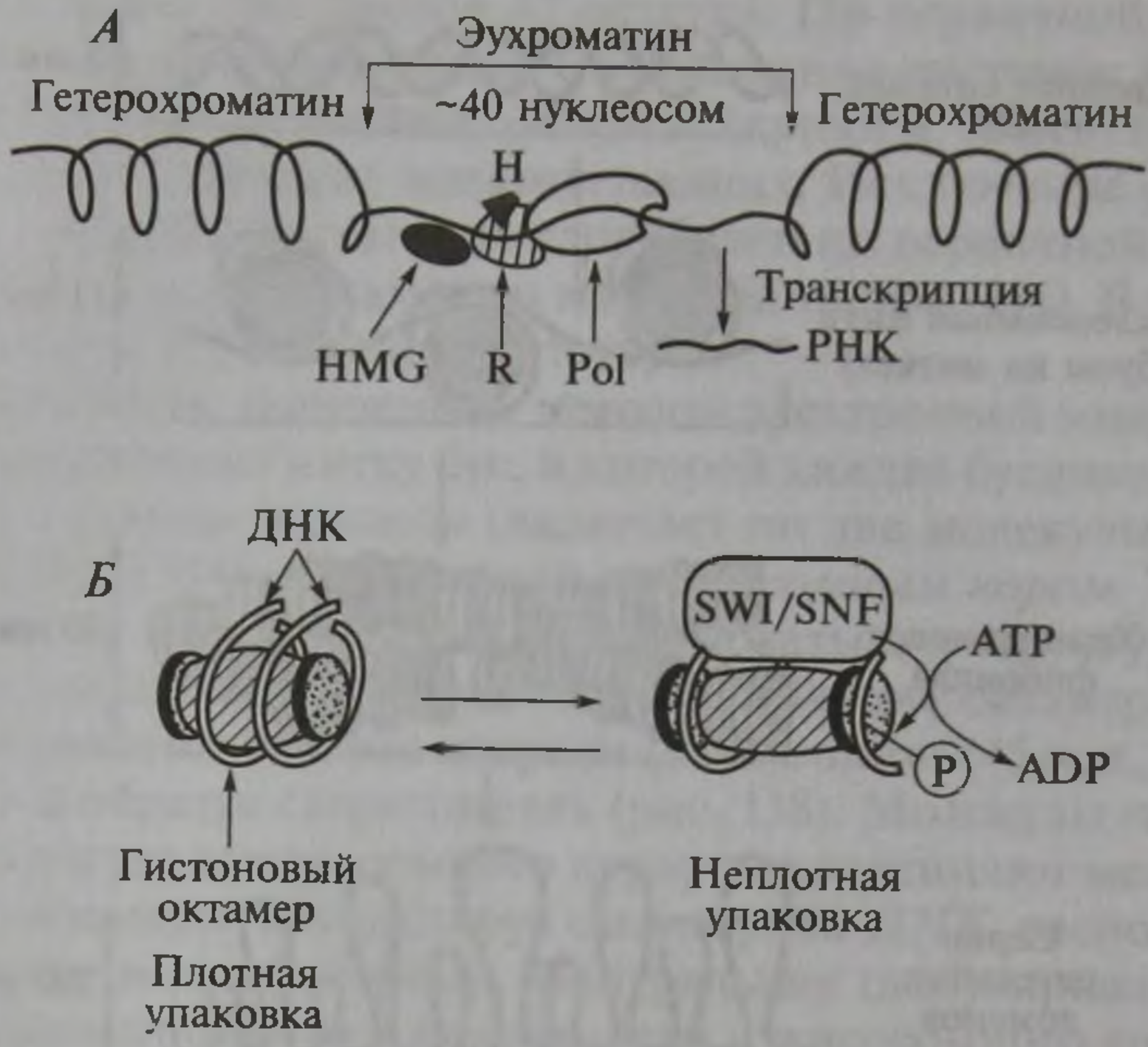


Рис. 140. Деконденсация хроматина в зоне активной транскрипции:

А — белки, связанные с ДНК в области эухроматина; *Б* — изменение степени (прочности) связывания гистонов с ДНК по АТР-зависимому механизму; R — рецептор стероидного гормона; H — стероидный гормон; Pol — РНК-полимеразный комплекс; SWI/SNF — белковый комплекс, фосфорилирующий гистоны

ставляют собой функционально независимые домены, содержащие один или несколько генов. Петли отделены друг от друга особыми последовательностями — *инсуляторами* (от англ. *insulate* — изолировать), с которыми взаимодействуют особые узнающие их белки. Инсуляторы препятствуют влиянию регуляторных последовательностей одной петли на экпрессию генов другой петли (домена) и способствуют целенаправленному воздействию энхансеров на транскрипцию генов определенного домена. Некоторые подвижные генетические элементы (элемент gypsy дрозофилы) содержат в себе нуклеотидные последовательности инсуляторов, и их внедрение в определенные участки хроматина влияет на экспрессию генов. Перемещение таких подвижных элементов может изменять границы доменов, экспрессия которых контролируется определенным энхансером, а внедрение между промотором и оператором может вообще блокировать транскрипцию.

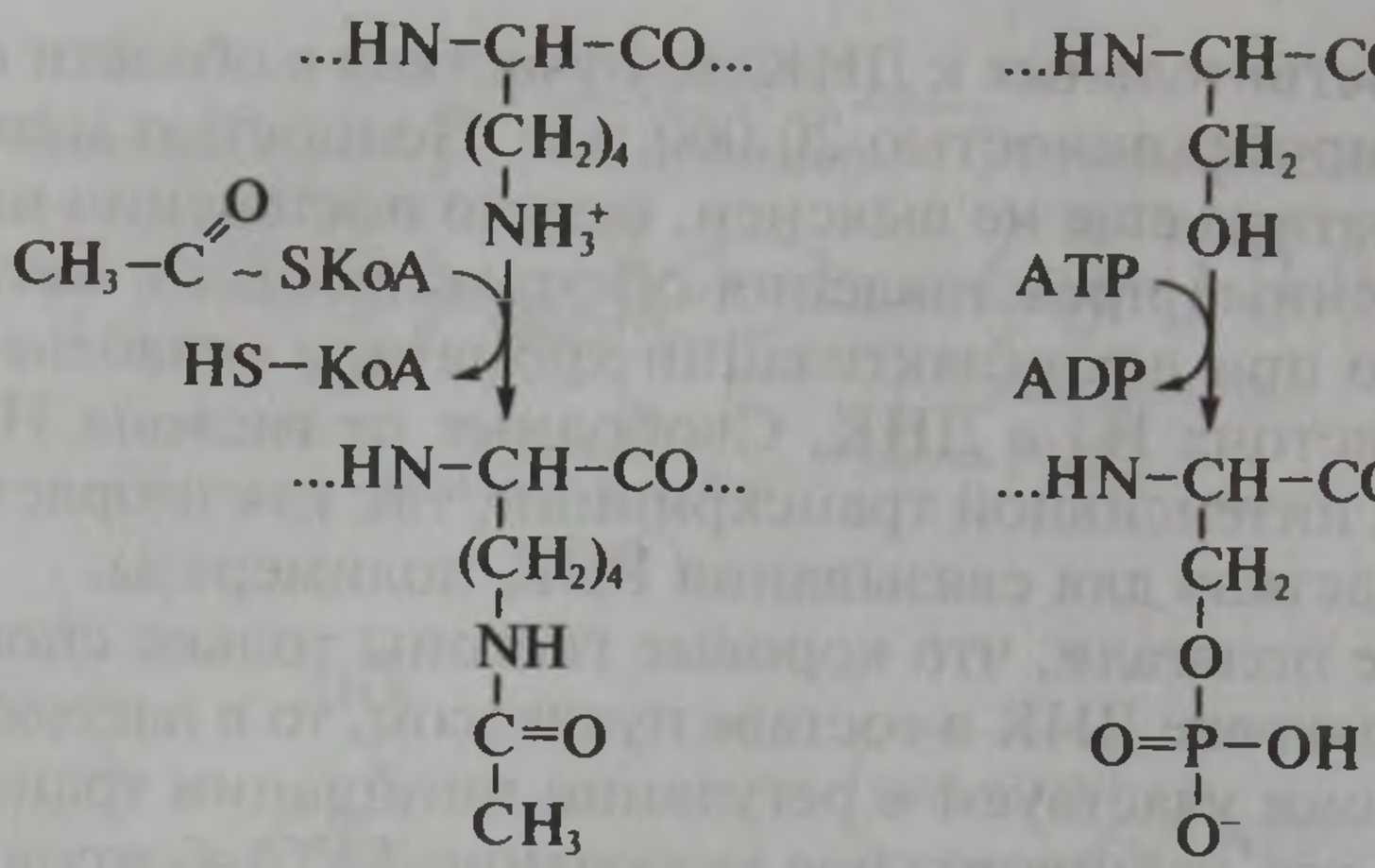
При активации эукариотических генов происходит избирательная деконденсация хроматина. Это продемонстрировано в экспериментах с ДНКазой I — ферментом, способным атаковать ДНК вне суперспирализованных участков. Обработка ДНКазой I вызывает деградацию только тех участков ДНК, которые составляют транскрибируемую часть генома эукариот (~10%). В гене лизоцима курицы выявлено

семь гиперчувствительных к ДНКазе I участков в области активного эухроматина протяженностью 20 000 н.п. Полностью механизм активации хроматина еще не выяснен, однако постепенно накапливаются определенные представления об этом процессе. Установлено, например, что при декомпактизации хроматина ослабевает прежде всего связь гистона H1 с ДНК. Свободные от гистона H1 области подвергаются интенсивной транскрипции, так как возрастает число доступных участков для связывания РНК-полимеразы.

Если ранее полагали, что коровые гистоны только способствуют регулярной упаковке ДНК в составе нуклеосом, то в настоящее время считают, что они участвуют в регуляции инициации транскрипции. Установлено, что избирательное включение ТАТА-боксов в нуклеосомы блокирует работу РНК-полимеразы, однако если образование нуклеосом происходит в кодирующей области генов, то транскрипция не блокируется, хотя интенсивность ее все же снижается. Предполагают, что гистоны предотвращают связывание основных белковых факторов транскрипции (так называемых *базальных факторов*) с ДНК. Так, нуклеосомы блокируют присоединение фактора TFIID и TBP к ДНК. В то же время известно, что рецепторы глюкокортикоидов способны присоединяться как к свободной, так и организованной в нуклеосомы ДНК.

Белки-активаторы транскрипции конкурируют с гистонами за связь с ДНК. Ослабление связи гистонов с ДНК происходит скорее всего в результате взаимодействия между белком-активатором и одним из доменов гистона H4. Этот домен, видимо, не участвует в сборке нуклеосом, имеет нитевидную (неспирализованную) форму и способен к взаимодействию с окружающими молекулами. Ацетилирование этого (N-концевого) домена гистона H4 ведет к снижению его положительного заряда, изменению его конформации и, как следствие, — к временному отсоединению от гистонов H2A и H2B. Далее происходит раскручивание ДНК вокруг оставшихся гистонов и ДНК становится доступной для базальных факторов транскрипции. Нитевидные хвосты гистонов H4 и H3 могут участвовать и в репрессии транскрипции, что показано, в частности, для генов, кодирующих типы спаривания у дрожжей. N-концевые участки («хвосты») гистонов H3 и H4 наиболее консервативны в эволюционном плане и представляют собой объекты ряда ковалентных модификаций, которые могут приводить к изменению взаимодействия этих гистонов с другими компонентами нуклеосом, а также с самой ДНК и ДНК-связывающими белковыми факторами транскрипции.

Среди ковалентных модификаций *ацетилирование* гистонов, по всей видимости, наиболее важно для тотальной регуляции транскрипции генов эукариот. Этот процесс катализируется гистонацетилтрансферазой (HATS)-ферментом, способным переносить ацетильные остатки с молекулы ацетил-коэнзима А на остатки лизина в молекулах гистонов и тем самым снижать их положительный заряд (рис. 141).



А *Б*

Рис. 141. Ковалентные модификации гистонов, ослабляющие их связь с ДНК:

А — ацетилирование, снимающее положительный заряд с остатков лизина; *Б* — фосфорилирование радикалов серина с образованием отрицательно заряженного остатка фосфосерина

Установлена тесная взаимосвязь между ацетилированием гистонов, их взаимодействием с ДНК и транскрипцией. Один из компонентов белкового комплекса TFIID у человека содержит белок TAFII250, обладающий ацетилтрансферазной активностью, что указывает на необходимость ацетилирования для начала транскрипции. Ацетилирование гистонов имеет обратимый характер: для их деацетилирования, вероятно, необходим особый белковый комплекс деацетилирования. Само ацетилирование также стимулируется особым белковым комплексом (НАТ-комплекс), связывающимся с линкерной ДНК вблизи сайта присоединения РНК-полимеразы.

Фосфорилирование гистонов представляет собой еще один путь активации транскрипции. Этот АТР-зависимый механизм, в основе которого, очевидно, лежит фосфорилирование радикалов серина коровых гистонов (см. рис. 141).

В фосфорилировании гистонов принимает участие специфический белковый комплекс, итог работы которого состоит в ослаблении связи гистонов с ДНК. Менее плотная упаковка ДНК в составе нуклеосом после фосфорилирования гистонов создает условия для успешной инициации и элонгации транскрипции. В ходе элонгации транскрипции, осуществляемой РНК-полимеразой II, нуклеосомы располагаются перед и после работающей полимеразы, что указывает на возможность перемещения нуклеосом либо их разрушение (разворачивание) и реконструкцию после прохождения фермента (рис. 142). Белковый комплекс SWI/SNF, задействованный в фосфорилировании гистонов (см. рис. 140), участвует в преодолении РНК-полимеразой нуклеосомного барьера (т. е. в элонгации транскрипции).

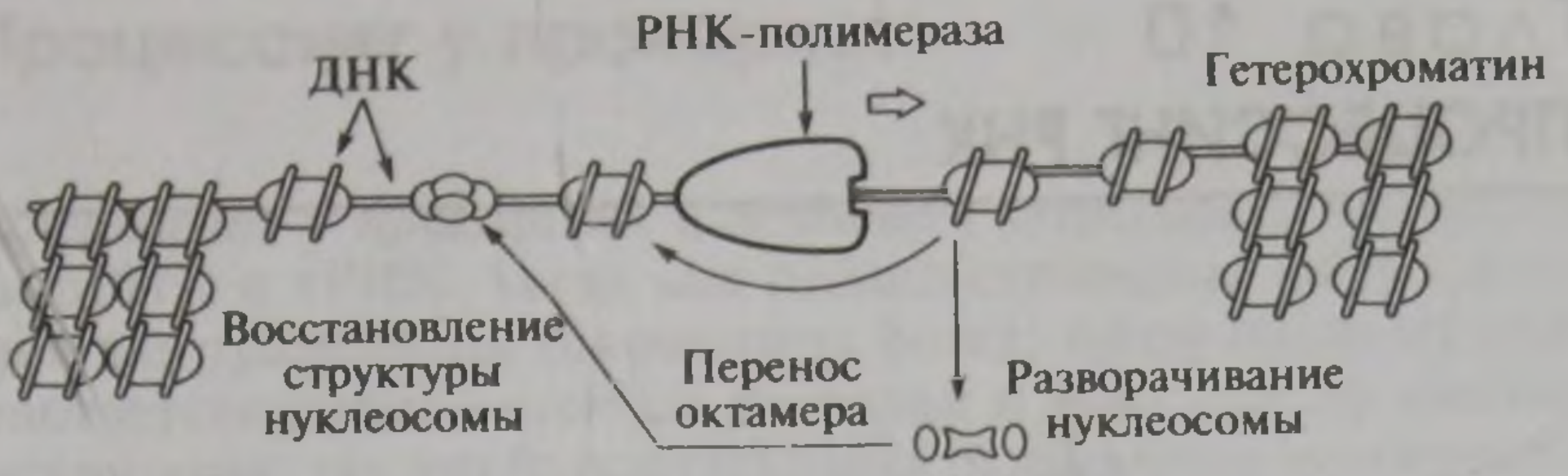


Рис. 142. Гипотетическая модель, иллюстрирующая изменения в нуклеосомной организации хроматина в период транскрипции. Пояснения см. в тексте

В последние годы установлено, что у дрожжей, растений и млекопитающих модификации гистонов могут регулироваться с участием малых интерферирующих РНК. Эти РНК способны также направлять процесс метилирования ДНК, что приводит к образованию гетерохроматина и репрессии транскрипции определенных генов.

Глава 10
ПРОЦЕССИНГ РНК

Образующиеся в результате транскрипции первичные РНК-транскрипты поначалу далеко не всегда представляют собой функционально активные молекулы РНК. Прежде чем стать таковыми, они, как правило, должны претерпеть ряд модификаций, которые и превращают их в зрелые РНК, пригодные к выполнению соответствующих функций. Этот процесс посттранскрипционной модификации первичных транскриптов (РНК-предшественников) называется *процессингом* и носит достаточно специфический характер в отношении разных видов РНК у про- и эукариот.

В связи с известными вышеотмеченными различиями в этапах реализации генетической информации у прокариот и эукариот процессинг у прокариот имеет весьма ограниченный характер, тогда как у эукариот он представляет собой достаточно тонко организованный процесс, непосредственно влияющий на регуляцию экспрессии генетического материала в дифференцированных клетках высших организмов. Наиболее детально изучен процессинг мРНК эукариот, который включает три главных момента: сплайсинг, кэпирование 5′-конца и полиаденилирование 3′-конца первичных транскриптов.

Сплайсинг (от англ. *splice* — соединять концы) состоит в вырезании из пре-мРНК некодирующих областей — *интронов* — и сшивании кодирующих структуру белка участков — *экзонов*. Различают несколько различных механизмов сплайсинга (см. ниже). *Кэпирование* представляет собой образование на 5′-конце мРНК особой структуры — кэпа (шапочки) (см. гл. 3). Кэпирование происходит вскоре после начала синтеза мРНК и осуществляется с участием GTP, из состава которого GMP переносится на 5′-дифосфат первого нуклеотида мРНК. *Полиаденилирование* осуществляется ферментом поли(А)-полимеразой и приводит к образованию на 3′-конце олиго(А)-фрагмента, содержащего 100—200 остатков адениловой кислоты подряд и называемого также «поли(А)-хвостом». Поли(А)-хвост определяет стабильность мРНК и время ее жизни в клетке. Кроме того, у эукариот поли(А)-хвост возможно способствует выходу мРНК из ядра в цитоплазму, а также существен для регуляции трансляции мРНК (см. гл. 11).

Процессинг у прокариот

Процессинг у прокариот в основном затрагивает предшественники рРНК и тРНК, тогда как полицистроновые мРНК используются для трансляции (биосинтеза белка) сразу после их синтеза на соответствующих оперонах бактерий и даже еще до окончания транскрипции, так что биосинтез белка оказывается непосредственно сопряжен с транскрипцией (см. рис. 122). Ранее считали, что полиаденилирование свойственно только мРНК эукариот. Однако в настоящее время установлено, что поли(А)-хвосты есть и у мРНК прокариот, но они короче, чем у мРНК эукариот (содержат от 14 до 60 остатков адениловой кислоты) и встречаются не у всех молекул мРНК. В отличие от эукариот полиаденилирование дестабилизирует структуру ряда мРНК бактерий.

Процессинг у бактерий необходим для образования зрелых молекул рРНК и тРНК. Первичный транскрипт, содержащий последовательности этих РНК, образуется на оперонах, включающих гены 16S рРНК, 23S рРНК, 5S рРНК и тРНК (рис. 143). Первичные транскрипты таких оперонов имеют длину ~5 000 нуклеотидов и помимо последовательностей рРНК содержат вставочные последовательности (*спейсеры*), существенные для процессинга. Вычленение индивидуальных рРНК происходит с участием эндонуклеаз, разрезающих первичные транскрипты в областях спейсеров. Так, РНКаза III узнает и расщепляет двунитевые участки (дуплексы), образующиеся за счет спейсеров и составляющие стебли петель (шпилек) в структуре предшественников. В петлях этих шпилек содержатся последовательности зрелых молекул 16S рРНК (~1 600 нуклеотидов) и 23S рРНК (~2 900 нуклеотидов). Другой фермент — РНКаза М5, идентифицированный у *Baccillus subtilis*, отщепляет от первичного транскрипта 5S рРНК. Образующиеся после первичного эндонуклеотического расщепления пре-тРНК подвергаются дальнейшему процессингу при участии двух ферментов: эндонуклеазы Р и экзонуклеазы D (рис. 144). При этом пре-тРНК, содержащая более 100 нуклеотидных остатков, укорачивается с обоих концов до стандартного размера (70—90 нуклеотидов). С 5′-конца на пре-тРНК действует эндонуклеаза Р, а с 3′-конца — экзонуклеаза D. Интересно, что молекула эндонуклеазы Р (РНКазы Р) сама содержит в своем составе РНК длиной до 400 нуклеотидов, которая обладает основной каталитической активностью, а белковая часть этой нуклеазы (мол. масса 17 000) только усиливает РНКазную активность. Ферменты этого типа выявлены как у бактерий, так и у эукариот. У *E. coli* длина РНК в составе РНКазы Р достигает 377 нуклеотидов, у дрожжей — 250 нуклеотидов, а у человека — 100 нуклеотидов.

Методами генетической инженерии получен вполне активный в процессинге тРНК гибридный фермент, содержащий РНК-компонент

РНКазы Р, выделенный из *E. coli*, и белковую часть — из соответствующего фермента человека. Эндонуклеазное действие РНКазы Р заключается в отщеплении 41 нуклеотида от 5′-концевой (лидерной) последовательности предшественника и сразу приводит к форми-

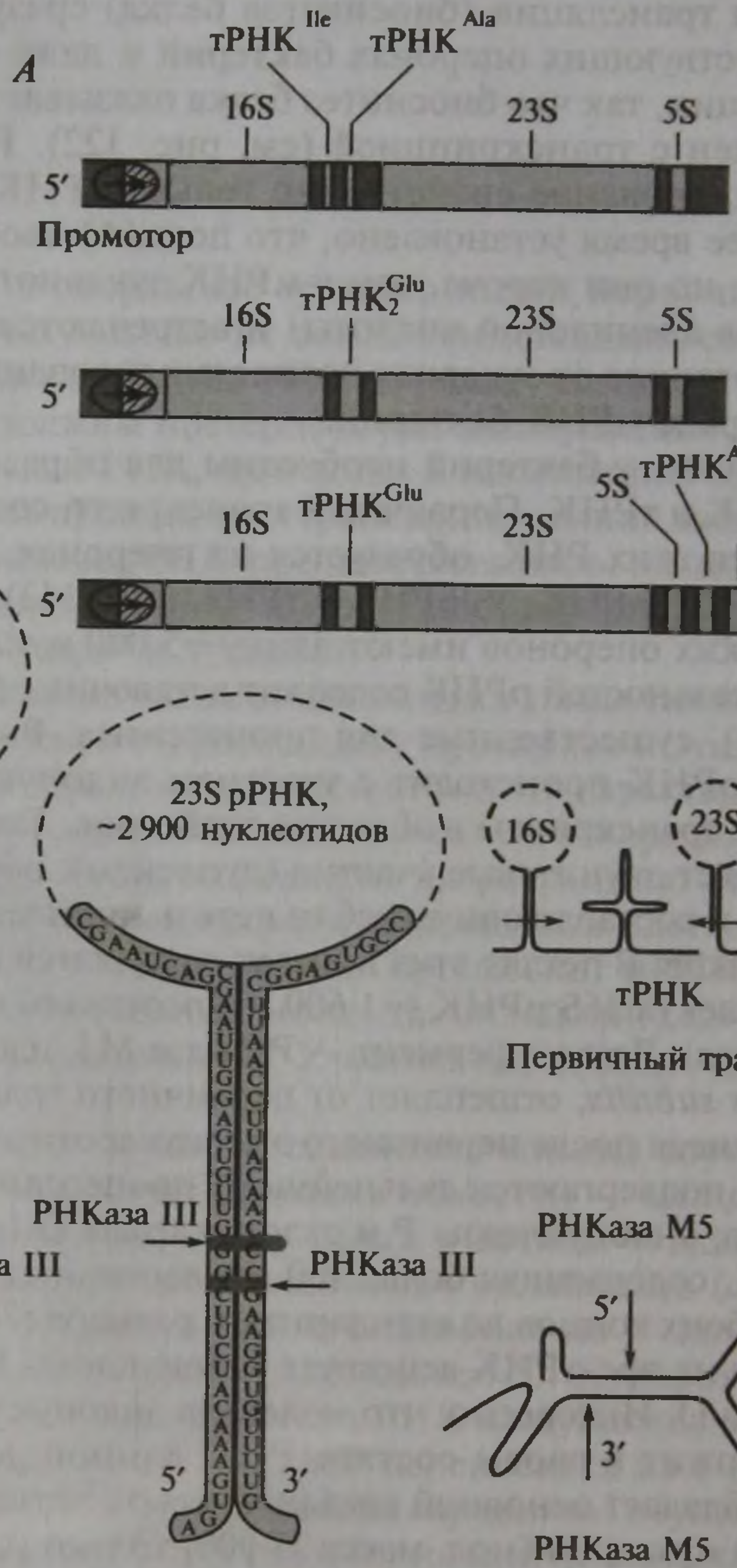


Рис. 143. Процессинг рРНК у прокариот (М. Сингер, П. Берг, 1998):

А — схема типичных единиц транскрипции рРНК *E. coli*. Между кодирующими участками для 16S, 23S и 5S рРНК и тРНК расположены спейсерные последовательности; *Б* — спейсерные последовательности, расположенные в предшественнике рРНК, образуют черешки (стебли), которые расщепляются РНКазой III; *В* — РНКаза М5 отщепляет от первичного транскрипта 5S РНК

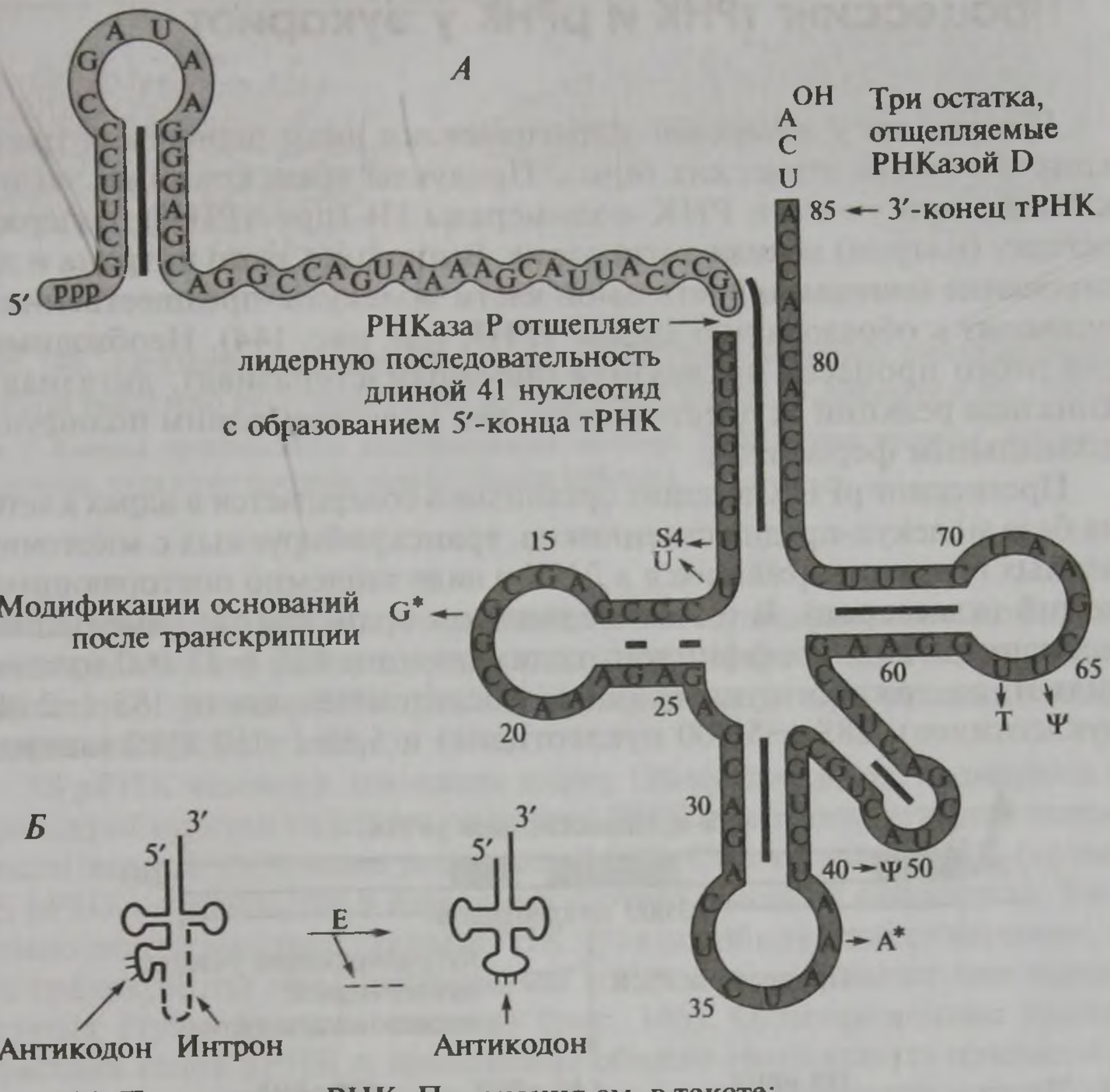


Рис. 144. Процессинг тРНК. Пояснения см. в тексте:

А — процессинг тРНК из молекулы-предшественника $тРНК^{Tyr}$ *E. coli* (М.Сингер, П.Берг, 1998); *Б* — удаление интрона из предшественника тРНК у эукариот; Е — полифункциональный фермент, обладающий киназной, лигазной и фосфодиэстеразной активностью. Звездочками обозначены модифицированные основания, U^{S4} — тиоурацил

рованию 5′-конца тРНК, представленного фосфорилированным гуанозином (5′-pG). Экзонуклеаза D, со своей стороны, отщепляет последовательно несколько нуклеотидов с 3′-конца молекулы-предшественника, останавливаясь на аденозине, соседствующем с двумя цитидиловыми нуклеотидами, и таким образом формируется 3′-конец (акцептирующий стебель) тРНК, содержащий стандартную последовательность CCA-OH. Одновременно с процессингом рРНК и тРНК осуществляется модифицирование ряда азотистых оснований, приводящее, в частности, к образованию характерных для тРНК дигидроурацила (UH2), псевдоуридина (Ψ) и тимина (T). Ферменты, осуществляющие эти модификации, до сих пор не охарактеризованы. Эти модификации называют *редактированием РНК* (см. с. 297).

Процессинг тРНК и рРНК у эукариот

Процессинг у эукариот затрагивает все виды первичных транскриптов эукариотических генов. Продукты транскрипции, возникающие при участии РНК-полимеразы III (пре-тРНК), содержат вставку (интрон) вблизи антикодона. Вырезание этого интрона и лигирование (сшивание) остальной части молекулы-предшественника приводит к образованию зрелой тРНК (см. рис. 144). Необходимые для этого процесса нуклеазная (фосфодиэстеразная), лигазная и киназная реакции осуществляются, как полагают, одним полифункциональным ферментом.

Процессинг рРНК высших организмов совершается в ядрах клеток на базе молекул-предшественников, транскрибируемых с многочисленных генов, содержащихся в ДНК в виде тандемно повторяющихся копий (кластеров). В одном первичном транскрипте, имеющем у млекопитающих коэффициент седиментации 45S (~13 000 нуклеотидов), содержатся нуклеотидные последовательности 18S (~2 000 нуклеотидов), 28S (~5 000 нуклеотидов) и 5,8S (~160 нуклеотидов)

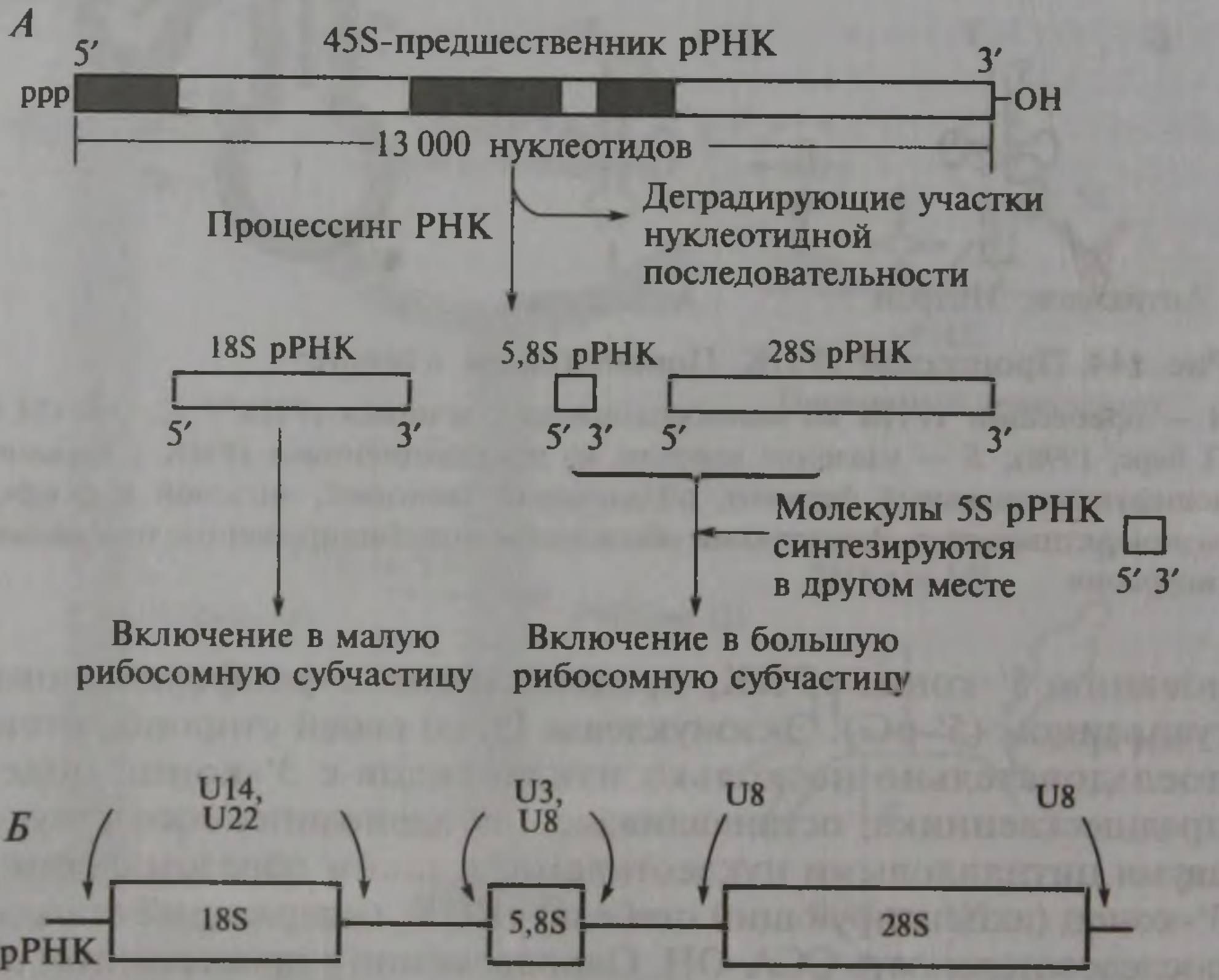


Рис. 145. Процессинг рРНК у эукариот:

А — процессинг 45S-предшественника рРНК (спейсерные последовательности на рисунке темные); *Б* — взаимодействие различных U-РНК с пограничными участками спейсерных последовательностей в молекуле-предшественнике рРНК у лягушек *Xenopus laevis*

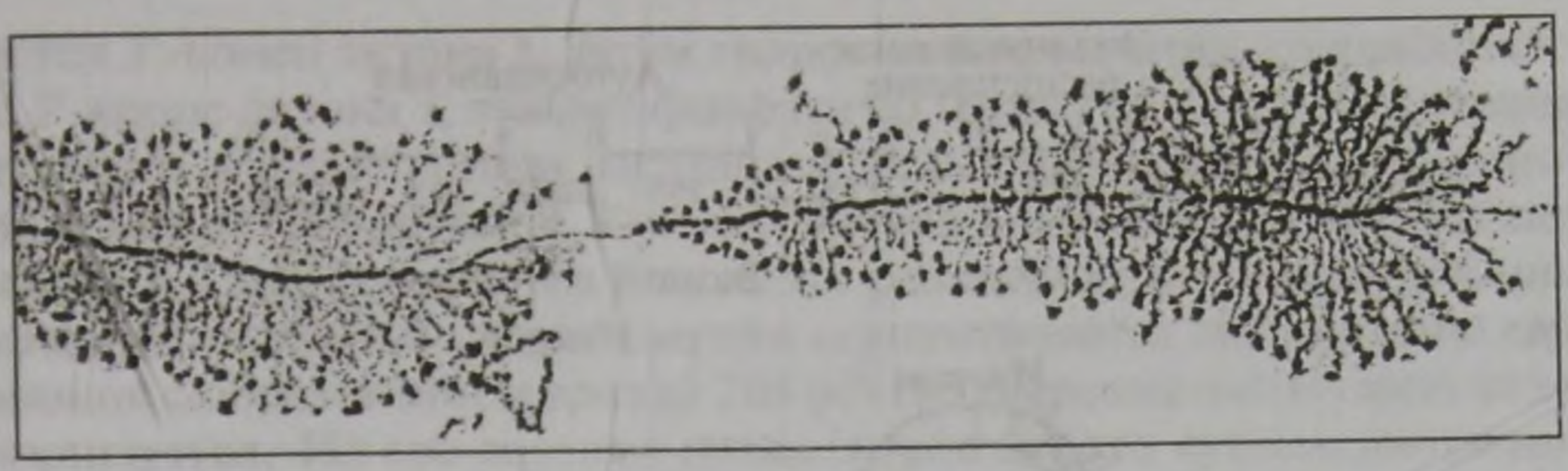


Рис. 146. Электронная микрофотография РНК-транскриптов, синтезированных на тандемно расположенных в ДНК генах рРНК (Б.Албертс и др., 1994)

На 5′-концах прогрессивно удлиняющихся молекул РНК видны частицы, которые, вероятно, отражают процесс начала сборки рибосом

рРНК, которые разделяются спейсерными последовательностями. Эндонуклеазное расщепление этого предшественника ведет к выщеплению зрелых рРНК (рис. 145) и происходит при участии малых ядерных РНК (см. ниже).

5S рРНК человека, имеющая длину 120 нуклеотидов, кодируется и транскрибируется отдельно от других РНК с соответствующих генов, число которых в геноме достигает 2 000. Синтез всех рРНК (кроме 5S рРНК) происходит в ядрышках с очень высокой скоростью. Тандемно повторяющиеся гены рРНК транскрибируются синхронно, и их транскрипты под электронным микроскопом имеют вид характерных структур типа «елочек» (рис. 146). Одновременная транскрипция генов рРНК и процессинг общего транскрипта приводят к синхронному образованию пропорциональных количеств всех рРНК, необходимых для сборки рибосом.

У ряда видов примитивных эукариот (*Tetrahymena thermophila*, *Physarum polycephalum* и др.) гены рРНК содержат особые интроны (интроны группы I), для которых характерен уникальный механизм сплайсинга. Такие интроны встречаются также в генах рРНК митохондрий, хлоропластов, дрожжей и грибов, однако они не выявлены в генах позвоночных животных. Изучение процессинга 26S рРНК тетрахимены (аналог 28S рРНК высших эукариот), выполненное Т. Чеком и сотрудниками, привело к открытию особого вида сплайсинга, осуществляемого без участия каких-либо белков и получившего название *аутосплайсинг* (сплайсинг типа I).

Оказалось, что содержащаяся внутри 26S рРНК тетрахимены вставка (интрон) длиной 400 нуклеотидов способна сама осуществлять вырезание этого интрона и сшивание экзонов в присутствии Mg^{2+} и свободного гуанозина (либо его фосфорилированных производных). Таким образом была открыта *аутокаталитическая функция* РНК и положено начало изучению *РНК-ферментов — рибозимов*. Схема аутосплайсинга рРНК у тетрахимены представлена на рис. 147. Ин-

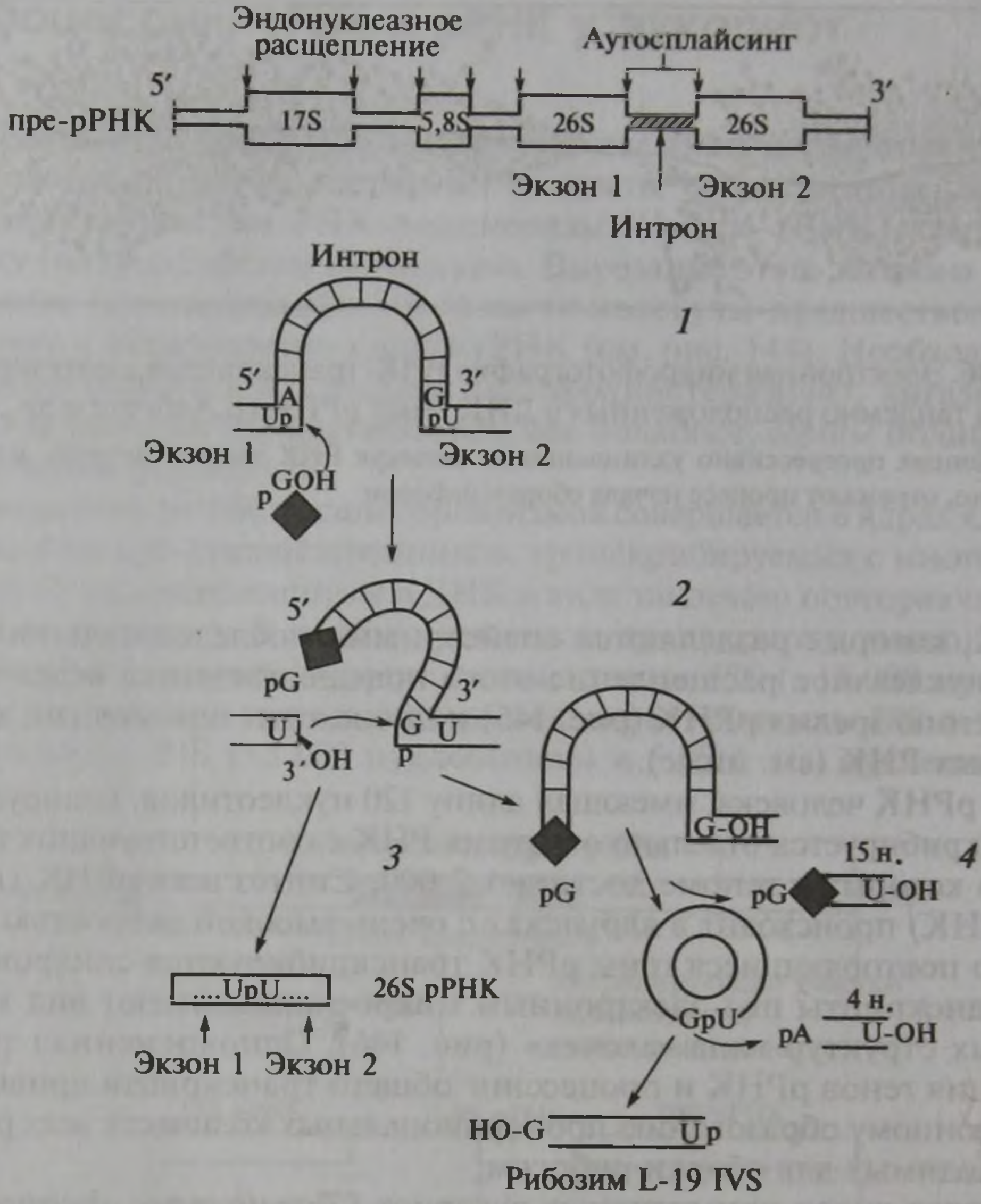


Рис. 147. Схема аутосплайсинга 26S рРНК у тетрахимены и процесс образования рибозима. В верхней части рисунка представлена структура предшественника рРНК. В последовательности 26S рРНК имеется вставка (интрон):

1 — гуанозин атакует 5′-конец интрона, в результате чего происходят разрыв межнуклеотидной связи и высвобождение 5′-конца интрона; *2* — гидроксильная группа 3′-конца экзона 1 атакует фосфодиэфирную связь на 3′-конце интрона, что ведет к вычленению интрона; *3* — замыкание фосфодиэфирной связи между ОН-группой 3′-конца экзона 1 и 5′-фосфатной группой экзона 2 приводит к сшиванию экзонов и образованию зрелой 26S рРНК; *4* — в результате двухэтапного аутокаталитического выщепления 19 нуклеотидов из состава интрона образуется рибозим, обладающий каталитическими свойствами (L –19 IVS)

трон 26S рРНК тетрахимены — IVS, состоит из 413 нуклеотидов. Аутосплайсинг индуцируется гуанозином, гидроксильная группа которого атакует фосфатную группу на 5′-конце интрона, в результате чего разрывается межнуклеотидная (фосфодиэфирная) связь и высвобож-

дается 3′-конец экзона 1. Затем гидроксильная группа, содержащаяся на 3′-конце экзона 1, атакует фосфатную группу на 3′-конце интрона, что ведет к вычленению интрона и замыканию фосфодиэфирной связи между ОН-группой 3′-конца экзона 1 и 5′-фосфатной группой экзона 2. Таким образом в результате реакции *трансэтерификации* без дополнительных затрат энергии осуществляется лигирование двух экзонов с образованием зрелой 26S рРНК. Вырезанный интрон затем циклизуется. Из его состава путем двухэтапного ауторасщепления освобождается фрагмент, содержащий 19 нуклеотидов, в результате чего образуется РНК длиной 376 нуклеотидов (L –19 IVS), которая и представляет собой истинный РНК-фермент (рибозим), обладающий каталитическими свойствами. Этот рибозим обладает устойчивой структурой, имеет эндонуклеазную активность, расщепляя длинные одноцепочечные РНК, и проявляет специфичность, распознавая в составе атакуемого субстрата тетрануклеотиды CUCU. Оказалось также, что рибозим L –19 IVS помимо нуклеазной обладает in vitro нуклеотидилтрансферазной (полимеразной) активностью и способен катализировать синтез олигонуклеотидов (олиго-С). Это указывает на возможность *аутокаталитической репликации* РНК и является одним из важных свидетельств в пользу существования «мира РНК» (см. гл. 3). В структуре интронов типа I выявлены характерные внутренние олигопуриновые последовательности (у тетрахимены это последовательность GGAGGG), называемые *адапторными* последовательностями, которые участвуют в образовании активного центра РНК-ферментов и выполняют важнейшую роль в каталитическом расщеплении РНК.

Детальные исследования природных РНК-ферментов послужили мощным стимулом к моделированию и синтезу рибозимов заданного строения. Эти работы позволили установить, что каталитической активностью обладают не только крупные РНК (~400 нуклеотидов у тетрахимены и РНКазы Р), но и короткие 13—20-членные олигонуклеотиды, которые могут быть синтезированы in vitro. Такие рибозимы стали называть *минизимами*. Одна из детально исследованных моделей функционирования таких рибозимов получила название «головка молотка» (рис. 148). Третичная структура «головки молотка» стабилизируется ионами двухвалентных металлов, которые нейтрализуют отрицательно заряженные атомы кислорода фосфодиэфирных связей и одновременно соединяют фосфатные группы ковалентными связями, что существенно для образования стабильного переходного состояния (фермент-субстратного комплекса). Как и в случае катализа, осуществляемого ферментами белковой природы, рибозимы и атакуемый субстрат (природные или синтетически полученные молекулы РНК) образуют фермент-субстратный комплекс, а затем — фермент-продуктный комплекс (см. рис. 148). Специально сконструированные рибозимы, содержащие необходимую для катализа адаптерную последовательность, способны, в частности, расщеплять

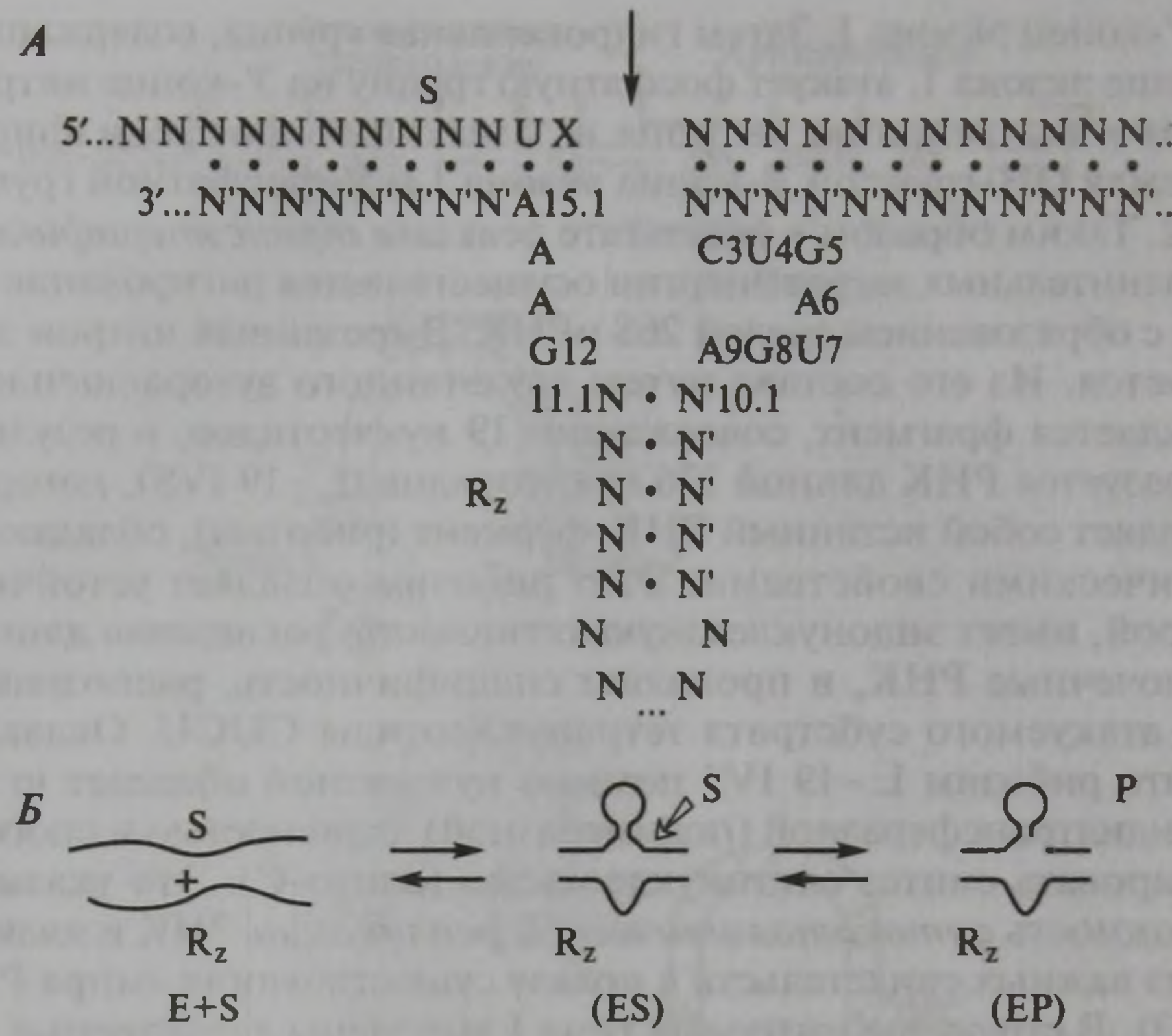


Рис. 148. Модель «головка молотка» (С. М. Женодарова, 1998):

А — общий вид комплекса, образующегося между рибозимом (R_z) и расщепляемым полирибонуклеотидным субстратом (S): N, N′ — любые комплементарные нуклеотиды; X — любой нуклеотид, кроме G. Стрелкой обозначено положение атакуемой связи в субстрате; *Б* — этапы реакции, катализируемой рибозимом, аналогичные стадиям реакции, ускоряемой ферментами белковой природы: ES — фермент-субстратный комплекс; EP — фермент-продуктный комплекс; *В* — специально сконструированный рибозим, расщепляющий РНК вируса иммунодефицита человека в области гена *gag*

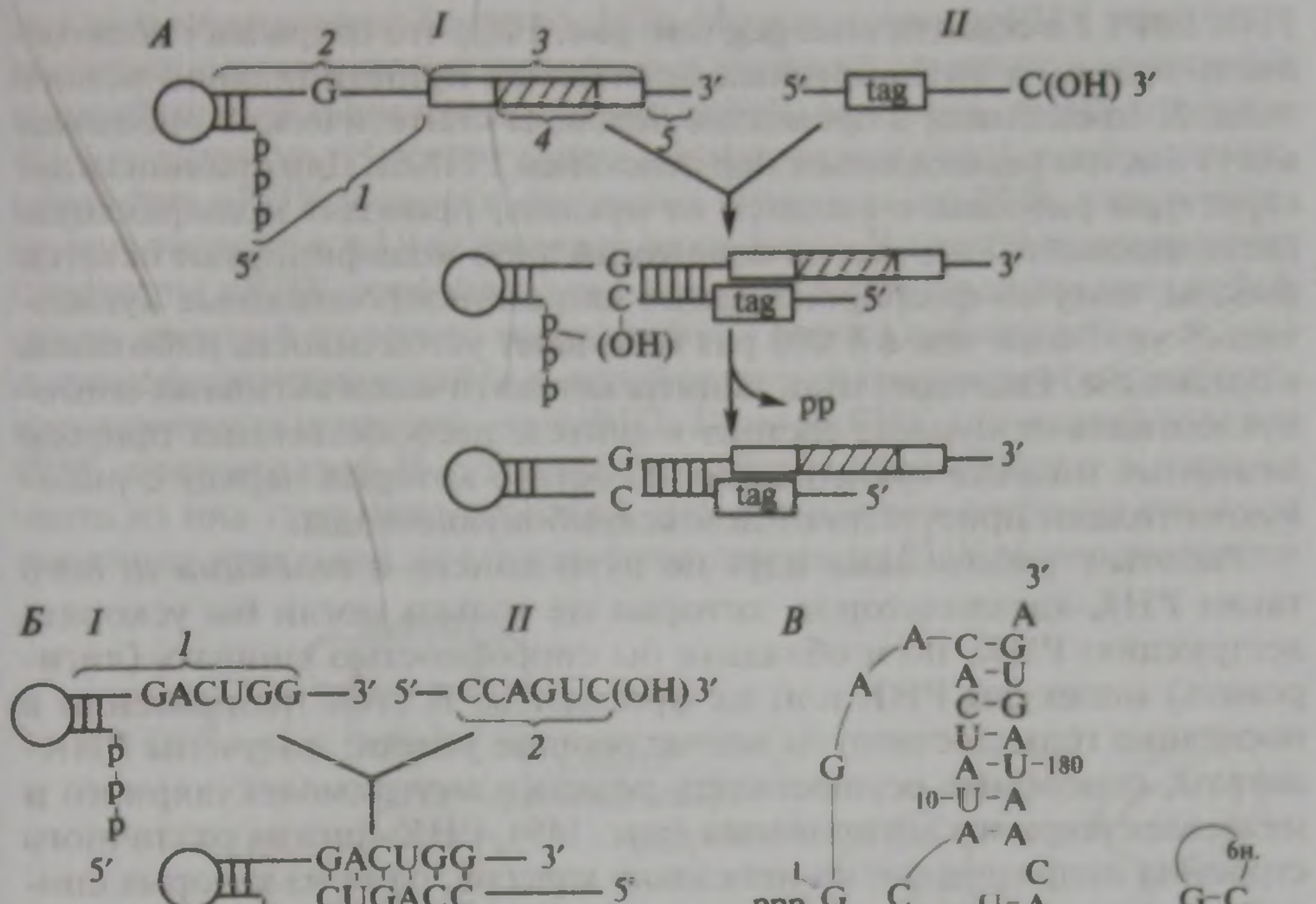


Рис. 149. Схема реакций лигирования РНК и структура рибозима класса 1, способного осуществлять лигирование с образованием 3′—5′-фосфодиэфирной связи (*Женодарова С.М., Клягина В.П.* // Молекулярная биология. — 1998. — Т. 32. — № 6):

А — схема реакции внутримолекулярного лигирования; *I* — РНК: *1* — константная 5′-область; *2* — матричный участок; *3* — область хаотической последовательности; *4* — каталитический домен; *II* — олигонуклеотидный субстрат: tag — определенная последовательность нуклеотидов, используемая как метчик; pp — пирофосфат; *Б* — схема реакции межмолекулярного лигирования; *I* и *II* — РНК-субстраты: *I* — матричный участок; *2* — праймер; *III* — комплекс субстратов; *IV* — каталитический домен РНК-лигазы; *V* — продукты лигирования; *В* — модель вторичной структуры лигазы класса 1. Линии после U-19 указывают место, где молекула рибозима может быть расщеплена для превращения в рибозим, катализирующий межмолекулярное лигирование; P1 — дуплекс между матричным участком и олигонуклеотидным субстратом

РНК ВИЧ 1 в области гена *gag* (см. рис. 148), что открывает возможность создания антивирусных препаратов принципиально нового типа. К сожалению, в организме человека синтетические рибозимы могут быстро расщепляться под действием РНКаз. Для стабилизации структуры рибозимов (защиты от нуклеаз) проводят модификацию (метилирование) азотистых оснований либо модифицируют остатки рибозы, получая фторпроизводные или аминопроизводные нуклеотидов, что более чем в 1 000 раз повышает устойчивость рибозимов в организме. Еще один путь защиты каталитически активных олигонуклеотидов от нуклеаз состоит в синтезе несвойственных природе химерных молекул *нуклеозимов*, в составе которых наряду с рибонуклеотидами присутствуют дезоксирибонуклеотиды.

Работы с рибозимами идут по пути поиска и *селекции in vitro* таких РНК-катализаторов, которые не только могли бы ускорять деструкцию РНК, но и обладали бы способностью сшивать (лигировать) молекулы РНК или их фрагменты. В этом направлении в последние годы достигнуты впечатляющие успехи: получены РНК-лигазы, способные осуществлять реакции внутримолекулярного и межмолекулярного лигирования (рис. 149). РНК-лигазы различного строения подразделяют на несколько классов, одни из которых способны катализировать образование 3′—5′-фосфодиэфирных связей (лигазы класса 1), а другие — 2′—5′-фосфодиэфирных связей. Все эти ферменты получают путем транскрипции, отбора (селекции) и амплификации специально сконструированных молекул ДНК. Эти молекулы ДНК имеют центральную область длиной ~200 нуклеотидов, расположенных случайным (хаотическим) образом, и служат для синтеза in vitro набора (пула) различных по структуре РНК, среди которых ведется отбор рибозимов.

Процессинг мРНК у эукариот

Посттранскрипционные модификации пре-мРНК особенно существенны для эукариот, что связано с *мозаичным строением их генов*, содержащих некодирующие последовательности — интроны. Число интронов в генах эукариот может достигать нескольких десятков, что диктует необходимость существования четкой системы сплайсинга первичных транскриптов. В сплайсинге пре-мРНК у высших эукариот задействован ряд белков, а также РНК особого вида — *малые ядерные РНК* (мяРНК). Малые ядерные РНК имеют последовательности протяженностью от 65 до 1000 и более нуклеотидов (10S—90S), богатые уридиловыми нуклеотидами, и поэтому называются также uРНК (U1, U2 и т.д.). У дрожжей выявлено 25 различных мяРНК, у позвоночных животных — 15. У шпорцевых лягушек *Xenopus laevis* ряд мяРНК (U3, U8, U14 и U22) участвуют в процессинге рибосомальных РНК, связываясь с пограничными участками спейсерных

последовательностей (см. рис. 145). Малые ядерные РНК выявлены не только у позвоночных животных и дрожжей, но также у насекомых и архибактерий. Они представляют собой, вероятно, очень древнюю группу молекул. Нуклеотидная последовательность всех соответствующих uРНК эукариот совпадает более чем на 90%, что, в частности, относится к U1 человека и дрозофилы. Высокий консерватизм структуры uРНК говорит о том, что сплайсинг представляет собой очень древний процесс, начавшийся с аутосплайсинга (см. выше) и трансформировавшийся в сплайсинг с участием особых рибонуклеопротеидных частиц — мяРНП. Гены мяРНК транскрибируются РНК-полимеразой II и имеют различную локализацию в геноме: часть из них представляет собой дискретные независимые гены, не имеющие интронов, тогда как гены других мяРНК *располагаются*

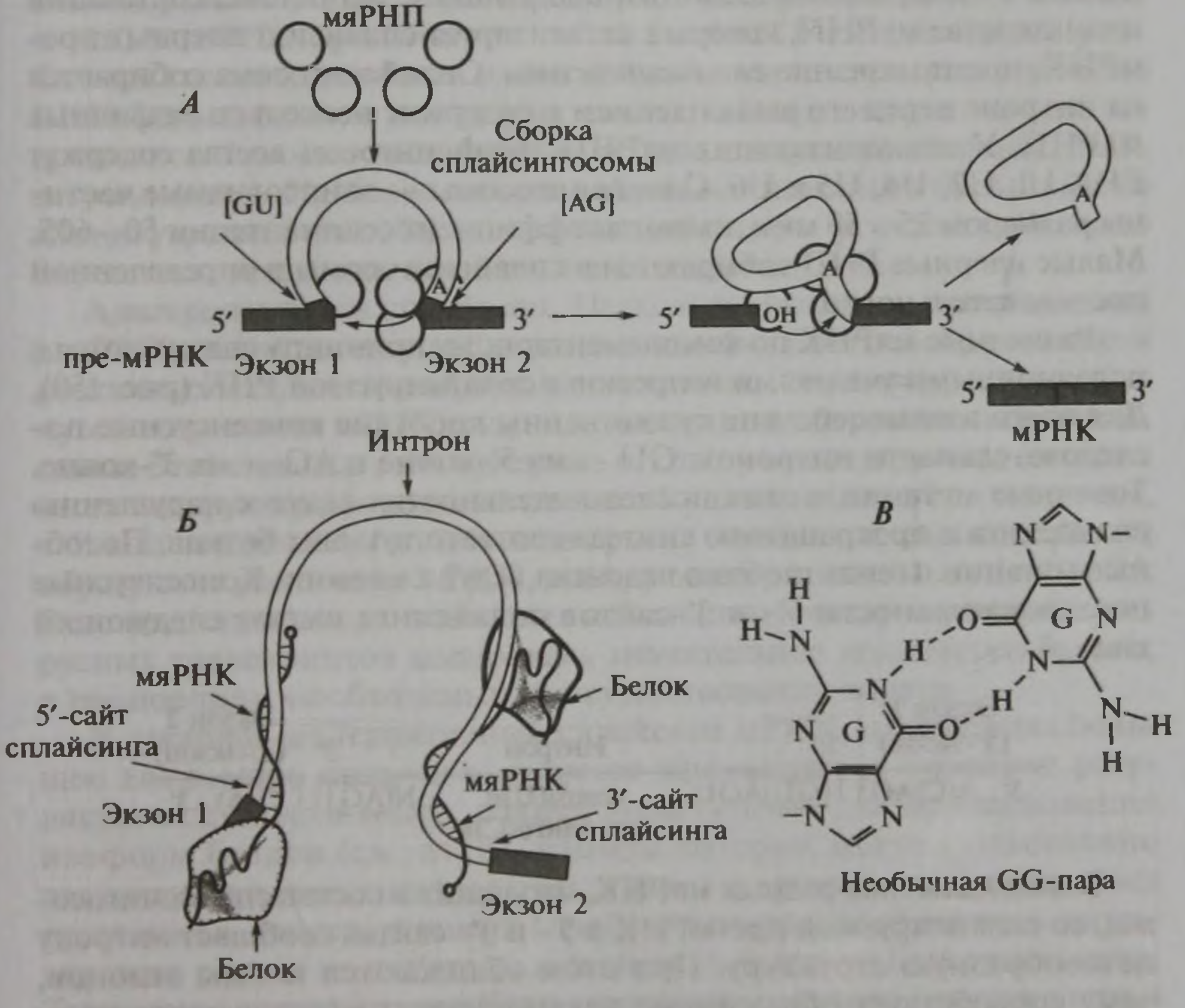


Рис. 150. Сплайсинг предшественника мРНК с участием мяРНП:

А — этапы сплайсинга с участием сплайсингосом. Пояснения см. в тексте. Кругами обозначены отдельные мяРНП; GU и AG — консенсусные пары нуклеотидов интрона, расположенные в 5′- и 3′-сайте сплайсинга соответственно; *Б* — схема комплементарного взаимодействия мяРНК, входящих в состав мяРНП, с пограничными участками интрона; *В* — необычные G-G-взаимодействия, способствующие сближению 5′- и 3′-концов интрона

внутри интронов генов, кодирующих белки. Так, у *Xenopus* U13 кодируется тремя уникальными последовательностями, находящимися в интронах 5, 6 и 8 генов белков теплового шока, а ген U16 находится внутри интрона рибосомального белка L1. Последнее обстоятельство имеет важное значение, так как показывает, что процессинг рРНК и процессинг мРНК белков рибосом может быть скоординирован при участии мяРНК. Кроме того, предполагают, что мяРНК способны служить *РНК-шаперонами*, участвуя в *фолдинге рРНК*, т.е. помогая ей принять необходимую структуру в пространстве.

Малые ядерные РНК присутствуют в ядрах в комплексах с белками, получившими название *малые рибонуклеопротеиновые частицы* (мяРНП). Стабильным компонентом мяРНП является белок фибриллярин — очень консервативный по структуре белок с молекулярной массой 34 кДа, локализованный в ядрышках. Комплекс, состоящий из множества мяРНП, который катализирует сплайсинг ядерных промРНК, носит название *сплайсингосомы*. Сплайсингосома собирается на интроне перед его выщеплением и содержит несколько различных мяРНП. У млекопитающих мяРНП сплайсингосом всегда содержат РНК U1, U2, U4, U5 и U6. Сплайсингосомы — элипсовидные частицы размером 25 × 50 мкм, имеют коэффициент седиментации 50—60S. Малые ядерные РНП собираются в сплайсингосомы в определенной последовательности.

Различные мяРНК по комплементарному принципу связываются с пограничными участками интронов в сплайсируемой РНК (рис. 150). Для этого взаимодействия существенны короткие консенсусные последовательности интронов: GU — на 5′-конце и AG — на 3′-конце. Точечные мутации в этих последовательностях ведут к нарушению сплайсинга и прекращению синтеза соответствующих белков. Подобные мутации в генах глобина человека ведут к анемии. Консенсусные последовательности 5′- и 3′-сайтов сплайсинга имеют следующий вид:

Экзон 1 (5′-экзон)	5′	Интрон	3′	Экзон 2 (3′-экзон)
5′...A(C)AG	GU AGU...	олиго(U)*n* или олиго(C)*n*	...N AG	G(A)...3′

Взаимодействие разных мяРНК, входящих в состав сплайсингосомы, со сплайсируемой пре-мРНК в 5′- и 3′- сайтах сообщает интрону петлеобразную структуру. При этом сближаются концы экзонов, чему способствует образование неканонических (отличающихся от уотсон-криковских пар) водородных связей между двумя гуанинами, содержащимися в 5′- и 3′-сайтах сплайсинга (см. рис. 150). Сближение экзонов создает условие для атаки 5′-конца интрона адениловым нуклеотидом, расположенным вблизи 3′-конца. В результате разрыва фосфодиэфирной связи между экзоном 1 и 5′-концом интрона последний взаимодействует с адениловым нуклеотидом и образованием

в интроне петли типа «лассо» (см. рис. 150). Вслед за этим освободившийся 3′-ОН-конец экзона 1 разрезает 3′-сайт сплайсинга, вырезает интрон и, соединяясь с экзоном 2, образует в итоге зрелую (сплайсированную) молекулу мРНК.

Полагают, что сплайсинг, катализируемый сплайсингосомами, возник из аутосплайсинга. Об этом свидетельствует сходство последовательности реакций, приводящих к вырезанию интронов: образование петлеобразной структуры в результате сближения двух экзонов; высвобождение 5′-конца интрона (под действием активного гуанилового нуклеотида в сплайсинге интронов типа I, и то же — под действием аденилового нуклеотида в сплайсинге интронов типа II); атака высвободившегося 5′-конца экзона фосфодиэфирной связи в 3′-сайте сплайсинга; высвобождение 3′-конца интрона и его существование в петлеобразной форме. С учетом широко распространенного мнения о том, что РНК представляла собой главный биополимер в первобытных клетках, сплайсинг в этих клетках, вероятно, катализировался РНК-ферментами, а некоторые самосплайсирующие интроны сохранились и до настоящего времени (как, например, интроны рРНК тетрахимены). На более поздних ступенях эволюции большинство клеток утратили способность к аутосплайсингу и функции рибозимов перешли к отдельным компонентам сплайсингосом.

Альтернативный сплайсинг. Несколько интронов, содержащихся в мРНК, могут сшиваться в разных комбинациях с образованием различных матричных последовательностей. Такая дифференциация путей созревания мРНК получила название альтернативного сплайсинга. Этот сплайсинг был впервые открыт у аденовирусов, у которых часть кодирующих последовательностей удаляется подобно интронам, а 5′-кэп соединяется с любым из других экзонов, и, таким образом, образуются различные мРНК, кодирующие разные белки (рис. 151). Этот вид сплайсинга позволяет небольшому числу первичных вирусных транскриптов кодировать значительное количество белков, в пропорциях, необходимых для существования вируса.

У эукариот альтернативный сплайсинг мРНК, содержащих большое количество интронов, является эффективным способом регуляции активности генов, создает возможность для возникновения изоформ белков (см. гл. 2), наборы которых могут существенно различаться в различных клетках, тканях и органах многоклеточных организмов. Альтернативный сплайсинг позволяет организму синтезировать разные по структуре и свойствам белки на базе одного гена. Такие гены кодируют семейства родственных белков, участвующих в мышечных сокращениях, формировании цитоскелета, нервных волокон, молекул иммуноглобулинов, пептидных гормонов и т. д.

В формировании альтернативных мРНК могут быть задействованы три основные механизма. Первый из них состоит в том, что для образования различных мРНК могут использоваться разные промоторы. В этом случае образуются транскрипты, имеющие разные

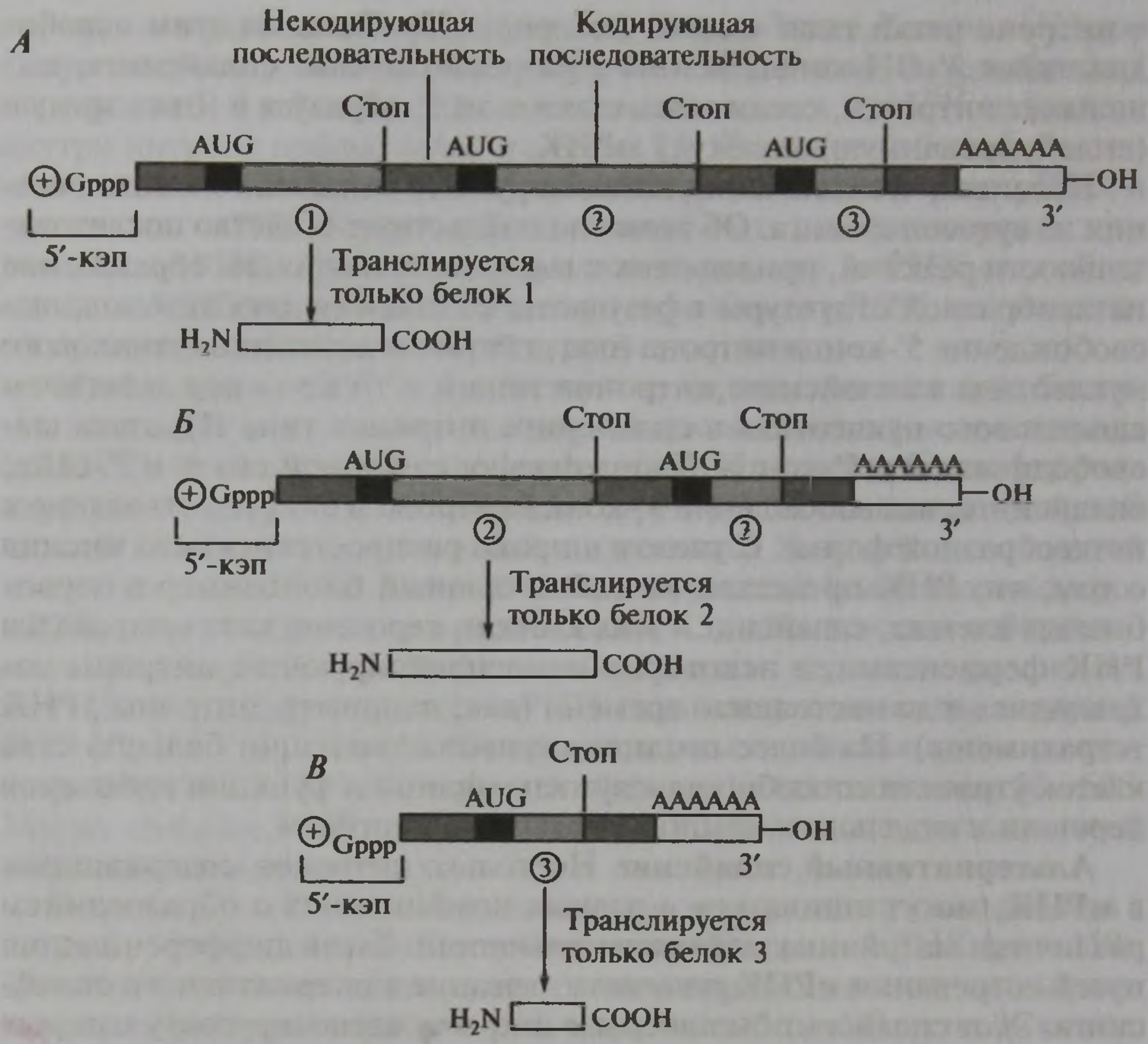


Рис. 151. Три различные мРНК, возникающие из одной молекулы-предшественника у аденовирусов (по Б. Албертсу и др., 1994). Пояснения см. в тексте

по длине 5′-концы и разное количество экзонов. Такой механизм сплайсинга выявлен для пре-мРНК легкой цепи миозина позвоночных животных.

Второй тип альтернативного сплайсинга имеет место при изменении сайта полиаденилирования первичного транскрипта. В этом случае изменяются размеры и структура 3′-концевого участка пре-мРНК (рис. 152). Таким способом образуются два вида мРНК тяжелой цепи иммуноглобулинов.

Третий тип альтернативного сплайсинга включает выбор различных экзонов из одинаковых пре-мРНК. При этом для формирования зрелых РНК могут использоваться различные экзоны, а часть из них не включается в сплайсинг. Таким образом происходит сплайсинг пре-мРНК тропонина Т скелетных мышц млекопитающих, содержащий 18 экзонов. При этом выбор экзонов зависит от стадии развития организма: экзон 16 присутствует в мРНК тропонина Т у взрослых,

а экзон 17 — в мРНК этого белка у эмбрионов. Тонкие механизмы выбора сайта сплайсинга еще неизвестны. Однако одной из причин возникновения альтернативных продуктов могут служить мутации, приводящие к разрушению нормальных сайтов сплайсинга и возникновению новых сайтов (рис. 153). Такой механизм альтернативного сплайсинга выявлен у больных талассемией (группа наследственных заболеваний, при которых резко падает содержание гемоглобина). Потеря в результате мутации одного из сайтов сплайсинга еще не означает отмену сплайсинга в целом, так как оставшийся другой сайт «ищет» подходящий участок и может соединяться с таким скрытым, ранее не реализовавшимся сайтом сплайсинга с образованием новой мРНК. Таким образом, мутации в мозаичных генах эукариот могут стимулировать образование новых белков, а альтернативный сплайсинг может играть решающую роль в эволюции высших организмов.

Процесс созревания РНК кроме вышеохарактеризованных сплайсинга, кэпирования и полиаденилирования включают еще ряд модификаций первичной структуры, называемых *редактированием РНК*. К этим реакциям относятся модификации азотистых оснований (дезаминирование, метилирование, восстановление и др.), в результате которых образуются необычные для РНК минорные основания (инозин, тимин, дегидроурацил и др.). К редактированию относятся и вставки нуклеотидов (чаще всего уридиловых) внутрь транскриби-

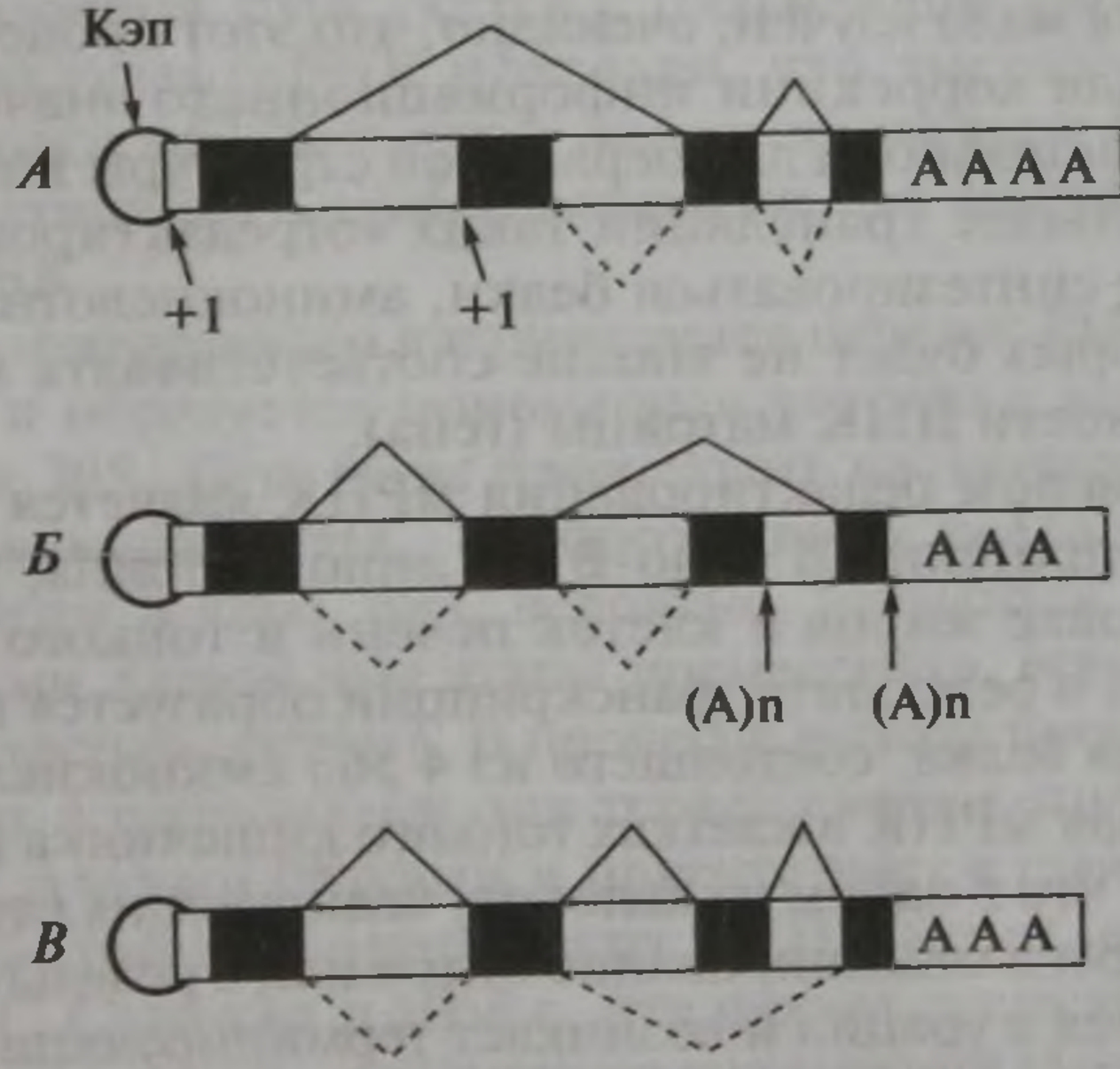


Рис. 152. Типы альтернативного сплайсинга пре-мРНК:

А — альтернативные промоторы (тип I); *Б* — альтернативные сайты полиаденилирования (тип II); *В* — выбор различных экзонов (тип III). Темные участки — экзоны, светлые — интроны; сплошными и пунктирными линиями соединены 5′- и 3′-сайты сплайсинга в зависимости от того, имеет ли пре-мРНК разные 5′-концы, разные 3′-концы или одинаковые 5′- и 3′-концы

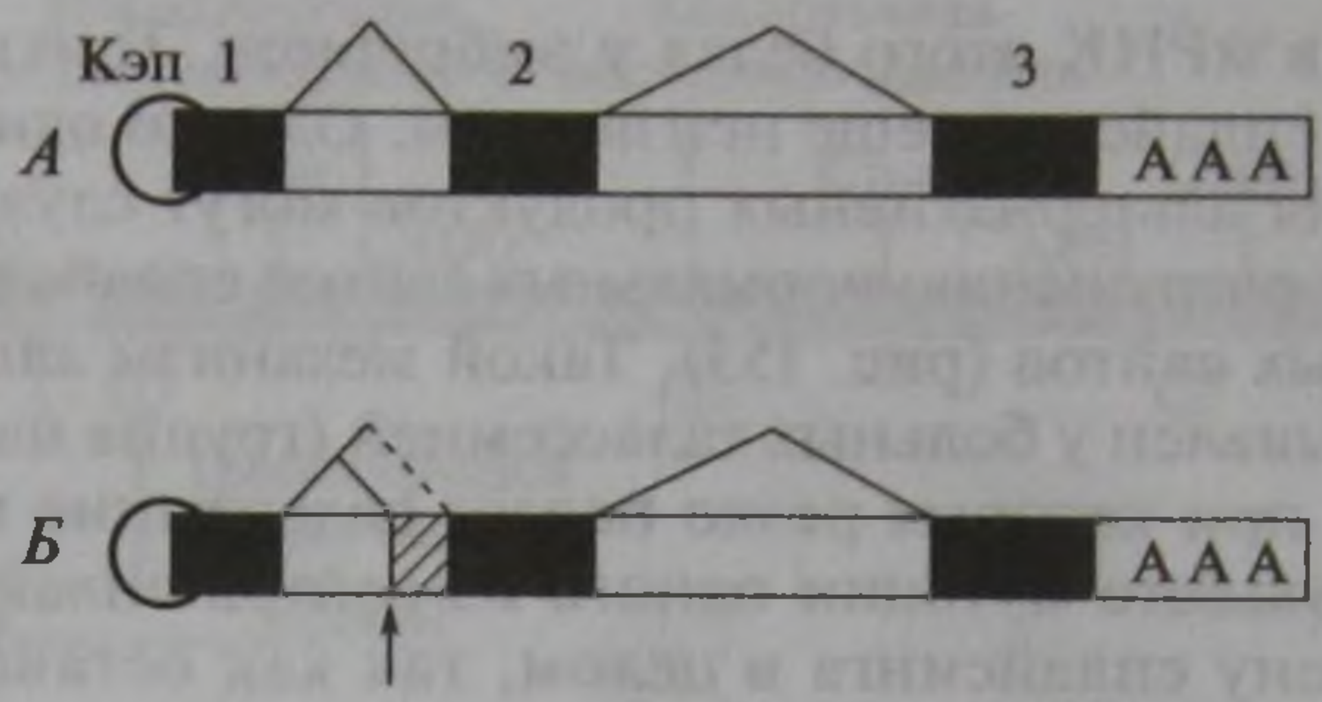


Рис. 153. Аномальный процессинг мРНК β-глобина у больных талассемией:

А — первичный транскрипт мРНК нормального β-глобина (зрелая мРНК состоит из трех экзонов); *Б* — мутация, приводящая к появлению нового сайта сплайсинга и удлинению экзона 2; *В* — мутация, приводящая к появлению дополнительного экзона между экзонами 2 и 3. Мутантные сайты обозначены стрелками. Темные участки — экзоны, светлые — интроны; заштрихованные участки — новые последовательности нуклеотидов, которые включаются в зрелую мРНК в результате мутации; линии соединяют 5′- и 3′- сайты сплайсинга

рованной цепи РНК. Несмотря на то что ферментативный аппарат редактирования мало изучен, очевидно, что этот процесс может быть очень важен для коррекции информационного значения кодонов мРНК, а следовательно, и для первичной структуры кодируемых ими белков. В результате трансляции таких «отредактированных» РНК в клетке могут синтезироваться белки, аминокислотная последовательность которых будет не вполне соответствовать нуклеотидной последовательности ДНК матрицы (гена).

Ярким примером редактирования мРНК является редактирование мРНК апопротеина В (Апо-В) — липопротеина, участвующего в транспортировке жиров и клеток печени и тонкого кишечника в кровь. В печени в результате транскрипции образуется мРНК, служащая для синтеза белка, состоящего из 4 563 аминокислотных остатков. Аналогичная мРНК в клетках тонкого кишечника редактируется таким образом, что в одном из кодонов, кодирующих глутамин (кодон CAA), происходит дезаминирование цитозина, в результате чего цитозин превращается в урацил и возникает терминирующий кодон UAA. Этот кодон прекращает трансляцию мРНК в середине ее молекулы, и поэтому в тонком кишечнике синтезируется укороченный белок, состоящий из 2 152 аминокислотных остатков (рис. 154).

Синтезированная в ядрах эукариотических клеток гетерогенная ядерная РНК (гяРНК) сразу после синтеза связывается с белками и образует гяРНП-частицы. В состав гяРНП-частиц входит не менее

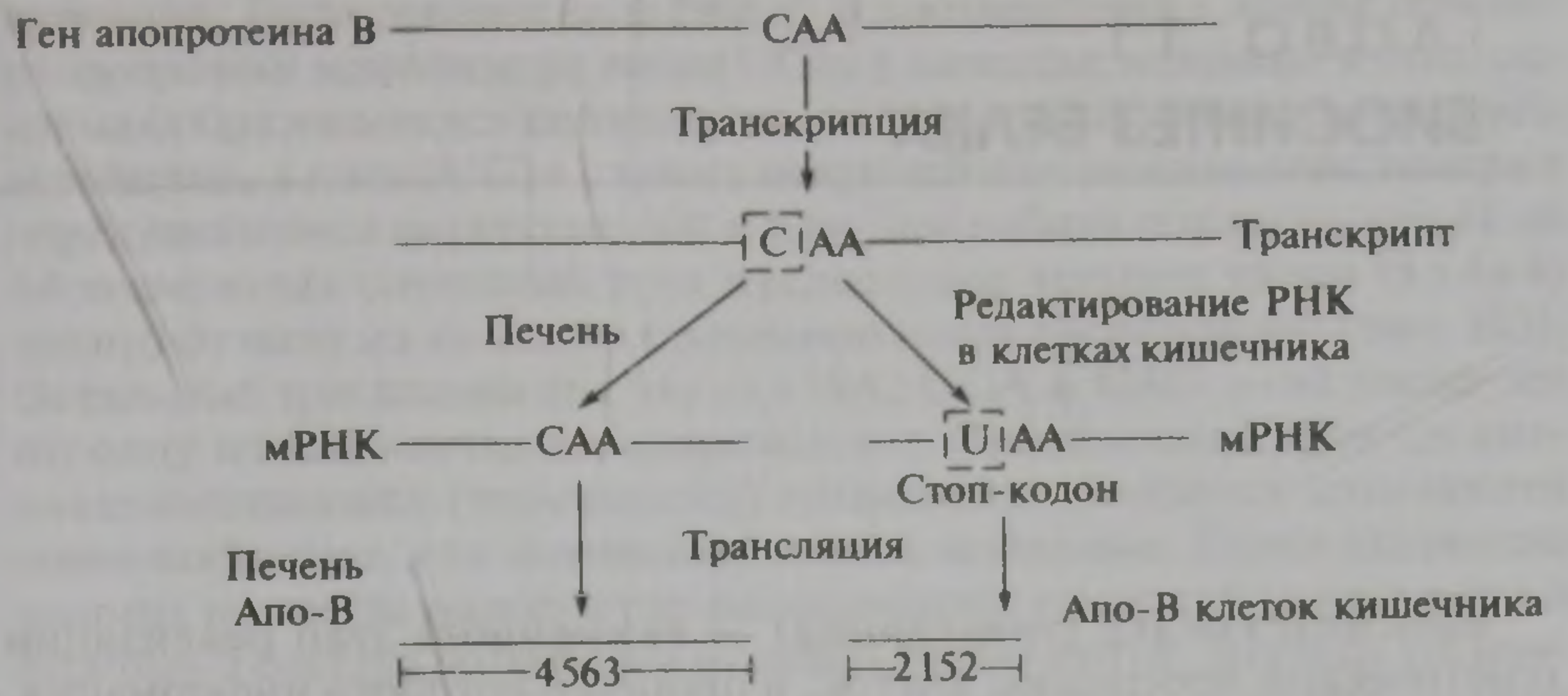


Рис. 154. Редактирование мРНК апопротеина В.

В гене апопротеина показан триплет (CAA) на нематричной цепи ДНК. Остальные пояснения см. в тексте

восьми разных белков. Все эти РНК-связывающие белки содержат в своем составе высококонсервативную последовательность: NH_2... Lys(Arg)-Gly-Phe(Tyr)-Ala(Gly)-Phe(Tyr)-Val-X-Phe(Tyr)...COOH. Гетерогенная ядерная РНК, открытая О. П. Самариной и Г. П. Георгиевым в 1960 г., представляет собой пре-мРНК, названную авторами ДНК-подобной РНК (дРНК). Дальнейшие работы в области изучения структуры гяРНК показали, что высокомолекулярные предшественники мРНК в ядрах ассоциированы с глобулярными белковыми частицами — *информоферами*, т. е. носителями информационной РНК.

На каждый информофер накручивается отрезок РНК длиной ~600 нуклеотидов, и образуется мономерная частица с коэффициентом седиментации 30S. При этом происходит частичное «плавление» вторичной структуры РНК. На протяжении всего периода, пока гяРНК находится в ядре, она, вероятно, находится в комплексе с информоферами. Отдельные этапы процессинга, вероятно, требуют изменения структуры гяРНП. В процессе выхода процессированной мРНК из ядра в цитоплазму она теряет связь с информоферами, связывается с другими белками и присутствует в цитоплазме в виде *информосом* — цитоплазматических РНП-частиц, открытых в лаборатории А. С. Спирина в 1964 г. Эта форма существования мРНК имеет прямое отношение к регуляции трансляции (биосинтеза белка) в рибосомальном аппарате клеток.

Глава 11
БИОСИНТЕЗ БЕЛКА

Биосинтез белка (трансляция) — важнейший этап реализации генетической программы клеток, в процессе которого информация, закодированная в первичной структуре нуклеиновых кислот, переводится в аминокислотную последовательность синтезируемых белков. Иными словами, трансляция — это перевод четырехбуквенного (по числу нуклеотидов) «языка» нуклеиновых кислот на двадцатибуквенный (по числу протеиногенных аминокислот) «язык» белков. Перевод осуществляется в соответствии с правилами *генетического кода*. Трансляция происходит с участием специализированных внутриклеточных частиц — *рибосом*, и в ее осуществлении принимает участие три главных класса РНК (мРНК, рРНК и тРНК), а также большая группа особых *белковых факторов трансляции*. Особенности молекулярного аппарата и механизм трансляции изучены в основном на прокариотических объектах (бактериях и бактериофагах), однако есть все основания считать, что основные принципы трансляции реализуются и в эукариотических клетках, в которых в то же время более развиты механизмы регуляции белоксинтезирующей системы.

Генетический код

Изучение трансляции и прежде всего расшифровка генетического кода стали одной из наиболее ярких страниц молекулярной биологии второй половины XX в. Еще в начале 50-х годов XX в. Г. Гамов предположил, что генетический код является триплетным: три соседних нуклеотида в полинуклеотидной цепи программируют включение одной аминокислоты в полипептидную цепь белка. В середине 60-х годов XX в. в серии оригинальных экспериментов Ф. Крик, С. Бреннер, Г. Виттман и другие исследователи действительно установили, что код является *триплетным* и *непрерывным* (не содержит «запятых»), т. е. в процессе синтеза белка последовательность мРНК считывается последовательно группами по три нуклеотида. Полная расшифровка генетического кода, проведенная М. Ниренбергом, С. Очоа и Н. Г. Корана с использованием бесклеточных систем, содержащих рибосомы и специальные синтетически полученные матрицы определенного

строения, была закончена в 1966 г. В соответствии с кодом при использовании кополимера поли(UC)n в качестве матрицы в этих системах образовывался полипептид, построенный из остатков серина и лейцина, а поли(UG)n служил матрицей для синтеза сополимера с чередующимися остатками Val и Cys. Эта работа показала, что 61 из 64 возможных сочетаний трех нуклеотидов четырех типов (4×4×4) кодируют одну из двадцати протеиногенных аминокислот (рис. 155). Остальные три кодона (из 64) — UAA, UGA и UAG — не кодируют ни одну из канонических аминокислот. Эти кодоны являются сигналами остановки (терминации) трансляции и поэтому называются *стоп-кодонами*, или *терминирующими кодонами*. Терминирующие кодоны не всегда однозначно распознаются системой трансляции и поэтому в составе мРНК они нередко дублируются. Первым (основным) стоп-кодоном обычно является кодон UAA, а на небольшом расстоянии следом за ним располагается один из других терминирующих триплетов (UGA или UAG). Следует также учитывать, что при синтезе ряда белков кодон UGA используется для включения в белок аминокислотного остатка селеноцистеина (см. ниже).

Поскольку число кодирующих триплетов (61) в три раза больше числа аминокислотных остатков, обычно присутствующих в белках, генетический код сильно *вырожден*, так что многие аминокислоты кодируются двумя и более кодонами. Только две аминокислоты

Вторая буква

Первая буква (5′-конец)	U			C			A			G			Третья буква (3′-конец)
U	UUU	Phe		UCU	Ser		UAU	Tyr		UGU	Cys		U
	UUC			UCC			UAC			UGC			C
	UUA	Leu		UCA			UAA	стоп		UGA			A
	UUG			UCG			UAG	стоп		UGG	Trp		G
C	CUU	Leu		CCU	Pro		CAU	His		CGU	Arg		U
	CUC			CCC			CAC			CGC			C
	CUA			CCA			CAA	Gln		CGA			A
	CUG			CCG			CAG			CGG			G
A	AUU	Ile		ACU	Thr		AAU	Asn		AGU	Ser		U
	AUC			ACC			AAC			AGC			C
	AUA			ACA			AAA	Lys		AGA	Arg		A
	AUG	Met		ACG			AAG			AGG			G
G	GUU	Val		GCU	Ala		GAU	Asp		GGU	Gly		U
	GUC			GCC			GAC			GGC			C
	GUA			GCA			GAA	Glu		GGA			A
	GUG			GCG			GAG			GGG			G

Рис. 155. Генетический код.

При синтезе белка триплеты нуклеотидов мРНК (кодоны) транслируются в соответствующие им аминокислоты. Например, кодоны AUG и GUG детерминируют включение в белок метионина и валина соответственно. Кодоны UAA, UAG и UGA являются терминирующими кодонами (стоп-кодоны)

(Met и Trp) кодируются единичными кодонами (AUG и UGG соответственно), поэтому, очевидно, встречаются в белках реже других. Вырожденность генетического кода проявляется в том, что для каждой аминокислоты существует более одной тРНК, и одна тРНК может взаимодействовать более чем с одним кодоном мРНК. В трехбуквенном генетическом коде наиболее важны первые две буквы, тогда как третья буква часто бывает разной. Так, например, глицин кодируется четырьмя *синонимическими кодонами*: GGA, GGC, GGG и GGU. В связи с преобладающей ролью первых двух букв кодонов (считая с 5′-конца триплета мРНК) генетический код иногда называют *квазидуплетным* (псевдодуплетным). Эта особенность кода позволяет использовать меньшее число тРНК: для взаимодействия с 61 кодоном достаточно 31 тРНК в цитоплазме и всего 22 тРНК в белоксинтезирующей системе митохондрий животных.

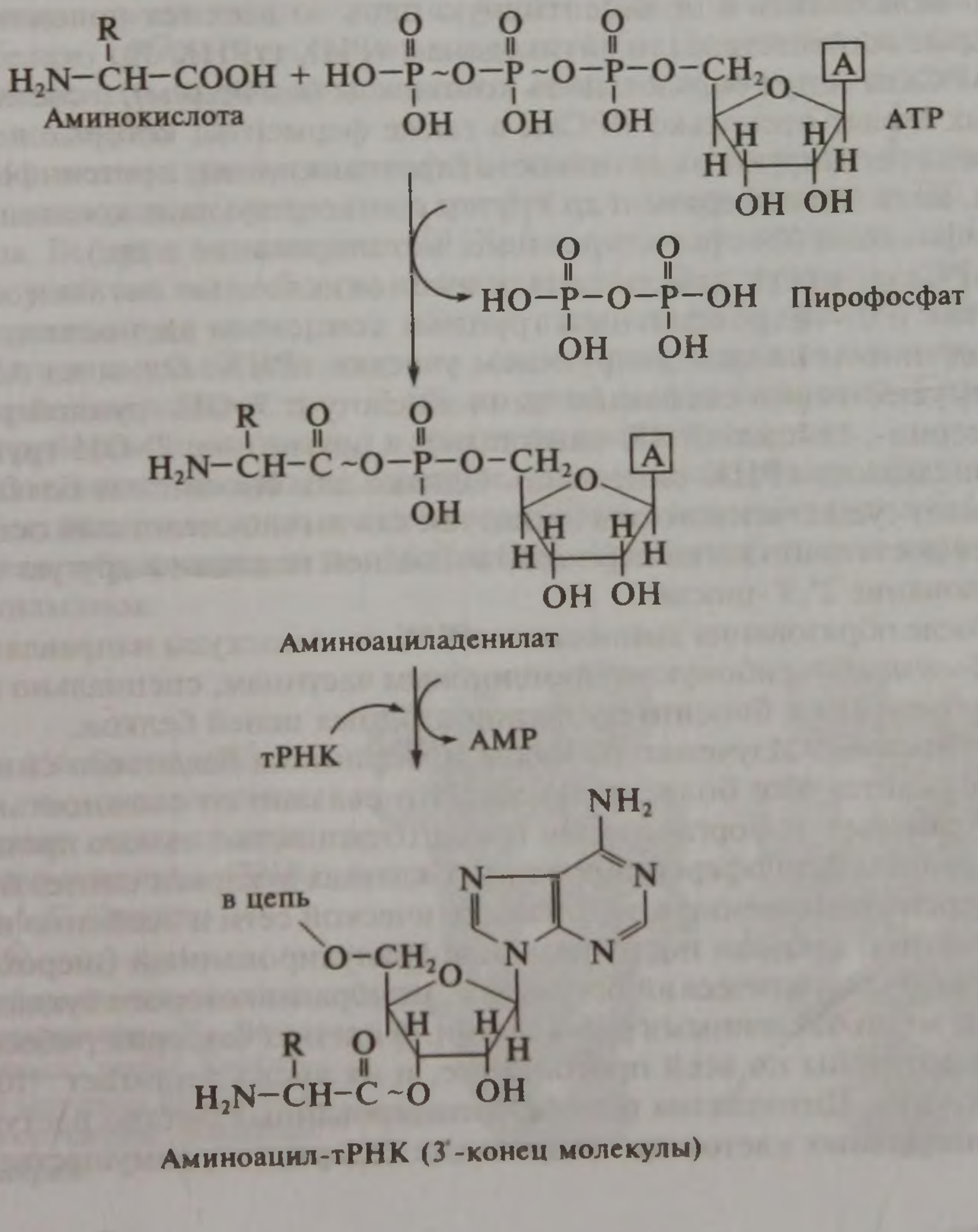
Генетический код почти всегда *универсален*, т. е. един для всех живущих на Земле организмов — от бактерий до человека. Небольшие отличия имеются, однако, в генетическом коде митохондрий и хлоропластов. Нет никаких данных о том, что когда-либо существовали организмы с другим кодом или другими аминокислотами. Очевидно, генетический код тщательно сохраняется в эволюции и изменения в коде, а также в рибосомальном аппарате клеток следует признать сильно заторможенными.

Активация аминокислот

Перед началом трансляции синтезированные в результате разнообразных биохимических реакций или полученные с пищей протеиногенные аминокислоты должны пройти *стадию активации* и присоединиться к тРНК, осуществляющими их доставку к рибосомам. Во всех клетках имеется набор тРНК, которые служат *адапторами* при переводе нуклеотидных последовательностей мРНК в аминокислотные последовательности белков. В структуре тРНК содержится несколько функционально важных участков (петель), главными из которых для адаптерной функции являются *антикодон* и *акцептирующий конец* (см. рис. 56).

Антикодон служит для взаимодействия с соответствующим (комплементарным) кодоном мРНК, а акцептирующий конец (расположенная на 3′-конце молекулы тРНК последовательность ССА-ОН) — для присоединения аминокислоты. Наличие антикодона и акцептирующие свойства тРНК позволяют им выполнять адаптерную функцию: связывая и перенося аминокислотный остаток и присоединяясь за счет антикодона к соответствующему кодону мРНК, они позволяют аминокислотным остаткам выстраиваться в порядке, диктуемом последовательностью нуклеотидов в матричной молекуле РНК, и таким образом способствуют переводу последовательности нуклеотидов в последовательность аминокислотных остатков

синтезируемого белка в соответствии с правилами генетического кода. Каждая тРНК может переносить только одну из аминокислот, вовлекаемых в биосинтез белка. Для большей части аминокислот имеется несколько тРНК, которые называются *изоакцепторными* и обозначаются соответственно $тРНК_1^{Gly}$, $тРНК_2^{Gly}$ и т. д. Существование изоакцепторных тРНК связано с вырожденностью генетического кода. В результате присоединения аминокислот к 3′-концу молекулы тРНК аминокислоты активируются — между карбоксильным концом молекулы и концевым аденозином акцептирующего конца тРНК возникает *макроэргическая связь*, энергия которой используется далее для синтеза пептидной связи в ходе биосинтеза белка на рибосомах (см. рис. 161). В результате специфического взаимодействия тРНК и соответствующей аминокислоты возникает *аминоацил*-тРНК — молекула, содержащая активированный аминокислотный остаток и соответствующий антикодон и являющаяся *истинным субстратом* для реакции синтеза полипептидной цепи белка:



Для каждой из аминокислот существует своя особая аминоацил-тРНК-синтетаза, которая осуществляет специфическое узнавание и связывание аминокислот и тРНК, в результате чего и возникает аминоацил-тРНК. Аминоацил-тРНК-синтетазы (АРСазы) обладают исключительно высокой субстратной специфичностью, активируя только белковые аминокислоты L-ряда и присоединяя их к имеющей определенный антикодон тРНК. Таким образом, они как бы сообщают аминокислоте определенный шифр в виде трех нуклеотидов антикодона тРНК и поэтому эти ферменты иногда называют *кодазами* или *шифразами*. Выдающаяся роль АРСаз в безошибочном синтезе полипептидных цепей белков была доказана в специальных экспериментах с использованием бесклеточных систем белкового синтеза. В этих опытах осуществляли химическое превращение аминокислотных остатков в составе аминоацил-тРНК, заменяя, например, остаток цистеина на остаток аланина. При использовании таких молекул в ходе биосинтеза неправильные аминокислотные остатки (Ala) включались в полипептидную цепь во всех тех положениях, которые соответствовали антикодонам тРНК ($тРНК^{Cys}$).

АРСазы могут образовывать комплексы (*кодосомы*), в состав которых входят несколько АРСаз, а также ферменты, которые в свою очередь регулируют их активность (протеинкиназы, протеинфосфатазы, метилтрансферазы и др.) путем соответствующих ковалентных модификаций (фосфорилирование, метилирование и др.).

АРСазы могут присоединять аминокислотные остатки как к 3′-, так и 2′-гидроксильным группам концевого аденозина, расположенного на акцептирующем участке тРНК. Одни из АРСаз преимущественно связывают аминокислоту с 3′-ОН группой рибозы (серил-, глицил-тРНК-синтетазы), а другие — с 2′-ОН-группой (фенилаланил-тРНК-синтетаза). Однако для биосинтеза белка это не имеет существенного значения, так как аминокислотный остаток может достаточно легко переходить из одной позиции в другую через образование 2′,3′-цикла.

После образования аминоацил-тРНК эти молекулы направляются к *рибосомам* — рибонуклеопротеиновым частицам, специально приспособленным к биосинтезу полипептидных цепей белков.

Рибосомы. Изучение рибосом и механизма белкового синтеза продолжается уже более полувека, что связано со сложностью их пространственной организации и многоэтапностью самого процесса трансляции. В дифференцированных клетках эукариот синтез белка идет преимущественно в эндоплазматической сети и особенно в той ее фракции, которая носит название гранулированный (шероховатый) эндоплазматический ретикулум, мембраны которого буквально усеяны многочисленными рибосомами. В клетках бактерий рибосомы рассредоточены по всей протоплазме, и их число достигает ~10^4 на одну клетку. Цитоплазма недифференцированных быстро растущих эмбриональных клеток эукариот также содержит преимущественно

свободные рибосомы. Определенное число белков синтезируется также в ядре, в митохондриях и хлоропластах растений. В митохондриях выявлены рибосомы, которые несколько отличаются от обычных (цитоплазматических) рибосом по размеру (коэффициенту седиментации), что, вероятно, связано с особенностями происхождения этих органелл, спецификой их геномов и генетического кода. Так, митохондриальные рибосомы грибов имеют коэффициент седиментации 75S, а митохондриальные «минирибосомы» млекопитающих — 55S. Цитоплазматические рибосомы эукариот и рибосомы прокариот очень сходны по структуре. Каждая их них состоит из двух *субчастиц* — большой и малой, комплекс которых и представляет собой собственно рибосому с молекулярной массой 2,5 млн (70S) у прокариот или 4,2 млн (80S) — у эукариот. Малая субчастица (30—40S) служит для связывания мРНК и тРНК, а большая участвует в образовании пептидной связи. Основу (каркас) каждой субчастицы составляют молекулы рибосомальных РНК, вокруг которых в определенном числе и порядке группируются белки малой (S-белки) и большой (L-белки) субчастицы рибосомы (рис. 156). Сравнение структурных компонентов прокариотических и эукариотических рибосом приведено на рис. 157.

У всех прокариотических форм жизни (эубактерии, актиномицеты, синезеленые водоросли и архебактерии) присутствуют 70S рибосомы. Весовое соотношение РНК: белок у них составляет 2:1. Около 2% сухой массы рибосом составляет Mg. В препаратах рибосом, выделенных из прокариот, обычно присутствуют органические поликатионы, такие, как спермин, спермидин, кадаверин и путресцин (суммарное количество ~2,5% от сухой массы рибосом). Строение 70S рибосом практически идентично у всех прокариот.

В цитоплазме эукариот содержатся несколько более крупные 80S рибосомы, у которых соотношение РНК : белок близко к 1:1. В их составе обнаружены ионы Mg^{2+} и Ca^{2+}, а также небольшое количество полиаминов.

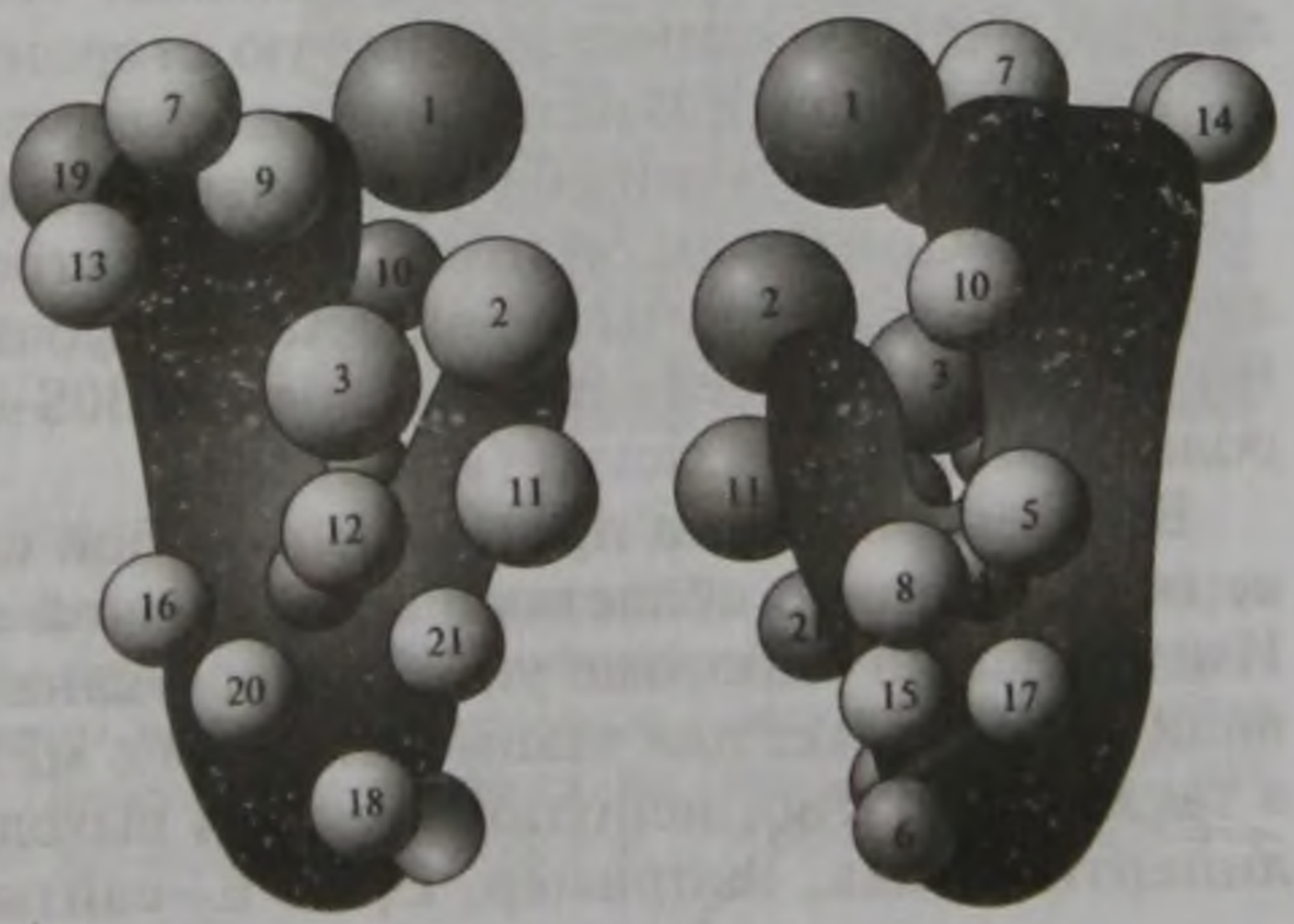


Рис. 156. Белки рибосомы в определенном порядке группируются вокруг молекулы рРНК (по А. С. Спирину, 1986).

Изображены две проекции расположения 21 белка (S-белки) вокруг Y-образной молекулы 16S рРНК в 30S-субчастице рибосомы бактерий

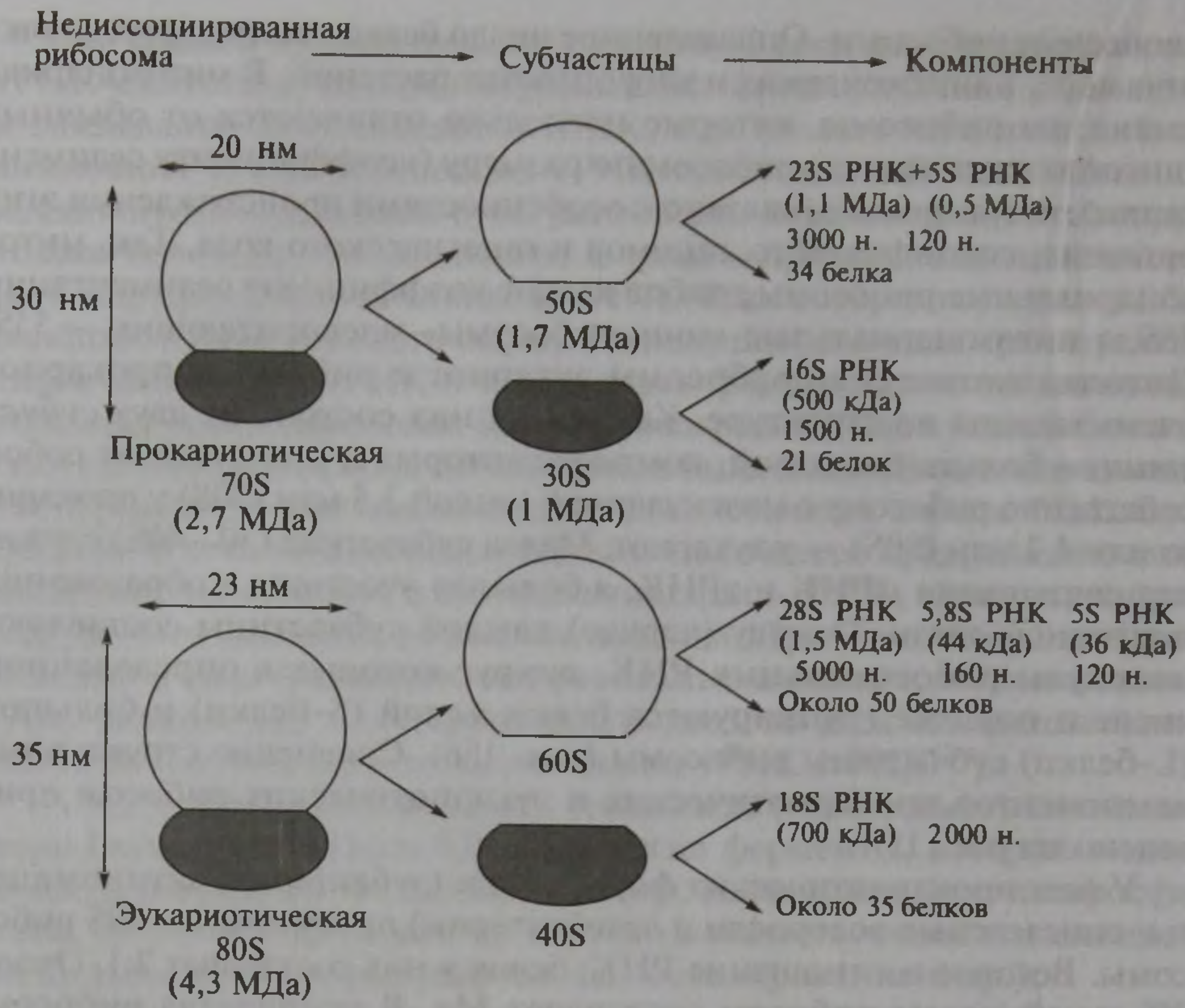


Рис. 157. Сравнение состава прокариотических и эукариотических рибосом. Пояснения см. в тексте

Наиболее детально изучена структура 70S рибосом бактерий (*E. coli*). В составе малой (30S) субчастицы рибосом различают определенные доли, называемые «головкой», «телом» и боковой выступ, или платформа (рис. 158). Большая (50S) субчастица содержит три выступа, или протуберанца: центральный протуберанец (головка) и два боковых, среди которых различают «L7/L12 — стержень» («палец») и расположенные по другую сторону от головки «L1-ребро» (или боковая доля). В названии этих выступов отражена локализация в них соответствующих белков большой субчастицы (L-белков).

В последние годы методом криоэлектронной микроскопии получены более детальные изображения субчастиц рибосом бактерий. На них кроме вышеназванных долей в 30S-субчастице идентифицированы «канал» и выступ («шпора»).

В целом рибосома представляет собой сложное по архитектонике образование с областями повышенной электронной плотности. Имеющиеся в рибосоме углубления («канал», «туннель» и др.), по-видимому, служат для взаимодействия с мРНК и аминоацил-тРНК, а также, вероятно, используются для выхода синтезированного полипептида (как, например, E_1- и E_2-сайты). В ходе трансляции в

процессе продвижения по мРНК рибосомальные субчастицы, по всей видимости, динамично контактируют друг с другом и с мРНК.

Так, в лаборатории академика А. С. Спирина с использованием оригинального метода тритиевой бомбардировки установлено, в частности, что стержень («палец») большой субчастицы обладает способностью «обнимать» малую субчастицу после присоединения GTP-связывающих белковых факторов трансляции. Структура всех белков бактериальных рибосом досконально изучена, и в настоящее время практически завершена работа по выяснению полной пространственной структуры рибосом прокариот, включая детали взаимодействия всех составляющих ее белков и рРНК.

В 2011 г. завершена работа по определению пространственной структуры 40S-субчастицы эукариотической рибосомы (рис. 159).

Рибосомальные РНК, составляющие более половины массы рибосомы, имеют весьма сложную пространственную организацию и выполняют, по современным представлениям, ведущую роль в биосинтезе белка. При некоторых отличиях структуры больших (16—18S

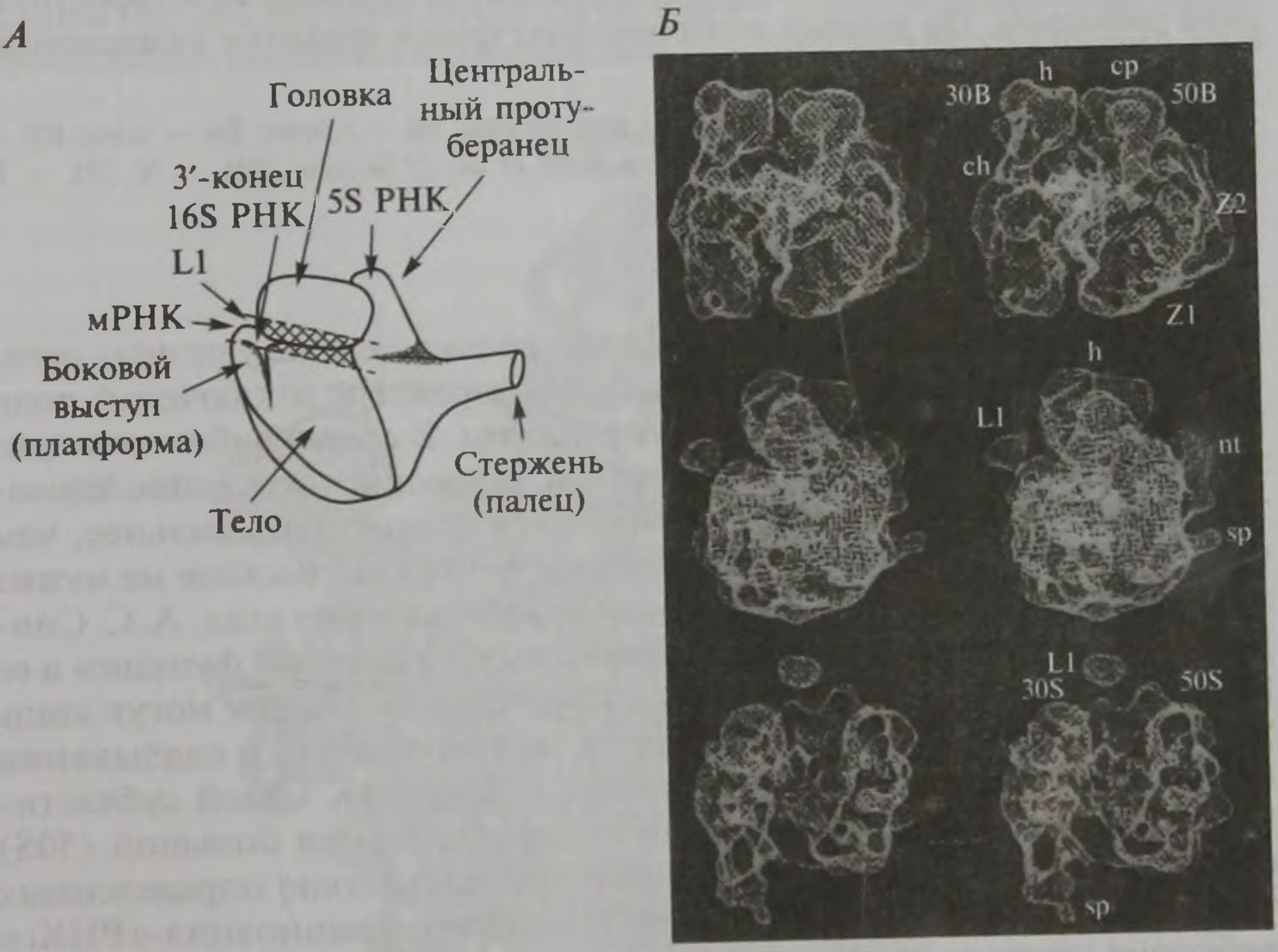


Рис. 158. Схема строения рибосомы бактерий (по А. С. Спирину, 1986):

А — заштрихованная область показывает расположение мРНК-связывающего участка на 30S-субчастице (изображена на первом плане). 50S-субчастица изображена на втором плане; *Б* — изображение рибосомы прокариот, полученное с использованием метода криоэлектронной микроскопии. Представлены три проекции полной рибосомы *E. coli*. Слева — 30S-субчастица, справа — 50S-субчастица (*Frank* et al. // Nature — 1995. — V. 376. — P. 441—444.)

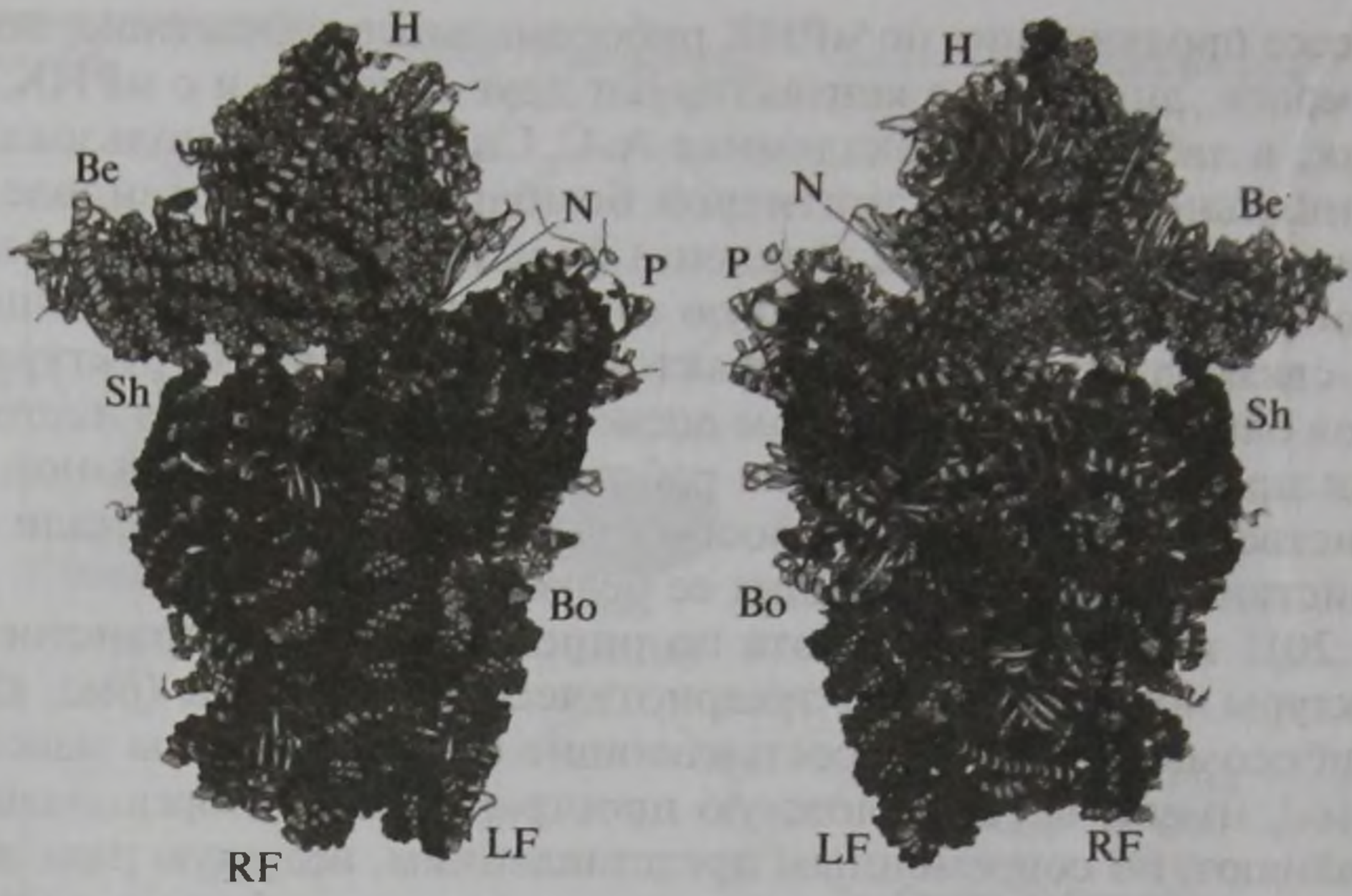


Рис. 159. Модель пространственной структуры 40S-субчастицы эукариотической рибосомы. По данным рентгеноструктурного анализа с разрешением 3,9 Å:

H — головка; Be — клюв; N — шея; P — платформа; Sh — плечо; Bo — тело; RF — правая ступня; LF — левая ступня (*Julius Rabi et al.* // Scince. 2011. — V. 331. — P. 730—736.)

и 23—28S) и малых (5S) рРНК в целом весьма сходны у разных организмов, что говорит об универсальном характере и ключевой роли этих молекул в функционировании рибосом. В состав рибосом входит также значительное количество белков, аминокислотные последовательности которых в эволюции менялись значительно сильнее, чем структура рРНК. Часть из этих белков, возможно, вообще не нужна для работы рибосом. Так, по образному выражению акад. А. С. Спирина, белки только «декорируют рибосому», а главные функции в ее работе, несомненно, выполняют РНК, тогда как белки могут лишь усиливать эти функции. Например, ведущую роль в связывании мРНК выполняет 3′-концевой участок 16S рРНК малой субчастицы рибосомы прокариот. В основе формирования большой (50S) субчастицы лежит комплементарное взаимодействие определенных участков 23S и 5S рРНК. Отбор и связывание аминоацил-тРНК в рибосоме также осуществляются с помощью 16S рРНК, тогда как сам синтез пептидной связи требует участия 23S рРНК. Разнообразные функции рРНК в осуществлении трансляции на рибосомах, вероятно, отражают определяющее значение РНК на первичных (добелковых) этапах биологической эволюции, когда сами РНК могли формировать каталитически активные поверхности и выполнять функции биокатализаторов (рибозимов).

Как прокариотические, так и эукариотические рибосомы содержат по одной высокополимерной РНК на каждую субчастицу и одну низкомолекулярную (5S РНК) в составе большой субчастицы, головку (центральный протуберанец) которой она и образует. В эукариотических рибосомах содержится также низкомолекулярная 5,8S РНК, которая представляет собой гомолог 5′-концевого участка высокополимерной РНК большой субчастицы рибосом прокариот. Эта РНК и аналогичная ей 4,5S рРНК хлоропластов высших растений представляют собой результат процессинга предшественника высокополимерной РНК большой субчастицы и участвуют в ее формировании. Молекулярные массы и число нуклеотидов различных рРНК представлены на рис. 160.

Нуклеотидная последовательность всех видов рРНК прокариот расшифрована и полностью изучена их пространственная структура (см. гл. 3). В составе этих молекул выявлены определенные участки (домены), изучено их структурно-функциональное назначение. В целом 16S рРНК бактерий имеет Y-образную структуру, отдельные концы (ветви) которой формируют головку, тело и платформу малой субчастицы. Наиболее примечателен тот факт, что общая форма этой

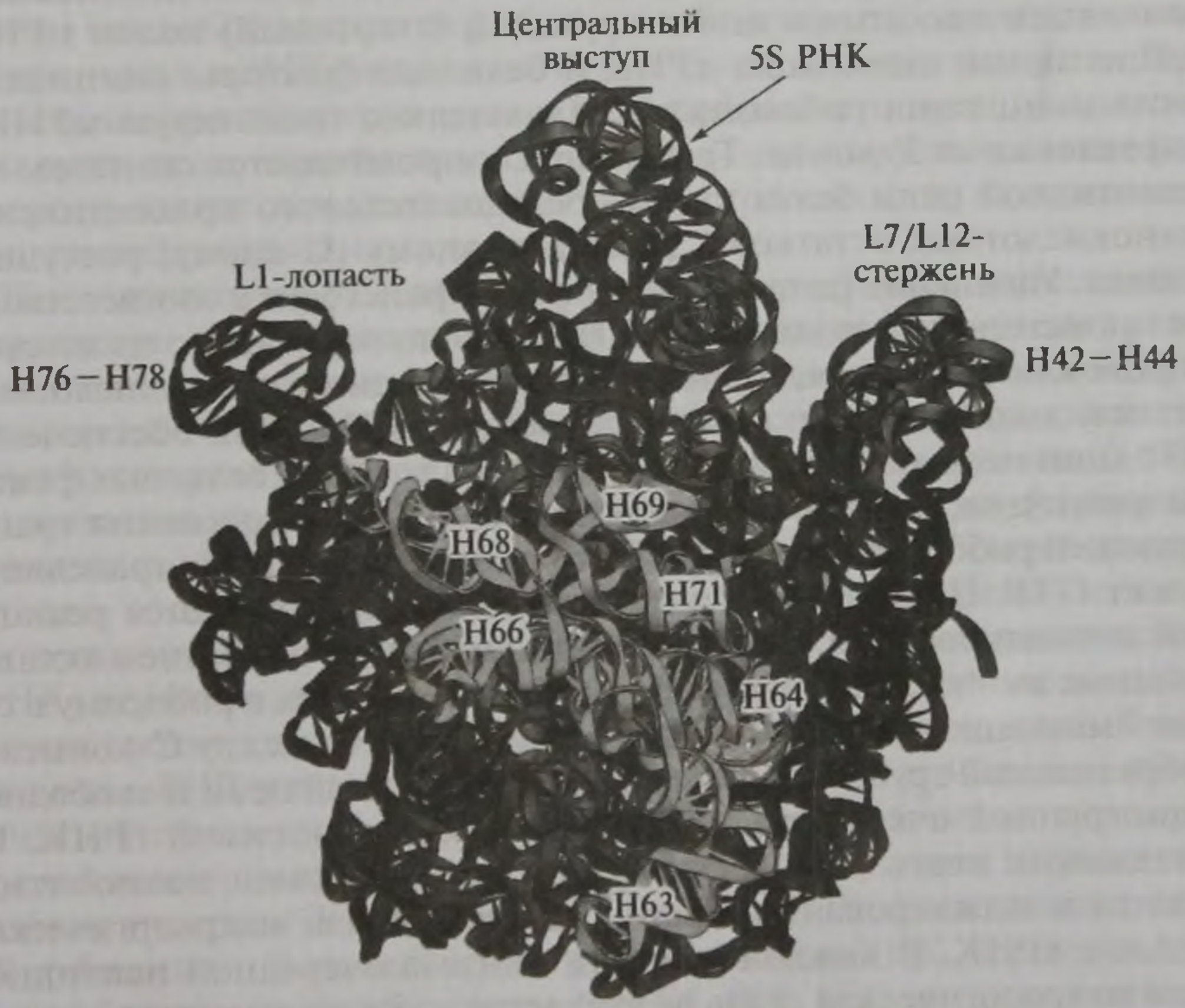


Рис. 160. Модель третичной структуры 23S рРНК архебактерии *Haloarcula marismortui* по данным рентгеноструктурного анализа. Н — петли (спиральные участки молекулы РНК) (по А. С. Спирину , 2011)

РНК напоминает форму самой 30S-субчастицы. Именно способность 16S рРНК к самопроизвольной укладке в пространстве (без участия белков) определяет общую форму и размеры рибосомальной субчастицы, в которой 16S рРНК служит каркасом (своего рода «арматурой») для ассоциации S-белков. Высокомолекулярная 23S рРНК бактерий имеет в пространстве более округлую форму, чем 16S рРНК, что связано с особенностями ее вторичной структуры: 5′-конец и 3′-конец 23S рРНК сближены (спарены) и образуют устойчивую спираль из 8 нуклеотидных пар. В определенной области (петле) вторичной структуры 23S рРНК происходит комплементарное взаимодействие с 5S рРНК. За исключением нескольких редуцированных выступов (протуберанцев) контуры 23S рРНК совпадают с контурами 50S-субчастицы, что опять-таки говорит о лидирующей роли РНК в формировании пространственной структуры рибосомы (рис. 160).

Последовательность событий в процессе биосинтеза белков на рибосомах обычно подразделяют на три больших этапа: *инициацию*, *элонганию* и *терминацию*.

Процесс прочтения (трансляции) мРНК стартует с определенного его участка, где начинается кодирующая (транслируемая) последовательность нуклеотидов. Для обеспечения начала (инициации) трансляции необходим инициирующий (стартовый) кодон мРНК, инициаторная аминоацил-тРНК и белковые факторы инициации. После инициации рибосома последовательно транслирует мРНК в направлении ее 3′-конца. Трансляция сопровождается синтезом полипептидной цепи белка путем последовательного присоединения аминокислотных остатков к карбоксильному (С-концу) растущего пептида. Удлинение (элонгация) пептида происходит в соответствии с последовательностью кодонов мРНК, прочтение каждого из которых сопровождается присоединением к пептиду одного аминокислотного остатка, доставляемого соответствующей тРНК. Для обеспечения элонгации необходимо наличие аминоацил-тРНК, белковых факторов элонгации, а также источник энергии для продвижения транслирующей рибосомы по мРНК. Источником энергии для трансляции служит GTP. Центральным моментом элонгации является реакция транспептидирования — синтез пептидной связи с участием активированных аминокислотных остатков, доставляемых в рибосому в составе аминоацил-тРНК. Реакция осуществляется между С-концевой карбоксильной группой растущей полипептидной цепи и свободной аминогруппой очередной аминокислоты, переносимой тРНК. На протяжении всего процесса карбоксильный конец полипептида остается в активированной форме, будучи связан макроэргической связью с тРНК. В каждой реакции синтеза очередной пептидной связи макроэргическая связь разрывается и тут же замещается такой же связью, образованной следующим аминокислотным остатком, доставленным аминоацил-тРНК (рис. 161). Терминация трансляции происходит тогда, когда рибосома достигает терминирующего кодона

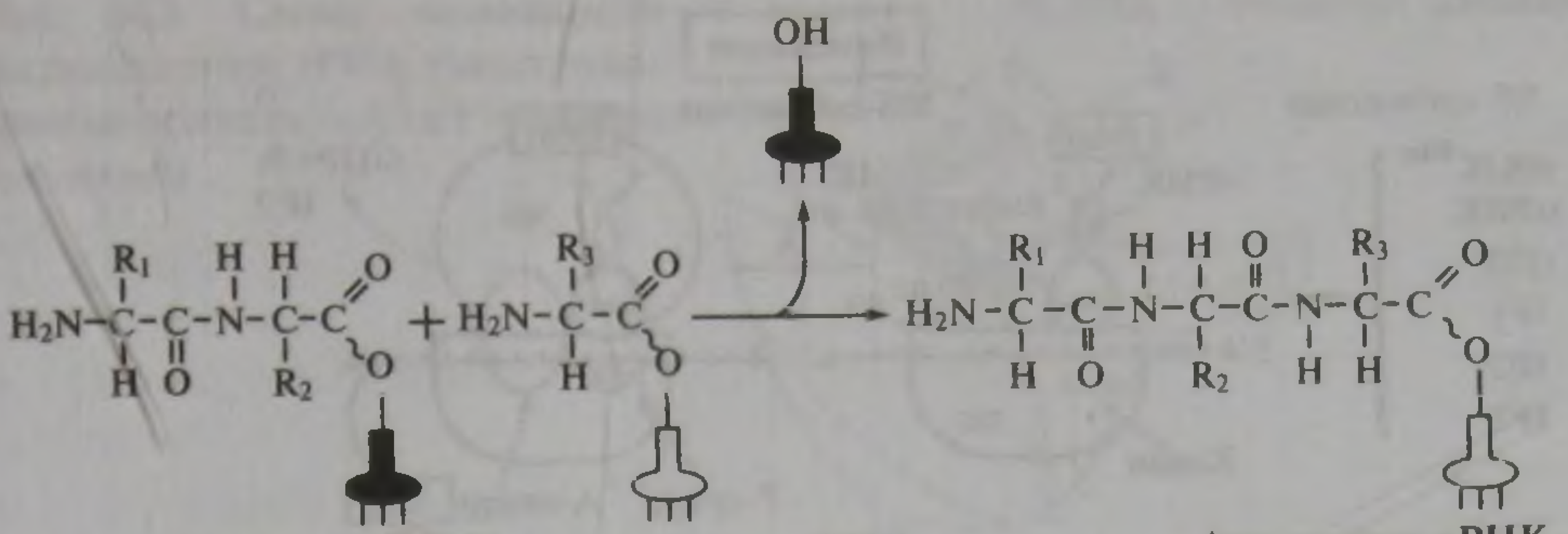


Рис. 161. Реакция транспептидирования. Пояснения см. в тексте

мРНК. С данным кодоном обычно не взаимодействует ни одна из аминоацил-тРНК, вместо этого к мРНК присоединяются белковые факторы терминации, под действием которых синтезированный пептид высвобождается из рибосомы. В случае трансляции моноцистроновых мРНК биосинтез белка на этом заканчивается. При трансляции полицистроновых мРНК рибосома может продолжить движение, пока не дойдет до очередного инициирующего кодона, с которого начнется новый полный цикл трансляции, включающий инициацию, элонгацию и терминацию.

Последовательность событий, происходящих в ходе одного цикла трансляции, представлена на рис. 162. Рибосома в процессе трансляции выполняет ряд функций: связывает и удерживает мРНК, связывает аминоацил-тРНК, осуществляет синтез пептидной связи, удерживает растущий полипептид, участвует в гидролизе GTP и продвигается по мРНК, взаимодействует с белковыми факторами трансляции. Для выполнения всех этих функций рибосома должна обладать соответствующими функциональными центрами, из которых только часть в настоящее время четко идентифицирована. Так, мРНК, вероятно, связывается с 30S-субчастицей в районе желобка, отделяющего головку от боковой лопасти и головку от тела. В связывании мРНК у прокариот главную роль играет комплементарное взаимодействие определенного пуринбогатого участка ее молекулы (*последовательность Шайна—Дальгарно*) с коротким пиримидин-богатым участком 3′-конца молекулы 16S рРНК, локализованной в 30S-субчастице. В связывании и удерживании мРНК, вероятно, участвуют белки малой субчастицы (S3, S5, S10, S14), а также стержень (палец) большой субчастицы (белковый комплекс L7/L12).

Связывание аминоацил-тРНК происходит в рибосоме в двух специальных центрах, которые формируются в результате объединения

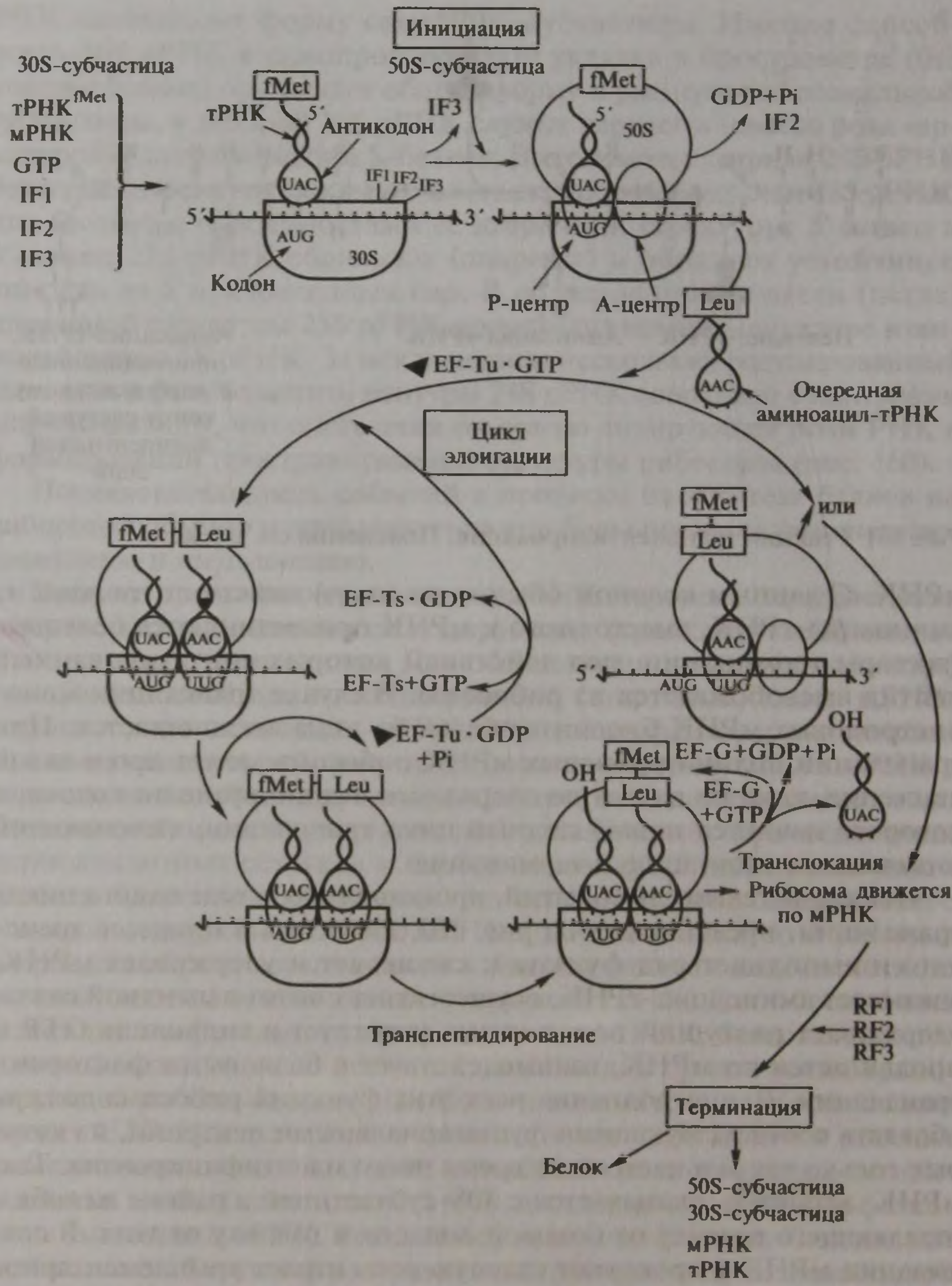
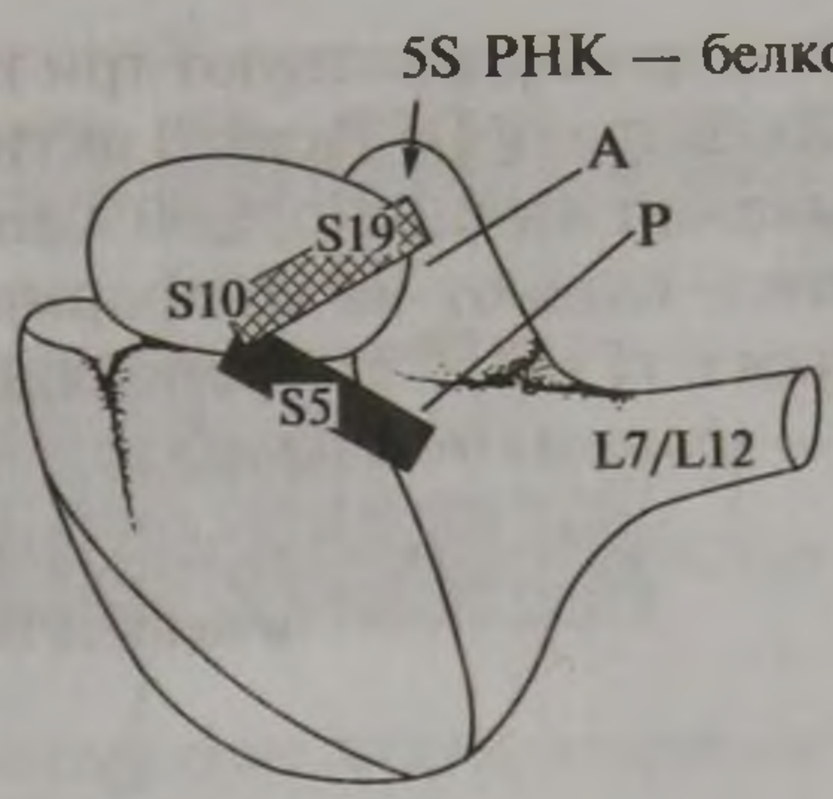


Рис. 162. Этапы трансляции у бактерий. Пояснения см. в тексте

большой и малой субчастиц — P-центре (пептидильный центр) и A-центре (аминоацильный центр). Оба центра достаточно компактны и расположены рядом друг с другом вблизи головки 30S-субчастицы и большого протуберанца 50S-субчастицы (рис. 163). Здесь происходит взаимодействие акцептирующих концов аминоацил-тРНК и пептидил-тРНК с образованием пептидной связи (реакция транс-

Рис. 163. Схема возможного расположения тРНК (заштрихованные области) в А- и Р-центрах рибосомы



пептидирования). Для формирования А- и Р-центров, вероятно, необходимо взаимодействие всех рРНК рибосомы, а также целого ряда белков большой и малой субчастиц.

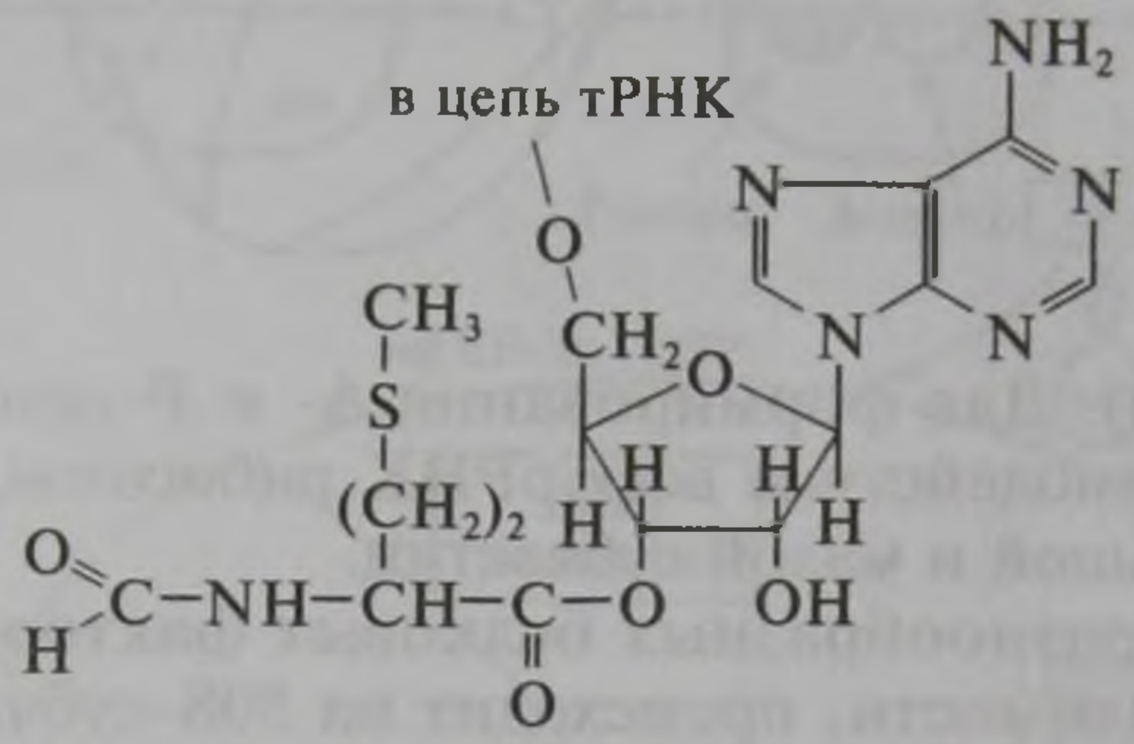
Связывание разнообразных белковых факторов элонгации и GTP, по всей видимости, происходит на 50S-субчастице в районе L7/L12- стержня. К связыванию GTP способны специфические белковые факторы трансляции: EF-Tu, EF-G (факторы элонгации), IF-2 (фактор инициации) и RF-2 (фактор терминации). Эти факторы всегда взаимодействуют с рибосомой в комплексе с GTP, т.е. представляют собой GTP-связывающие белки (G-белки). При связывании с рибосомой активируется GTPазная активность того или иного из этих факторов, а высвободившаяся в результате гидролиза GTP энергия используется на соответствующих энергоемких стадиях трансляции.

Пептидилтрансферазный центр, в котором идет синтез пептидной связи, локализован в 50S-субчастице и расположен вблизи центрального протуберанца на вогнутой поверхности, контактирующей с 30S-субчастицей. В организации этого центра, вероятно, принимают участие один из доменов 23S рРНК и некоторые белки 50S-субчастицы. Многочисленные попытки идентифицировать пептидилтрансферазу среди белков рибосом не увенчались успехом, и складывается впечатление, что за синтез пептидной связи ответственна 23S рРНК. Комплекс перемещений (транслокация) молекулы мРНК, деацилированной тРНК и пептидил-тРНК, осуществляемый по GTP-зависимому механизму, может, вероятно, происходить в результате взаимодействия субчастиц рибосом (их размыкания и смыкания) при участии стержня большой субчастицы. Такая динамичная модель работы рибосомы детально разработана академиком А. С. Спириным.

Этапы трансляции

Инициация белкового синтеза состоит из нескольких стадий и обслуживается рядом *белковых факторов инициации* (IF — Initiation

Factors). У бактерий существуют три таких фактора: IF-1, IF-2 и IF-3. Фактор IF-2 обладает GTPазной активностью и играет центральную роль в связывании инициаторной аминоацил-тРНК. Два других фактора, вероятно, влияют на конформацию 30S-субчастицы, помогая ей связываться с $тРНК^{fMet}$, переносящей модифицированный остаток метионина — *формилметионил*:



Формилметионил-тРНК
(3′-конец молекулы)

Именно этой тРНК предстоит связаться с инициирующим кодоном мРНК — кодоном AUG. В некоторых случаях первый аминокислотный остаток включается не на метиониновом кодоне AUG, а на кодоне GUG, который кодирует валин. Это связано с тем, что для эффективной инициации у прокариот стартовый кодон должен располагаться на вершине шпилечной структуры мРНК, перед которой на расстоянии от 3 до 10 нуклеотидов должна располагаться пурин-богатая последовательность Шайна — Дальгарно. Комплементарное спаривание антикодона $тРНК^{fMet}$ с инициирующим кодоном имеет принципиально важное значение, так как именно это взаимодействие определяет *рамку считывания* (трансляции) нуклеотидной последовательности мРНК, которая в принципе может декодироваться с трех различных позиций (по числу нуклеотидов в кодоне), и в результате могут быть синтезированы три различные полипептидные цепи (рис. 164).

На следующей стадии инициации происходит присоединение 30S-субчастицы к мРНК и узнавание инициирующего кодона AUG среди других аналогичных кодонов, кодирующих метионин. Инициаторная $тРНК^{fMet}$ при этом связывается с 30S-субчастицей в виде комплекса с факторами инициации и GTP (см. рис. 162). Выбор правильной точки инициации на молекуле мРНК определяется, очевидно, 30S-субчастицей. Выбор стартового кодона у бактерий происходит с участием короткой (от 5 до 8 нуклеотидов) пурин-богатой последовательности Шайна — Дальгарно в мРНК, которая комплементарно связывается с пиримидин-богатым участком вблизи 3′-конца 16S рРНК, локализованной в 30S-субчастице рибосомы. Таким образом, взаимодействие трех видов РНК (мРНК, 16S рРНК

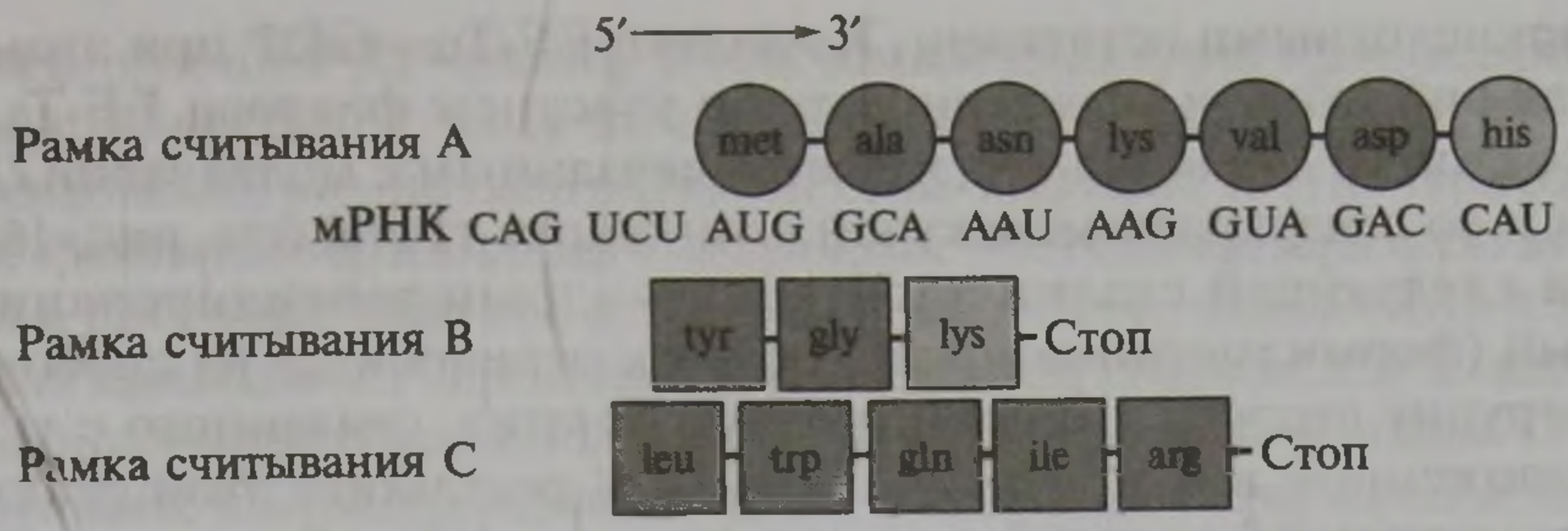


Рис. 164. Различные рамки считывания (трансляции) нуклеотидной последовательности мРНК.

Трансляция может осуществляться с помощью одной из трех рамок считывания. В приведенном примере рамки считывания B и C прерываются стоп-кодонами. «Открытой» является только рамка считывания A, начинающаяся с кодона AUG

и тРНК^{fMet}) устанавливает рамку считывания и предопределяет аминокислотную последовательность синтезируемого белка.

После образования комплекса 30S-субчастицы с мРНК и тРНК^{fMet} белковые факторы инициации IF-3 и IF-1 покидают 30S-субчастицу, уступая место 50S-субчастице, присоединение которой завершает процесс сборки полной (70S) рибосомы прокариот. Объединение субчастиц приводит к формированию двух центров связывания тРНК: Р-центра и А-центра, назначение и локализация которых в рибосоме были отмечены нами ранее. Далее под действием фактора IF-2 происходит гидролиз GTP, а высвободившаяся при этом энергия, вероятно, расходуется на стабилизацию (закрепление) тРНК^{fMet} в Р-центре рибосомы. Одновременно формируется А-центр рибосомы, предназначенный для присоединения следующей аминоацил-тРНК и, таким образом, заканчивается процесс сборки активной (транслирующей) рибосомы, осуществляющей элонгацию белкового синтеза.

Элонгация трансляции у бактерий обслуживается тремя белковыми факторами (EF-Tu, EF-Ts и EF-G) и может быть подразделена на три стадии. На первой стадии к А-центру рибосомы присоединяется аминоацил-тРНК, приносящая второй (после f-Met) аминокислотный остаток, кодируемый вторым (после триплета AUG) кодоном мРНК. Таким образом, создаются условия для образования пептидной связи между аминокислотными остатками, связанными с соответствующими тРНК закрепленными в А- и Р-центрах рибосомы. Поступая в А-центр, аминоацил-тРНК закрепляется в нем в комплексе с белковым фактором EF-Tu и GTP. В связывании аминоацил-тРНК, вероятно, принимают участие не только антикодоны, но и другие участки, поскольку посттранскрипционные модификации ее молекулы существенно влияют на связывание с рибосомой. При участии фактора EF-Tu осуществляется гидролиз GTP до GDP и фосфата, а выделяющаяся энергия расходуется на сближение двух акцептирующих концов молекул тРНК с прикрепленными к ним

аминокислотными остатками. Комплекс EF-Tu—GDP при этом покидает рибосому и регенерируется с участием фактора EF-Ts, так что фактор EF-Tu вновь оказывается связанным с молекулой GTP, а затем и со следующей молекулой аминоацил-тРНК (см. рис. 162).

На следующей стадии элонгации — транспептидировании — первый (формилметиониновый) остаток переносится на свободную NH_2-группу второго аминокислотного остатка, связанного с тРНК, расположенной в А-центре рибосомы. В результате этой реакции, химизм которой был разобран выше (см. рис. 161), образуется первая пептидная связь в молекуле синтезированного белка и возникает дипептидил-тРНК. Стоящим первым на N-конце остаток fMet (у эукариот им является остаток Met) впоследствии отщепляется при участии аминопептидаз. Отщепление метионина существенно потому, что N-концевой аминокислотный остаток может определять время жизни белка, воздействуя на убиквитин-зависимый путь его деградации (см. гл. 2).

Третья стадия элонгации — *транслокация* — заключается в том, что дипептидил-тРНК, оставаясь связанной с соответствующим кодоном мРНК, перемещается из А-центра в Р-центр рибосомы, одновременно с продвижением транслируемой мРНК через рибосому. При этом освободившаяся от аминокислотного остатка в результате реакции транспептидирования $тРНК^{fMet}$ удаляется из Р-центра, перемещаясь в *Е-центр рибосомы*, и далее выводится из нее. Одновременно освобождается А-центр, необходимый для связывания следующей аминоацил-тРНК. Транслокация, которую можно рассматривать как «шаг» рибосомы по мРНК в направлении 5′→3′, происходит строго на расстояние, соответствующее размеру одного триплета в мРНК. В результате транслокации напротив освободившегося А-центра располагается очередной кодон мРНК, определяющий присоединение соответствующей аминоацил-тРНК. Транслокация у бактерий идет с участием фактора EF-G и сопровождается гидролизом еще одной молекулы GTP.

Многократное последовательное воспроизведение всех вышеотмеченных стадий элонгации обеспечивает удлинение полипептидной цепи в соответствии с кодом белкового синтеза и происходит с весьма высокой скоростью: в бактериальной клетке продолжительность одного цикла элонгации составляет ~1/20 с, а весь синтез белка среднего размера (~400 аминокислотных остатков) занимает около 20 с.

Терминация трансляции у бактерий связана с функционированием трех белковых факторов: RF-1, RF-2 и RF-3 (от англ. *recognize* — узнавать), распознающих «бессмысленные» стоп-кодоны в мРНК. Фактор RF-1 узнает кодоны UAG и UAA, а фактор RF-2 — кодоны UAA и UGA. Фактор RF-3 выполняет вспомогательную роль, стимулируя работу RF-1 и RF-2. При поступлении в рибосому одного из терминирующих кодонов с ним немедленно связывается

соответствующий RF-фактор и тем самым блокирует присоединение аминоацил-тРНК. К тому же в клетках обычно отсутствуют тРНК с антикодонами, комплементарными стоп-кодонам тРНК. Присоединение факторов терминации стимулирует пептидилэстеразную активность в рибосоме (она свойственна, в частности, белкам L11 и L16), которая приводит к гидролизу сложноэфирной связи между С-концом синтезированного полипептида и акцептирующим концом тРНК. В результате синтезированный белок отделяется от рибосомы; одновременно отделяются тРНК и мРНК, а сама рибосома диссоциирует на 30S- и 50S-субчастицы. В терминации транскрипции принимает участие молекула GTP, которая, вероятно, служит аллостерическим регулятором активности белковых факторов терминации.

Поскольку мРНК бактерий, как правило, полицистроновые (т.е. несут информацию о структуре нескольких белков), они содержат несколько инициирующих и несколько терминирующих кодонов. Хотя рибосома бактерий распознает терминирующие кодоны, она может пройти по мРНК и дальше, начав трансляцию следующего участка в полицистроновой мРНК. Кроме этого, трансляция может стартовать не с первого, а со второго или последующего инициирующего кодона, и таким образом на одной мРНК могут синтезироваться разные белки.

В клетках эукариот этапы биосинтеза белка в основном соответствуют таковым у прокариот, однако представляют собой более тонко организованный процесс, связанный с особенностью организации их мРНК и иным набором белковых факторов трансляции. Как известно, зрелые мРНК эукариот содержат на 5′-конце особую структуру — кэп (см. гл. 3), который, как установлено, необходим для связывания мРНК с малой (40S) субчастицей рибосомы. У большинства моноцистроновых мРНК эукариот в качестве стартового используется ближайший к 5′-концу триплет AUG, и ни один из других триплетов AUG, расположенных в кодирующей области мРНК, не может быть использован в качестве инициирующего. Инициация трансляции с кодона AUG происходит только тогда, когда он находится в составе соответствующего нуклеотидного окружения: за два нуклеотида до него должен располагаться пуриновый нуклеотид (A или G), а сразу после него — G:

...NNN A(G)NN **AUG** GNNN...,

где N — пиримидиновый нуклеотид.

Перед началом трансляции у эукариот последовательность мРНК как бы сканируется (просматривается) начиная с 5′-конца (кэпа) для поиска AUG, расположенного в вышеотмеченном оптимальном нуклеотидном окружении. Инициация трансляции у эукариот связана с функционированием большего числа белковых факторов инициации, которые называются eIF-факторами (eukaryotic Initiation Factors). К собственно факторам инициации трансляции у эукариот

относят белки, обозначаемые eIF1 — eIF5. Эти факторы условно подразделяются на две группы. К первой группе относятся те из них, которые взаимодействуют с субчастицами рибосомы и тем самым облегчают присоединение инициаторной аминоацил-тРНК (тРНК^{Met}) и мРНК. Факторы второй группы участвуют в подготовке 5′-концевой области мРНК к инициации трансляции. К первой группе принадлежат факторы eIF-1, eIF-2 и eIF-3 (аналоги факторов IF-1 — IF-3 прокариот), а также ряд дополнительных факторов (eIF-2B и eIF-S). Вторая группа факторов специфична для эукариот и включает мРНК-связывающие и мРНК-расплетающие белки. Важнейшими из них являются факторы из группы eIF-4 (eIF-4A, eIF-4B, eIF-4E и eIF-4F). Фактор инициации eIF-4F представляет собой белковый комплекс, который связывается с кэпом эукариотической мРНК. Он состоит из трех субъединиц (α, β и γ), одна из которых (субъединица β) обладает РНК-зависимой АТРазной активностью. Активность факторов eIF-4E регулируется путем их фосфорилирования/дефосфорилирования.

После связывания кэпа для инициации трансляции необходимо расплетание вторичной структуры мРНК в 5′-нетранслируемой области. Этот АТР-зависимый процесс у эукариот осуществляется за счет РНК-*хеликазной* активности белковых факторов eIF- 4A и eIF-4B. Фактор eIF-4A существует как в свободном виде (белок с массой 45 кДа), так и в комплексе с фактором eIF-4F, и именно ему, по-видимому, свойственна АТРазная и РНК-хеликазная активность.

В целом большинство белковых факторов инициации эукариот функционируют в виде крупных белковых комплексов инициации трансляции (коэффициенты седиментации 43S и 48S). Связывание и продвижение этих комплексов по мРНК (вместе с малой субчастицей рибосомы) блокируется особым *белком-репрессором* инициации. Репрессор взаимодействует с элементами вторичной структуры (шпильками) мРНК, расположенными в 5′-нетранслируемом районе мРНК (либо в начале кодирующей области), стабилизирует соответствующие шпильки и тем самым предотвращает функционирование комплекса факторов инициации.

В элонгации трансляции у эукариот участвуют два фактора (eEF-1 и eEF-2), которые являются аналогами факторов EF-Tu и EF-G бактерий. Фактор eEF-1 образует комплекс с аминоацил-тРНК и GTP, и этот комплекс присоединяется к А-центру рибосомы. После предварительного «узнавания» (кодоно-антикодонового взаимодействия) фактор элонгации расщепляет GTP и отходит от рибосомы вместе с GDP, освобождая тем самым акцептирующий конец пептидил-тРНК и открывая возможность для удлинения пептида в последующей реакции транспептидирования. Небольшая пауза между узнаванием и гидролизом GTP достаточна для отделения от мРНК несущих неправильные (не вполне комплементарные) антикодоны аминоацил-тРНК. Таким образом этот фактор снижает вероятность включения несоответствующих аминокислотных остатков в полипептидную цепь

белка. Факторы элонгации эукариот в отличие от соответствующих факторов прокариот обладают выраженной *РНК-связывающей способностью*. Эти белки, как и аминоацил-тРНК-синтетазы, взаимодействуют с кодирующим участком мРНК, к которому присоединяется еще ряд ферментов, имеющих специфическое сродство к РНК (глицеральдегид-3-фосфатдегидрогеназа, лактатдегидрогеназа и др.). Предполагают, что эти NAD^+/NADH-зависимые ферменты участвуют в энергетическом обеспечении биосинтеза белка, способствуя созданию более высокой концентрации ATP и GTP в белоксинтезирующей системе.

В терминации трансляции эукариот участвует один белковый фактор — R, функция которого аналогична белковым факторам терминации бактерий.

Регуляция трансляции

Интенсивность и эффективность биосинтеза белков в клетках про- и эукариот определяются как общими для всех типов клеток механизмами, так и специфическими элементами контроля, в большей мере свойственными белоксинтезирующей системе эукариот.

Процессы синтеза и созревания различных видов РНК и роль белковых факторов рассмотрены нами выше. В этом разделе мы остановимся на других, ранее не отмеченных моментах регуляции трансляции, и рассмотрим отдельные примеры регуляции, выявленные у разных объектов, уделив дополнительное внимание роли мРНК в регуляции трансляции у эукариот.

Первый этап регуляции трансляции имеет место еще до начала биосинтеза белка на рибосомах и связан с деятельностью аминоацил-тРНК-синтетаз. Установлено, что АРСазы взаимодействуют со всеми тРНК, но проявляют избирательность к определенным изоакцепторным тРНК. Ошибки, допускаемые АРСазами при синтезе аминоацил-тРНК, очень редки и происходят не чаще чем 1 раз на 10 000 реакций. Однако, даже в случае ошибки, АРСазы способны к самокоррекции: они гидролизуют сложноэфирную связь между тРНК и неправильными присоединенными аминокислотными остатками.

Инициация трансляции на рибосомах является ключевым моментом регуляции всего процесса и включает несколько механизмов. У прокариот, мРНК которых не содержат кэпа и отсутствует кэп-зависимая регуляция инициации, ведущую роль в запуске трансляции играет *«сила» инициирующих последовательностей мРНК*. Чем длиннее эта последовательность и чем выше ее комплементарность 16S рРНК, тем быстрее достигается инициация трансляции. Кроме этого, у прокариот (равно как и эукариот) имеет место регуляция инициации за счет пространственной структуры инициирующего участка мРНК: сворачивание этого участка в стабильную вторич-

ную и третичную структуру может блокировать инициацию. Для продолжения трансляции необходимо, чтобы рибосома «расплела» эту структуру либо сработали особые дестабилизирующие белковые факторы. И, наконец, *избирательная негативная регуляция* трансляции может достигаться связыванием белков-репрессоров с инициирующим районом мРНК.

В качестве регуляторных белков могут выступать специализированные белки (вышеохарактеризованные белковые факторы трансляции, а также другие белки трансляционного контроля, о которых мы скажем ниже) либо белки, имеющие в клетке иные функции, но способные при определенных условиях выступать в роли регуляторов трансляции. Так, например, трансляция мРНК фермента треонил-тРНК-синтетазы у бактерий контролируется самим продуктом трансляции. Избыточный синтез этого фермента приводит к тому, что часть его молекул связывается с 5′-нетранслируемой областью, свойственной мРНК, в результате чего в этой области формируется шпилечная структура, препятствующая связыванию рибосом с инициирующим триплетом. В результате такой *авторегуляции* инициация трансляции мРНК треонил-тРНК-синтетазы останавливается. Такой способ регуляции, называемый также *регуляцией трансляции по принципу обратной связи*, характерен и для трансляции рибосомных белков бактерий. Обычно бактериальная клетка не допускает сверхпродукции рибосомных белков, синтезируя их в том количестве, которое достаточно для сборки рибосом. Скоординированный синтез этих белков происходит несмотря на то, что их гены представлены независимыми оперонами, рассеянными по всей ДНК, на которых синтезируются полицистроновые мРНК, кодирующие группы рибосомных белков. Координация синтеза достигается тем, что среди синтезированных на мРНК белков есть один белок, который может выступать в качестве репрессора трансляции своей мРНК. Так, например, один из продуктов трансляции (белок S7) полицистроновой мРНК (она кодирует также белок S12 и факторы элонгации EF-Tu и EF-G) при повышении его концентрации в клетке способен прикрепляться к определенному участку мРНК и блокировать трансляцию всех кодируемых ею белков. Таким образом, полагают, что регуляция синтеза рибосомных белков идет по простому и элегантному механизму: репрессия синтеза этих белков наступает тогда, когда исчерпывается количество рРНК, необходимое для сборки рибосом (пока идет процесс сборки, белки рибосом в силу их высокого сродства к рРНК используются для формирования рибосом, а по мере исчерпания рРНК избыток рибосомных белков подавляет трансляцию своих собственных мРНК).

В клетках эукариот в качестве белка-репрессора начального этапа трансляции определенных белков может выступать один из ферментов углеводного обмена — аконитат-гидратаза. Биохимикам было давно известно, что этот фермент катализирует взаимопревращение

лимонной и изолимонной кислот с образованием в качестве промежуточного соединения *цис*-аконитовой кислоты. Однако оказалось, что этот фермент задействован в регуляции трансляции мРНК ферритина — белка, связывающего токсичные для клеток ионы железа. Регуляция трансляции этой мРНК зависит от определенной последовательности, образующей шпилечную структуру в области 5′-нетранслируемого участка ее молекулы. Эта шпилька в отсутствие железа связывается со специфическим белком-репрессором — аконитат-гидратазой, которая стабилизирует структуру шпильки и тем самым препятствует продвижению (сканированию) рибосом по 5′-концу мРНК, подавляя таким образом инициацию трансляции мРНК-ферритина. Репрессор, имеющий сродство к ионам железа, при повышении их концентрации в клетке отходит от мРНК, что и вызывает инициацию трансляции и синтез ферритина. Синтезированный ферритин, в свою очередь, способен отнимать железо у белка-репрессора, который вновь может остановить трансляцию по вышеуказанному механизму (рис. 165).

Элонгация трансляции также подвержена регуляции, на что указывает непостоянная скорость трансляции, выявленная при

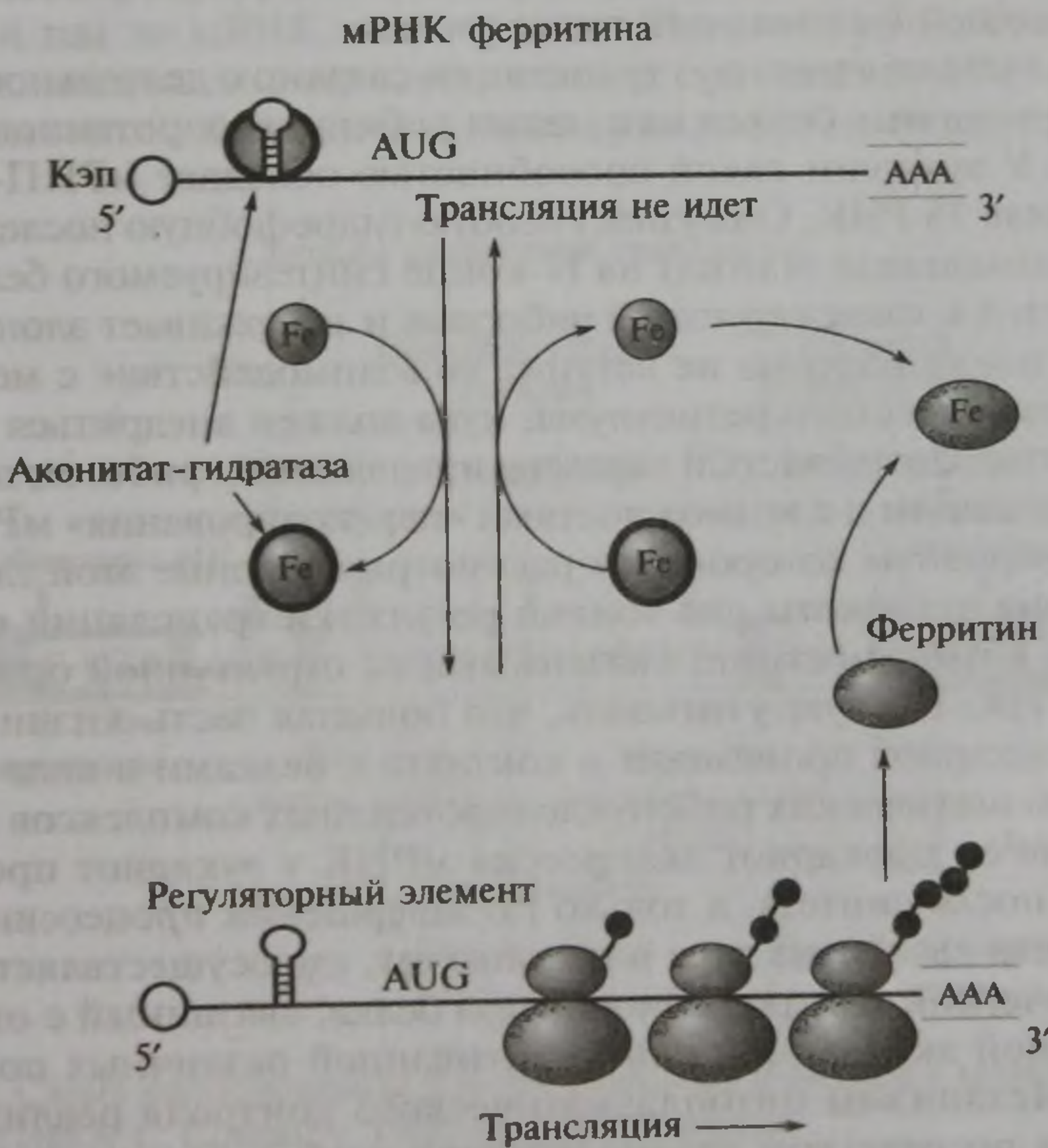


Рис. 165. Регуляция трансляции мРНК ферритина с участием аконитат-гидратазы и ионов железа. Пояснения см. в тексте

«прочтении» рибосомами разных мРНК. Такая прерывистая, содержащая паузы трансляция выявлена при синтезе глобиновых цепей в ретикулоцитах кролика, в интактных клетках и бесклеточных системах синтеза для других белков: фиброина шелка, вителлогенина и сывороточного альбумина позвоночных, белков вируса табачной мозаики и белка оболочки фага MS2.

Существует три объяснения пауз трансляции. Первое объяснение состоит в том, что рибосомы могут задерживаться на кодонах, соответствующих минорным (т. е. присутствующих в клетках в малых количествах) изоакцепторным тРНК. Кодоны минорных тРНК называют *модулирующими*, так как они могут модулировать скорость трансляции. Чем больше этих кодонов в мРНК, тем ниже скорость трансляции.

Второе объяснение прерывистого хода трансляции состоит в неодинаковой доступности разных участков мРНК для рибосом: скорость трансляции неизбежно снижается при преодолении рибосомами стабильных элементов вторичной и третичной структуры различных участков мРНК. Преодолению «трудных» участков мРНК у эукариот способствует фактор элонгации eEF-2, обладающий в отличие от фактора EF-G прокариот РНК-связывающей и РНК-расплетающей (хеликазной) активностью.

Третье объяснение пауз трансляции связано с деятельностью особых регуляторных белков или малых рибонуклеопротеинов (мРНП-частиц). У эукариот такой способностью обладает мРНП-частица, содержащая 7S РНК. Она узнает особую гидрофобную последовательность (сигнальный пептид) на N-конце синтезируемого белка, присоединяется к транслирующей рибосоме и задерживает элонгацию до тех пор, пока рибосома не вступит во взаимодействие с мембраной эндоплазматического ретикулума, куда должен внедриться сигнальный пептид. Прерывистый характер продвижения рибосом по мРНК, вероятно, связан и с возможностями «перекодирования» мРНК, различные варианты которого мы рассмотрим в конце этой главы.

Высшие эукариоты для тонкой регуляции трансляции обладают особыми возможностями, связанными со структурной организацией их мРНК. Следует учитывать, что большая часть жизни мРНК в клетках эукариот происходит в контакте с белками в виде ядерных и цитоплазматических рибонуклеопротеидных комплексов (мРНП). В отличие от прокариот экспрессия мРНК у эукариот происходит не сразу после синтеза, а только по завершении процессинга и последующего выхода из ядра в цитоплазму, где осуществляется цитоплазматический контроль биосинтеза белка, связанный с отбором и селективной активацией или инактивацией различных популяций мРНК. Механизмы цитоплазматического контроля реализуются в основном посредством взаимодействия особых белков — факторов трансляционного контроля — с определенными участками молекулы мРНК (рис. 166).

В эукариотических мРНК сразу за кэпом располагается 5′-нетранслирующая область (5′-UTR), с которой связываются белки, играющие важнейшую роль в регуляции инициации трансляции. К их числу относятся собственно белковые факторы инициации (eIF), а также белки-репрессоры трансляции. В инициации трансляции принимают участие также 3′-нетранслируемая область (3′-UTR) и поли(A)-хвосты мРНК. В 3′-UTR имеются и особые последовательности: *сигналы полиаденилирования и сигнал внутриклеточной локализации*. С 3′-UTR способны связываться *маскирующие белки* (они делают мРНК недоступной для трансляции рибосомами) и дестабилизирующие белки, детерминирующие доступность мРНК для атаки нуклеаз и таким образом определяющие время жизни мРНК в клетке. В достаточно протяженной области 3′-UTR присоединяется также белок мРНП.

Кодирующий участок (транслируемая область) мРНК способен связывать не только вышеохарактеризованные факторы элонгации (eEF), АРСазы и рибосомы, но и *коровый белок* мРНП-*частиц*, а также определенные ферменты, необходимые для процесса трансляции. В кодирующей области мРНК расположены также сигналы *перекодирования* (репрограммирования) трансляции, благодаря которым с одной и той же мРНК могут транслироваться разные белки.

Функциональная роль кэпа, а также значение 5′-UTR и связывающихся с этой областью факторов трансляционного контроля рас-

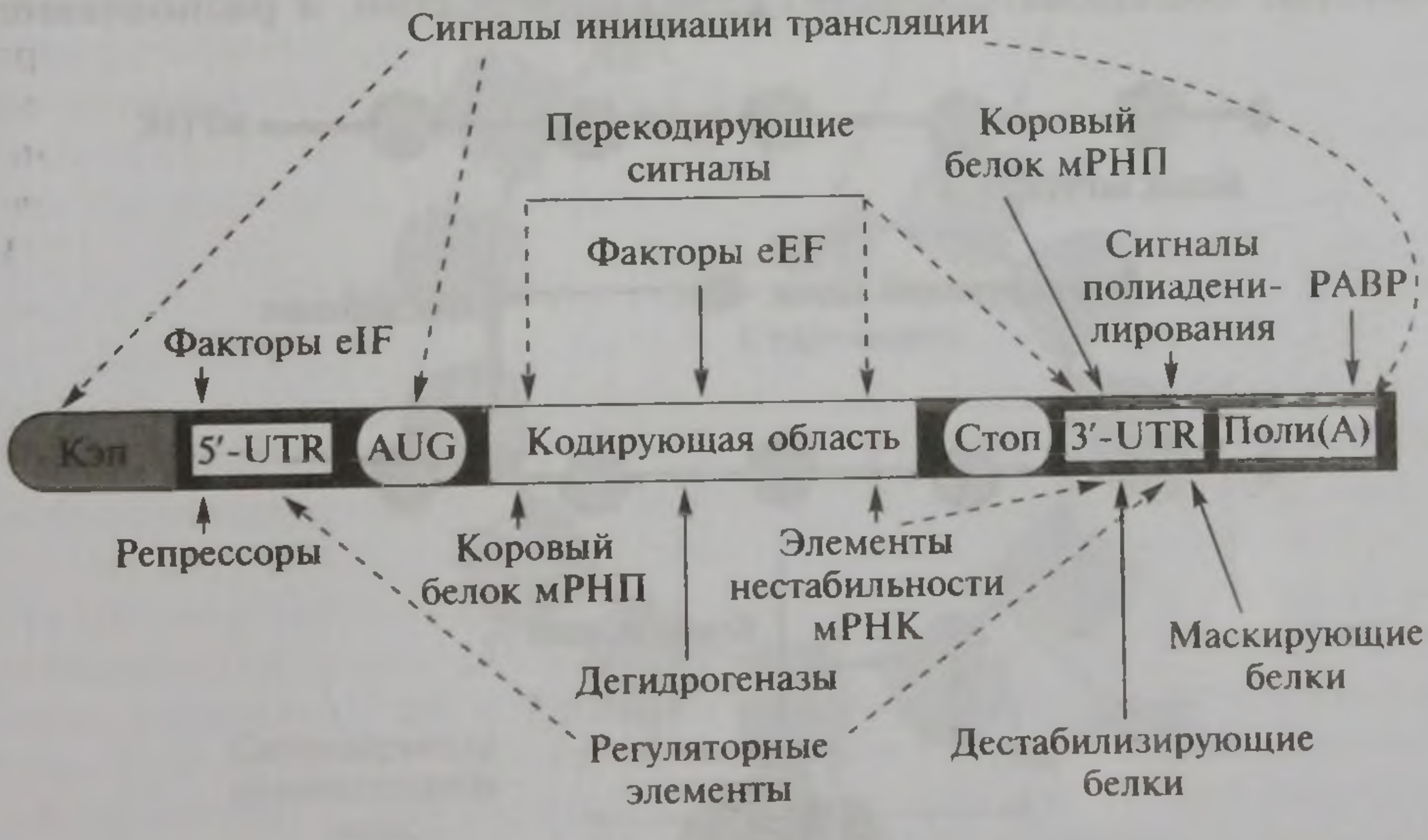


Рис. 166. Схема расположения функциональных участков в молекуле мРНК эукариот (по А. С. Спирину, 1996, с дополнениями).

Сплошные стрелки — области присоединения белков, участвующих в регуляции трансляции, пунктирные стрелки — расположение тех или иных регуляторных последовательностей (сигналов) в различных участках мРНК

смотрены выше. Здесь кратко остановимся на механизмах контроля трансляции, реализуемых с участием других районов мРНК.

3′-UTR играет, вероятно, решающую роль в активации (или инактивации) и деградации мРНК, внутриклеточной локализации и структурной организации мРНП. Эта область отвечает за маскирование мРНК и усиление трансляции. *Маскирование* осуществляется белками, которые делают мРНК недоступной ни рибосомам, ни ферментам деградации (нуклеазам). Такая мРНК приобретает стабильную и неактивную в смысле трансляции форму и называется *запасной* мРНК. Эта РНК свойственна ооцитам и сперматозоидам, а в отдельных случаях и соматическим клеткам.

3′-UTR и связанные с ней белки участвуют в блокировании инициации трансляции, которая, как известно, происходит на противоположном конце мРНК (5′-UTR). Предполагают, что 3′- и 5′-UTR контактируют друг с другом либо с помощью белка-посредника, либо благодаря прямому сближению, происходящему с изменением структуры мРНП. В этом случае под действием маскирующего белка релаксированная (вытянутая) форма мРНП, вероятно, может переходить в конденсированную (компактную) форму, благодаря чему 5′- и 3′-UTR сближаются (рис. 167).

К 3′-UTR способны присоединяться *дестабилизирующие* белки, которые стимулируют быструю деградацию определенных видов мРНК. Такие короткоживущие мРНК выявлены у цитокининов и протоонкогенов (*c-myc*, *c-foc* и др.). Они содержат в 3′-UTR AU-богатые последовательности, которые, вероятно, и распознаются



Рис. 167. Предполагаемый механизм маскирования мРНК (по А. С. Спирину, 1996). Пояснения см. в тексте

дестабилизирующими белками. Напротив, существуют белки, стабилизирующие мРНК, которые взаимодействуют с особыми структурными элементами в 3′-UTR и защищают мРНК от атаки нуклеаз. К этим белкам относится вышеупомянутая аконитат-гидратаза, регулирующая трансляцию мРНК ферритина.

Весьма существенную роль в регуляции трансляции играет и поли(А) последовательность на 3′-конце мРНК. Часть мРНК эукариот полиаденилируется в ходе процессинга в ядре клеток (см. гл. 10), а сигналом для ядерного полиаденилирования служит последовательность AAUAAA, расположенная вблизи 3′-конца молекулы. В ряде случаев процессированные мРНК не проходят полиаденилирования в ядре и не используются в течение определенного времени при биосинтезе белка. Такие (запасные) формы мРНК выявлены в ооцитах амфибий и сухих зародышах пшеницы. Активация этих мРНК сопровождается их полиаденилированием в цитоплазме, для чего служит определенный сигнал в 3′-UTR. У амфибий таким сигналом является последовательность $(U)_6$ AUAAAG. Поли(А)-хвост мРНК находится

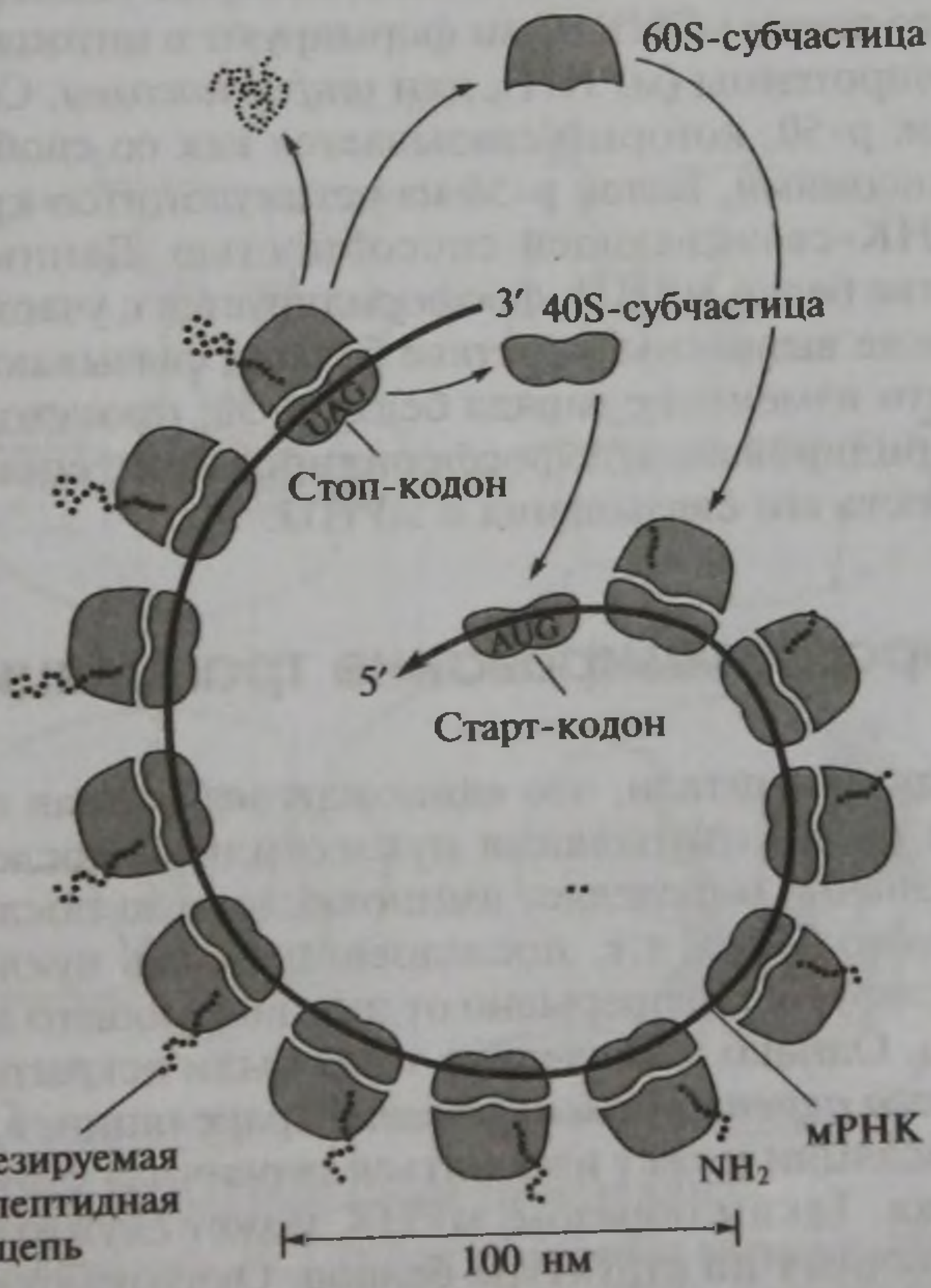


Рис. 168. Схематическое изображение полирибосомы, демонстрирующее возможность сближения 5′- и 3′-концов молекулы мРНК, где расположены транскэп, 5′-UTR и поли(А)-хвост, принимающие участие в инициации трансляции (по Б. Албертсу и др., 1994)

часто в комплексе с поли(А)-связывающим белком (PABP), молекулы которого выполняют функцию стабилизаторов мРНК. PABP содержит четыре РНК-связывающих домена, и его комплекс с поли(А)-хвостом не только предохраняет мРНК от деградации, но и выступает в качестве *энхансера инициации трансляции*. Предполагают, что PABP облегчает присоединение 60S-субчастицы рибосомы эукариот к инициирующему комплексу 40S-субчастицы. Этот процесс, естественно, требует сближения кэпа и 5′-UTR с поли(А)-хвостом, что, вероятно, достигается в ходе трансляции эукариотических мРНК при образовании *полирибосом* (полисом) — структур, состоящих из одной молекулы мРНК и целого ряда связанных с ней рибосом, осуществляющих последовательную интенсивную трансляцию этой мРНК (рис. 168).

Методом электронной микроскопии установлено, что полирибосомы в клетках эукариот имеют замкнутые структуры типа «кругов», «розеток» и т. п. Наряду с белками, распознающими определенные участки в мРНК, эукариотам свойственна особая группа *неспецифических мРНК-связывающих белков*, которые универсально взаимодействуют со всеми мРНК. Они формируют в цитоплазме матричные рибонуклеопротеины (мРНП), или *информосомы*. Среди них преобладает белок р-50, который связывается как со свободными мРНК, так и с полисомами. Белок р-50 из ретикулоцитов кролика обладает сильной РНК-связывающей способностью. Данный белок, как и многие другие белки мРНП, фосфорилируется с участием ферментов, которые также выявлены в составе белков, связывающихся с мРНК. Понятно, что изменение заряда белка р-50, происходящее в результате фосфорилирования/дефосфорилирования, способно влиять на эффективность его связывания с мРНП.

Перепрограммирование трансляции

Долгое время считали, что единожды избранная при инициации трансляции рамка считывания нуклеотидной последовательности мРНК одназначно определяет аминокислотную последовательность синтезируемого белка, т. е. последовательность нуклеотидов мРНК всегда транслируется непрерывно от инициирующего до терминирующего кодона. Однако в последние годы были вскрыты возможности *динамического перепрограммирования* трансляции, в результате чего по ходу трансляции может изменяться первичная структура синтезируемого белка. Таким образом, мРНК может служить матрицей для синтеза различных по структуре белков. Оказывается, что клетки и поражающие их вирусы с малым числом генов способны «экономно» использовать свой генетический потенциал не только путем альтернативного сплайсинга (см. выше), но и за счет перепрограммирования (перекодирования) мРНК.

Изменение рамки считывания (т. е. перепрограммирование трансляции в отношении порядка считывания кодонов мРНК) обычно происходит на один — два нуклеотида в разных направлениях: в сторону 5′-конца мРНК (–1 нуклеотид) или в сторону 3′-конца (+1, +2 нуклеотида). Так происходит в том случае, когда кодирующие (транслируемые) области некоторых мРНК оказываются «испорчены» присутствием в их составе терминирующих кодонов. Однако такие мРНК все же могут транслироваться за счет изменения рамки

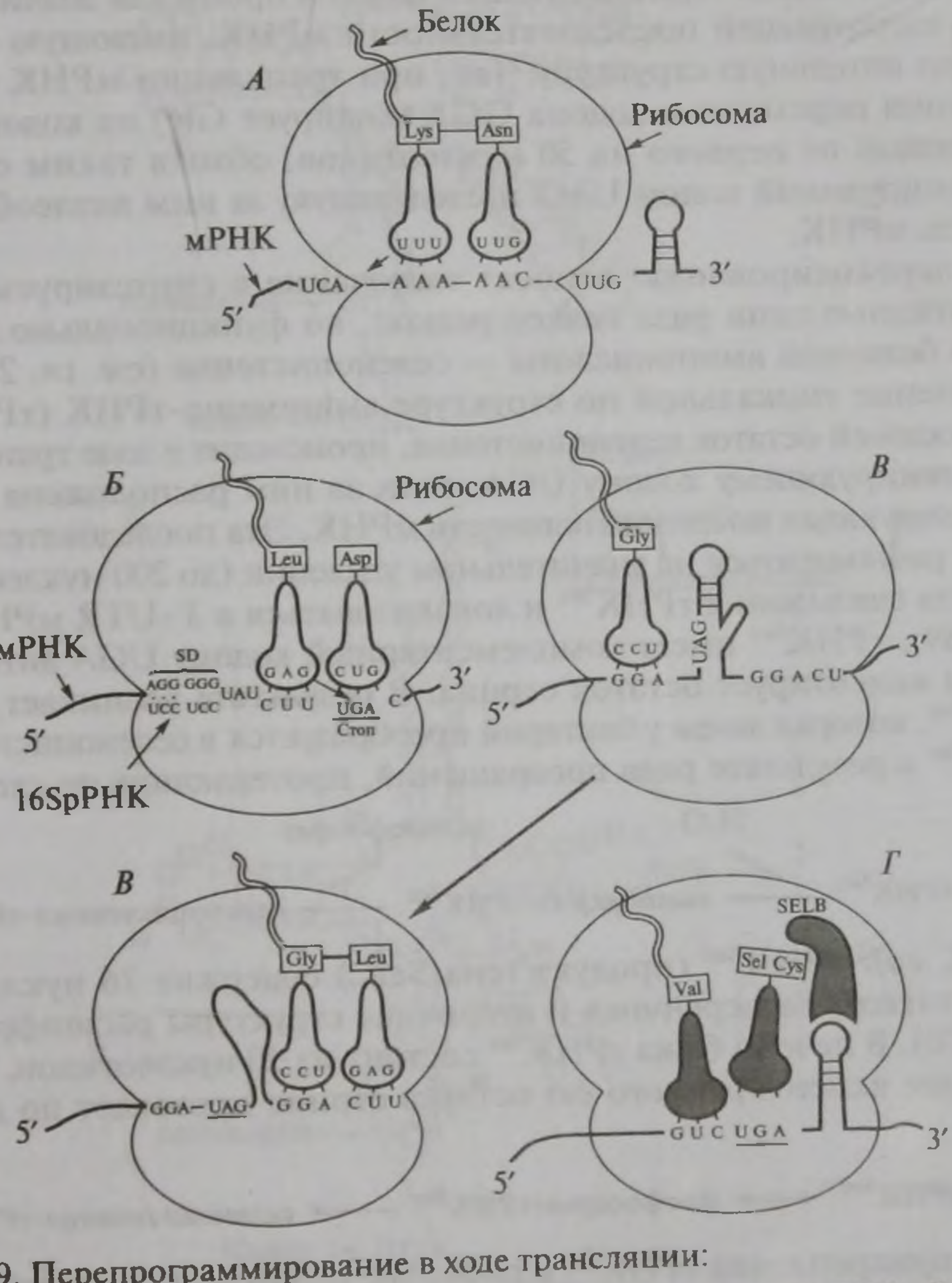


Рис. 169. Перепрограммирование в ходе трансляции:

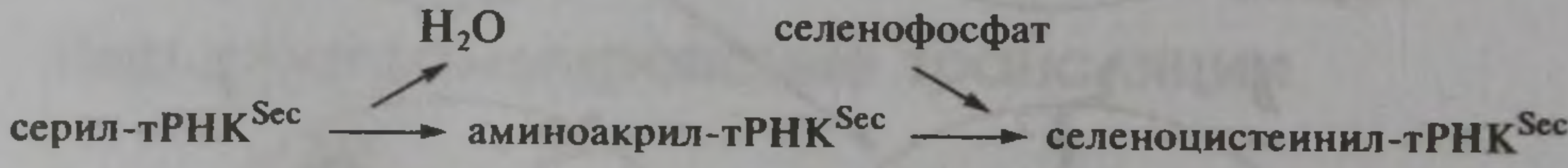
А — сдвиг рамки считывания на (–1) при трансляции вирусной РНК; *Б* — сдвиг рамки считывания на (+1) при трансляции мРНК бактериального фактора терминации RF-2 (SD — последовательность Шайна — Дальгарно); *В* — «прыжок» рибосомы на 50 нуклеотидов при трансляции мРНК одного из белков бактериофага Т4; *Г* — прочитывание терминирующего кодона UGA как кодона аминокислоты селеноцистенина (SelSyc) при биосинтезе формиатдегидрогеназы у *E. coli*

считывания, для чего необходимо наличие определенных сигналов в транслируемой области.

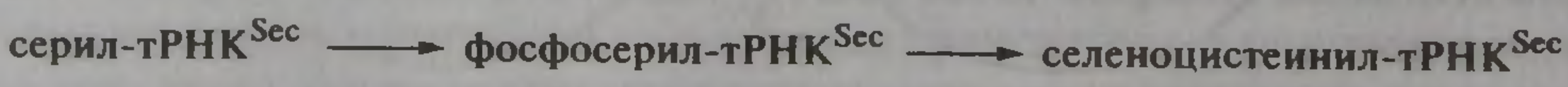
Сдвиг рамки считывания на (–1) при трансляции РНК ретровирусов происходит в области особой гептануклеотидной последовательности перед шпилечной структурой (рис. 169). У бактерий при трансляции мРНК, кодирующей белковый фактор терминации RF2, сдвиг рамки на +1 также зависит от окружения терминирующего кодона UGA.

В других случаях рибосомы в ходе трансляции могут совершать «прыжки», обходя терминирующий кодон и пропуская значительную часть кодирующей последовательности мРНК, имеющую обычно особую петельную структуру. Так, при трансляции мРНК фага T_4 рибосома переходит с кодона GGA (кодирует Gly) на кодон GGA, отстоящий от первого на 50 нуклеотидов, обходя таким образом терминирующий кодон UAG и следующую за ним петлеобразную область мРНК.

К перекодированию относят включение в синтезируемые полипептидные цепи ряда белков редкой, но функционально важной (21-й) белковой аминокислоты — селеноцистеина (см. гл. 2). Присоединение уникальной по структуре аминоацил-тРНК (тРНК^{Sec}), переносящей остаток селеноцистеина, происходит в ходе трансляции к терминирующему кодону UGA, если за ним расположена особая стимулирующая последовательность мРНК. Эта последовательность может располагаться на значительном удалении (до 200 нуклеотидов) от места связывания тРНК^{Sec} и локализоваться в 3′-UTR мРНК (см. рис. 169). тРНК^{Sec} имеет комплементарный кодону UGA антикодон ACU и акцептирует остаток серина. В результате возникает серил-тРНК^{Sec}, которая затем у бактерий преобразуется в селеноцистеинил-тРНК^{Sec} в результате ряда превращений, протекающих по схеме:



У *E. coli* тРНК^{Sec} (продукт гена SelC) содержит 76 нуклеотидных остатков. Ее первичная и вторичная структуры расшифрованы (рис. 170). В печени быка тРНК^{Sec} состоит из 90 нуклеотидов, и превращение акцептируемого ею остатка серина протекает по другой схеме:



Для прикрепления тРНК^{Sec} к рибосоме необходим специфический белковый фактор SEL B (молекулярная масса 68 кДа), который представляет собой аналог фактора элонгации EF-Tu с N-концевым доменом, связывающим GTP. В мРНК формиатдегидрогеназы (одного из селеноцистеинилсодержащих белков *E. coli*) идентифицирована особая зона узнавания длиной 40 нуклеотидов в области кодона

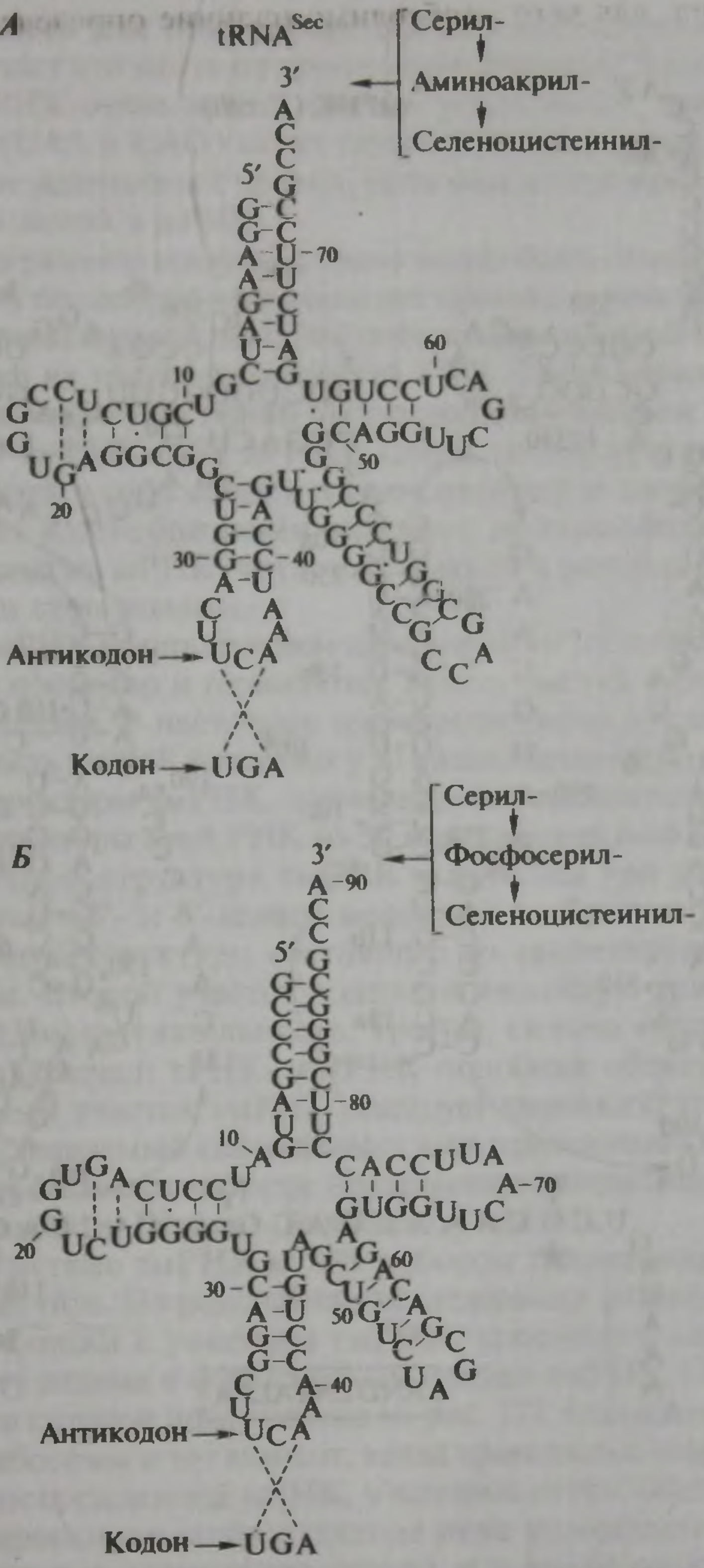


Рис. 170. Структура тРНК^{Sec} (*Stadtman T.C.* // Ann. Rev. Biochem. — 1996. — V. 65. — P. 83):

А — структура тРНК^{Sec} прокариот. У *E. coli* тРНК^{Sec} акцептирует серин, а затем преобразуется в селеноцистеинил-тРНК; *Б* — структура тРНК^{Sec} эукариот. В печени быка тРНК^{Sec} также акцептирует серин, а затем превращается в селеноцистеинил-тРНК

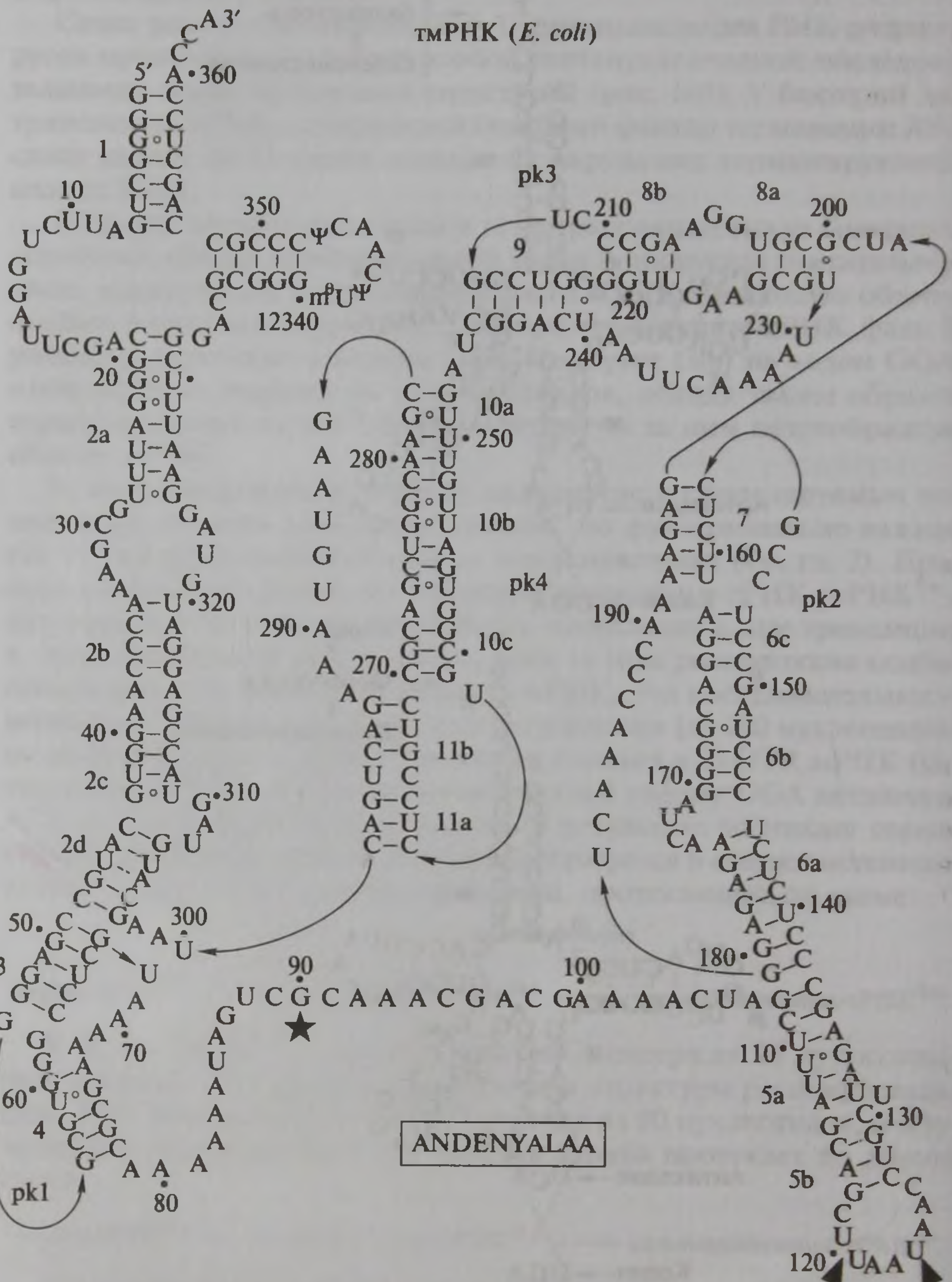


Рис. 171. Модель вторичной структуры тмРНК *E. coli* (М.Э.Зверева и др., 2000).

Звездочкой обозначено начало транслируемого района; ▲ — стоп-кодоны в рамке считывания; pk1—pk4 — псевдоузлы. В рамке выделена аминокислотная последовательность tag-пептида

UGA, служащего для присоединения селеноцистеинил-тРНК^{Sec}, которая отличает это место от других стоп-кодонов. Данный элемент структуры мРНК очень важен, так как установлено, что два других стоп-кодона (UAA и UAG) могут служить для включения селеноцистеина в синтезированный пептид, если они ассоциированы с такой «узнаваемой» зоной в мРНК.

К перепрограммированию условно может быть отнесен и активно исследуемый в последние несколько лет процесс *транс-трансляции*, когда синтез полипептида, начатый рибосомой на одной мРНК, затем переключается на трансляцию другой РНК. Такой вариант трансляции был выявлен у эубактерий. Он проходит с участием необычного вида РНК — тмРНК (10S аРНК). Столь небольшая по размерам РНК совмещает в себе свойства транспортной и матричной РНК. Она участвует в освобождении рибосом, не способных закончить биосинтез белка на мРНК, которые утратили в результате случайной модификации стоп-кодоны.

У *E. coli* тмРНК кодируется моноцистроновым геном ssrA, имеющим собственный промотор и терминатор. Размер тмРНК варьирует от 349 до 411 нуклеотидов. В настоящее время определена нуклеотидная последовательность тмРНК более чем у 50 видов бактерий. Создана модель вторичной структуры тмРНК, основанная на сравнительном анализе первичной структуры этой РНК из 50 микроорганизмов (рис. 171).

Во вторичной структуре тмРНК выражены три участка: один из них включает 3′- и 5′-концы молекулы и образует характерную тРНК-подобную структуру, состоящую из акцептирующего стебля и ТΨС-петли. Второй участок содержит линейную транслируемую (матричную) последовательность. Третий, сильно структурированный участок включает тРНК- и мРНК-подобные области молекулы. Транслируемый участок тмРНК кодирует короткий полипептид — tag-пептид, узнаваемый специфическими протеазами. Кодирующей tag-пептид последовательности предшествует типичная адапторная последовательность Шайна—Дальгарно (AAGG).

Взаимодействие тмРНК и 70S рибосом исследовано в системе трансляции in vitro. По результатам исследования разработана модель *транс*-трансляции с участием тмРНК, способной акцептировать аминокислоту аланин с образованием аланил-тмРНК. Согласно этой модели, схема которой представлена на рис. 172, аланил-тмРНК входит в А-центр рибосомы в тот момент, когда трансляция останавливается на 3′-конце поврежденной мРНК, у которой отсутствует стоп-кодон. Недосинтезированная полипептидная цепь переносится на аланил-тмРНК (реакция транспептидирования), и далее рибосома продолжает трансляцию, используя матричный участок тмРНК. Синтез продолжается до поступления в А-центр стоп-кодона тмРНК (UAA), после чего вступает в действие фактор терминации и трансляция завершается. После терминации трансляции от рибосомы отсоединяется *гибридный белок*, содержащий на С-конце tag-пептид, который «распознается»

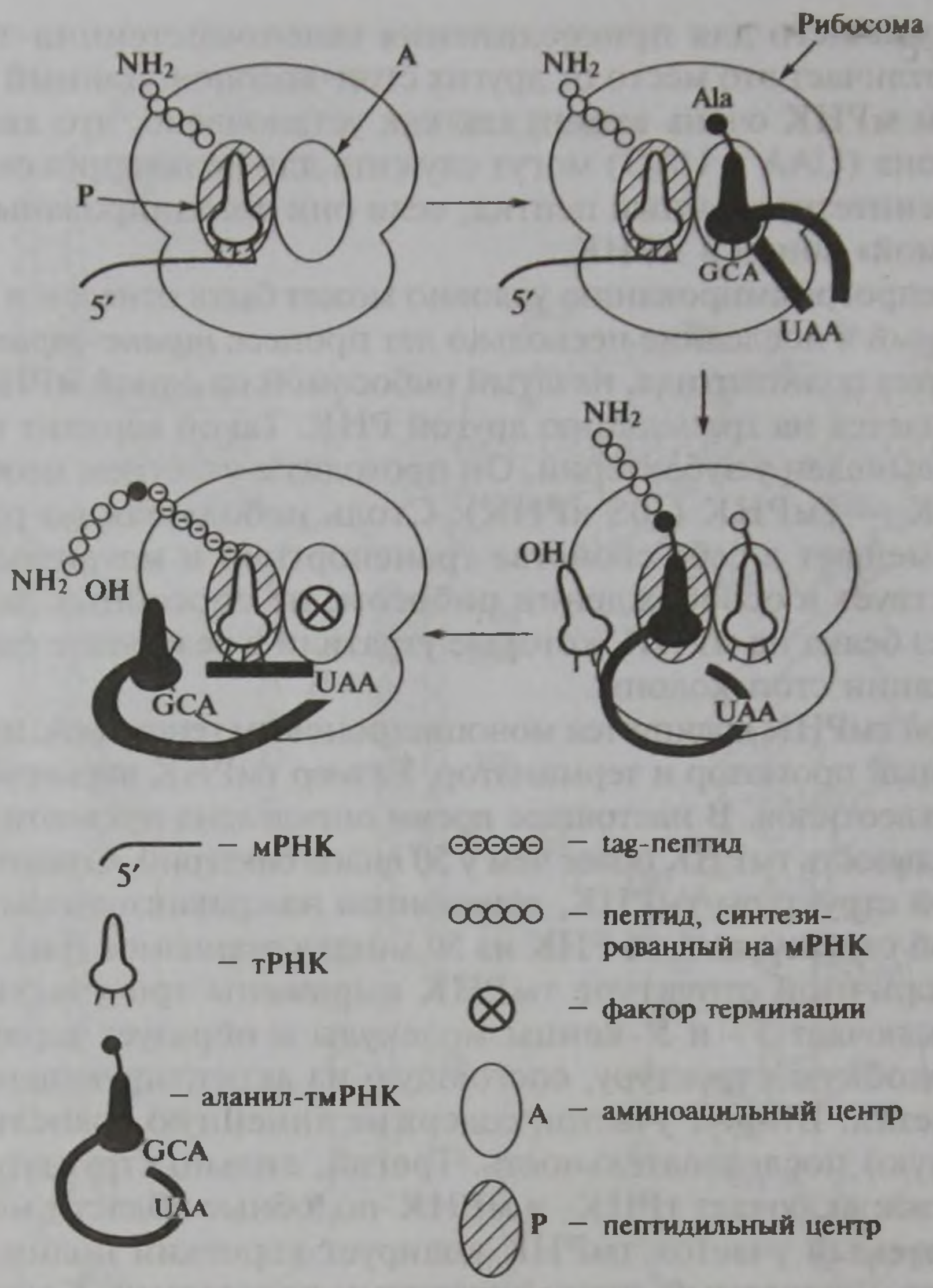


Рис. 172. Схема *транс*-трансляции с участием аланил-тмРНК (М.Э.Зверева и др., 2000). Пояснения см. в тексте.

протеазами, атакующими и разрушающими необычный продукт трансляции. Терминация освобождает ранее связанную рибосому, которая далее может участвовать в синтезе других белков.

Таким образом, тмРНК выполняет две существенные функции: освобождает рибосомы от связи с дефектными мРНК и одновременно способствует быстрой деградации «неправильных» белков, присоединяя к ним узнаваемый протеазами tag-пептид. Поскольку тмРНК отсутствует у высших организмов, не исключена возможность ее использования в качестве мишени при создании новых антибактериальных препаратов, которые могли бы блокировать функции тмРНК у бактерий и привести к гибели патогена, не затрагивая синтез белков человека.

Глава 12
РЕПАРАЦИЯ ДНК

Передача наследственной информации в неискаженном виде — важнейшее условие выживания как отдельного организма, так и вида в целом. Большинство изменений в структуре ДНК совершенно недопустимы: они либо ведут к вредным мутациям, либо блокируют репликацию ДНК и вызывают гибель клеток. Между тем ДНК постоянно подвергается химическим изменениям в результате воздействия как спонтанных, так и индуцированных факторов среды: УФ-облучением, ионизирующей радиацией, химическими мутагенами, температурой и др. К спонтанным повреждениям относятся: ошибки репликации (в результате появляются некомплементарные пары нуклеотидов — *мисмэтчи*); апуринизация (отщепление азотистых оснований от сахаро-фосфатного остова — образование *АР-сайтов*) и дезаминирование (отщепление аминогруппы от азотистого основания).

К индуцируемым повреждениям принято относить: димеризацию (сшивание соседних пиримидиновых оснований с образованием димера); размыкание пуринового кольца; однонитевые и двунитевые разрывы в ДНК; сшивки между цепями ДНК. В ходе эволюции выработалась система, позволяющая исправлять нарушения в ДНК, вызванные ошибками репликации или повреждающими агентами внешней среды, — система репарации ДНК. В результате ее активности на 1 000 повреждений в ДНК только одно приводит к мутации. Нарушения в системе репарации могут приводить к преждевременному старению, развитию онкологических заболеваний, болезням аутоиммунной системы и целому ряду других генетически обусловленных дефектов. Так, у людей, страдающих *пигментной ксеродермой* (наследственное заболевание, выражающееся в очагах рака кожи), нарушение системы репарации приводит к тому, что они не могут бывать на солнечном свету в силу полной незащищенности от УФ-облучения. Частота заболевания раком увеличивается с возрастом, вероятно, за счет накоплений повреждений в ДНК, связанных с ослаблением системы репарации. В специальных определениях установлено, что в геноме зародышевой линии клеток млекопитающих и человека происходит в среднем 6 нуклеотидных замен в год. Вероятно, и в соматических клетках происходит такое же количество мутаций. Их накопление с возрастом повышает вероят-

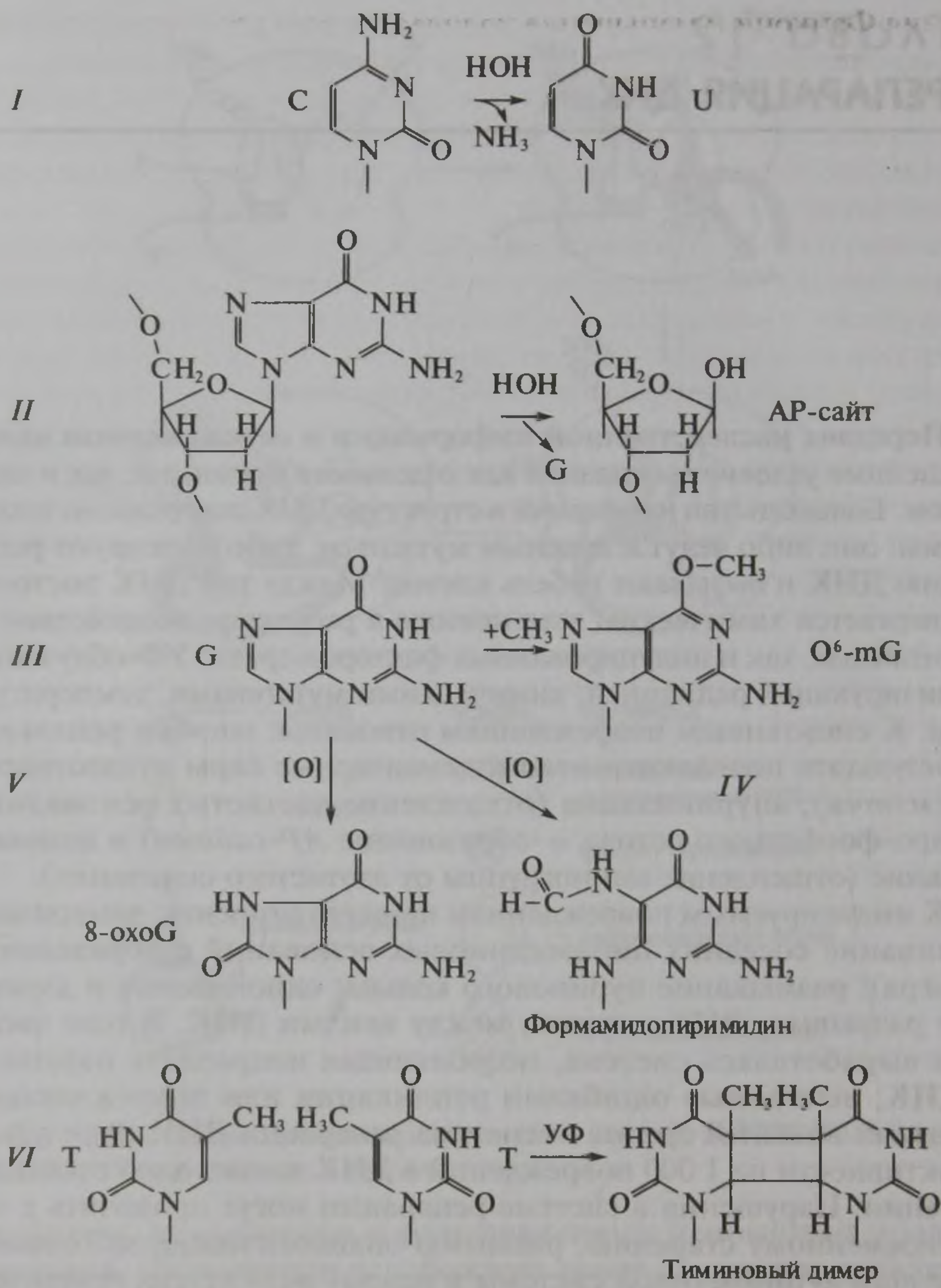


Рис. 173. Наиболее часто выявляемые повреждения в нуклеотидах ДНК:
I — дезаминирование; *II* — апуринизация; *III* — метилирование; *IV* — размыкание кольца; *V* — окисление; *VI* — образование тиминовых димеров

ность ракового перерождения клеток. В целом полагают, что 80—90% всех раковых заболеваний связаны с отсутствием репарации ДНК. Многие другие наследственные болезни человека также связаны с дефектами в работе системы репарации. К ним, в частности, относят триходистрофию (нехватка серы в клетках волос, ведущая к их ломкости; аномалии кожи и зубов; дефекты полового развития), синдром Кокэйна (карликовость, глухота, атрофия зрения и др.);

анемия Фанкони (уменьшение количества всех клеточных элементов крови, скелетные нарушения, микроцефалия, потеря слуха).

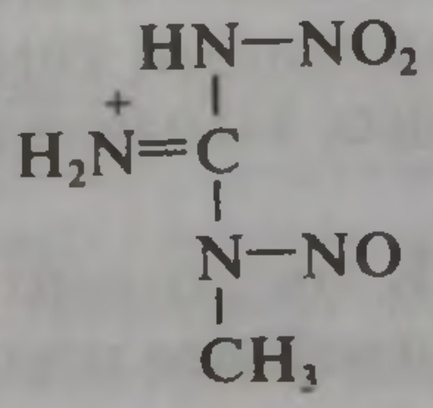
Известны четыре основных типа повреждений в ДНК: повреждение одиночных нуклеотидов; повреждение пары нуклеотидов; разрыв цепей ДНК; образование поперечных сшивок между основаниями одной цепи или разных цепей ДНК. Несомненно, что система репарации способна противостоять всем типам повреждений в ДНК, однако механизмы репарации наиболее изучены только в отношении первых двух типов, в основе которых лежат изменения структуры гетероциклических азотистых оснований. Устранение разрывов в цепях ДНК, вероятно, достигается прямым лигированием с участием ДНК-лигаз либо в процессе рекомбинации молекул ДНК (см. гл. 8), а механизмы устранения поперечных сшивок пока не изучены.

Наиболее часто наблюдаемые повреждения нуклеотидов включает *окисление*, *дезаминирование* и *алкилирование азотистых оснований*, образование *пиримидиновых димеров*, а также вставку нуклеотидов или в*ключение оснований-аналогов* (образование неканонических пар оснований), гидролиз N-гликозидных связей между основанием и дезоксирибозой (*апуринизация*). Схемы этих повреждений представлены на рис. 173.

Так, спонтанное (или индуцированное повышением температуры) дезаминирование азотистых оснований может изменять структуру ДНК, превращая цитозин в урацил— азотистое основание, не свойственное ДНК и распознаваемое системой репарации (рис. 173, *I*).

Апуринизация — гидролитическое выщепление азотистых (преимущественно пуриновых) оснований из полинуклеотидной цепи ДНК — приводит к образованию так называемых АР-сайтов (рис. 173, *II*), что связано с термолабильностью N-гликозидной связи пуриновых нуклеотидов. Однако в целом термин «АР-сайт» объединяет все случаи выщепления оснований с образованием и апуриновых, и апиримидиновых сайтов. Полагают, что ежедневно ДНК каждой клетки человека теряет от 5 000 до 10 000 пуриновых оснований, которые должны быть замещены (вставлены) ферментами репарации.

Алкилирование ДНК происходит при участии алкилирующих реагентов (мутагенов), большинство из которых являются канцерогенами. Один из продуктов такой модификации — O^6-метилгуанин — представлен на рис. 173, *III*. К числу мощных алкилирующих (метилирующих) агентов принадлежат *афлатоксин* низших грибов рода *Aspergillus* (поражают растения арахиса), а также разнообразные *нитрозоамины*, как, например, N-метил-N-нитро-N-нитрозогуанидин:



Нитрозоамины могут возникать из вторичных аминов и азотистой кислоты и ее солей (нитритов), которые в свою очередь образуются в организме человека в процессе восстановления нитратов, поступающих с пищей.

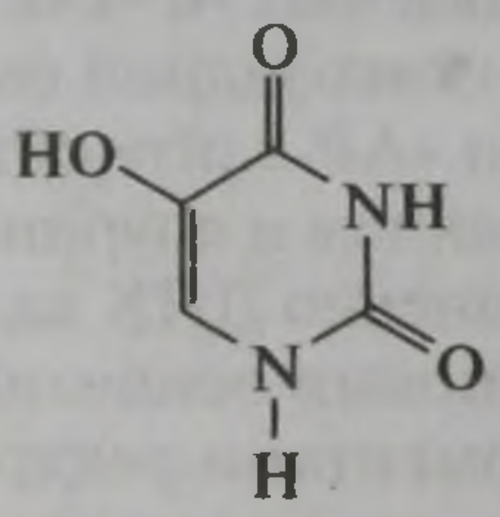
К числу наиболее мощных алкилирующих реагентов относятся иприт и его производные:

$$ClCH_2CH_2—S—CH_2CH_2Cl$$

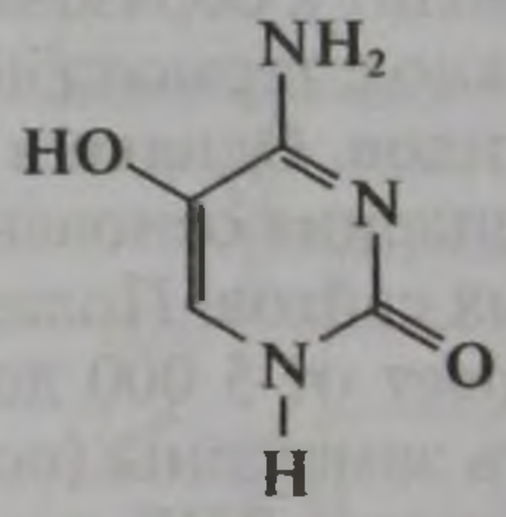
иприт (β, β′-дихлордиэтилсульфид)

Производные иприта способны вызывать многочисленные поперечные сшивки в молекуле ДНК, которые ведут к летальному исходу.

Окисление азотистых оснований в ДНК вызывается различными *активными формами кислорода* (АФК). Среди них наиболее выраженными окислительными свойствами обладают *супероксиданион-радикал* ($O_2^{\bullet}$) и *одноэлектронный гидроксил* ($^{\bullet}OH$). Оба эти соединения могут вызывать окислительные повреждения в ДНК: размыкание пуринового кольца с образованием формамидопиримидинов (рис. 173, *IV*) либо окисление пуринов (рис. 173, *V*). В последнем случае наиболее часто образуется 8-оксогуанин (8-охо-G) — производное гуанина, которое является мутагеном и способно образовывать водородные связи не с цитозином, а с аденином и таким образом приводит к образованию мутации (*трансверсии*) в процессе репликации ДНК (GC-пара заменяется на AT-пару). Наиболее стабильными продуктами окисления цитозина являются 5-гидроксицитозин (5—OH—C) и 5-гидроксиурацил (5—OH—U):



5-Гидроксиурацил
(5—OH—U)



5-Гидроксицитозин
(5—OH—C)

Наличие таких оснований в матричной цепи ДНК приводит к тому, что в процессе репликации образуются неканонические пары оснований: (5—OH—C:A; 5—OH—U:C), что ведет к трансверсиям C → T и C → G.

Тиминовые димеры возникают в ДНК за счет образования ковалентных связей между соседними тиминами, расположенными в одной цепи ДНК. Этот процесс вызывается УФ-облучением (рис. 173, *VI*).

Системы репарации ДНК достаточно консервативны в эволюции от бактерий до человека и наиболее изучены у *E. coli*.

Известны два типа репарации: *прямая* и *эксцизионная* (от англ. *excision* — вырезание). Прямая репарация — наиболее простой путь устранения повреждений в ДНК, в котором обычно задействованы специфические ферменты, способные быстро (как правило, в одну стадию) устранять соответствующее повреждение, восстанавливая исходную структуру нуклеотидов. Так действует, например, O^6-метилгуанин-ДНК-метилтрансфераза, которая снимает метильную группу с азотистого основания на один из собственных остатков цистеина (рис. 174, *А*). У *E. coli* может в 1 мин синтезироваться до 100 молекул этого белка. Аналогичный по функциям белок высших эукариот, очевидно, играет важную роль в защите от рака, вызываемого внутренними и внешними алкилирующими факторами. АР-сайты могут репарироваться путем прямой вставки пуринов при участии ферментов, называемых ДНК-инсертазами (от англ. *insert* — вставлять).

Тиминовые димеры «расшиваются» путем прямой репарации при участии *фотолиаз*, осуществляющих соответствующее фотохимическое превращение (рис. 174, *Б*). ДНК-фотолиазы представляют собой группу ферментов, активируемых светом, с длиной волны 300 — 600 нм (видимая область), для чего в их структуре имеется особый светочувствительный центр. Они широко распространены в природе и обнаружены у бактерий, дрожжей, насекомых, рептилий, земноводных и человека. Эти ферменты нуждаются в разнообразных кофакторах (FADH, тетрагидрофолиевая кислота и др.), участвующих в фотохимической активации фермента. Фотолиаза *E. coli* представляет собой белок с молекулярной массой 35 кДа, прочно связанный

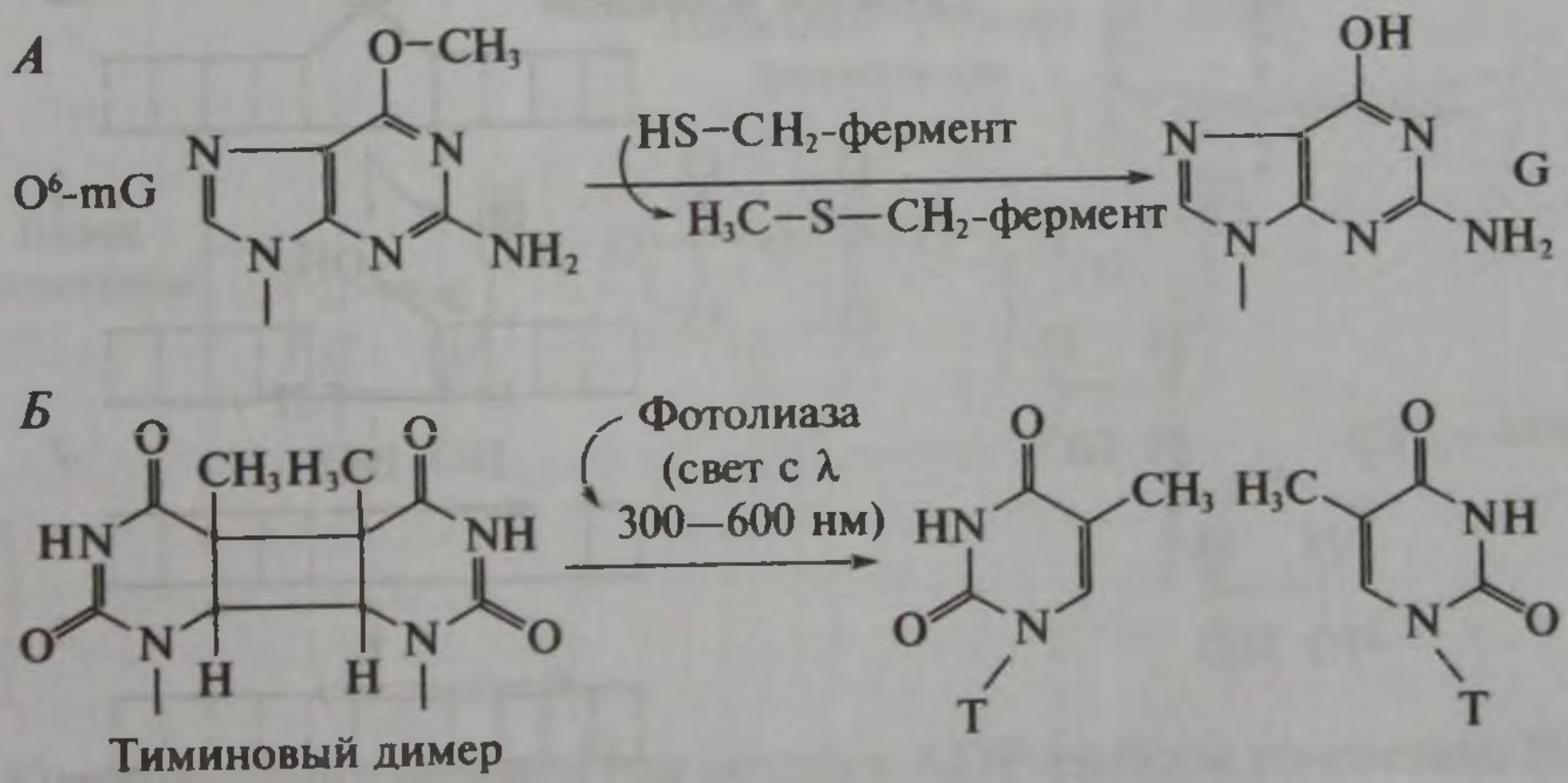


Рис. 174. Примеры прямой репарации повреждений в ДНК:

А — метилированное основание (O^6-mG) демитилируется ферментом метилтрансферазой, которая переносит метильную группу на один из своих остатков цистеина; *Б* — фотолиаза присоединяется к тиминовому димеру, и после облучения этого комплекса видимым светом (300 — 600 нм) димер расшивается

с олигорибонуклеотидом длиной 10 — 15 нуклеотидов, необходимым для активности фермента.

Эксцизионная репарация включает удаление поврежденных азотистых оснований из ДНК и последующее восстановление нормальной структуры молекулы.

В эксцизионной репарации обычно принимают участие несколько ферментов, а сам процесс затрагивает не только поврежденный, но и соседние с ним нуклеотиды. Кроме этого, для эксцизионной репарации необходима вторая (комплементарная) цепь ДНК. Общая упрощенная схема эксцизионной репарации представлена на рис. 175.

Первым этапом эксцизионной репарации является вырезание аномальных азотистых оснований. Его катализируют ДНК-N-гликозилазы — ферменты, расщепляющие гликозидную связь между дезоксирибозой и азотистым основанием. У человека ДНК-

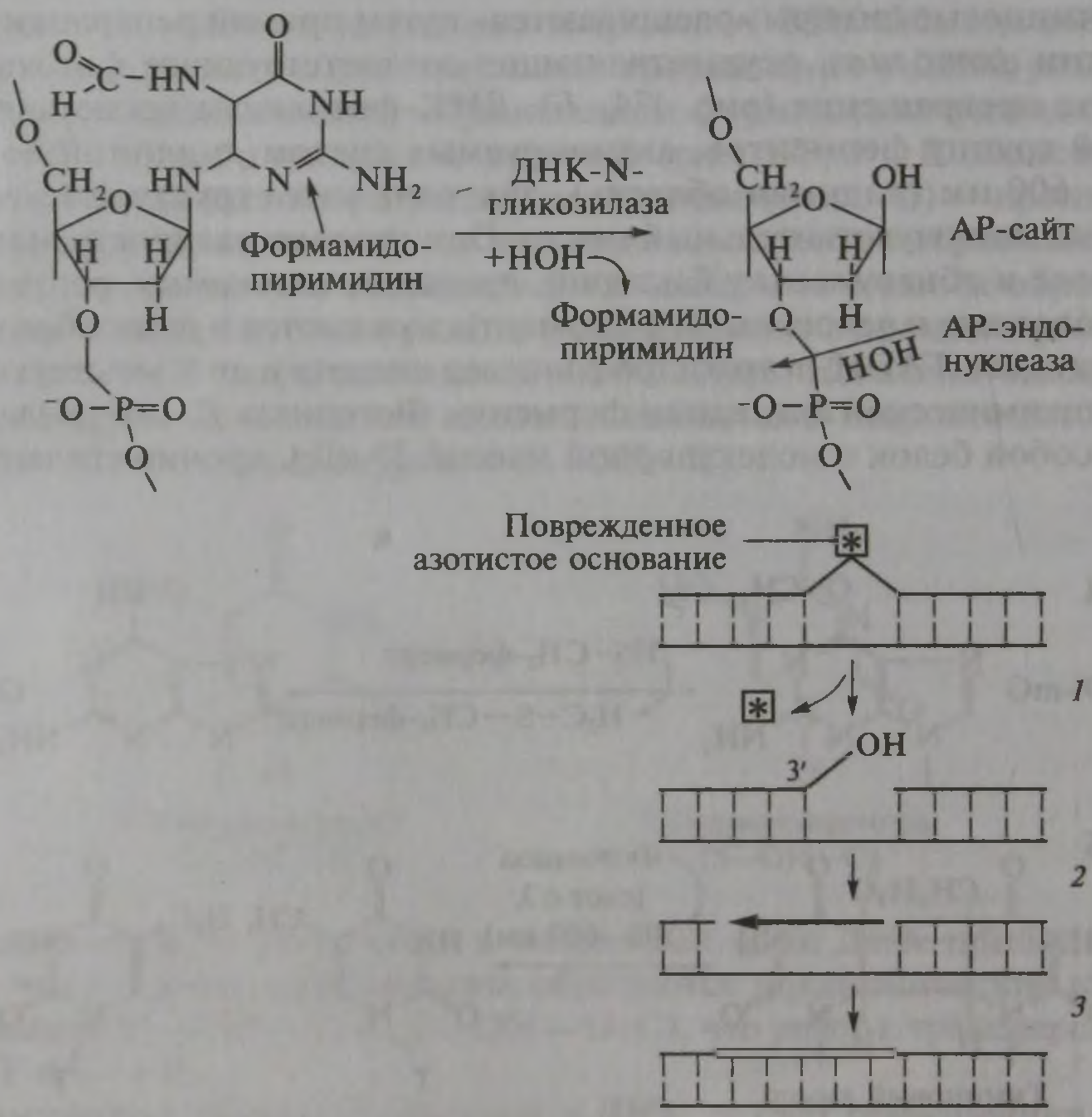


Рис. 175. Схема эксцизионной репарации:

1 — ДНК-N-гликозилаза удаляет поврежденное основание, AP-эндонуклеаза вносит разрыв в цепь ДНК; *2* — экзонуклеаза удаляет ряд нуклеотидов; *3* — ДНК-полимераза заполняет освободившийся участок комплементарными нуклеотидами, ДНК-лигаза сшивает репарированную цепь ДНК. Вверху детализирована стадия *1*

N-гликозилазы обладают высокой субстратной специфичностью: разные ферменты этого семейства распознают и вырезают различные аномальные основания (8-оксогуанин, урацил, метилпурины и др.). В результате действия ДНК-N-гликозилазы образуется АР-сайт, который атакуется ферментом АР-эндонуклеазой. Она разрывает сахарофосфатный остов молекулы ДНК в АР-сайте и тем самым создает условия для работы следующего фермента — экзонуклеазы, которая последовательно отщепляет несколько нуклеотидов от поврежденного участка одной цепи ДНК. У бактерий освобожденное место далее заполняется соответствующими нуклеотидами при участии ДНК-полимеразы I, ориентирующейся на вторую (комплементарную) цепь ДНК. Поскольку ДНК-полимераза I способна удлинять 3′-конец одной из цепей в месте разрыва в двуцепочечной ДНК и удалять нуклеотиды с 5′-конца того же разрыва, т. е. осуществлять *ник-трансляцию* (см. гл. 14), этот фермент играет ключевую роль в репарации ДНК. Окончательное сшивание репарированных участков осуществляет ДНК-лигаза.

Эксцизионная репарация ДНК в клетках млекопитающих сопровождается резким всплеском активности еще одного фермента — поли(ADP-рибозо)-полимеразы. При этом происходит ADP-рибозилирование белков хроматина (гистонов и негистоновых белков), что ведет к ослаблению их связи с ДНК и открывает доступ ферментам репарации. Донором ADP-рибозы в этих реакциях выступает NAD^+, запасы которого сильно истощаются при эксцизионной репарации повреждений, вызываемых рентгеновским облучением:

Отрицательно заряженные остатки ADP-рибозы из состава NAD^+ присоединяются через радикал глутаминовой кислоты или фосфосерина к белкам хроматина, что ведет к нейтрализации положительных зарядов этих белков и ослаблению их контакта с ДНК.

ДНК-гликозилазы, участвующие в устранении окислительных повреждений ДНК в клетках прокариот и эукариот, весьма раз-

нообразны и отличаются по субстратной специфичности, пространственной структуре и способам взаимодействия с ДНК. К наиболее изученным ДНК-гликозилазам относятся эндонуклеаза III (EndoIII), формамидопиримидин-ДНК-гликозилаза (Fpg), Mut T и Mut Y кишечной палочки.

Эндонуклеаза III *E. coli* «узнает» и специфически выщепляет из ДНК окисленные пиримидиновые основания. Этот фермент представляет собой мономерный глобулярный белок, состоящий из 211 аминокислотных остатков (мол. масса 23,4 кДа). Ген, кодирующий EndoIII, секвенирован, установлена его нуклеотидная последовательность. EndoIII представляет собой железосерный белок [$(4Fe–4S)^{2+}$-белок], обладающий элементом надвторичной структуры типа «греческий ключ» (спираль — шпилька — спираль), служащим для связывания с ДНК. Ферменты с аналогичной субстратной специфичностью и сходной аминокислотной последовательностью выделены также из клеток быка и человека.

Формамидопиридин-ДНК-гликозилаза *E. coli* «узнает» и выщепляет из ДНК окисленные гетероциклические основания пуринового ряда.

Mut T — небольшой белок с молекулярной массой 15 кДа, обладающий нуклеозидтрифосфатазной активностью, который преимущественно гидролизует dGTP до dGMP и пирофосфата. Биологическая роль Mut T заключается в предотвращении образования во время репликации неканонических пар A:G и A:8-oxo-G. Такие пары могут появляться в том случае, когда окисленная форма dGTP (8-oxo-dGTP) становится субстратом ДНК-полимеразы. Mut T гидролизует 8-oxo-dGTP в 10 раз быстрее, чем dGTP. Это делает 8-oxo-dGTP наиболее предпочтительным субстратом Mut Y и объясняет его функциональную роль.

Mut Y представляет собой специфическую аденин-ДНК-гликозилазу, расщепляющую N-гликозидную связь между аденином и дезоксирибозой аденозина, образующего неканоническую пару с гуанином. Функциональная роль этого фермента заключается в предотвращении мутации T:A→G:A путем отщепления неповрежденного остатка аденина из пары оснований A:8-oxo-G.

В последнее время в эксцизионной репарации особое внимание уделяют ATP-зависимому механизму удаления повреждений из ДНК. Этот вид эксцизионной репарации получил название ***нуклеотидная эксцизионная репарация*** (nucleotide excision repair; NER). Она включает в себя удаление из ДНК олигонуклеотидных фрагментов, содержащих повреждение, и последующую реконструкцию цепи ДНК с участием комплекса ферментов (нуклеаз, ДНК-полимеразы, ДНК-лигазы и др.). Удаление фрагмента ДНК происходит по обе стороны поврежденного нуклеотида. Длина удаляемых олигонуклеотидных фрагментов отличается у прокариот и эукариот. Так, у *E. coli*, *B. subtilus*, *Micrococcus luteus* вырезается фрагмент длиной 12—13

нуклеотидов, а у дрожжей, земноводных и человека — фрагмент, состоящий из 24—32 нуклеотидов. Выщепление фрагмента ДНК осуществляется ферментом *эксинуклеазой* (excinuclease). У *E. coli* этот фермент состоит из 3 различных протомеров — *uvr A*, *uvr B* и *uvr C*, каждый из которых выполняет определенную функцию в ходе эксцизионного выщепления фрагмента ДНК. Название этих белков дано по первым буквам слов «ultra violet repair». Протомер uvr A обладает АТРазной активностью, связывается с ДНК в виде димера, осуществляя первичное распознавание повреждения и связывание uvr B. Протомер uvr B обладает латентной АТР-азной и латентной хеликазной активностью, необходимой для изменения конформаций и расплетания двойной спирали ДНК; он же обладает эндонуклеазной активностью, расщепляя межнуклеотидную (фосфодиэфирную) связь со стороны 3′-конца выщепляемого фрагмента. Протомер uvr C действует как нуклеаза, вносящая разрыв в репарируемую цепь ДНК с 5′-конца вырезаемого фрагмента. Таким образом, протомеры uvr A, uvr B, uvr C взаимодействуют с ДНК в определенной последовательности, осуществляя АТР-зависимую реакцию выщепления олигонуклеотидного фрагмента из репарируемой цепи ДНК. Образовавшаяся брешь в молекуле ДНК реставрируется при участии ДНК-полимеразы I и ДНК-лигазы.

Эксцизионные репарации у человека также имеют АТР-зависимый характер и включают три основных этапа: узнавание повреждения, двойное разрезание цепи ДНК, восстановительный синтез и лигирование репарируемой цепи. Однако в эксцизионной репарации ДНК человека принимают участие 25 различных полипептидов, 16 из которых участвуют в выщеплении олигонуклеотидного фрагмента, являясь протомерами эксинуклеазы, а остальные осуществляют синтез репарируемого участка молекулы. В репарационной системе ДНК у человека весьма существенную роль выполняют белки транскрипции — РНК-полимераза II и TF IIH — один из шести основных факторов транскрипции эукариот. Следует отметить, что эксцизионная репарация у прокариот, как и у эукариот, зависит от функционального состояния ДНК: транскрибируемая ДНК репарируется быстрее, чем транскрипционно неактивная. Этот феномен объясняется следующими факторами: структурой хроматина, гомологией цепей транскрибируемых участков ДНК, эффектом повреждения цепей и его влиянием на РНК-полимеразу. Известно, что такие крупные повреждения, как образование тиминовых димеров, блокируют транскрипцию как у бактерий, так и у человека, если они происходят на матричной цепи ДНК (повреждения на кодирующей цепи не влияют на транскрипционный комплекс). РНК-полимераза останавливается в месте повреждения ДНК и блокирует работу транскрипционного комплекса. У *E. coli* усиление репарации при транскрипции опосредуется одним специальным белком — транскрипционно-репарационным фактором сцепления

(TRCF). Этот белок способствует отсоединению РНК-полимеразы от ДНК и одновременно стимулирует образование комплекса белков, осуществляющих репарацию поврежденного участка. По окончании репарации РНК-полимераза встает на место и транскрипция продолжается.

Репарация ошибок репликации ДНК. Ошибки спаривания азотистых оснований во время репликации ДНК происходят достаточно часто (у бактерий один раз на 10 тыс. нуклеотидов), в результате которых в дочернюю цепь ДНК включаются некомплементарные нуклеотидам материнской цепи нуклеотиды — мисмэтчи (от англ. *mismatch*). Несмотря на то что ДНК-полимераза I прокариот обладает способностью к самокоррекции (см. гл. 14), ее усилия по устранению ошибочно присоединенных нуклеотидов иногда оказываются недостаточны, и тогда в ДНК остаются некоторые неправильные (некомплементарные) пары. В этом случае репарация происходит с использованием определенной системы, связанной с метилированием ДНК. Действие этой системы репарации основано на том, что после репликации через определенное время (несколько минут) ДНК подвергается метилированию. У *E. coli* метилируется в основном аденин с образованием N^6-метил-аденина (N^6-mA). До этого вновь синтезированная (дочерняя) цепь остается неметилированной. Если в такой цепи есть неспаренные нуклеотиды, то она подвергается репарации: метилирование как бы метит ДНК и включает систему исправления ошибок репликации. В этой системе репарации узнаются особые структуры: последовательность G—N^6-mA—T—C и следующая за ней деформация в двойной спирали в месте отсутствия комплементарности (рис. 176). В устранении неспаренных нуклеотидов в полуметилированной молекуле ДНК принимает участие достаточно сложный комплекс ферментов репарации, который сканирует поверхность молекулы ДНК, вырезает участок дочерней цепи, прилегающей к мисмэтчам, а затем создает условия для застраивания его нужными (комплементарными) нуклеотидами. Различные компоненты этого комплекса обладают разными активностями (нуклеазной, хеликазной, АТРазной), необходимыми для внесения разрывов в ДНК и выщепления нуклеотидов, расплетания двойной спирали ДНК и энергетического обеспечения движения комплекса вдоль репарируемой части молекулы. Сходный по структуре и функциям комплекс ферментов репарации выявлен и у человека.

Рекомбинантная (пострепликативная) репарация. В тех случаях, когда по тем или иным причинам вышерассмотренные системы репарации оказываются нарушенными, в цепях ДНК могут образовываться бреши (недорепарированные участки), имеющие иногда весьма существенные размеры, что чревато нарушением системы репликации и может привести к гибели клеток. В этом случае клетка в состоянии использовать для репарации одной молекулы ДНК другую полученную при репликации молекулу ДНК, т. е. привлечь для этой

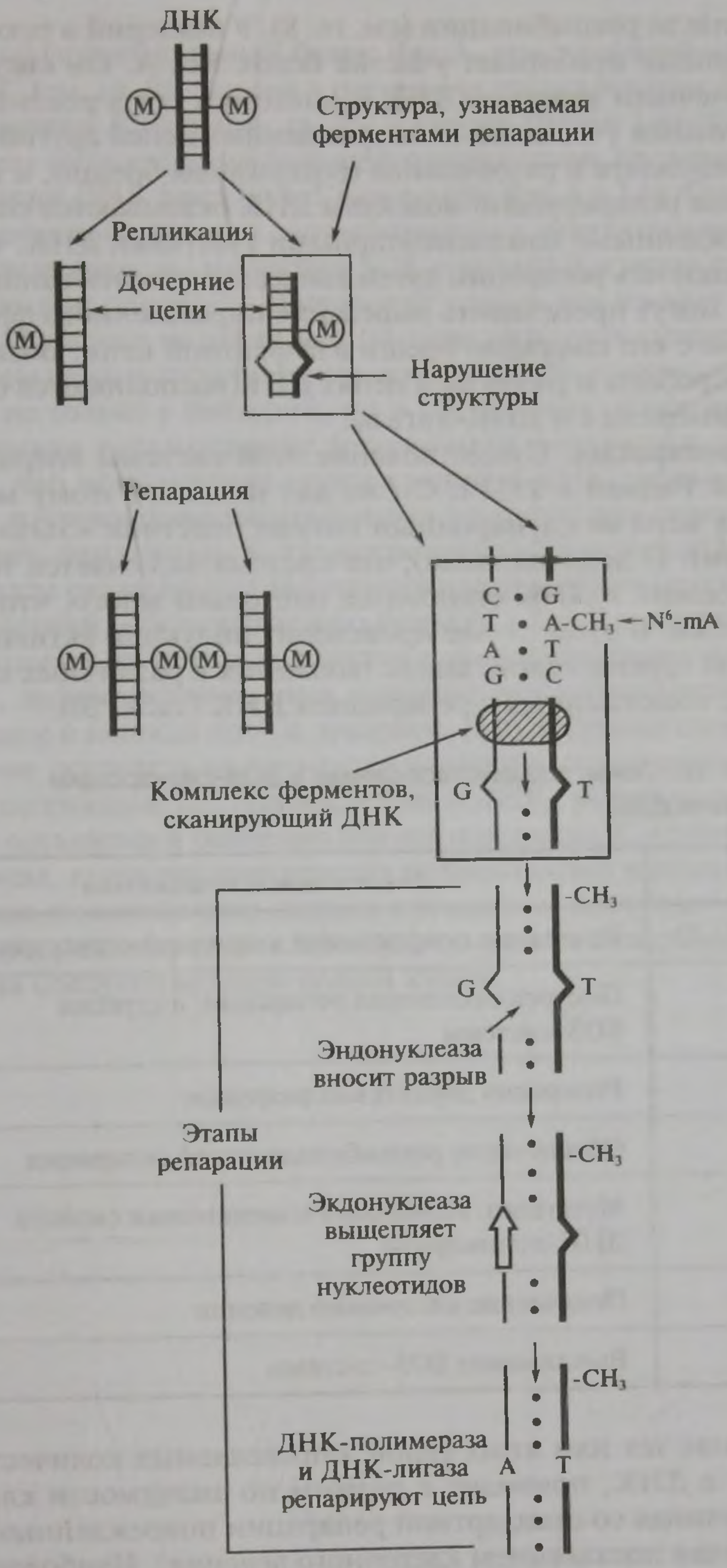


Рис. 176. Система репарации, связанная с метилированием ДНК. Пояснения см. в тексте

цели механизм рекомбинации (см. гл. 8). У бактерий в рекомбинантной репарации принимает участие белок Rec A. Он связывается с одноцепочечным участком ДНК и вовлекает его в рекомбинацию с гомологичными участками неповрежденных цепей другой молекулы ДНК. В результате и разорванная (содержащая бреши), и неповрежденная цепи репарируемой молекулы ДНК оказываются спаренными с неповрежденными комплементарными участками ДНК, что открывает возможность репарации путем вышеохарактеризованных систем. При этом могут происходить вырезание определенного фрагмента и заполнение с его помощью бреши в дефектной цепи. Возникающие при этом пробелы и разрывы в цепях ДНК восполняются с участием ДНК-полимеразы I и ДНК-лигазы.

SOS-репарация. Существование этой системы впервые постулировал М. Радман в 1974 г. Он же дал название этому механизму, включив в него международный сигнал бедствия «SOS» (спасите наши души). И действительно, эта система включается тогда, когда повреждений в ДНК становится настолько много, что угрожает жизни клетки. В этом случае происходит индукция активности разнообразной группы генов, задействованных в различных клеточных процессах, сопряженных с репарацией ДНК (табл. 20).

Таблица 20. **Гены, задействованные в SOS-репарации повреждений ДНК**

Гены	Последствия активации гена
uvr A, B, C, D	Репарация повреждений вторичной структуры ДНК
rec A	Постреплikативная репарация, индукция SOS-системы
rec N, ruv	Репарация двунитевых разрывов
ssb	Обеспечение рекомбинационной репарации
umu C, D	Мутагенез, вызванный изменениями свойств ДНК-полимеразы
sul A	Подавление клеточного деления
lex A	Выключение SOS-системы

Включение тех или иных генов, определяемых количеством повреждений в ДНК, приводит к разным по значимости клеточным ответам (начиная со стандартной репарации поврежденных нуклеотидов и кончая подавлением клеточного деления). Наиболее изучена SOS-репарация у *E. coli*, главными участниками которой являются белки, кодируемые генами rec A и lex A. Первый из них представляет

собой полифункциональный белок Rec A, участвующий в рекомбинации ДНК (см. гл. 8), а также в регуляции транскрипции генов фага λ, поражающего *E. coli* (см. гл. 9), а второй (белок Lex A) является репрессором транскрипции большой группы генов, предназначенных для репарации ДНК бактерий. Связывание Rec A с Lex A приводит к расщеплению последнего и соответственно к активации генов репарации. В свою очередь, индукция SOS-системы бактерии служит для фага λ сигналом опасности и приводит к тому, что профаг переключается с пассивного на активный (литический) путь существования, вызывая тем самым гибель клетки-хозяина. SOS-система репарации выявлена не только у бактерий, но и у животных, и человека.

Заканчивая рассмотрение механизмов репарации, нельзя не отметить, что исправление повреждений в ДНК тесным образом связано с другими фундаментальными молекулярно-генетическими процессами: репликацией, транскрипцией и рекомбинацией. Все эти процессы оказываются переплетенными в общую систему (сеть) взаимодействий, обслуживаемую большим числом разнообразных белков, многие из которых являются полифункциональными молекулами, задействованными в контроле реализации генетической информации в клетках про- и эукариот. В то же время очевидно, что природа «не скупится» на элементах контроля, создавая сложнейшие системы коррекции тех повреждений в ДНК, которые несут опасность для организма и особенно для его потомства. С другой стороны, в тех случаях, когда репарационных возможностей недостаточно для сохранения генетического статуса организма, наступает необходимость в программируемой клеточной смерти — апоптозе, которому посвящена следующая глава нашей книги.

Глава 13
ПРОГРАММИРУЕМАЯ КЛЕТОЧНАЯ СМЕРТЬ (АПОПТОЗ)

У всех многоклеточных организмов, включая человека, генетически заложена программа гибели клеток. В тех случаях, когда программа самоликвидации активируется под влиянием определенных внешних или внутренних стимулов (вирусов, токсинов, клеточных онкогенов, ДНК-повреждающих агентов и др.), вступает в действие форма клеточной гибели, называемая *апоптозом*, или *запрограммированной клеточной смертью* (programmed cell death, PCD). Наглядным примером апоптоза служит осенний листопад. Апоптоз — необходимое общебиологическое явление, позволяющее биологическим системам освободиться от стареющих или больных, появившихся в результате мутаций и ставших потенциально опасными для организма клеток. Такие клетки должны быть умерщвлены для того, чтобы, во-первых, обеспечивался баланс между пролиферацией и смертью части клеток организма, а во-вторых, чтобы отягощенные грузом мутаций клетки не исказили наследственную программу развития.

Клеточная смерть имеет, таким образом, двойственное предназначение: с одной стороны, это инструмент гомеостаза, в поддержании постоянного числа клеток и необходимый элемент дифференцировки клеток, а с другой — оружие самозащиты организма от грозящих ему и его потомству опасностей. Программируемая смерть клеток обнаруживается на ранних стадиях эмбриогенеза при формировании органов, замене одних тканей на другие, резорбции временных органов и при других процессах, сопутствующих нормальной дифференцировке и развитию. Во взрослом организме апоптоз становится комплементарен митозу, участвует в регуляции клеточных популяций, в частности, служит для уничтожения временно возникающего избытка лимфоцитов, вызванного инфицированием организма патогенами. Одновременно апоптоз необходим для иллиминации клеток, которые претерпели генетические повреждения или встали на путь ракового перерождения. Потеря контроля над апоптозом вызывает патологические изменения в организме, чему могут способствовать вирусы, обладающие генами, кодирующими белки с антиапоптозной активностью. Таким образом, клеточные и молекулярные аспекты апоптоза тесно переплетены с процессами онкогенеза и вирусного патогенеза и имеют непосредственное отношение к изучению тонких

механизмов регуляции клеточного деления. Поэтому в последние годы наблюдается бурный всплеск работ в области изучения программируемой клеточной смерти.

Цитологически апоптоз существенно отличается от некроза — другой формы гибели клеток, которая рассматривается как метаболическая катастрофа, вызванная тяжелыми внезапными повреждениями клетки. Некротическую гибель клетки могут вызывать физические или химические повреждения (ожоги, обморожения, отравления, гипоксия, вирусные поражения). При некрозе происходит набухание клетки, наблюдаются изменения в структуре цитоплазматической и внутриклеточных мембран, сопровождающиеся высвобождением лизосомальных (литических) ферментов, и наступает лизис клетки (рис. 177). Некроз сопровождается воспалительным процессом, который полностью отсутствует при апоптозе.

В отличие от некроза апоптоз проявляется в деструкции хроматина, уменьшении объема клетки и фрагментации клеточного ядра. Апоптоз заканчивается фрагментацией клетки с образованием апоптозных телец, которые затем быстро разрушаются макрофагами. Для апоптоза характерна ферментативная межнуклеосомная фрагментация ДНК, которая предшествует разрушению клеточной мембраны.

В молекулярном плане апоптоз представляет собой многоэтапный (каскадный) процесс, который начинается с приема определенного сигнала и заканчивается деструкцией клеточных молекул и структур под действием литических ферментов (протеаз, нуклеаз и др.). Сигнал воспринимается определенными рецепторами, передается молекулам-посредникам (адаптерам, мессенджерам) и достигает ядра клетки, где происходит экспрессия генов, детерминирующих развитие апоптоза.

В настоящее время идентифицированы четыре основные функциональные группы молекул, вовлеченные в процесс апоптоза: особые протеолитические ферменты (каспазы); адаптерные белки, контролирующие активацию каспаз; белки суперсемейства рецепторов факторов некроза опухолей (TNF) и белки семейства Bcl-2.

В клетках животных и человека апоптоз связан с активацией *каспаз* — группы эволюционно консервативных цистеиновых протеиназ, расщепляющих белки после остатков аспарагиновой кислоты. У млекопитающих идентифицировано 14 каспаз. Они распознают определенные тетрапептидные блоки (мотивы) в полипептидных цепях белков и расщепляют пептидную связь, образованную карбоксильной группой остатка Asp. Каспазы синтезируются в виде профермен тов с очень низкой активностью. Зрелые (процессированные) каспазы представляют собой тетрамеры и состоят из двух субъединиц по 20 кДа и двух по 10 кДа. Механизм образования активных каспаз реализуется путем протеолитического саморасщепления или перекрестного расщепления (ауто- или транспроцессинга). В ходе этого

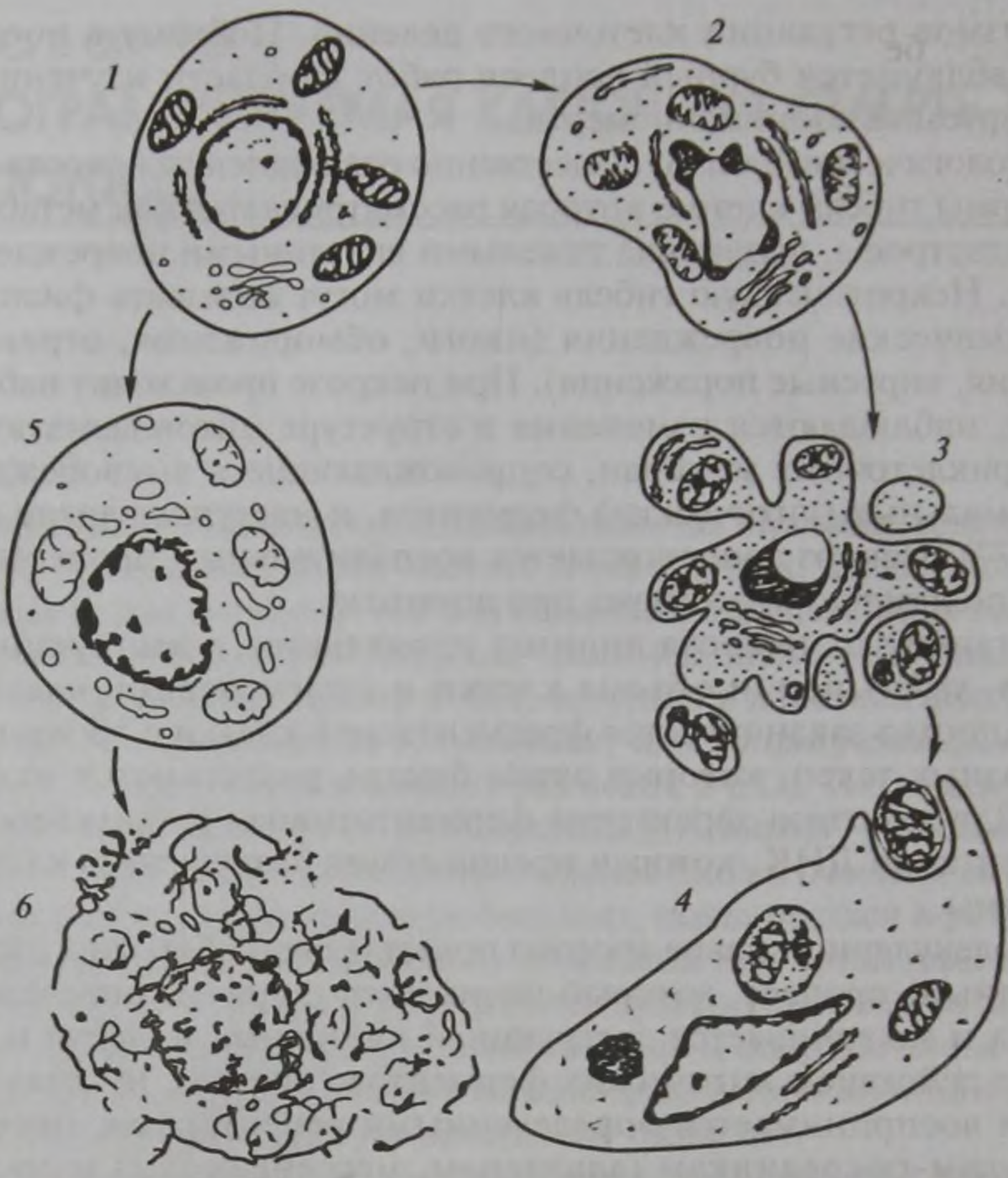


Рис. 177. Структурные изменения в клетке в процессе некроза (*слева*) и апоптоза (*справа*) (В.Д. Самуилов и др., 2000):

1 — интактная клетка; *2* — сегрегация хроматина в ядре; *3* — фрагментация ядра и образование апоптозных телец; *4* — фагоцитоз апоптозных телец соседней клеткой; *5* — конденсация хроматина и деградация цитоплазматических структур; *6* — разрушение мембран и гибель клетки

процесса от прокаспазы (белка-предшественника с молекулярной массой 30—50 кДа) отщепляется N-концевой домен, а оставшаяся часть молекулы разделяется на большую и малую субъединицы, которые затем ассоциируются с образованием гетеродимеров и тетрамеров (рис. 178). В функциональном плане среди каспаз различают инициирующие каспазы (8 и 9) и активируемые или эффекторные каспазы (3, 6, 7). Так, например, инициирующая каспаза 8 активирует прокаспазу 3, из которой возникает активная эффекторная каспаза 3. Активация эффекторных каспаз сообщает процессу апоптоза необратимый характер. Каспазы, особенно эффекторные (их еще называют каспазами второго эшелона) расщепляют и инактивируют определенные жизненно важные для клетки белки, такие, как

ферменты репарации ДНК, белок-ингибитор белка р-53 (подробнее о белке р-53 см. ниже), протеинкиназа С и др. Одной из мишеней каспаз является поли(ADP-рибозо)-полимераза — фермент, связанный с репарацией и транскрипцией ДНК (см. гл. 12). Активируемая каспазами ДНКаза CAD (caspase-activated DNAase) в норме находится в неактивном состоянии в комплексе с ингибитором DFF (DNA fragmentation factor). В ходе апоптоза DFF расщепляется каспазами, что приводит к освобождению активной эндонуклеазы CAD, которая осуществляет межнуклеосомальные разрывы хроматина в области линкерной ДНК. В результате образуются фрагменты ДНК, соответствующие по размерам нуклеосомным частицам (180 — 200 н. п.).



Рис. 178. Активация прокаспаз путем протеолиза на субъединицы и их ассоциации.
Большие субъединицы каспаз на рисунке темные, малые субъединицы — светлые; заштрихована область N-концевого домена в молекуле прокаспазы

В клетках существуют особые рецепторы апоптоза, воспринимающие сигнал от соответствующего лиганда (индуктора), стимулирующего программу клеточной смерти. Так, рецептор Fas, взаимодействуя с соответствующим лигандом (лиганд FasL — трансмембранный лиганд цитотоксических клеток, Т-киллеров), активируется и запускает программу смерти клеток, инфицированных вирусами. FasL относится к семейству лигандов фактора некроза опухолей TNF (tumor necrosis factor).

Рецепторы TNF представляют собой трансмембранные белки, которые своими внеклеточными доменами взаимодействуют с тримерными молекулами лигандов — индукторов апоптоза, а их цитоплазматические участки служат для взаимодействия с внутриклеточными участниками апоптозного каскада (рис. 179).

Так, рецептор Fas, называемый также CD-95 (от англ. *cell death* — клеточная смерть), содержит цитоплазматический участок (домен), необходимый для взаимодействия с молекулой адаптора — внутриклеточного белка, служащего передатчиком (мессенджером) сигнала от рецептора к эффекторным белкам — каспазам. При активации прокаспазы 8 образуются агрегаты FasL-Fas-FADD-прокаспаза 8, которые носят название *апоптосом*, или *апоптозных шаперонов*.

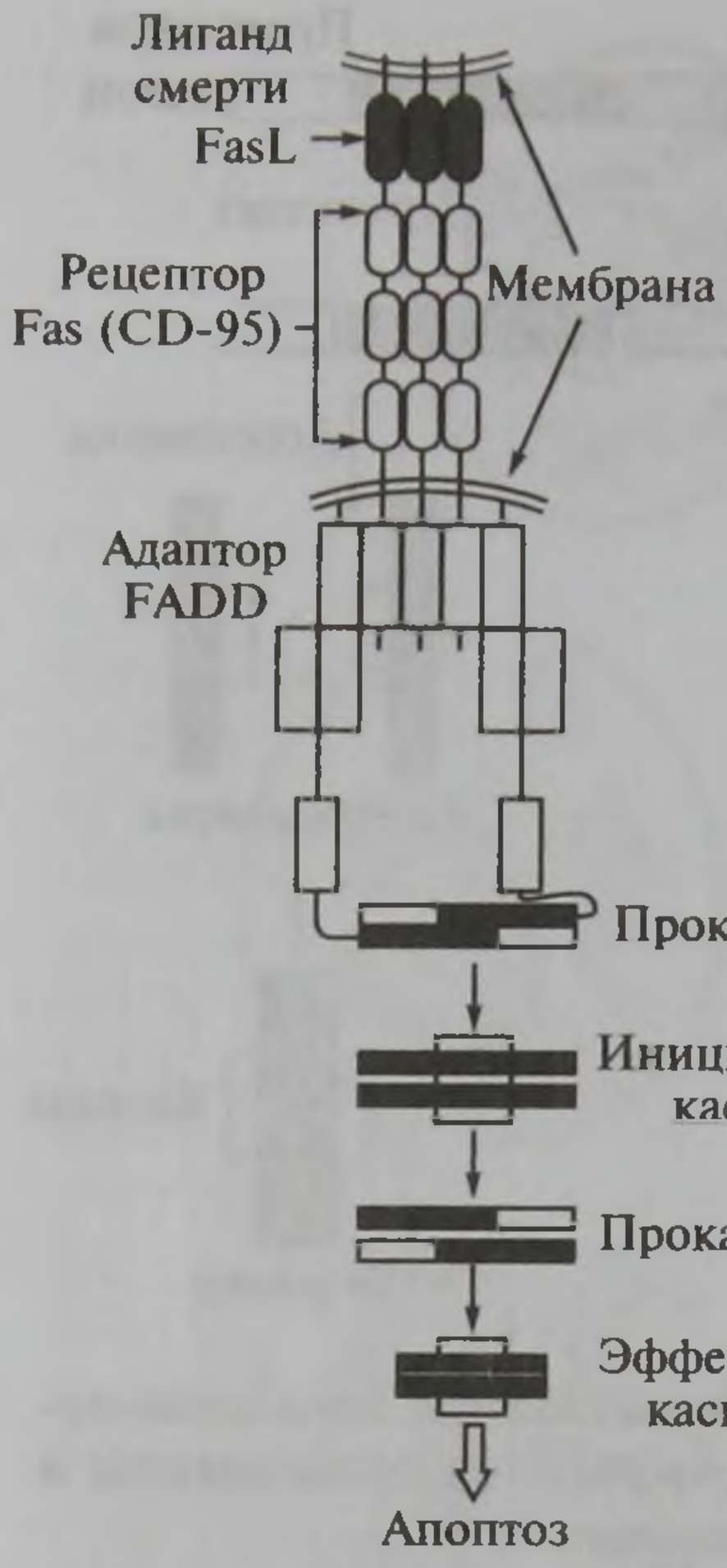


Рис. 179. Каскадная система индукции апоптоза, включающая функционирование лиганда FasL, рецептора CD-95, адаптора FADD, инициаторной каспазы 8 и эффекторной каспазы 3

Таким образом, в регуляцию апоптоза оказываются вовлечены продукты целого ряда генов. Впервые прямые доказательства участия специфических генов в апоптозе были получены при изучении программируемой гибели клеток нематод *Caenorhabditis elegans*, у которых 131 из 1 090 соматических клеток погибают на определенной стадии развития. У этого вида червей для клеточной гибели необходимы гены ced 3 и ced 4. Еще один ген — ced 9 кодирует белок, предотвращающий гибель клеток, индуцированную ced 3 и ced 4. Гомологами ced 9 у млекопитающих являются гены семейства bcl-2. В это семейство входят гены, кодирующие два типа структурно-родственных белков. К первому типу относятся белки млекопитающих, препятствующие апоптозу (Bcl-x1, Bcl-w и др.), а также множество вирусных гомологов Bcl-2, содержащих сходные с Bcl-2 участки первичной структуры. Второй тип включает белки, усиливающие апоптоз (Bax, Bak и Bok). Активность проапоптозных белков Bcl-2 контролируется на стадии посттрансляционной модификации: фосфорилированием (при участии протеинкиназ) и ограниченным протеолизом (при участии каспаз).

В клетках, подвергавшихся воздействию индукторов апоптоза, происходят разительные изменения структуры митохондрий — образование в их внутренних мембранах гигантских пор. Следствием этого являются набухание матрикса митохондрий и разрыв их наружных мембран. К факторам, вызывающим раскрытие пор, относятся активные формы кислорода ($O_2^{\cdot-}$, $^{\cdot}OH$ и др.), которые в целом являются мощными индукторами апоптоза. При разрушении митохондрий из них высвобождается ряд апоптогенных факторов: цитохром *c*, белок флавопротеин AIF (apoptosis inducing factor), а также прокаспазы 2, 3 и 9. Таким образом, по образному выражению крупнейшего специалиста в области биохимии и молекулярной биологии митохондрий академика В. П. Скулачева, митохондрии «подобны термоядерным

станциям», способным не только снабжать клетку энергией, но и при определенных обстоятельствах приводить к ее гибели.

В ряде случаев апоптоз реализуется с участием двух систем — рецепторов плазматической мембраны и митохондриальных апоптогенных факторов.

Повреждение ДНК ведет к накоплению в клетке белка p-53, который может останавливать деление клетки и/или индуцировать апоптоз. Этот белок расценивается как главный антиопухолевый белок клеток млекопитающих. У более чем 50% изученных видов опухолевых клеток ген белка p-53 инактивирован и у них нарушена p-53-зависимая регуляция клеточного гомеостаза. Обработка клеток ДНК-повреждающими агентами приводит к увеличению содержания белка p-53, индукции p-53-зависимых генов и задержке или блоку клеточного цикла на фазе G1. Таким образом, при воздействии мутагенов белок p-53 не только увеличивает время репарации ДНК, но и защищает организм от пролиферации клеток с опасными мутациями. *Белок p-53 является фактором транскрипции*, регулирующим активность ряда генов. Существенно, что он является индуктором транскрипции гена белка-ингибитора циклинзависимой протеинкиназы (CDK) — фермента, необходимого для функционирования группы белков, задействованных в митотическом делении клеток (рис. 180). Активная форма киназы (CDK) образуется в клетке в результате взаимодействия молекулы фермента с *циклином* — белком, выполняющим роль регуляторной субъединицы CDK. Весь период митоза CDK остается в активном состоянии, а затем инактивируется после разрушения циклина в результате убиквитинзависимого механизма протеолиза (см. гл. 2). При завершении митоза киназа CDK переходит в неактивное состояние в результате присоединения к ее молекуле специфического белка-ингибитора. Образование комплекса циклин — киназа — белок-ингибитор останавливает митоз. Таким образом, индукция синтеза белка-ингибитора, осуществляемая благодаря белку p-53, является ключевым моментом остановки митотического деления клетки. Предполагают, что ответная реакция клетки на образование белка p-53 зависит от степени мутаций в ДНК. При умеренном нарушении генома происходит остановка клеточного деления, осуществляется репарация ДНК и клетка может продолжать существование, перейдя к очередному туру репликации ДНК, а затем к митозу. При массовых нарушениях структуры генома, когда репарация уже неэффективна, включаются апоптозные каскады активации каспаз и наступает апоптоз.

К факторам, препятствующим нормальному функционированию белка p-53, относятся белки онкогенных вирусов, в частности белок E6 вируса папиломы человека (HPVs), а также белки Bcl-2 аденовирусов. Кроме того, обнаружены вирусные гены — ингибиторы каспаз, задействованных в CD95-индуцированном апоптозе.

Таким образом, раковое перерождение клеток под действием онкогенных вирусов в последних исследованиях все в большей мере

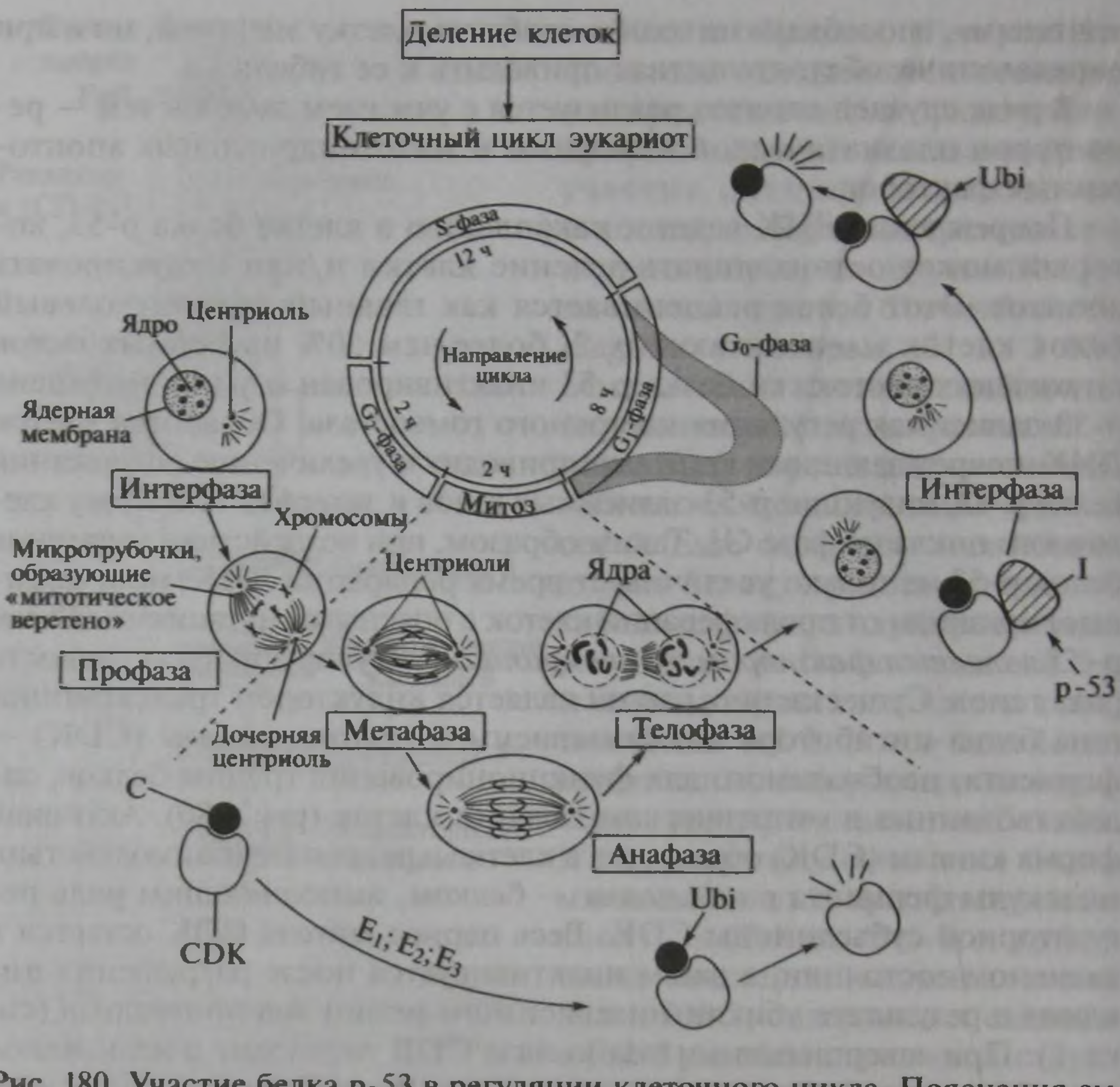


Рис. 180. Участие белка p-53 в регуляции клеточного цикла. Пояснения см. в тексте:

C — циклин; CDK — циклинзависимая протеинкиназа; I — ингибитор; E_1, E_2, E_3 — ферменты, обслуживающие митотическое деление клетки; Ubi — убиквитин. Пунктирной стрелкой обозначена роль белка p-53 как индуктора синтеза белка-ингибитора киназы CDK

связывают с нарушением системы апоптоза, в результате чего клетки переходят к нелимитированному делению. Роль нарушений апоптоза в развитии многих болезней выглядит уже бесспорной. Так, ген bcl-2 был изначально идентифицирован как протоонкоген, а позднее установлена его роль в апоптозе. Сам по себе ген bcl-2 представляет собой слабый онкоген, однако он может выступать синергистом таких онкогенов, как c-myc и pin-1. Другие протоапоптозные белки семейства Bcl-2 могут выступать как супрессоры опухолей. Многие из создаваемых в настоящее время противоопухолевых средств нацелены на стимулирование апоптоза.

Большое значение имеют исследования апоптоза у ВИЧ-инфицированных клеток. Установлено, что при поражении этим

РНК-содержащим ретровирусом (подробнее о ВИЧ см. гл. 4) значительную часть погибающих клеток иммунной системы составляют здоровые неинфицированные Т-лимфоциты. Ряд белков ВИЧ оказывают усиливающее действие на апоптоз. Так, Tat-белок вируса может индуцировать апоптоз Т-лимфоцитов. В ходе заболевания растет число Т-клеток, экспрессирующих рецептор Fas.

Экспрессия FasL также индуцируется белками Tat и gp-120 вируса, приводя в итоге к усилению Fas-зависимого пути развития апоптоза. Вирусный белок gp-120, связываясь с рецептором CD-4 Т-клеток, также инициирует апоптоз этих клеток. В то же время некоторые белки ВИЧ (в частности, белок Vpr) участвуют в ингибировании апоптоза. В целом продолжающаяся программируемая гибель Т-клеток при ВИЧ-инфекции приводит к истощению иммунной системы организма.

Глава 14
ГЕНЕТИЧЕСКАЯ ИНЖЕНЕРИЯ

Методы генетической инженерии
(технология получения рекомбинантных ДНК)

По определению академика А. А. Баева, долгое время возглавлявшего работы в области генетической инженерии в нашей стране, «генетическая инженерия представляет собой конструирование in vitro функционально активных генетических структур (рекомбинантных ДНК) или иначе — создание искусственных генетических программ».

Формально датой рождения генетической инженерии можно считать 1972 г., когда П. Берг и сотрудники получили первую рекомбинантную ДНК, состоящую из ДНК вируса SV40 и бактериофага. Создание генетической инженерии имело огромное значение для молекулярной биологии в целом. Благодаря ее методам практически структура и функции любого гена и продуктов его экспрессии (РНК и белки) стали доступны для исследования.

Появилась возможность не только детального изучения гигантских геномов высших организмов, но и целенаправленного изменения их структуры. Создались условия для создания генетически трансформированных видов микроорганизмов, растений и животных, способных стать продуцентами биологически активных соединений (ферментов, гормонов, антибиотиков и т. п.), т. е. была создана база для развития новых биотехнологий.

Триумф генетической инженерии в конце XX в. связан с развитием в ее недрах новых, быстро совершенствовавшихся методов исследований. Среди этих методов следует выделить основные пять:

• ***расщепление ДНК*** (***рестрикция***), что необходимо для выделения генов и манипуляций с ними;

• ***гибридизация нуклеиновых кислот***, которая благодаря их способности связываться друг с другом по комплементарному принципу, позволяет выявлять специфические последовательности ДНК и РНК, а также совмещать различные генетические элементы; используется в цепной полимеразной реакции для амплификации ДНК in vitro;

• ***клонирование ДНК***, осуществляемое путем введения фрагментов ДНК или их групп в быстро реплицирующиеся генетические элементы (плазмиды или вирусы), что дает возможность размножать гены в клетках бактерий, дрожжей или эукариот;

• ***определение нуклеотидных последовательностей*** **(*секвенирование*)** в клонируемом фрагменте ДНК, что позволяет определить структуру генов и аминокислотную последовательность кодируемых ими белков;

• ***химико-ферментативный синтез полинуклеотидов***, часто необходимый для целенаправленной модификации генов и облегчения манипуляции с ними.

Следует отметить, что возникновение генетической инженерии было напрямую связано с открытием целой серии ферментов, оказавшихся незаменимыми инструментами для манипулирования с генами, ввиду практически полного отсутствия соответствующих химических методов, позволяющих получать, модифицировать и амплифицировать рекомбинантные молекулы ДНК. Ниже, в разделе, посвященном рестрикции, приведен перечень наиболее важных ферментов, используемых в генетической инженерии.

Рестрикция ДНК

Рестрикция ДНК достигается с помощью специфических ферментов — *рестриктаз*. Они представляют собой эндонуклеазы бактериального происхождения, предназначенные для защиты клеток бактерий от чужеродной (вирусной) ДНК.

Рестриктазы способны осуществлять разрывы межнуклеотидных (фосфодиэфирных) связей внутри полинуклеотидных цепей ДНК. Впервые рестриктазы были выявлены в клетках кишечной палочки (*E. coli*), зараженных бактериофагом λ. При этом оказалось, что фаговое потомство, выращенное на двух различных штаммах этой бактерии, с различной интенсивностью размножается в клетках противоположных штаммов.

Так, фаг, выращенный на штамме С, плохо размножался в клетках бактерии штамма К12. Это ограничение развития (рестрикция) бактериофага, как оказалось, связано с эндонуклеазной деградацией его ДНК. Сама клеточная ДНК защищается от рестриктаз штаммоспецифической модификацией — метилированием части нуклеотидов, которое осуществляется особым ферментом — ДНК-метилазой. Чаще всего продуктами метилирования являются 6-метиладенин и 5-метилцитозин (см. гл. 3).

Таким образом, присутствие в клетках бактерий двух ферментов (рестриктазы и ДНК-метилазы) обеспечивает комплексную защиту ее ДНК. Эта система рестрикции — модификации (R-M-система), вероятно, препятствует скрещиванию между разными видами и штаммами бактерий и тем самым обеспечивает сохранность их видов в эволюции.

В генетической инженерии рестриктазы используются для фрагментации молекул ДНК при создании рекомбинантных геномов. Начиная с первых работ Д. Натанса и Г. Смита (1962), число открытых рестриктаз быстро росло, и в 1973 г. Смит и Натанс предложили номенклатуру рестриктаз, которая включает следующие пункты:

1) название рестриктазы формируется из родового и видового названия микроорганизма-хозяина, содержащего R-M-систему; при этом к первой заглавной букве родового названия добавляются две первые строчные буквы вида, например: *Escherichia coli — Eco*, *Haemophilus influenzae — Hin*;

2) за родо-видовым названием, в случае необходимости, приводится обозначение штамма (*Eco* B) или серотипа (*Hin* d);

3) различные R-M-системы, присущие одному виду бактерий, обозначаются римскими цифрами: *Hin*d I, *Hin*d II, *Hin*d III;

4) если R-M-система локализована в геноме плазмиды, то после родо-видового названия указывается символ плазмиды: *Eco* RI.

Основой для классификации рестриктаз служат характер их воздействия на ДНК и зависимость от кофакторов. Среди нескольких тысяч известных к настоящему времени рестриктаз выделяют 3 класса (типа).

Рестриктазы типа I содержат три неидентичных субъединицы с мол. массами 135, 60 и 50 кДа (α, β и γ) и массой эпимолекулы около 500 кДа. Эти рестриктазы специфичны к неметилированной двуцепочечной ДНК и требуют в качестве кофакторов S-аденозилметионин (донор метильных групп в реакции метилирования), ATP и Mg^{2+}.

Субъединицы рестриктаз типа I имеют различные функции: α-субъединицы проводят расщепление ДНК, β-субъединицы способны метилировать ДНК, γ-субъединицы осуществляют узнавание специфических последовательностей ДНК.

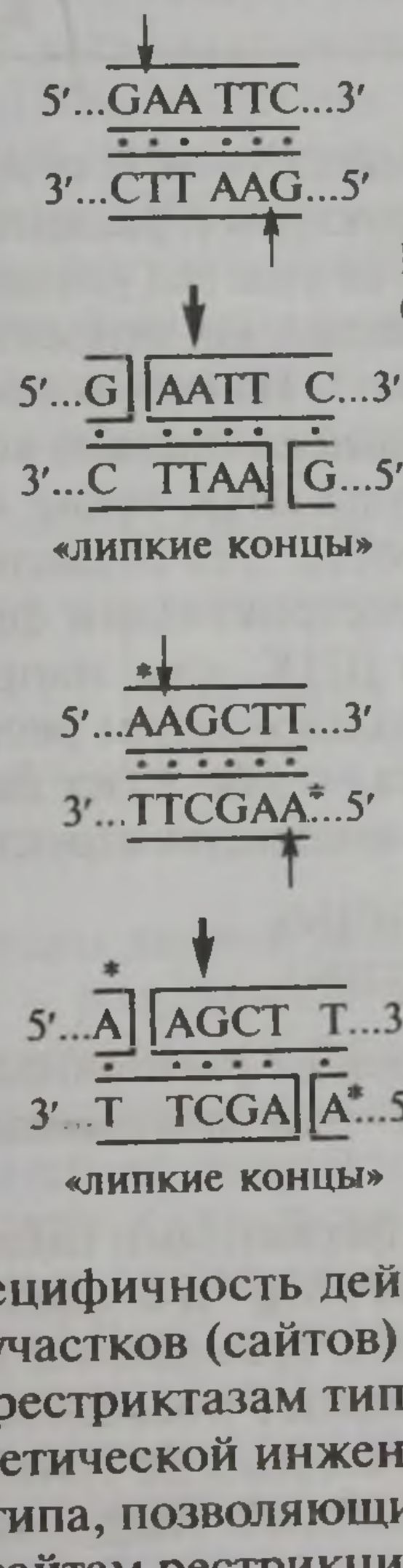
В то же время разрывы в цепях ДНК под действием этих рестриктаз происходят случайным образом на значительном расстоянии от участка узнавания, а продукты расщепления оказываются весьма гетерогенными, что затрудняет их использование в генетической инженерии.

Рестриктазы типа II состоят из двух субъединиц одного вида, имеют массу эпимолекулы около 100 кДа и нуждаются для осуществления нуклеазной активности только в ионах Mg^{2+} (метилазы этой системы являются самостоятельными ферментами).

Рестриктазы этого типа узнают в ДНК последовательности из 4—8 нуклеотидов, имеющие ось симметрии второго порядка (палиндромы).

Они гидролизуют фосфодиэфирные связи в двух цепях ДНК, осуществляя либо симметричные, либо несимметричные разрывы в зоне палиндрома. В первом случае образуются так называемые «тупые концы»:

Во втором случае действие рестриктаз приводит к образованию в ДНК так называемых «липких концов»:



Высокая специфичность действия, приводящая к образованию определенных участков (сайтов) рестрикции, определила повышенный интерес к рестриктазам типа II, которые стали важнейшим инструментом генетической инженерии. Целенаправленный поиск рестриктаз этого типа, позволяющих проводить *картирование геномов* по различным сайтам рестрикции, а также выщеплять определенные участки, расположенные между этими сайтами (рестрикционные фрагменты), и далее комбинировать эти участки, привел к тому, что число известных рестриктаз типа II стало быстро нарастать. Если в 1976 г. были открыты 84 таких рестриктазы, то к 1995 г. стало из-

вестно 2 400 рестриктаз этого типа, имеющих 188 специфических сайтов рестрикции. Выделенные из разных источников, но имеющие одинаковую специфичность рестриктазы стали называть *изошимеразами*. Ниже приведены названия некоторых широко используемых рестриктаз и обозначения соответствующих сайтов рестрикции:

Рестриктаза	Сайт рестрикции
Alu I	AI
Bam HI	B
Eco RI	E
Eco RII	E_2
*Hin*d III	HIII
Hpa I	HI
Sau 3A	S_3

В последнее время среди рестриктаз типа II открыты ферменты, распознающие непалиндромные структуры и расщепляющие ДНК на строго фиксированном расстоянии от участка узнавания с образованием как тупых, так и «липких концов». Их относят к подклассу IIS (от англ. *shift* — перестановка, сдвиг). Из них особенно интересны те, которые образуют достаточно длинные «липкие концы». В каждом месте гидролиза изучаемой молекулы ДНК такие «липкие концы» имеют уникальную последовательность. Это позволяет направленно объединять образованные этими рестриктазами фрагменты с восстановлением исходной структуры ДНК, как, например, это было сделано в отношении фага f1 с использованием рестриктазы *Hpa* I. Еще один фермент этого подкласса — *Bbv* I (из *Bacillus brevis*) — узнает в ДНК последовательность, имеющую структуру

$$GCAGC(N)_8$$
$$CGTCG(N)_{12}$$

и расщепляет ДНК по одной из цепей на расстоянии 8 нуклеотидов от сайта узнавания, а по другой цепи — на расстоянии 12 нуклеотидов от узнаваемого участка.

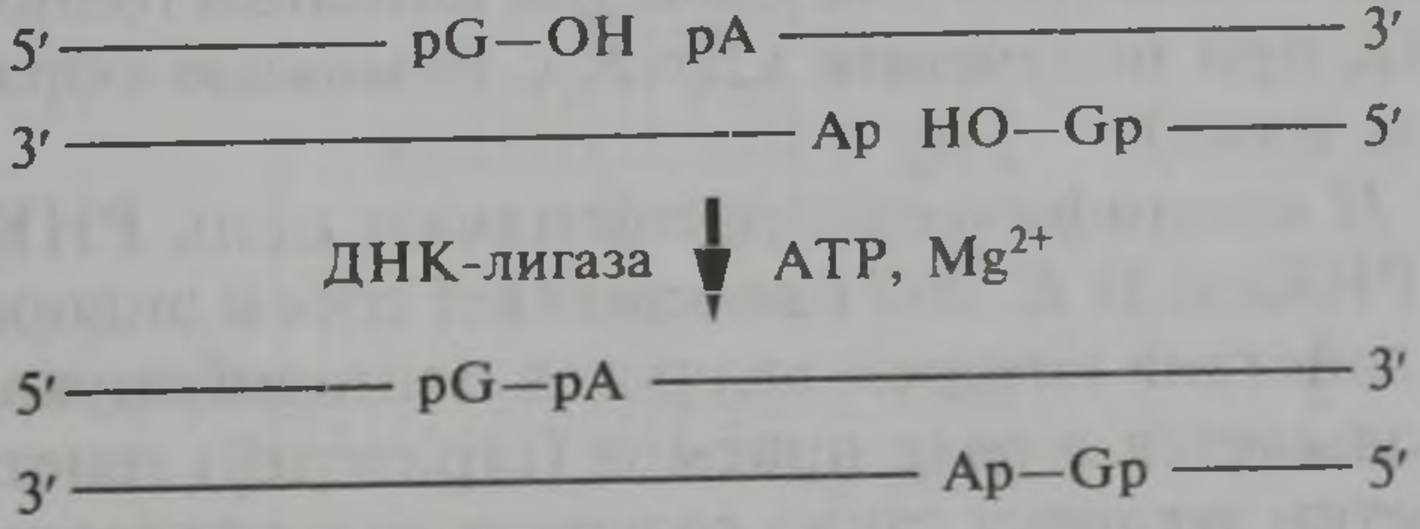
Рестриктазы типа III сходны с ферментами типа I. Для проявления активности они нуждаются в ATP, Mg^{2+} и S-аденозилметионине. Эти рестриктазы состоят из 2 типов субъединиц (рестриктазы и метилазы). Они узнают непалиндромные последовательности длиной 5—6 н.п. и расщепляют ДНК в стороне от сайтов узнавания на расстоянии 24—25 н.п. Исчерпывающего гидролиза ДНК они не производят.

Кроме бактерий рестриктазы выявлены у дрожжей *Sachramyces serevisiae* и одноклеточных водорослей *Chlamydomonas reinhardtii* и по свойствам напоминают рестриктазы типа III. У водорослей рестриктаза *Cre I* выявлена только в зиготах и активна в отношении

только хлоропластической ДНК одного из партнеров. В генетической инженерии рестриктазы (особенно типа II) используют для получения *библиотек*, или *банков геномной ДНК*. При обработке хромосомной ДНК рестриктазами (метод дробовика — «shotgun») получают огромное количество фрагментов ДНК (10^5—10^7 фрагментов из генома млекопитающих). Далее путем использования плазмид и их клонирования (см. ниже) получают миллионы колоний клеток бактерий или дрожжей, каждая из которых содержит плазмиду с различными фрагментами ДНК. Это и есть библиотека, или банк геномной ДНК. Соответствующий банк был использован и при изучении структуры генома человека.

Ниже приведены сведения о других ферментах, используемых в различных методах генетической инженерии.

ДНК-лигаза представляет собой еще один важнейший инструмент, способный катализировать синтез фосфодиэфирной связи в двуцепочечной ДНК. Наиболее часто в генетической инженерии используется ДНК-лигаза фага Т4, способная в присутствии АТР сшивать фрагменты ДНК как с липкими, так и с тупыми концами. Этот фермент состоит из одной полипептидной цепи (68 кДа) и ускоряет синтез фосфодиэфирной связи между 5′-фосфатным (5′-p) и 3′-гидроксильными концами (3′—OH) цепей ДНК. При лигировании липких концов реакция идет по схеме

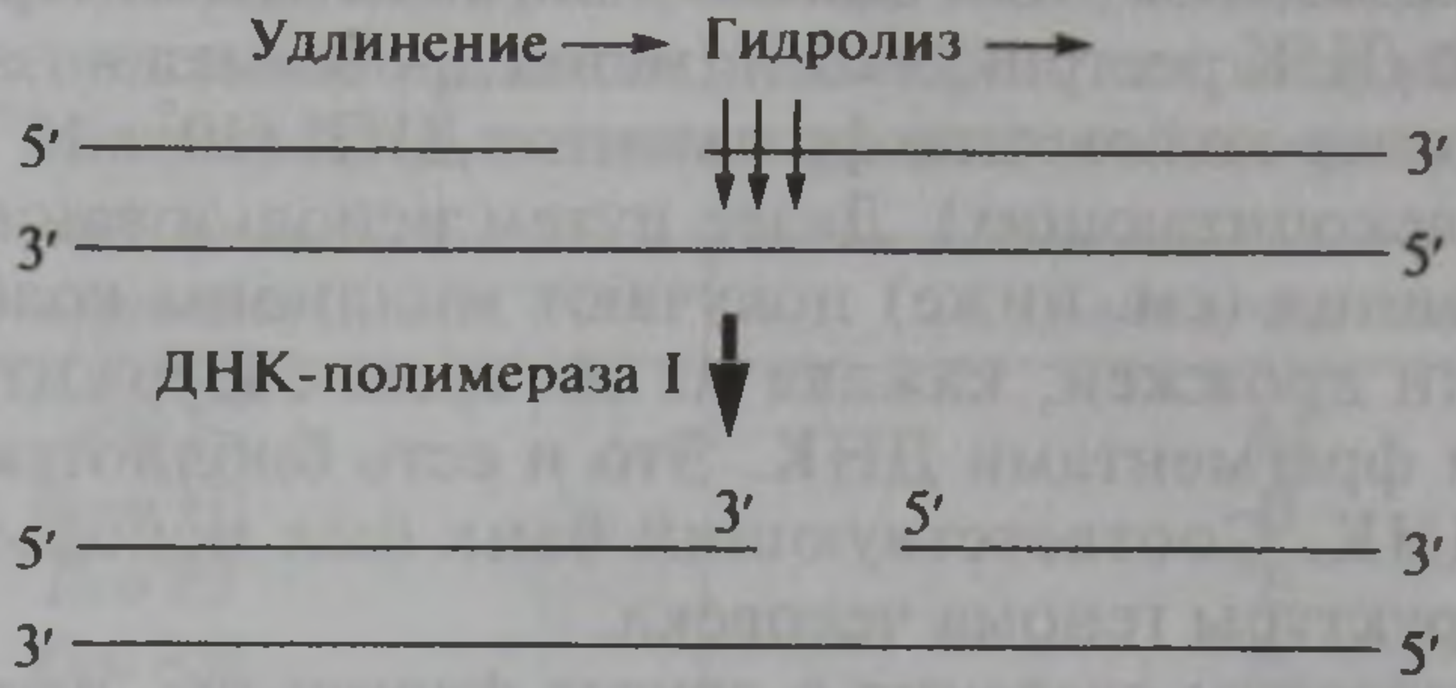


Частным случаем данной реакции является лигирование так называемого *ника* — разрыва в одной из цепей ДНК (см. ниже).

ДНК-полимераза E. coli была открыта А. Корнбергом и соавт. в 1958 г. Этот фермент, называемый в настоящее время ДНК-полимеразой I, состоит из одной полипептидной цепи (103 кДа) и имеет трехдоменную структуру. Каждый домен обладает определенной ферментативной активностью: N-концевой домен 5′→3′ — экзонуклеазной; С-концевой домен 5′→3′ — полимеразной (нуклеотидилтрансферазной); а средний домен 3′→5′ — экзонуклеазной. N-Концевой домен может быть отщеплен с использованием протеаз (трипсином, субтилизином и др.); остающаяся часть молекулы (68 кДа) — *фрагмент Кленова* — сохраняет присущие ей каталитические активности.

За счет 5′→3′— нуклеазной активности ДНК-полимераза I способна отщеплять нуклеотиды с 5′-конца одной цепи ДНК (в спаренных участках), что происходит в месте разрыва (ника). При этом одновременно идут две реакции: 5′→3′ экзонуклеазный гидролиз и 5′→3′

полимеразная реакция. В этом случае происходит процесс «переноса» одноцепочечного разрыва вдоль цепи ДНК — *ник-трансляция*:



Если использовать в приведенной реакции радиоактивные нуклеотиды, то можно получать меченую ДНК. Разрыв (ник) обычно проводят, добавляя к препарату ДНК следовые количества эндодезоксирибонуклеазы (ДНКазы I из поджелудочной железы быка), добиваясь исходной бреши в цепи, необходимой для присоединения фермента и запуска ник-трансляции. Эта реакция in vivo играет важную роль в репарации поврежденной ДНК.

Нуклеазы, специфичные в отношении одноцепочечной ДНК, выделены из разных объектов: плесневого гриба *Aspergillus ouyzae*, *Neurospora* и др. Нуклеаза SI, выделенная из *A. oryzae*, используется в генетической инженерии для удаления шпильки (одноцепочечного участка) ДНК при получении кДНК с помощью обратной транскриптазы (см. ниже).

РНКазы Н специфически расщепляют цепь РНК в дуплексе РНК:ДНК. РНКаза Н *E. coli* представляет собой эндорибонуклеазу, продуктами действия которой являются олигорибонуклеотиды. Она широко используется в ряде приемов (стратегий) генетической инженерии. Клетки эукариот также содержат эндонуклеазу такого типа. Эндонуклеазная активность РНКазы Н свойственна и обратным транскриптазам ретровирусов.

Обратная транскриптаза (*ревертаза*). В 1964 г. Г. Темин выдвинул гипотезу о существовании специфических для РНК-содержащих ретровирусов ферментов, способных синтезировать ДНК на матрице РНК, а в 1970 г. Г. Темин и С. Мизутани, а также независимо от них Д. Балтимор открыли этот фермент у вируса саркомы Рауса. Эта РНК-зависимая ДНК-полимераза получила название *обратная транскриптаза*, или *ревертаза*. Наиболее изучена ревертаза вирусов птиц, которая состоит из двух субъединиц — α (65 кДа) и β (95 кДа) и обладает по крайней мере тремя активностями: ДНК-полимеразной (может использовать в качестве матрицы как РНК, так и ДНК); активностью РНКазы Н (гидролизует РНК в составе гибрида РНК:ДНК, но не атакует свободную РНК); ДНК-эндонуклеазной активностью. Первые две активности используются для синтеза ДНК, комплементарной вирусной РНК, которая затем интегрируется в

геном клетки-хозяина. Эндонуклеазная активность, очевидно, используется для внесения разрывов в цепи ДНК-хозяина. Показано, что β-субъединицы данной ревертазы обладают всеми тремя активностями, α-субъединица — только активностями ДНК-полимеразы и РНКазы Н. Затравкой (праймером) для полимеразной реакции могут служить небольшие участки одноцепочечных молекул ДНК или РНК (а также тРНК).

В генетической инженерии ревертаза очень широко используется для целенаправленного синтеза на матричных РНК комплементарных молекул ДНК (кДНК). Для повышения эффективности этой процедуры дополнительно применяют и другие ферменты (рис. 181).

Если включить полученную таким образом кДНК в быстрореплицирующиеся генетические элементы (плазмиды), то в них можно вырастить клоны кДНК. С использованием различных мРНК таким способом получают *библиотеки кДНК*. Библиотека кДНК отличается от библиотеки геномной ДНК (см. выше), которую получают



Рис. 181. Схема получения кДНК с использованием ревертазы вируса птичьего миелобластоза и трех дополнительных ферментов: поли(А)-полимеразы, фрагмента Кленова ДНК-полимеразы I и нуклеазы S1 (разрушает характерную шпильку в молекуле кДНК, которую образует ревертаза)

после обработки ДНК рестриктазами. В библиотеке геномной ДНК содержатся миллионы клонов — колоний, в которых не обязательно присутствуют дискретные гены, так как рестриктазы разрезают ДНК и внутри генов. Обратная транскриптаза дает возможность получить более прогрессивную библиотеку *транскрибируемых генов*, в которой нетрудно идентифицировать клон, если ген кодирует определенный белок, по которому (биохимическими методами) и идентифицируют клон. ДНК в таких библиотеках не содержит интронов, и по ней легко определить аминокислотную последовательность белка, как сейчас нередко и поступают, определяя структуру белков, не имея их препаратов.

Нуклеаза Bal 31, выделяемая из псевдомонады *Alteromonas espejiana*, способна как экзонуклеаза удалять мононуклеотиды одновременно с 5′- и 3′-концов двуцепочечной ДНК, а также действовать как эндонуклеаза, специфичная к одноцепочечной ДНК. Таким образом, этот фермент позволяет «подтупить» несимметричные концы ДНК либо укоротить фрагменты ДНК, чтобы сблизить их функционально значимые элементы, а затем объединить их с помощью ДНК-лигазы.

Концевая (*терминальная*) *дезоксинуклеотидилтрансфераза из тимуса* теленка наращивает нуклеотиды на 3′-ОН-конец одно- или двуцепочечной молекулы ДНК. С ее помощью можно «подтупить» концы ДНК (если 3′-ОН-конец короче). Используя этот фермент, к фрагментам ДНК можно присоединять (наращивать) олигонуклеотидные «хвосты», состоящие из одного вида нуклеотидов. Если к другому фрагменту ДНК присоединить комплементарный ему другой хвост (конец), то они могут гибридизоваться и образовывать гибридные (рекомбинантные) молекулы ДНК (см. рис. 182).

Именно этот фермент и был использован в 1972 г. при получении первых гибридных ДНК. Кроме этого с его помощью можно вводить радиоактивные нуклеотиды в 3′-конец ДНК.

Поли(*А*)-*полимераза E. coli* катализирует присоединение к свободному 3′-концу одноцепочечной РНК поли(А)-последовательности. Этот фермент используется при подготовке молекул РНК, не содержащих поли(А)-хвоста, для синтеза на них кДНК с помощью ревертазы. Он может быть использован также для введения радиоактивной метки в 3′-конец РНК.

Полинуклеотидилкиназа была выделена из *E. coli*, инфицированной бактериофагом Т4. Фермент кодируется геномом бактериофага и переносит фосфорные остатки на 5′-концевые гидроксильные группы молекул ДНК и РНК, используя в качестве донора фосфата АТР. Этот фермент применяют для мечения 5′-конца полинуклеотидной цепи с помощью радиоактивного ^{32}P, содержащегося в АТР, меченной по γ-фосфату. Содержащийся в исходной молекуле немеченый 5′-фосфат предварительно удаляют с помощью другого фермента — *фосфомоноэстеразы* (*фосфатазы*).

Гибридизация нуклеиновых кислот

Метод основан на способности азотистых оснований образовывать комплементарные пары: A-T (U) и G-C. Реакция гибридизации используется в генетической инженерии для создания гибридных молекул ДНК, а также как весьма чувствительный метод выявления определенных последовательностей в ДНК и РНК. Если водный раствор ДНК нагреть до температуры 96—100 °С и сильно защелочить (pH > 13,0), то ДНК диссоциирует на отдельные цепи. Этот процесс *денатурации* ДНК обратим, поскольку если две изолированные цепи ДНК выдерживать определенное время при 65 °С, то они вновь спариваются, образуя двойную спираль, — что и называют *ренатурацией*, или *гибридизацией* (*отжигом*). Гибридизация может идти между одинарными цепями ДНК и/или РНК (конечно, если они имеют комплементарные последовательности нуклеотидов), что приводит к образованию двойных цепей (дуплексов) разного состава: ДНК:ДНК; РНК:РНК; ДНК:РНК.

При конструировании гибридных ДНК гибридизация нуклеиновых кислот реализуется, как правило, в двух основных стратегиях (методах) получения рекомбинантных ДНК: коннекторном и рестриктазно-лигазном.

Коннекторный метод создает условия для гибридизации продуктов рестрикции разных геномов путем наращивания на их концах комплементарных олигонуклеотидных участков (рис. 182). Гибридизация (отжиг) этих фрагментов ведет к образованию гибридных

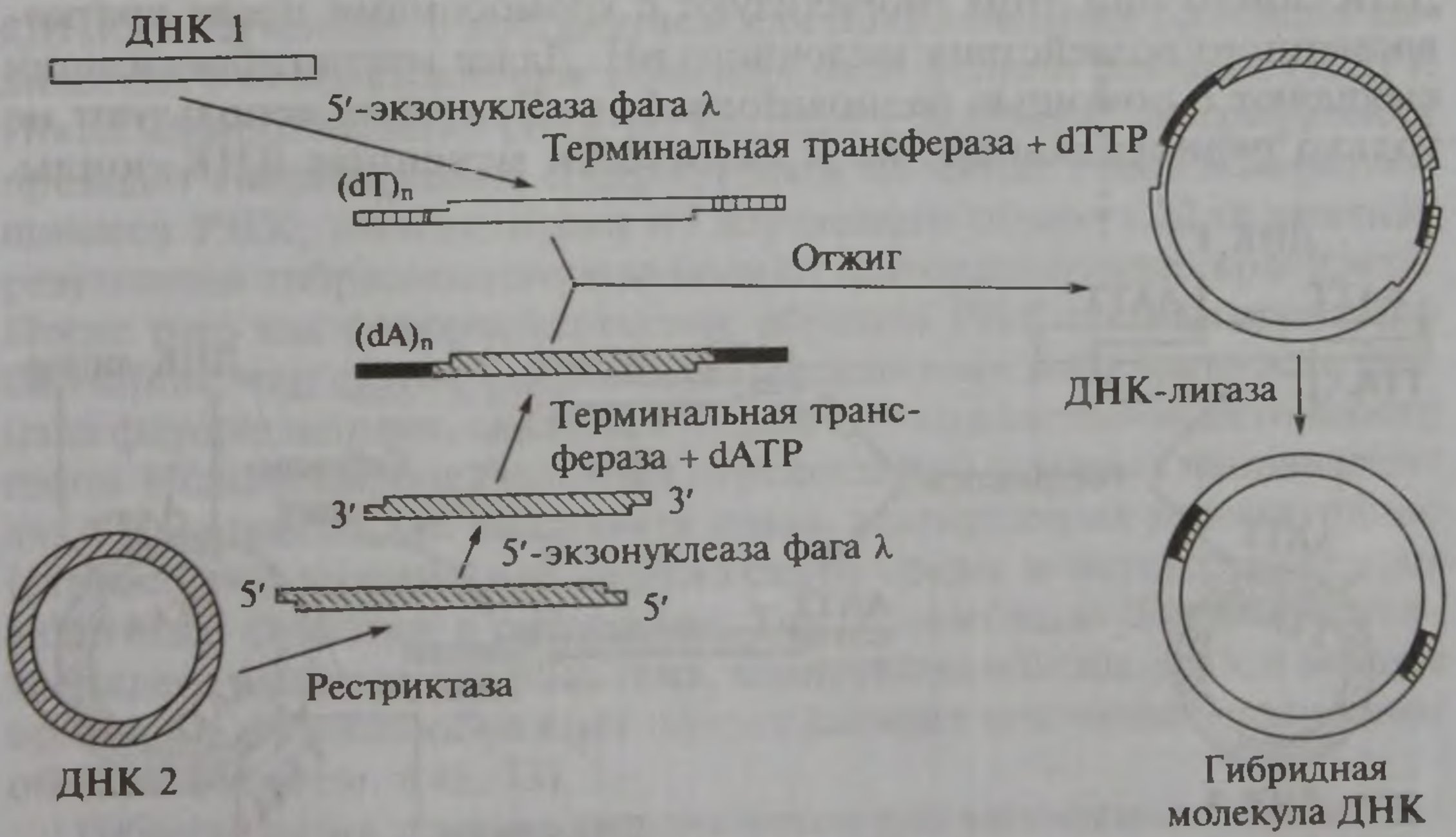


Рис. 182. Схема коннекторного метода получения гибридных молекул ДНК (по С. Н. Щелкунову, 1994)

молекул ДНК. В этом методе используются 3 фермента: 5′-экзонуклеза, терминальная нуклеотидилтрансфераза и ДНК-лигаза. Именно этим методом П. Берг и сотр. получили в 1972 г. первую гибридную молекулу ДНК.

Рестриктазно-лигазный метод, разработанный Коэном и сотр. в 1973 г., наиболее прост и популярен в генетической инженерии. В этом методе с использованием одной рестриктазы типа II, дающей фрагменты рестрикции с «липкими концами», гибридизация между фрагментами хромосомной ДНК и ДНК-плазмидой осуществляется без дополнительной процедуры наращивания комплементарных концов. После окончания гибридизации остается только сшить полинуклеотидные фрагменты с помощью ДНК-лигазы (рис. 183).

Гибридизация используется для нахождения числа определенных нуклеотидных последовательностей (генов) в ДНК. Для этой цели применяют *ДНК-зонды* — радиоактивные фрагменты ДНК с известной нуклеотидной последовательностью. Метку в зонд вводят путем ник-трансляции (см. выше). Поскольку скорость ренатурации (гибридизации) денатурированной нагреванием ДНК лимитируется вероятностью столкновения комплементарных последовательностей, скорость гибридизации с ДНК-зондом дает возможность определить число копий генов в ДНК. Это очень точный метод, который позволяет выявить один-единственный ген в клетке. Так выявляют отдельные (уникальные) гены, а также гены, представленные в геноме десятками или сотнями копий.

Для локализации специфичных последовательностей нуклеиновых кислот в хромосомах и клетках используют *гибридизацию in situ*. ДНК-зонды при этом гибридизуют с хромосомами после кратковременного воздействия щелочного pH. Далее места гибридизации выявляют с помощью радиоавтографии. При этом используют не только радиоактивные, но и химически меченные ДНК-зонды.

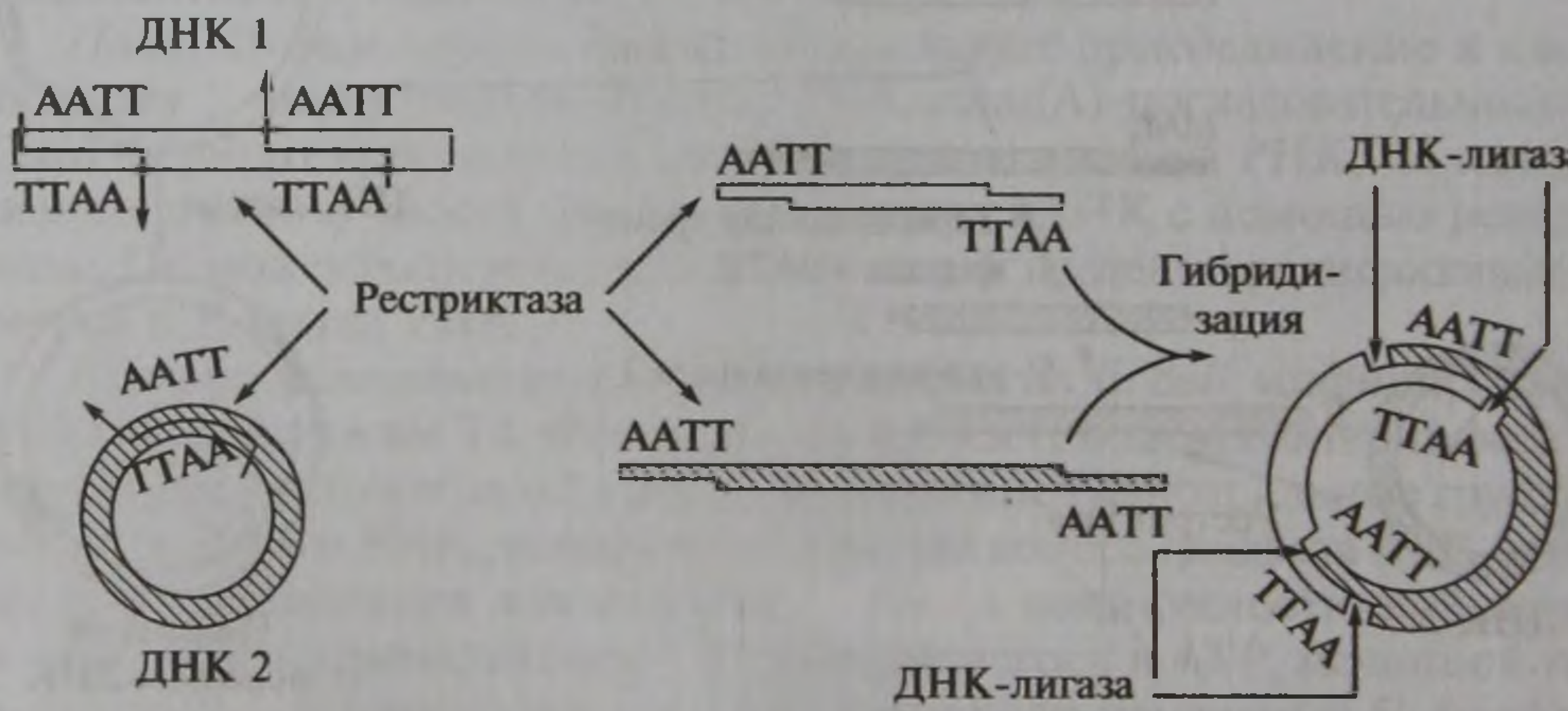


Рис. 183. Схема рестриктазно-лигазного метода получения гибридных молекул ДНК

Для этого при синтезе зонда используют нуклеотиды, содержащие боковую цепь *биотина*, которую после гибридизации окрашивают стрептовидином. С помощью гибридизации in situ можно выявить локализацию определенных видов мРНК в клетках и таким образом судить о дифференциальной активности генов в эмбриогенезе и процессе дифференцировки в клетках эмбриона. Популярной биологической моделью таких исследований являются эмбрионы плодовой мушки дрозофилы.

Гибридизация используется для выявления транскрибируемых и нетранскрибируемых последовательностей в ДНК. При этом анализируются продукты гибридизации ДНК и матричных РНК, что позволяет выявить активные (транскрибируемые) участки в молекулах ДНК. Именно анализ результатов гибридизации ДНК:РНК позволил У. Гилберту выявить *мозаичное строение генов* эукариот (см. гл. 6), что стало одним из крупнейших открытий в молекулярной биологии.

При высоком числе генов у эукариот (от 6 200 у дрожжей до 30 000 и более у человека) возникает необходимость в использовании специальной техники для одновременного получения экспериментальной информации об активности большого числа генов в клетке.

Современная экспериментальная техника позволяет создать матрицу — *биочип*, размером всего несколько сантиметров, с помощью которой можно получить данные о функциональной активности многих (если не всех) генов организма. В этом методе на специальную (стеклянную) подложку с помощью роботов наносят образцы ДНК, которые представляют собой либо отдельные гены, либо молекулы кДНК, полученные в достаточном для приготовления биочипов количестве с использованием полимеразной цепной реакции (ПЦР). После иммобилизации ДНК на матрице с помощью ультрафиолета проводят гибридизацию содержащихся на чипах генов с образующимися РНК, выделяемыми из изучаемого объекта. Для анализа результатов гибридизации используют флуоресцентные красители. После того как флуоресцирующие образцы РНК прореагировали с биочипом, чип сканируют лазером и оценивают интенсивность сигнала флуоресценции, служащей мерой функциональной активности генов. Использование биочипов перспективно в разных направлениях, и в частности для выявления генов, реагирующих на негативное (стрессовое) воздействие окружающей среды и осуществляющих защитные функции в организме. Так, с помощью биочипов подтверждена выдающаяся роль гена, кодирующего белок р-53, в защите организма от разнообразных повреждающих реагентов (подробнее об этом белке см. в гл. 13).

Гибридизация широко используется для выявления определенных нуклеотидных последовательностей в смеси рестрикционных фрагментов. Такой прием носит название «блоттинг» (от англ. *blot* — промокать). С этой целью рестрикционные фрагменты фрак-

ционируют методом электрофореза в 0,8 %-й агарозе, что позволяет разделить до 500 фрагментов, отличающихся по размеру всего на один нуклеотид. Затем гель совмещают с листом нитроцеллюлозной бумаги. В результате диффузии фрагменты частично переходят на этот лист (blot), и таким образом получается *реплика* (отпечаток) с геля. Затем методом радиоавтографии с использованием радиоактивного ДНК-зонда и рентгеновской пленки определяют на реплике положение фрагментов, которые гибридизуются с зондом. Такой прием выявления фрагментов ДНК с использованием ДНК-зондов, приводящий к образованию гибридов ДНК:ДНК, носит название *саузерн-блоттинг.*

Нозерн-блоттингом называют процедуру, при которой ДНК-зонды гибридизуют с репликами геля, в котором фракционируют РНК, что приводит к образованию гибридов ДНК:РНК.

Гибридизацию (блоттинг) зондов можно осуществлять и в отношении целых хромосом, которые удается фракционировать в крупнопористых агарозных гелях методом *пульс-электрофореза*, меняя направление электрического поля, что позволяет фракционировать огромные молекулы ДНК.

Способность ДНК к денатурации-ренатурации используется и в одном из самых распространенных в настоящее время методов амплификации фрагментов ДНК in vitro — полимеразной цепной реакции.

Полимеразная цепная реакция и другие методы амплификации нуклеиновых кислот

Полимеразная цепная реакция (*ПЦР*) — это метод амплификации фрагментов нуклеиновых кислот in vitro, с помощью которого можно достаточно быстро (в течение нескольких часов) получить миллионы копий определенных нуклеотидных последовательностей (генов). Метод был предложен в 1985 г. К. Мюллисом (биотехнологическая корпорация «Cetus», США) и получил широкое распространение с 1988 г., когда Р. Сайки с соавт. была опубликована основополагающая работа по теории ПЦР и ее оптимизации. Метод ПЦР, названный «изобретением века» и очень скоро, в 1993 г., отмеченный Нобелевской премией, ускорил реализацию программы «Геном человека» (см. гл. 6), а также способствовал внедрению в практику клинической диагностики многих заболеваний высокоэффективных диагностических наборов нового поколения.

В методе ПЦР для амплификации фрагментов ДНК используют термоустойчивую ДНК-полимеразу из термофильной бактерии *Thermus aquaticus* (Taq-полимеразу), которая в присутствии всех 4 видов дезоксирибонуклеозидтрифосфатов и коротких 20 — 30-член-

ных затравок (праймеров) осуществляет синтез комплементарных последовательностей ДНК.

Вначале осуществляют термическую денатурацию (расплетание) двуцепочечной молекулы (фрагмента) ДНК при 93—95 °С, что приводит к разделению комплементарных цепей, а затем пробы охлаждают до 60 °С, что дает возможность связаться праймерам и одноцепочечным дезоксиолигонуклеотидам, а также запускает работу Taq-полимеразы.

Праймеры обычно получают методом химического синтеза, определив предварительно нуклеотидную последовательность пограничных с амплифицируемым геном участков ДНК. В менее строгом варианте, когда не известны пограничные участки генов, используют случайные праймеры, которые могут связываться с ДНК. Праймеры подстраиваются сами (по комплементарному принципу) к денатурированным нагреванием гомологичным участкам одноцепочечных фрагментов исходной («материнской») ДНК. Они служат по существу для двух целей: во-первых, запускают работу ДНК-полимеразы, а во-вторых, ограничивают ее действие, как бы застопоривают фермент в рамках нужного участка копирования ДНК (рис. 184).

Как видно из схемы, ПЦР имеет циклический характер. В первом и частично во втором циклах образующиеся копии (*ампликоны*) ДНК не соответствуют границам амплифицируемого гена (все ампликоны в первом цикле и часть ампликонов во втором цикле получаются более протяженными в тех участках, где еще не произошло связывание второго — «антисмыслового» праймера). Однако начиная уже с третьего цикла длина ампликонов становится стандартной, т.е. соответствует числу пар нуклеотидов фрагментов ДНК-матрицы между 3′-концами праймеров. Ампликоны накапливаются в геометрической прогрессии и начинают доминировать среди продуктов амплификации. Процесс имеет цепной характер, так как синтезированные фрагменты ДНК в дальнейшем сами служат матрицей, на которой идет синтез.

При многократном повторении циклов синтеза (включающих нагревание и охлаждение проб в вышеназванных пределах температур) число копий специфических фрагментов ДНК экспоненциально увеличивается, что позволяет из небольшого числа имеющихся фрагментов получить большое количество копий, которые можно идентифицировать методом электрофореза.

Повторяя циклы амплификации 30—40 раз, за 1,5—3 ч получают миллионы копий фрагментов ДНК. Увеличение числа циклов неэффективно, так как ПЦР имеет экспоненциальный характер только на начальных этапах (20—25 циклов), после чего начинается выход на плато в силу истощения дезоксирибонуклеозидтрифосфатов, праймеров и нарастающего температурного повреждения Taq-полимеразы. Тем не менее результаты, достигаемые с помощью ПЦР, позволили этому методу занять ведущее место в молекулярной биологии как по популярности использования в научных исследованиях фун-

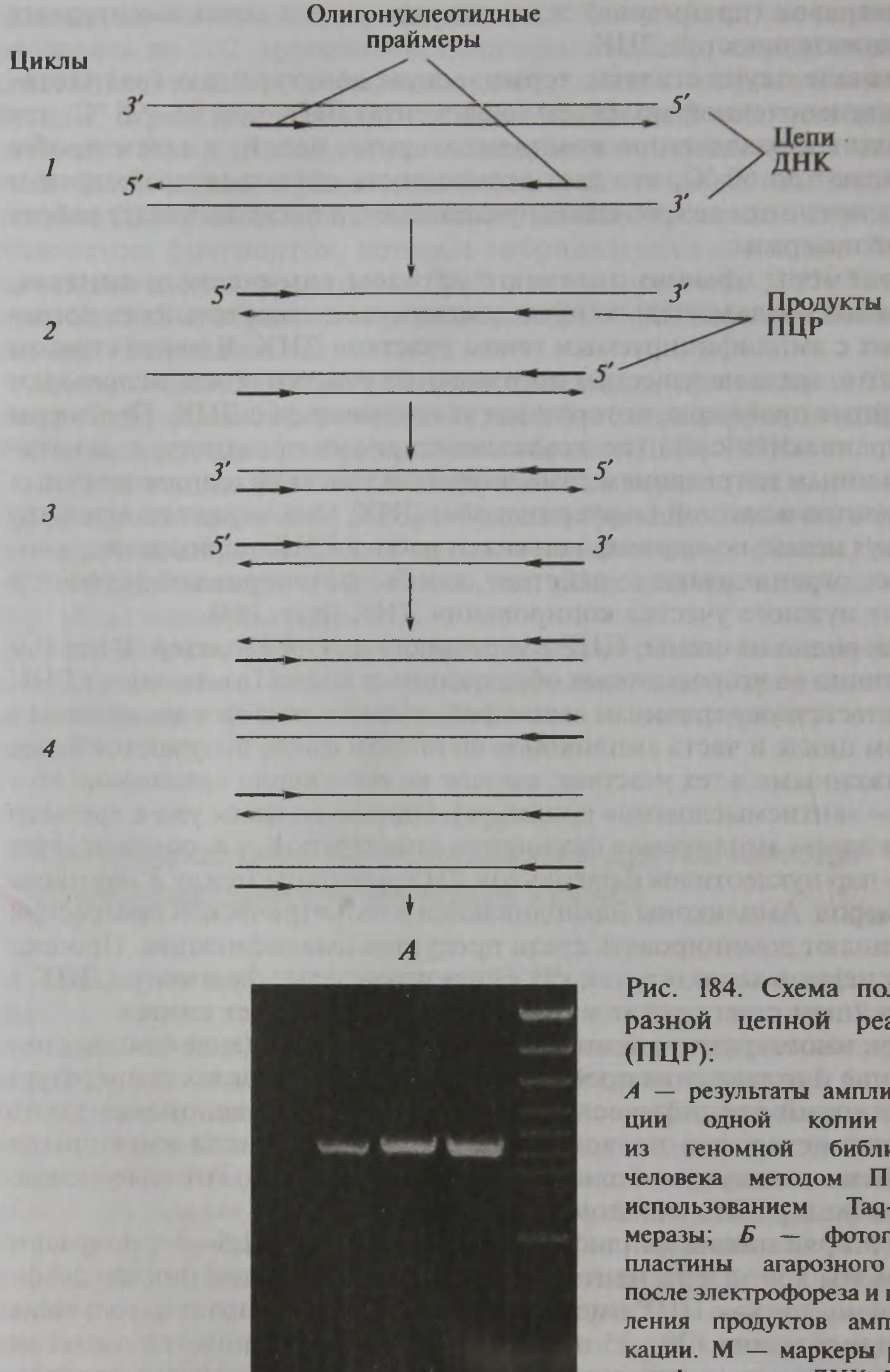


Рис. 184. Схема полимеразной цепной реакции (ПЦР):

А — результаты амплификации одной копии гена из геномной библиотеки человека методом ПЦР с использованием Taq-полимеразы; *Б* — фотография пластины агарозного геля после электрофореза и проявления продуктов амплификации. М — маркеры размеров фрагментов ДНК. Размер амплифицированного гена 960 н. п. (*Innis M. A. et al.* PCR Protocols and Applications: A Laboratory Manuals. — Academic. New York, 1989.)

даментального характера, так и по практическому применению в медицинских целях. В частности, этим методом удалось получить продукты амплификации из одной-единственной копии матричной последовательности гена глобина человека и из уникальных генов одного сперматозоида.

Метод ПЦР быстро прогрессирует, создаются его новые разновидности, нацеленные на решение конкретных задач. Ниже приведены краткие сведения по нескольким видам ПЦР.

• *ПЦР с обратной транскрипцией* (англ. *Reverse Transcription PCR, RT-PCR*) используется для амплификации, выделения или идентификации известной последовательности из библиотеки РНК. Перед обычной ПЦР проводят на матрице мРНК синтез одноцепочечной молекулы ДНК с помощью ревертазы и получают одноцепочечную кДНК, которая используется в качестве матрицы для ПЦР. Этим методом часто определяют, где и когда экспрессируются данные гены.

• *Асимметричная ПЦР* (англ. *Asymmetric PCR*) проводится когда нужно амплифицировать преимущественно одну из цепей исходной ДНК. Используется в некоторых методиках секвенирования и гибридизационного анализа. ПЦР проводится как обычно, за исключением того, что один из праймеров берется в большом избытке.

• *Количественная ПЦР* (англ. *Quantitative PCR, Q-PCR*) используется для быстрого измерения количества определенной ДНК, кДНК или РНК в пробе.

• *Количественная ПЦР в реальном времени* (англ. *Quantitative real-time PCR*). В этом методе используют флуоресцентно меченые реагенты для точного измерения количества продукта реакции по мере его накопления.

• *Touchdown* (*Stepdown*) *ПЦР* (англ. *Touchdown PCR*). С помощью этого метода уменьшают влияние неспецифического связывания праймеров на образование продукта. Первые циклы проводят при температуре выше температуры отжига, затем каждые несколько циклов температуру снижают. При определенной температуре система пройдет через полосу оптимальной специфичности праймеров к ДНК.

В отдельных видах ПЦР, кроме «классической» Taq-полимеразы, используются и другие термостабильные ДНК-полимеразы, несколько отличающиеся по свойствам. Так, например, Vent ДНК-полимеразы из *Thermococcus litoralis* и Pfu ДНК-полимеразы из *Pyrococcus furiosus*, отличающиеся высокой точностью копирования амплифицируемых участков ДНК, и Taq SE ДНК-полимеразы пригодны для амплификации особо протяженных участков ДНК (до 15 тыс. н.п.).

Лигазная цепная реакция проводится по принципу, аналогичному ПЦР, но вместо Taq-полимеразы и dNTP используется термостабильная ДНК-лигаза и четыре специфических олигонуклеотида, добав-

ляемых в реакционную смесь в избытке. Каждые 2 олигонуклеотида комплементарны к амплифицируемому фрагменту ДНК-матрицы и непосредственно примыкают друг к другу; одновременно они комплементарны и другой паре олигонуклеотидов. После денатурации ДНК-матрицы при 95 °С олигонуклеотиды связываются с соответствующими (комплементарными) фрагментами цепей ДНК-матрицы и сшиваются лигазой при 50 °С. Далее эти две стадии цикла повторяются много раз, в результате чего экспоненциально увеличивается число соединенных ковалентной связью олигонуклеотидов, и таким образом достигается амплификация определенных участков ДНК.

NASBA-метод (Nucleic Acid Seaquance — Base Amplification), разработанный в последние годы, является наиболее универсальным методом амплификации как ДНК, так и РНК. Этот метод, в отличие от ПЦР, является изотермальным и осуществляется при 41 °С. Основными компонентами NASBA-системы являются РНК-полимераза фага Т7, РНКаза Н и обратная транскриптаза вируса птичьего миелобластоза (об этих ферментах см. выше). В систему входят также нуклеозидтрифосфаты и два специфических праймера, один из которых содержит участок, представляющий собой последовательность (промотор), распознаваемую РНК-полимеразой. В качестве матриц могут быть использованы и двуцепочечные нуклеиновые кислоты, но после превращения их в одноцепочечные. NASBA-метод применяется для диагностики инфекционных заболеваний, генетических дефектов, диагностики и контроля терапии рака. Чувствительность NASBA при тестировании ВИЧ составляет три копии (РНК) вируса иммунодефицита человека в 100 мкл образца плазмы крови. Метод можно использовать для контроля пищевых продуктов, в ветеринарной диагностике и судебно-медицинской службе.

Клонирование ДНК

Методом клонирования фрагменты ДНК любого вида, полученные с помощью рестриктаз, можно ввести в плазмиду или в бактериофаг, получив, таким образом, вектор молекулярного клонирования, а затем размножить эти генетические элементы в клетках бактерий или дрожжей, увеличивая их число в миллионы раз. Встраивание фрагмента ДНК в плазмиду осуществляется in vitro, а затем рекомбинантная ДНК вводится в клетки бактерии-хозяина. Наиболее часто в роли хозяина выступает детально изученный штамм *E. coli* K12, а в роли вектора — плазмиды или фаги этой бактерии.

В целом бактерии могут приобретать новый генетический материал тремя путями: *трансформацией*, *конъюгацией* или *трансдукцией*. *Трансформация* представляет собой изменение генотипа клетки путем внесения в нее ДНК из культуральной среды. В молекулярной биологии этот термин обозначает стабильное изменение генотипа и

фенотипа клетки. *Конъюгация* обозначает прямой перенос ДНК из одной клетки (донора) в другую клетку (реципиент) через особый белковый мостик (канал), возникающий между клетками. У *E. coli* этот перенос занимает примерно 90 мин. Установив время, необходимое для переноса определенных генов в процессе конъюгации, удается построить *генетическую карту* бактерии. *Трансдукция* происходит с участием бактериофагов, которые вводят в ДНК клетки новые генетические элементы.

При использовании фагов в генетической инженерии носителем рекомбинантной ДНК является популяция фаговых частиц, и рекомбинантный фаговый геном можно клонировать непосредственно (в чашках Петри) на газоне клеток-хозяев, где образуются фаговые бляшки. Затем их можно ввести в клетки бактерий в виде изолированной ДНК (трансфекция) либо в виде реконструированных фаговых частиц (инфекция). Процедура получения рекомбинантных ДНК и их клонирования основана на тех же принципах, что и трансдукция. Помимо бактериофагов в ней используются плазмиды.

Плазмиды представляют собой кольцевые молекулы двуцепочечной ДНК, встречающиеся в клетках бактерий и дрожжей, где они реплицируются в процессе пролиферации клеток как самостоятельные единицы. С точки зрения генетической инженерии плазмиды обладают несколькими важными свойствами или функциями, к числу которых относятся: репликация, несовместимость, функция переноса и устойчивость к антибиотикам. Плазмиды, обладающие вышеперечисленными свойствами, наряду с бактериофагами используют в качестве векторов клонирования. *Вектор* — молекула ДНК, которая способна переносить в клетку чужеродную ДНК и обеспечивать там ее амплификацию. В генетической инженерии процесс введения молекул ДНК в живые клетки бактерий, называемый *трансфекция*, обычно осуществляется после обработки клеток $CaCl_2$, что делает их оболочки более проницаемыми, и введение ДНК занимает всего несколько минут.

Бактериальные плазмиды имеют блочную организацию. При этом гены переноса могут объединяться с *репликоном* (участком, ответственным за репликацию). Блочная организация плазмид дает возможность методами генетической инженерии специально конструировать векторы, обладающие необходимыми качествами. Одним из наиболее распространенных векторов является специально созданная плазмида pBR322 (рис. 185). Нуклеотидная последовательность этого вектора (4 362 н. п.) полностью изучена. В нем имеются гены устойчивости к ампициллину (Amp^r) и тетрациклину (Tet^r). Репликон в этом векторе взят из плазмиды Col EI. Эта природная плазмида *E. coli* содержит ген, кодирующий белок-токсин колицин, что и отражено в ее названии. Репликон обеспечивает инициацию репликации, включая синтез особого вида РНК (так называемой РНКI), которая контролирует интенсивность репликации (число ко-

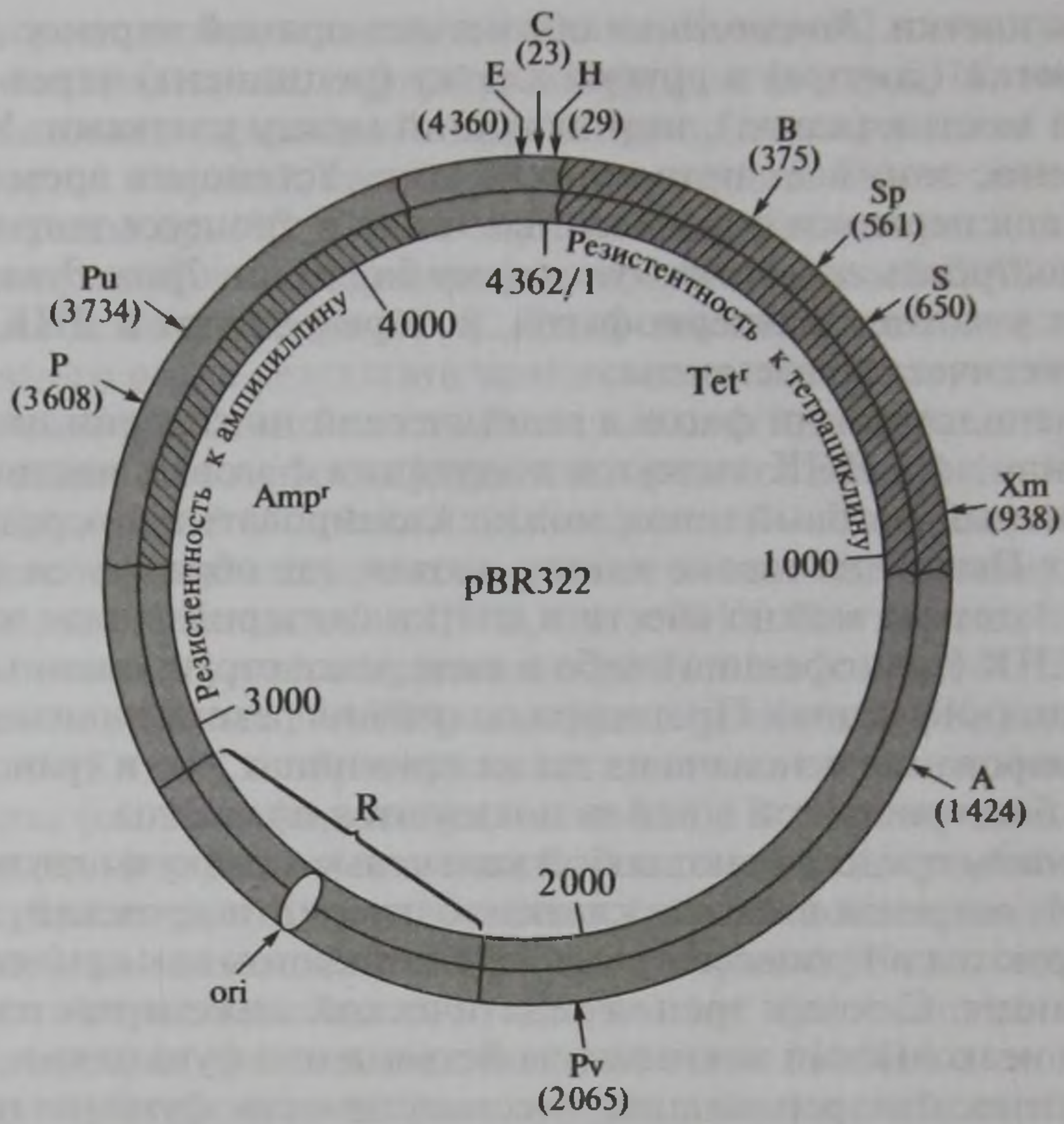


Рис. 185. Схема строения плазмидного вектора pBR322 (М. Сингер, П. Берг, 1998):

ori — точка начала репликации; Amp^r и Tet^r — гены устойчивости к ампициллину и тетрациклину; указаны сайты рестрикции для различных рестриктаз; цифры обозначают число нуклеотидных пар; R — репликон

пий) плазмиды в клетке. В норме в клетке *E. coli* может содержаться до 50 копий плазмиды pBR322. Клетки *E. coli*, содержащие данный вектор, выращивают в среде с ампициллином или тетрациклином, на которой клетки, не имеющие эту плазмиду, не растут, и таким образом на такой *селективной среде* накапливаются только те клетки, которые содержат рекомбинантную ДНК. Это свойство облегчает выделение больших количеств рекомбинантной плазмидной ДНК. Из-за своего небольшого размера плазмидная ДНК легко отделяется от хромосомной ДНК методом центрифугирования.

В генетической инженерии создаются плазмидные векторы, которые обеспечивают высокий уровень экспрессии определенных генов. Далее ген, который хотят использовать для усиленной наработки какого-либо белка, совмещают в плазмиде с сильным *промотором* — участком ДНК, обеспечивающим эффективное связывание РНК-полимеразы и высокую скорость синтеза матричной РНК. Кроме этого стараются использовать *регулируемые промоторы*, интенсивность действия которых можно изменять путем изменения

некоторых параметров (температуры, концентрации определенных ионов и т. д.). Одним из таких сильных регулируемых промоторов является промотор pL фага λ. Его активность возрастает при повышении температуры в связи с инактивацией белка репрессора (λ-репрессор). Примером вектора с промотором pL фага λ является и плазмида pI L203. Этот вектор был создан с использованием плазмиды pBR322, у которой регуляторная область была заменена на таковую фага λ (рис. 186). С помощью этой плазмиды были созданы бактериальные штаммы-суперпродуценты нескольких ферментов.

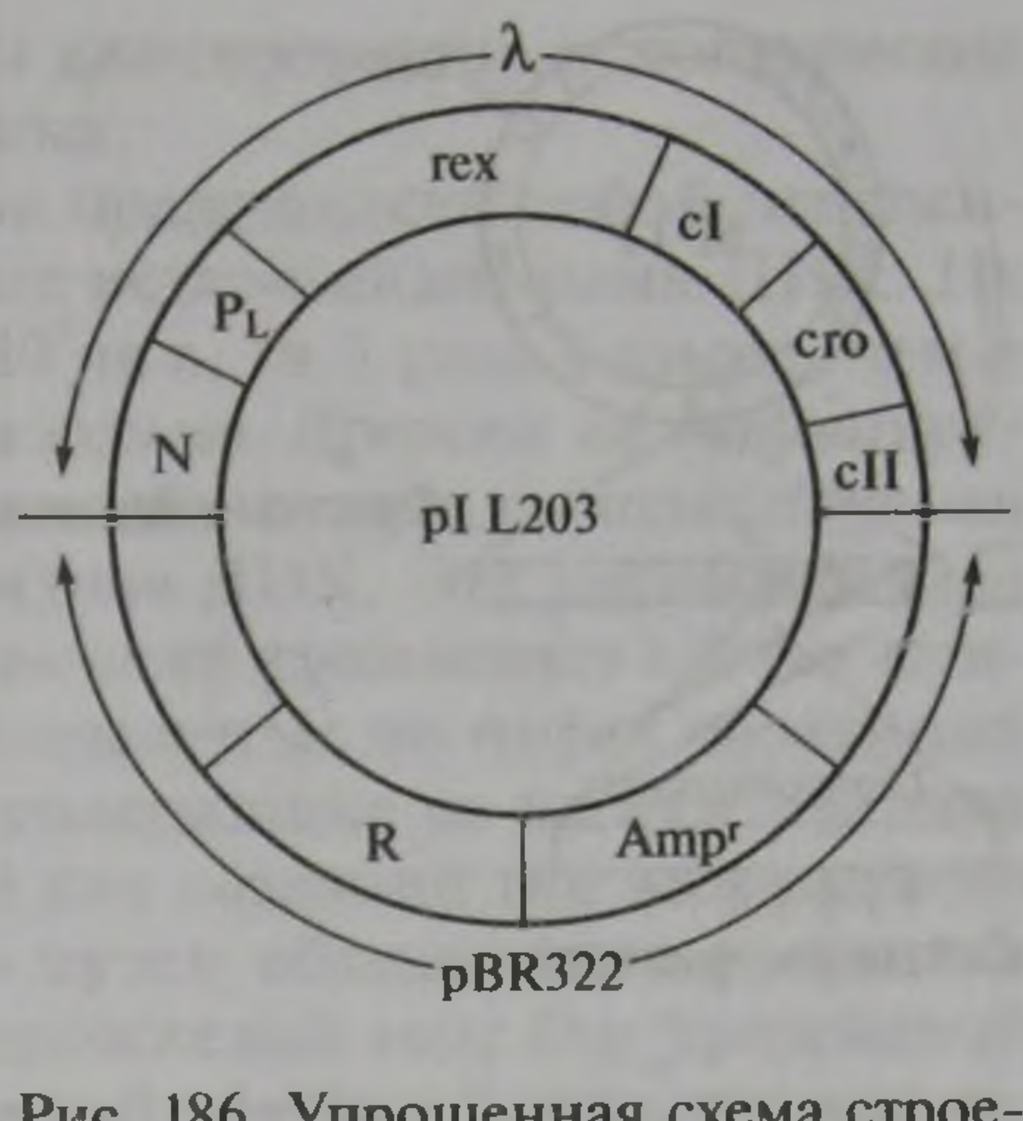


Рис. 186. Упрощенная схема строения плазмидного вектора pI L203:
N, rex, cro, cI, cII — гены фага λ; P_L — регулируемый промотор; Amp^r — ген устойчивости к ампициллину

Возможности плазмид для генетической инженерии не беспредельны, что связано с их небольшими размерами. Когда нужно клонировать крупные фрагменты ДНК, удобнее использовать векторы на основе бактериофага λ. Геном фага λ досконально изучен: он содержит область начала репликации (ori), гены белков головки и хвостового отростка, а также гены ферментов, необходимые для репликации фаговой ДНК, гены лизиса инфицируемых клеток и гены, детерминирующие лизогенный путь существования фага (подробнее о регуляции активности этих генов см. гл. 9). В центральной части линейного генома фага содержится область, необязательная для литической инфекции, которую и используют для вставки клонируемой ДНК.

ДНК фага λ разрезают с помощью рестриктазы *Eco* RI, удаляют необязательную центральную область и на ее место встраивают нужный фрагмент ДНК, получая, таким образом, *конкамер* — предшественник для упаковки фаговой ДНК в зрелые фаговые частицы. Фаг λ очень пластичен: без нарушения развития фага из него можно убрать до 25 % ДНК или пристроить до 6 % лишней ДНК. В этот фаговый вектор можно встраивать до 23 000 н. п.

В тех случаях, когда не хватает возможностей фага, используют еще более крупные векторы — *космиды*. В состав космид входят: ген (маркер) резистентности к антибиотикам, репликон плазмиды и фрагмент ДНК фага λ (так называемый cos-участок). Этот фрагмент представляет собой однонитевые комплементарные участки на концах фаговой ДНК, т. е. «липкие концы». Космиды по существу представляют собой подвижные «cos»-сайты, между «липкими концами» которых удается встроить фрагменты ДНК длиной до 45 000 н. п.



Фрагмент lac-оперона *E. coli*

Рис. 187. Вектор на основе фага M13:

1—9 — гены фага M13; z — ген lac z; P — промотор; O — оператор. Полилинкер (42 н. п.) содержит сайты рестрикции для различных рестриктаз (остальные пояснения см. в тексте)

Еще один вид векторов — *фазмиды*, искусственные гибриды между фагом и плазмидой. После встройки чужеродной ДНК они в одних условиях могут размножаться как фаги, а в других — как плазмиды. Фазмиды имеют определенные преимущества перед векторами других типов. Их можно «упаковывать» в оболочки фаговых белков, и они будут существовать как фаги, при температуре 42 °С, либо, используя температурочувствительный белок — репрессор фага λ, заставлять их размножаться как плазмиды (при более низкой температуре ~32 °С).

Векторы на основе нитевидных фагов применяют тогда, когда удобнее работать с одной цепью ДНК (например, для определения ее нуклеотидной последовательности — см. ниже). Так, фаги M13 и fd содержат кольцевую одноцепочечную ДНК с полностью изученными последовательностями нуклеотидов (6 407 нуклеотидов у фага M13) и их организацией. Эти фаги могут содержаться в количестве 200—300 копий на одну бактериальную клетку. Для размножения они переходят в двуцепочечную репликативную форму, в которую с помощью рестриктаз и лигаз легко встроить нужный фрагмент ДНК, где далее он будет размножаться и существовать в одноцепочечной форме. Типичный вектор на основе фага M13 изображен на рис. 187.

Для удобства отбора рекомбинантов геном дикого типа M13 модифицируют, встраивая в него часть лактозного оперона (lac-оперона) *E. coli*, включающую промотор, оператор и кодирующую область гена β-галактозидазы. В эту область дополнительно встраивают фрагмент ДНК, содержащий сайты узнавания для определенных рестриктаз, с тем, чтобы далее можно было их использовать для включения фрагментов ДНК с соответствующими «липкими концами». Такой сегмент называют *полилинкером*.

Все вышеизложенное имеет отношение к клонированию ДНК в наиболее часто используемых прокариотических системах хозяин-вектор. Однако в настоящее время начинают развиваться соответ-

ствующие эукариотические системы клонирования, использующие клетки дрожжей, растений и животных.

Дрожжи *Saccharomyces cerevisiae* представляют собой перспективные модели для экспериментов с рекомбинантными ДНК. Их гаплоидный геном содержит $1{,}4 \cdot 10^7$ н. п. (в 3 раза больше, чем у *E. coli*), распределенных по 17 хромосомам. Дрожжи не инфицируются вирусами, но в их клетках выявлена плазмида, используемая в качестве вектора — 2мкм-плазмидная ДНК. Эта двуцепочечная плазмида имеет длину в 318 н. п., на одну дрожжевую клетку приходится 50 копий. Дикий тип этой плазмиды не имеет маркерных генов, что не позволяет отбирать содержащие ее клетки. Однако 2мкм-плазмида была использована для создания так называемого *челночного вектора*, полученного путем объединения ее части с плазмидой pBR322, содержащей дрожжевой ген. Ген дрожжевой клетки кодирует фермент, необходимый для биосинтеза гистидина, благодаря чему удается проводить отбор клеток, растущих в отсутствие гистидина. Для введения вектора жесткую полисахаридную стенку дрожжевых клеток разрушают с помощью соответствующих ферментов либо обрабатывают их ацетатом лития, что открывает возможность для проникновения ДНК в клетки.

Важной особенностью дрожжевых систем является то обстоятельство, что клонируемые фрагменты ДНК способны к рекомбинации с гомологичными участками дрожжевого генома, что ведет к стабильной *сайт-специфической трансформации* последнего (независимо от сохранения или утраты вектора). Поэтому в последнее время созданы конструкции на основе дрожжевых хромосом, которые позволяют клонировать в клетках дрожжей фрагменты ДНК размером до 1 000 000 н. п.

В клетках растений нет природных плазмид и в качестве векторов могут быть использованы различные растительные вирусы. В то же время векторы для растительных клеток наиболее успешно создают на основе особых бактериальных Ti-плазмид, которые обладают способностью переносить сегменты ДНК в геномы растений (см. ниже).

При введении ДНК в клетки животных часто используют вирусы, в которые предварительно вводят необходимый сегмент ДНК. После образования вирионов вирусы проникают в клетки путем инфицирования. Упаковка происходит обычно in vivo, так как in vitro, в отличие от бактериофагов, этот процесс пока не осуществим. Можно проводить и прямую инъекцию ДНК в клетки и их ядра, но при этом трансформируется только малая часть исследуемых клеток (не более 10%). В ряде случаев используется электропорация — обработка смеси клеток и вводимой ДНК высоковольтным (2—4 тыс. В) электрическим разрядом, что приводит к образованию в мембранах клеток отверстий (пор), необходимых для проникновения ДНК. Векторы для трансформации клеток животных наиболее часто создают на основе вируса SV40 и вируса папилломы. На основе вируса SV40 аналогично

вышеупомянутой процедуре конструирования векторов на основе фага λ создано несколько типов векторов: трансдуцирующие, плазмидные и пассивные. *Трансдуцирующие векторы* реплицируются в клетках почек обезьяны и упаковываются в вирионы. *Плазмидные SV40-векторы* реплицируются, но не упаковываются в белковые оболочки. *Пассивные трансформирующие векторы* не могут ни реплицироваться, ни упаковываться, однако содержат участки генома вируса, способствующие экспрессии других генов. Стабильная трансформация в клетках животных происходит тогда, когда плазмидные векторы SV40 интегрируются в геном клетки-хозяина. Частота этой интеграции составляет не более 10^{-5}—10^{-3}.

Папилловирусы вызывают доброкачественные новообразования кожи (бородавки) у ряда видов животных. Геном папилломавируса крупного рогатого скота (BPV) используют для создания векторов животных. Он способен реплицироваться в виде стабильной плазмиды без образования вирионов. Встраивание ДНК в этот вектор производят с сохранением части генома вируса, которую называют «*трансформирующей*». При этом вектор не внедряется в геном клетки-хозяина, а остается в ней в виде плазмиды в количестве от 10 до 100 копий.

Определение нуклеотидных последовательностей

Для выяснения структуры, функций, происхождения, возможностей рекомбинации и амплификации (клонирования) молекул нуклеиновых кислот или их фрагментов необходимо установить их нуклеотидную последовательность. Работы в этой области начались с изучения первичной структуры РНК и были основаны на применении специфических ферментов — рибонуклеаз, расщепляющих молекулы РНК в местах расположения определенных нуклеотидов: пуриновых, распознаваемых РНКазой U2; пиримидиновых — панкреатической РНКазой I; гуаниловых — РНКазой T_1 и т. п. С помощью этого метода «разрезания молекулы», в значительной мере разработанного в нашей стране академиком А. А. Баевым и сотрудниками, в 60—70-е годы XX в. были расшифрованы нуклеотидные последовательности различных тРНК, 5S и 16S рРНК *E. coli*, РНК фагов R-17, MS2 и др. Этот метод трудно было автоматически перенести на изучение первичной структуры (секвенирование) ДНК, так как не известны были специфические эндонуклеазы, расщепляющие ДНК по определенным основаниям (типа РНКазы T_1). В середине 70-х г. в этой области произошел решительный перелом, связанный с открытием рестриктаз и усовершенствованием метода гель-электрофореза, когда стало возможным разделять достаточно протяженные фрагменты

ДНК, отличающиеся размером всего на один нуклеотид. С помощью рестриктаз ДНК стали разрезать на определенные блоки и определять в них позиции нуклеотидов химическими или энзиматическими методами.

Химическое секвенирование основано на специфической химической модификации пуриновых и пиримидиновых оснований с последующим выщеплением модифицированных нуклеотидов из полимерной цепи и анализом образовавшихся продуктов методом гель-электрофореза.

Этот метод, разработанный А. Максамом и У. Гилбертом в 1976 г., стал ярким примером использования химических приемов для решения задач молекулярной биологии. Однако, со временем он был постепенно вытеснен более «мягким» биохимическим (энзиматическим) методом секвенирования ДНК.

Энзиматический метод был разработан Сангером и Коулсоном в 1975 г. В этом методе анализируемый фрагмент ДНК используют в качестве матрицы в реакции полимеразного копирования (синтеза комплементарной цепи ДНК) с помощью ДНК-полимеразы I *E. coli*. Этот метод получил название «плюс-минус-метод», поскольку реакцию полимеризации в нем изначально проводили либо в отсутствие одного из 4 типов нуклеотидов («минус»-система), либо в присутствии только одного нуклеотида («плюс»-система), что ограничивает возможность наращивания полинуклеотидной цепи, т.е. останавливает (терминирует) ее синтез из-за недостатка соответствующего нуклеотида. Затем для остановки синтеза стали использовать специальные молекулы — терминаторы, в качестве которых берут дидезоксирибонуклеозидтрифосфаты (рис. 188). Для секвенирования этим методом используют одноцепочечную молекулу ДНК, а также короткий олигодезоксинуклеотидный праймер (затравку), комплементарный небольшому участку анализируемой ДНК. Одиночная (секвенируемая) цепь ДНК служит матрицей для работы ДНК-полимеразы, а новая (комплементарная) цепь наращивается, начиная с 3′-гидроксильной группы праймера путем присоединения соответствующих дезоксирибонуклеозидфосфатов. При секвенировании проводят четыре разных реакции копирования, для чего комплекс «матрица-праймер» инкубируют в присутствии всех четырех обычных дезоксинуклеозидтрифосфатов (dATP, dTTP, dGTP и dCTP) и какого-то одного определенного дидезоксинуклеозидтрифосфата, который и служит терминатором синтеза, после своего встраивания в растущую цепь ДНК: присоединение дидезоксинуклеозидфосфата, у которого отсутствует 3′-гидроксильная группа, делает невозможным присоединение следующего нуклеотида (рис. 188, *А*). Поскольку терминация синтеза происходит в разных местах, то образуется набор цепей (фрагментов) различной длины, начиная с 5′-конца праймера до места расположения соответствующего присоединенного дидезоксинуклеотида. Новосинтезированные фрагменты являются

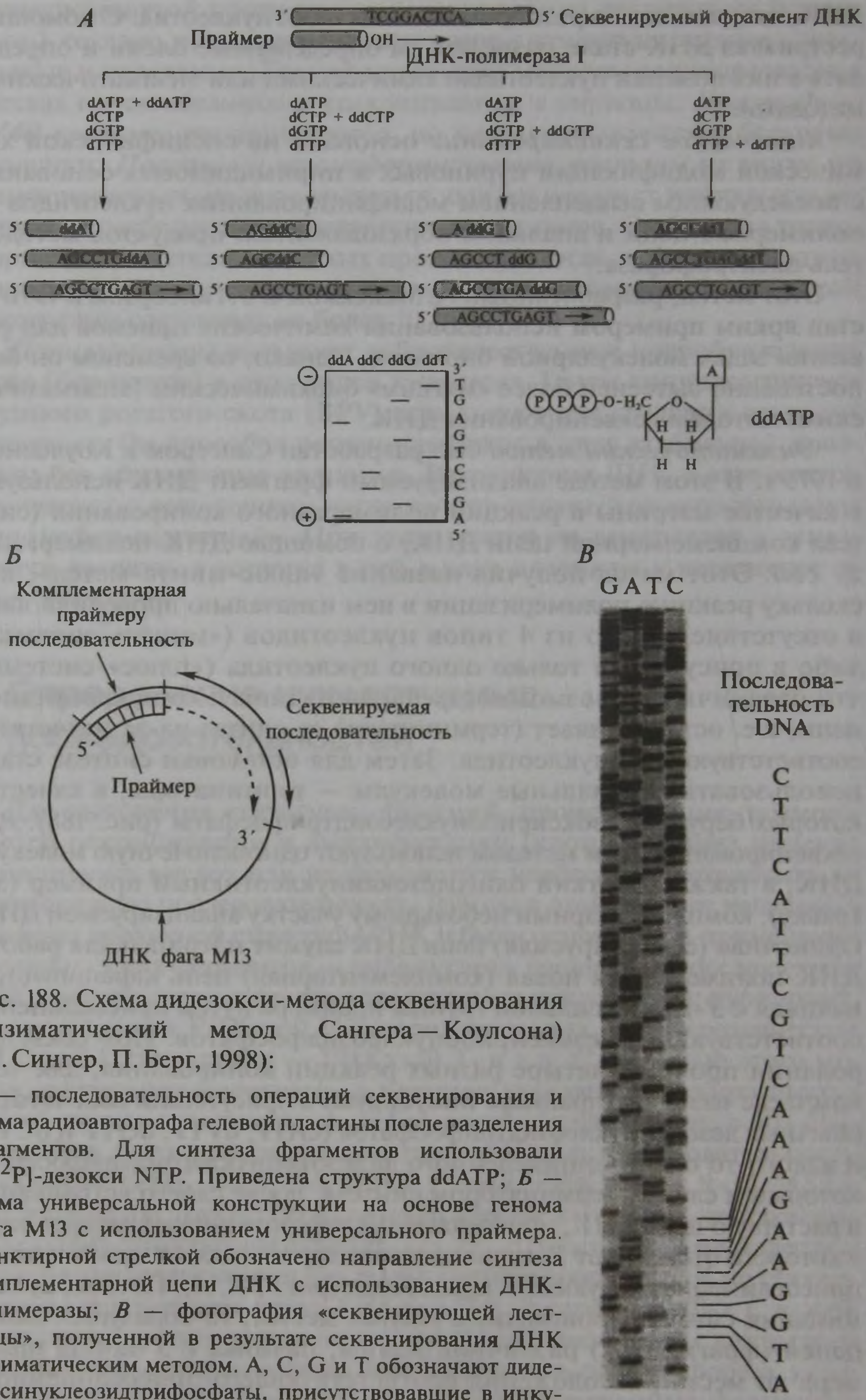


Рис. 188. Схема дидезокси-метода секвенирования (энзиматический метод Сангера — Коулсона) (М. Сингер, П. Берг, 1998):

А — последовательность операций секвенирования и схема радиоавтографа гелевой пластины после разделения фрагментов. Для синтеза фрагментов использовали α[^{32}P]-дезокси NTP. Приведена структура ddATP; *Б* — схема универсальной конструкции на основе генома фага M13 с использованием универсального праймера. Пунктирной стрелкой обозначено направление синтеза комплементарной цепи ДНК с использованием ДНК-полимеразы; *В* — фотография «секвенирующей лестницы», полученной в результате секвенирования ДНК энзиматическим методом. A, C, G и T обозначают дидезоксинуклеозидтрифосфаты, присутствовавшие в инкубационной смеси

радиоактивно меченными, так как для их синтеза используют по крайней мере один меченый дидезоксинуклеозидтрифосфат. Метка вводится в праймер или по α-фосфатной группе дезоксинуклеозидтрифосфата, используемого для наращивания цепи ДНК. При этом может быть использован не только атом фосфора ^{32}P, но и атом серы ^{35}S, заменяющий атом кислорода в фосфорном остатке нуклеотида. Продукты четырех реакций анализируют методом электрофореза, и секвенируемую последовательность считывают с радиоавтографа, так же как при анализе химического секвенирования. Длина выявленного при электрофорезе радиоактивного фрагмента указывает место расположения комплементарного нуклеотида в анализируемом фрагменте ДНК. Например, если при анализе на радиоавтографе пробы, полученной путем копирования в присутствии ddCTP, выявлены фрагменты длиной 3 и 4 нуклеотида, это значит, что в анализируемом фрагменте ДНК в соответствующих позициях расположен гуанин.

В настоящее время для секвенирования ДНК методом Сангера созданы универсальные конструкции (векторы) на основе генома фага M13, в которые встраивают изучаемый фрагмент ДНК (рис. 188, *Б*). В такой конструкции любая встроенная (клонируемая) последовательность всегда ограничена (фланкирована) одними и теми же последовательностями фаговой ДНК, что позволяет использовать универсальный праймер, который нетрудно синтезировать химическими методами (см. ниже). Таким образом, комбинируя клонирование разных фрагментов ДНК и их секвенирование с помошью дидезокситерминатора, были определены последовательности митохондриальной ДНК человека (16 569 н. п.) и ДНК фага λ (48 502 н. п.).

Работа по определению нуклеотидных последовательностей крупных геномов, состоящих из миллионов и более нуклеотидов, сильно облегчается использованием компьютерных программ, которые позволяют быстро находить перекрывающиеся последовательности в анализируемых фрагментах и затем составить полную карту генома.

Методы Максама — Гилберта и Сангера — Коулсона в настоящее время практически полностью автоматизированы. При работе наиболее распространенным методом Сангера в 5′-конец праймера вводят флуоресцентные метки, причем для каждого из 4 анализируемых нуклеотидов используют флуоресцирующие агенты с различными спектральными характеристиками. После электрофореза гель сканируют при различных длинах волн, и полученная информация обрабатывается на компьютере. При этом все биохимические операции также проводятся роботом, и скорость секвенирования превышает 100 000 нуклеотидов в сутки. Такой полностью автоматизированный прием был использован при определении структуры генома человека, что позволило существенно ускорить выполнение этого гигантского проекта.

Химический синтез гена

Синтез гена был впервые осуществлен в 1970 г. в США в лаборатории Х.-Г. Кораны. Этим геном стал цистрон, кодирующий дрожжевую $тРНК^{Ala}$, состоящий из 77 пар нуклеотидов. К тому времени структура $тРНК^{Ala}$ была расшифрована и созданы приемы органического синтеза, позволяющие синтезировать 20-членные олигонуклеотиды. Кораной вначале были синтезированы 15 одноцепочечных олигонуклеотидов, содержащих от 5 до 12 оснований. Синтез осуществляли ступенчато, присоединяя нуклеотиды один за другим. Полученные химическим синтезом фрагменты соединяли между собой с помощью ДНК-лигазы фага Т4. Так как для обеспечения работы этого фермента необходимо наличие ортофосфатной группы в 5′-положении на одном из сшиваемых концов олигонуклеотида, то при синтезе использовали еще один фермент — олигонуклеотидкиназу фага Т4, которая фосфорилировала 5′-ОН группу за счет АТР.

Поскольку лигаза способна сшивать фрагменты цепи ДНК на одной цепи только тогда, когда вторая цепь непрерывна, Х.-Г. Корана и сотрудники вводили синтезированные нуклеотиды в реакционную смесь попарно с таким расчетом, чтобы после взаимодействия комплементарных областей оставались концевые одноцепочечные участки («липкие концы»).

Так были получены три крупные блока олигонуклеотидов, которые с помощью ДНК-лигазы были соединены в целую молекулу гена. Позднее Корана синтезировал ген тирозиновой супрессорной РНК, состоящей из 207 нуклеотидов.

Химический синтез гена (который, строго говоря, является химико-ферментативным) стал крупнейшим достижением биоорганической химии и молекулярной биологии. Он показал и трудоемкость этой работы, однако открыл путь прямого получения необходимых генов для технологии рекомбинантных ДНК и стал важным условием прогресса генетической инженерии.

Метод, разработанный Х.-Г. Кораной, был затем применен для синтеза генов гормонов соматостатина и инсулина. Для синтеза гена соматостатина было использовано восемь (8), а для инсулина — 66(!) коротких одноцепочечных синтетических фрагментов ДНК, которые затем гибридизовали и получали двуцепочечные молекулы, стабилизированные водородными связями между комплементарными нуклеотидами.

Несмотря на то что синтезировать нуклеотидную последовательность целого гена очень трудно, значительную ценность представляет синтез даже их коротких участков длиной 15—20 нуклеотидов, которые пригодны в качестве специфических зондов для гибридизации. Такие зонды используют в саузерн- и нозерн-блоттинге, а также при скрининге рекомбинантных клонов ДНК. Синтезированные in vitro

короткие нуклеотидные последовательности применяют в качестве праймеров при определении нуклеотидных последовательностей по методу Сангера (см. выше). Синтетические сегменты — линкеры, содержащие сайты узнавания рестриктаз, лигируют с фрагментами ДНК, встраиваемыми в векторы, что существенно облегчает манипулирование с клонируемыми генами.

Химический синтез полидезоксирибонуклеотидов в настоящее время практически полностью автоматизирован и проводится на установках, называемых «синтезаторы ДНК».

Достижения и перспективы генетической инженерии

Получение биологически активных соединений. Генетическая инженерия не только открыла возможности для конструирования генов, но и легла в основу новых альтернативных технологий получения биологически активных соединений (гормонов, рилизинг-факторов, регуляторных пептидов, интерферонов и др.). Многие из этих соединений содержатся в биологических объектах в ничтожных количествах, и их препаративное выделение традиционными биохимическими методами оказывается слишком трудоемким, а иногда и невозможным.

Методы генетической инженерии позволяют преодолеть эти трудности путем амплификации соответствующих генов, их клонирования и экспрессии в клетках бактерий, дрожжей и быстро реплицирующихся клетках многоклеточных организмов и таким образом добиться получения значительных количеств физиологически активных молекул, пригодных для практического использования в качестве медицинских препаратов, пищевых продуктов и т.д. В качестве примеров практического использования методов генетической инженерии остановимся на генно-инженерных способах получения пептидных гормонов и интерферонов.

Одним из первых методами генетической инженерии был получен *гормон роста человека* — соматотропный гормон (HGH). В организме он синтезируется клетками гипофиза. Этот гормон состоит из 121 аминокислотного остатка, его мол. масса составляет 20 кДа.

Схема получения HGH с использованием генно-инженерных методов приведена на рис. 189. Отправной точкой для получения гормона роста являлось выделение соответствующей мРНК из клеток гипофиза. Далее осуществляли синтез кДНК, из которой в дальнейшем выщепляли область, кодирующую сигнальный пептид, а вместо нее встраивали триплет ATG, который кодирует аминокислоту метионин и служит сигналом начала трансляции в белоксинтезирующей системе клеток. К полученному таким образом гену присоединяли необходи-

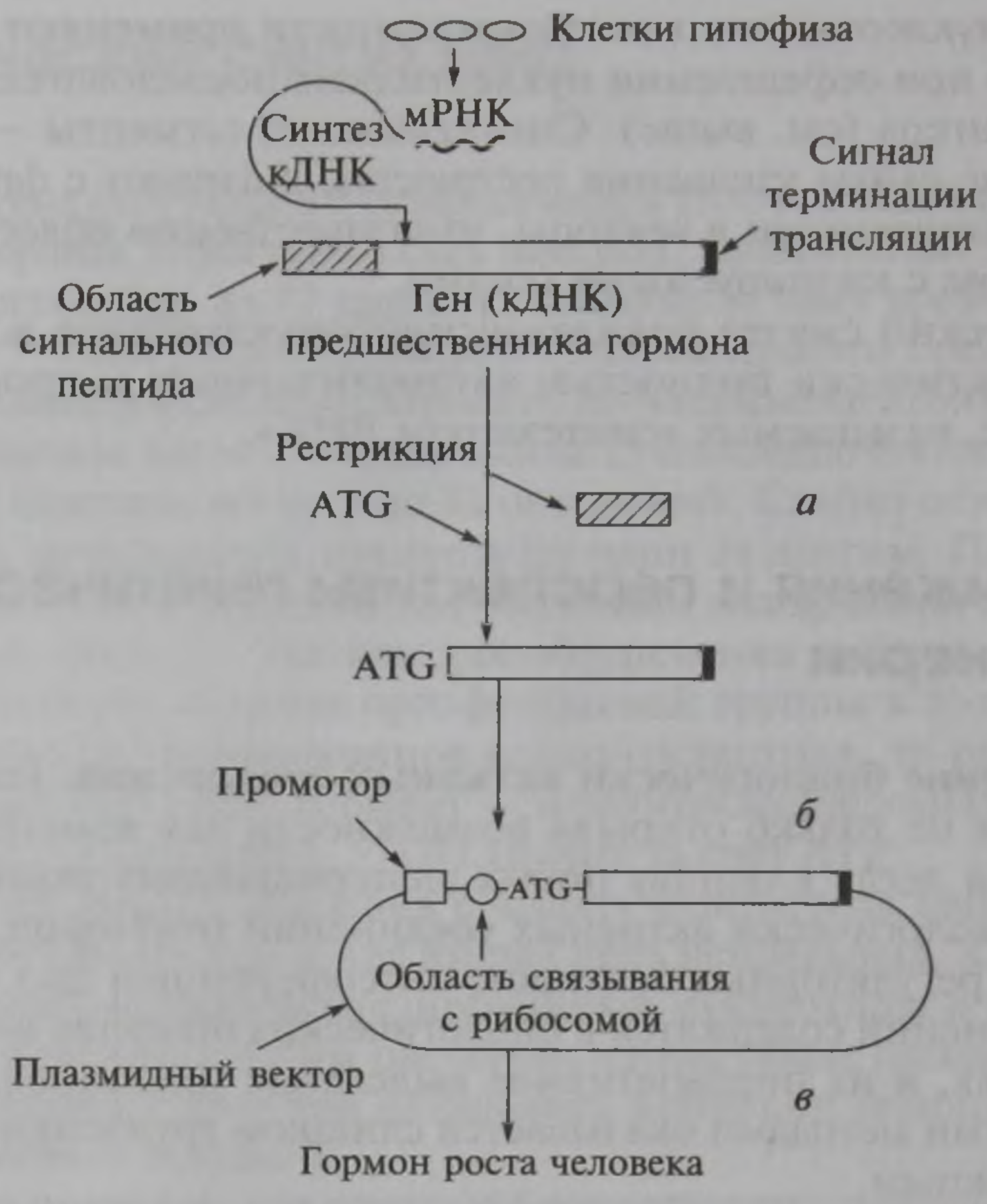


Рис. 189. Схема получения гормона роста человека:

а — выщепление рестриктазой области, кодирующей сигнальный пептид, и присоединение триплета ATG, кодирующего метионин; *б* — присоединение к гену промотора и области, необходимой для связывания мРНК с рибосомой *E. coli*, включение в вектор; *в* — экспрессия гена в клетках бактерии

мые для его эффективной экспрессии регуляторные элементы (промотор и лидерную последовательность, служащую для прикрепления мРНК к рибосоме) и затем включали всю эту конструкцию в соответствующую плазмиду, получая таким образом вектор молекулярного клонирования. Дальнейшая трансформация и клонирование этого вектора в бактериальных клетках, где осуществляется интенсивная репликация, транскрипция и трансляция, позволяет получить необходимое количество гормона, который затем используют не только в медицинских целях, но и в практическом животноводстве, повышая с его помощью интенсивность роста животных.

Получение *соматостатина* представляет собой интересный пример целенаправленного конструирования белков, выполняемого с помощью методов генетической инженерии. Соматостатин представляет собой короткий пептид (содержит 14 аминокислотных остатков), который синтезируется в желудочно-кишечном тракте и тормозит высвобождение из гипофиза гормона роста. Получить соматостатин

в клетках бактерий в значительных количествах первое время не удавалось, так как он быстро разрушался протеолитическими ферментами. Чтобы «обмануть» внутриклеточные бактериальные протеазы, пришлось сконструировать *химерный белок* — искусственный предшественник соматостатина. В этом белке в качестве N-концевого участка использовали легко синтезируемый белок бактерии, а к нему через аминокислоту-связку присоединили сам соматостатин. Все это было сделано на уровне генетического материала (ДНК), который затем клонировали в соответствующем векторе (рис. 190). Для синтеза соматостатина его ген был получен методом химического синтеза, а потом совмещен с бактериальным геном β-галактозидазы, к которому дополнительно подстроили фрагмент бактериального оперона, содер-

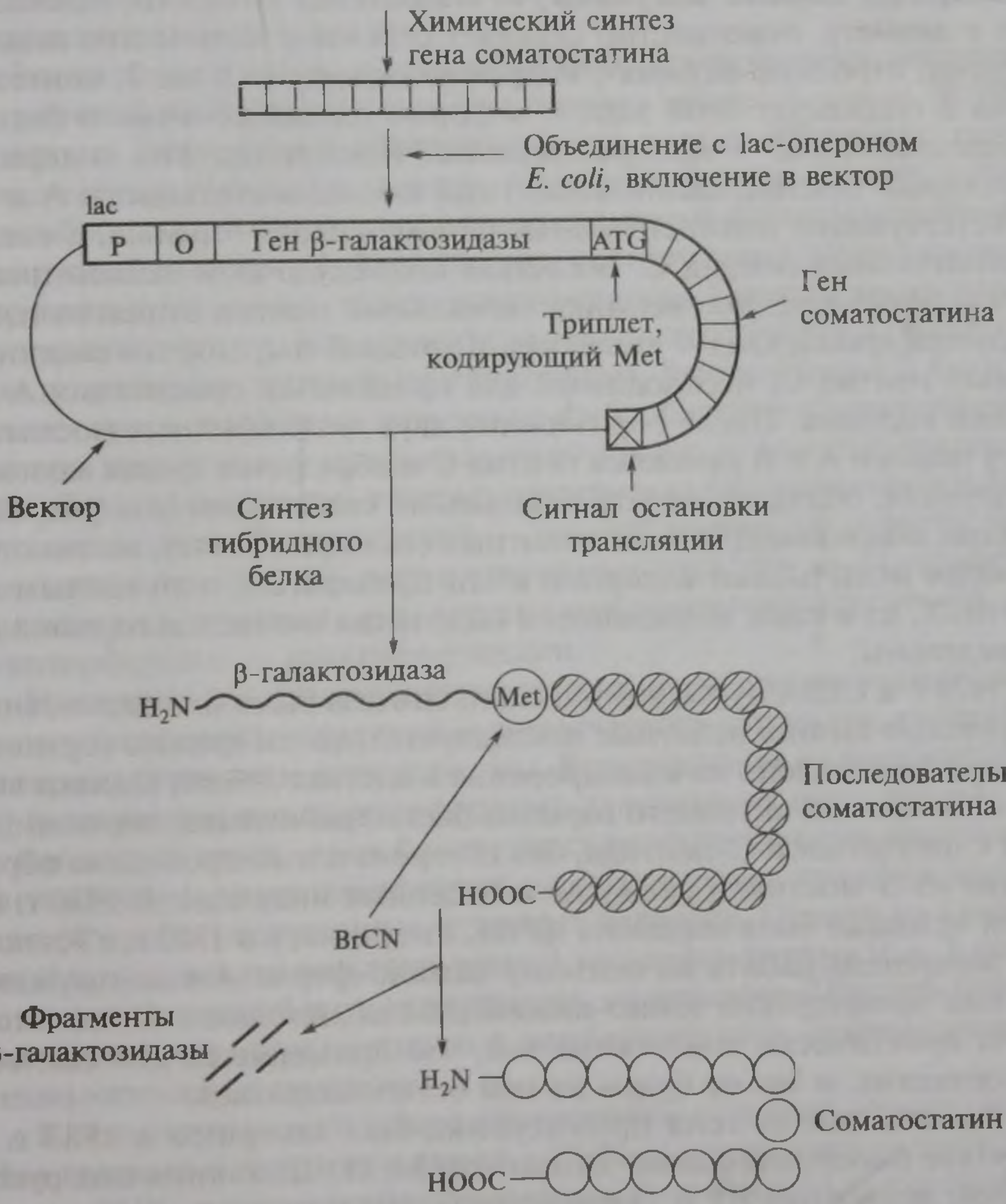


Рис. 190. Схема получения соматостатина. Пояснения см. в тексте

жащий промотор и оператор. В качестве связующего элемента в этой конструкции был использован триплет ATG. Будучи трансформирован в составе плазмидного вектора в бактериальные клетки, этот ген служил источником для синтеза в них химерного белка, содержащего последовательности β-галактозидазы и соматостатина, соединенные через остаток метионина. После выделения из клеток бактерии этот белок обрабатывали бромцианом, в результате чего происходило его расщепление по остаткам метионина и таким образом достигалось вычленение из его состава физиологически активного пептида — соматостатина. Успех работы обеспечивался предварительным анализом первичной структуры соматостатина, в составе которого отсутствуют остатки метионина.

Особое значение для практической медицины имел генно-инженерный ***синтез инсулина***, от недостатка которого, приводящего к диабету, повсеместно страдает огромное количество людей. Инсулин, строение которого подробно освещено в гл. 2, синтезируется в поджелудочной железе в форме одноцепочечного белка-предшественника — *препроинсулина*. Препроинсулин содержит сигнальный пептид, аминокислотные последовательности А и В, соответствующие полипептидным цепям зрелого гормона, а также соединительный пептид С. В клетках поджелудочной железы после синтеза белка-предшественника сигнальный пептид отщепляется и образуется *проинсулин* — молекула, в которой сохраняется соединительный пептид С, необходимый для правильной ориентации А- и В-цепей гормона. После образования двух дисульфидных мостиков между цепями А и В удаляется пептид С и образуется зрелая молекула инсулина, обладающая гормональными свойствами (см. рис. 30). Инсулин можно выделять из животных (свиней и телят), но некоторые люди испытывают аллергию к его препаратам, полученным из животных, да и сами возможности выделения и очистки гормона не беспредельны.

В 1978 г. в США путем химического синтеза были получены гены, кодирующие аминокислотные последовательности зрелого гормона, и далее осуществлено их клонирование в клетках *E. coli*. Однако выход биологически активного гормона был незначителен, вероятно, в связи с отсутствием С-пептида, без которого плохо проходило образование -S-S-мостиков между А- и В-цепями инсулина. В 1980 г. из тканей человека была выделена мРНК инсулина, а в 1982 г. в Канаде была завершена работа по полному химико-ферментативному синтезу гена проинсулина генно-инженерными методами. Схема этой работы практически аналогична той, что применяется для синтеза соматостатина, и мы не будем на ней останавливаться.

В России синтез гена проинсулина был завершен в 1983 г. в Институте биоорганической химии им. М. М. Шемякина под руководством академика Ю. А. Овчинникова. Лекарственные препараты инсулина человека, полученные методами генетической инженерии,

впервые поступили в продажу в Англии (компания *Eli Lilly*). Инсулин стал первым препаратом для врачебной практики, созданным с помощью технологии рекомбинантных ДНК. Бурное развитие этих работ привело к возникновению *генно-инженерной эндокринологии*, в результате чего было осуществлено клонирование генов *рилизинг-факторов* гормона роста, кортикотропина и соматокринина. Введение этих генов в генетический аппарат мышей позволяет регулировать рост животных. Таким образом открываются перспективы *генотерапии* — исправления наследственных дефектов путем введения в геном полноценных генов.

Большие успехи достигнуты в области получения методами генетической инженерии биологически активных пептидов. Так, в нашей стране созданы штаммы *E. coli*, синтезирующие в составе химерных белков нейропептид лей-энкефалин, а также пептидные гормоны брадикинин и ангиотензин.

Значительное место в практическом использовании методов генетической инженерии занимает ***синтез интерферонов***. В 1957 г. Исаакс и Линдерманн обнаружили, что клетки животных, подвергающиеся воздействию вирусов, выделяют в среду фактор, который способен придавать свежим (не зараженным) клеткам устойчивость к вирусной инфекции. Этот фактор, препятствующий размножению вирусов в клетке, был назван интерфероном. Позднее выяснили, что существуют три группы интерферонов: α-интерфероны (образуются при воздействии вирусов на лейкоциты), β-интерфероны (синтезируются при воздействии вирусов на фибробласты) и γ-интерфероны, или так называемые *иммунные интерфероны*, которые продуцируются Т-лимфоцитами в ответ на воздействие бактериальных и вирусных антигенов. Интерфероны — это небольшие по размеру белки, состоящие из 146—166 аминокислотных остатков. α–Интерфероны являются протеинами, не содержащими простетических групп, а β- и γ-интерфероны — гликопротеинами.

Интерфероны представляют высокую ценность для лечения целого ряда заболеваний: вирусного и хронического гепатитов, рассеянного склероза, миеломы, остеосаркомы. Есть указания на положительный эффект применения интерферонов для лечения рака легких, мозга и гортани. В то же время биологические исследования показали, что интерфероны видоспецифичны, т. е. для лечения людей необходимы интерфероны, полученные из клеток человека. Однако из 1 л донорской крови можно получить всего 1 мкг интерферона (т. е. 1 дозу для инъекции), и такой путь, естественно, не приемлем. Решить проблему оказалось возможным только с использованием современных биотехнологических подходов.

Клонирование интерферонов человека — задача весьма сложная. Это связано с тем, что в смеси мРНК, кодирующих разные белки, доля мРНК, кодирующей интерфероны и необходимой для получения кДНК, не превышает 0,1%. Поэтому трудность выделения достаточ-

ных количеств мРНК — первая (но не единственная) трудность при производстве интерферонов. Тем не менее со временем удалось клонировать гены всех трех типов интерферонов, что облегчило параллельно задачу изучения структуры этих белков. Первичные структуры интерферонов были вначале определены на основе анализа нуклеотидных последовательностей их генов, а затем подтверждены прямым секвенированием аминокислотных последовательностей. Затем в ряде стран, в том числе и у нас, были осуществлены работы по получению α-интерферона человека в клетках *E. coli*. Физико-химические свойства интерферона, выделенного из бактерий, оказались близки к свойствам интерферона, выделенного из крови доноров. Далее удалось получить такие штаммы бактерий, из 1 л бактериальной суспезии которых можно выделить 5 мг интерферона (т.е. в 5 000 раз больше, чем из 1 л крови человека). Это позволило наладить в ряде стран, включая Россию, промышленное производство интерферонов с использованием *E. coli*. Гены интерферонов клонировали в клетках дрожжей и высших эукариот, которые способны осуществлять гликозилирование этих белков, необходимое для точного воспроизведения структуры β- и γ-интерферонов.

Японскими исследователями разработана интересная система получения α-интерферона из крови (гемолимфы) гусениц тутового шелкопряда, пораженных вирусом ядерного полиэдроза. В этой работе ген белка вируса (полиэдрина) клонировали в составе плазмиды, а затем генно-инженерными методами конструировали рекомбинантную плазмиду, несущую ген интерферона, примыкающий к промотору полиэдринового гена. Эту плазмиду использовали для заражения культуры клеток тутового шелкопряда. Далее из клеток выделяли вирусные частицы, содержащие ген интерферона, и реплицировали их в личинках тутового шелкопряда, у которых по мере развития вирусной инфекции интерферон накапливался в гемолимфе (~50 мкг на 1 мл гемолимфы). Ниже приведены некоторые белки человека, полученные генно-инженерными методами (Б. Глик, Дж. Пастернак, 2002):

Белок	**Применение в медицине**
Адренокортикотропный гормон	Ревматизм
α_1-Антитрипсин	Эмфизема
Гемоглобин	Анемия
Гормон роста (соматотропин)	Задержка роста
Инсулин	Сахарный диабет
Интерлейкины	Злокачественное новообразование, имунные заболевания
Интерфероны (α, β, γ)	Вирусные заболевания, злокачественное новообразование
Кальцитонин	Остеомаляция
Лимфотоксин	Злокачественное новообразование

Релаксин	Роды
Рецептор интерлейкина-1	Астма, ревматоидный артрит
Соматолиберин	Задержка роста
Соматомедин С	Задержка роста
Сывороточный альбумин	Дефицит белков плазмы
Тиреотропный гормон	Рак щитовидной железы
Тканевой активатор плазминогена	Тромбообразование
Тромбоцитарный фактор роста	Атеросклероз
Урогастрон	Язвы
Урокиназа	Тромбообразование
Фактор, активирующий макрофаги	Злокачественное новообразование
Фактор некроза опухоли	Злокачественное новообразование
Фактор роста нервов	Повреждение нервной ткани
Фактор роста эпидермиса	Ожоги
Фактор VIII	Гемофилия
Фактор IX	Гемофилия
Факторы роста В-лимфоцитов	Иммунные заболевания
Хорионический гонадотропин	Женское бесплодие
Эндорфины и энкефалины	Боль
Эритропоэтин	Анемия, заболевания почек

Генетическая трансформация. Методами генетической инженерии удается не только получать отдельные генные продукты, но и направленно конструировать клетки и даже целые организмы животных. Первая такая попытка была сделана американским исследователем Дж. Гордоном в 1980 г.: методом микроинъекции в оплодотворенную яйцеклетку мыши была введена рекомбинантная плазмида pBR322, которая содержала ген тимидинкиназы вируса герпеса и фрагмент генома вируса SV40, вызывающего опухолевую трансформацию. В дальнейшем в геном мыши и других животных удалось ввести гены интерферонов, гормона роста человека и ряд других. Такие животные получили название *трансгенных*. Однако получение трансгенных животных (*генетическая трансформация*) представляет собой сложный и неоднозначный процесс, прежде всего потому, что интеграция чужеродных генов носит случайный характер и, как следствие, могут повреждаться соседние гены реципиента, расположенные вблизи участка внедрения чужеродной ДНК.

Регуляция экспрессии встраиваемых генов также представляет собой сложную проблему. Обычно экспрессия таких генов имеет неоднозначный характер (гены могут «бесконтрольно» включаться в разных тканях на разных стадиях развития трансгенного организма). Чтобы этого избежать, вводимые гены совмещают со специфическими *регулируемыми промоторами*, которые помогают по мере необходимости запускать экспрессию (транскрипцию) этих генов. К числу таких промоторов относится промотор гена металлотионеина 1 — белка, способного связывать ионы тяжелых металлов (см. гл. 2). Индукция биосинтеза металлотионеинов происходит в

тканях животных при поступлении с пищей повышенного количества солей тяжелых металлов. Учитывая это, в опытах по получению трансгенных животных промотор металлотионеина ($P_{мт-1}$) совмещали с геном гормона роста и вводили в плазмиду pBR322, которую затем инъецировали в оплодотворенную яйцеклетку мыши. Если затем родившимся трансгенным мышам добавляли в пищу цинк, то вследствие индукции транскрипции на промоторе металлотионеина начинался усиленный синтез гормона роста, что приводило к резкому (двукратному) увеличению размеров и массы тела мышей.

Введение соответствующих генетических конструкций в ДНК оплодотворенных яйцеклеток сельскохозяйственных животных (коров, коз, овец, свиней) и птиц стимулирует их рост и повышает выход продуктов питания. Получены популяции трансгенных рыб (лососей), в которых масса тела годовалых особей в 10 раз превышает таковую у обычных рыб. Кроме того, генетически трансформированные животные могут быть использованы как «биофабрики» по производству белков человека. Молоко трансгенных коров может служить, например, источником для получения белков человека, необходимых для лечения ряда заболеваний. Количество молока, полученное в течение года от 20 трансгенных коров, достаточно для полного удовлетворения потребностей медицины в белке С, предотвращающем тромбообразование. С другой стороны, одной трансгенной коровы будет достаточно для выделения из ее молока необходимого количества фактора IX (фактор Кристмаса) — компонента каскадного механизма свертывания крови, который вводят больным гемофилией. Получены линии трансгенных свиней, в крови которых накапливается гемоглобин человека. Созданы трансгенные свиньи, у которых продукты экспрессии генов ингибируют процесс отторжения органов животных при их пересадке приматам, что открывает перспективы успешной трансплантации органов животных в организм человека.

Методы генетической инженерии можно использовать для исправления дефектных генов в *обратной генетике*. В этом направлении начинают работу с выделения из клеток нужного белка. Затем получают ген этого белка, амплифицируют его, определяют нуклеотидную последовательность и видоизменяют ее, создавая мутантный ген. Если полученный таким образом in vitro ген ввести в оплодотворенную яйцеклетку, то можно получить мутантный (трансгенный) организм, что стало обычной процедурой в отношении не только мышей, но и других многоклеточных тест-объектов: дрозофилы, кроликов и др. Введенный ген может встроиться в хромосому в результате гомологической рекомбинации и стать, таким образом, постоянным элементом генома. Если этот ген транскрибируется, то в клетках будет синтезироваться измененный белок. При введении мутантных генов в клетки бактерий и дрожжей, которые, как правило, гаплоидны, эти гены достаточно часто рекомбинируют с нормальными гомологами,

и далее все клетки синтезируют мутантный белок. У высших (диплоидных) организмов вставка (*инсерция*) имеет случайный характер и обычно в геноме имеются два гена (нормальный и мутантный), а замена гена обычно происходит не чаще 1 раза на 1 000 опытов. Чтобы «вывести из строя» той или иной ген (в принципе любой ген, в том числе и тот, экспрессия которого по какой-то причине не желательна), применяют целенаправленно полученные искусственные гены, кодирующие *антисмысловые РНК*. С помощью таких генов можно создавать доминантные мутации. При одновременной экспрессии двух генов — «нормального» и мутантного — будет происходить синтез двух видов мРНК — нормальной и антисмысловой РНК. В результате их гибридизации биосинтез белка будет блокирован и клетка перестанет испытывать воздействие «нежелательного» гена.

Получение трансгенных растений. Этот метод имеет огромное научно-практическое значение прежде всего для увеличения пищевых ресурсов. По данным ЮНЕСКО, к 2050 г. население Земли может удвоиться и будет составлять 11 млрд. Традиционное сельскохозяйственное производство не в состоянии будет обеспечить соответствующий рост продовольствия, поэтому и основные надежды в этом отношении связаны с новыми генно-инженерными биотехнологиями. Традиционные методы селекции и методика слияния протопластов, хотя и позволяют получать новые генотипы растений, но требуют много времени и далеко не всегда приводят к улучшению хозяйственно-ценных признаков. В то же время важным преимуществом растений по сравнению с животными является способность их клеток и протопластов развиваться в целое растений — *тотипотентность*, что, естественно, используется в работах по получению трансгенных растений.

С использованием приемов генетической инженерии генетическая трансформация растений может осуществляться *векторным способом* (с использованием агробактерий и вирусов) либо путем *прямого переноса генов*. Начиная с середины 80-х г. XX в. создаются векторы на базе грамотрицательных почвенных бактерий *Agrobacterium tumefaciens*. Агробактерии обладают способностью интегрировать свой генетический материал в клетки двудольных растений. Они содержат крупные (~200 кб) Ti-плазмиды (от англ. *tumor inducing* — вызывающая опухоль), несущие так называемую Т-ДНК (от англ. *transferred* — передаваемая). Т-ДНК представляет собой определенную область плазмидной ДНК (~20 кб), способную включаться в ядерный геном растительной клетки. В Т-ДНК содержатся гены фитогормонов, продукты экспрессии которых индуцируют у трансформированных растений неконтролируемый рост недифференцированных клеток и вызывают образование опухолей — корончатых галлов. В Т-ДНК содержатся также гены, кодирующие опины, — необычные пептиды, которые используются бактериями в качестве источника азота и углерода. Гены, ответственные за ин-

дукцию опухоли, синтез опинов и подавление дифференцировки клеток поражаемого растения, расположены близко друг от друга в Т-ДНК-плазмиде. В то же время гены, ответственные за перенос и интеграцию Т-ДНК в хромосомы растений, находятся в другой области Ti-плазмиды — vir-области (рис. 191, *А*).

Ti-плазмиды представляют собой очень удобные естественные векторы, поскольку их Т-ДНК ограничена прямыми повторяющимися последовательностями длиной 25 н. п., между которыми можно встраивать нужный ген, и он подобно Т-ДНК будет переноситься в клетки растений. Трансгенные растения получают, используя модифицированную Т-ДНК, в которой онкогены заменены на нужный ген.

В связи с тем что большие размеры Ti-плазмид затрудняют клонирование, их используют в *бинарной системе*, добиваясь реком-

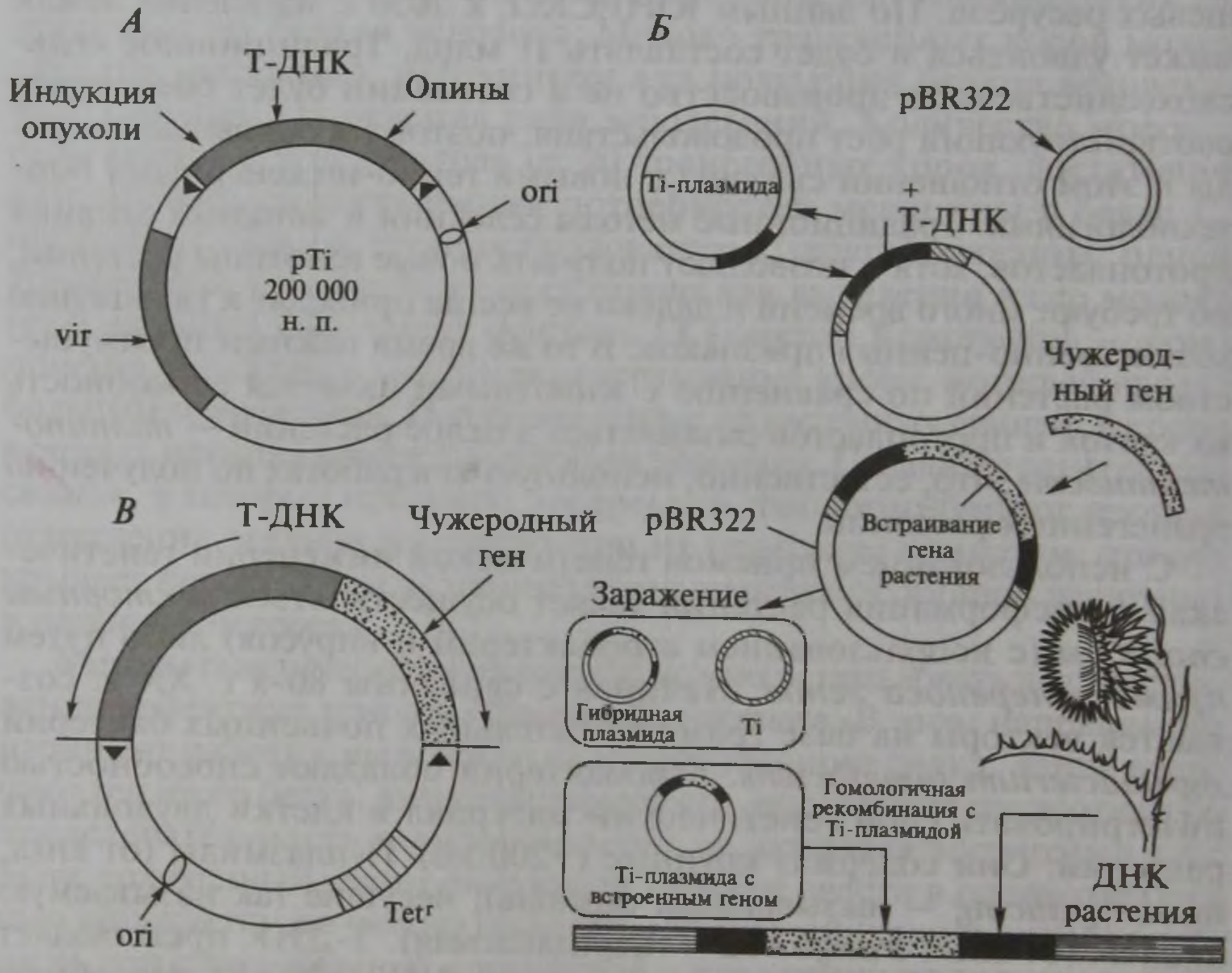


Рис. 191. Использование Ti-плазмиды в качестве вектора для переноса генов в растительные клетки (по Я. И. Бурьянову, 1999):

А — строение Ti-плазмиды: vir — область вирулентности; треугольниками показаны прямые повторы длиной 25 н.п., между которыми можно вставлять нужный ген; ori — точка начала репликации; *Б* — получение челночного вектора, служащего для переноса генов в растительные клетки; *В* — челночный вектор на основе Т-ДНК и К-плазмиды *E. coli*; Tet^r — ген устойчивости к тетрациклину; ori — точка начала репликации плазмиды с широким кругом хозяев

бинации области Т-ДНК с геномами других, небольших плазмид (типа pBR322), способных реплицироваться как в агробактериях, так и в клетках *E. coli*. Таким образом получают векторы, способные эффективно трансформировать геном клеток высших растений (рис. 191, *Б*, *В*).

Вирусные векторы получают на основе фитовирусов, большинство которых содержат одноцепочечную РНК, кодирующую белки, необходимые для репликации и транспорта вируса в растения. В генной инженерии часто используют сильные вирусные промоторы (например, 35S-промотор вируса мозаики цветной капусты). Сильные промоторы растений можно использовать для экспрессии чужеродных генов и в других векторных системах. Вышеназванный промотор вируса мозаики цветной капусты не обладает тканевой специфичностью и активен не только в клетках крестоцветных, но и в клетках растений других семейств.

Безвекторные системы дают возможность прямого переноса чужеродной ДНК в растительные протопласты. В этом случае используют методы электропорации, микроинъекции и баллистический метод.

Электропорация — кратковременные электрические разряды (1—100 мкс при напряжении 1 000—10 000 В/см) увеличивают проницаемость мембран протопластов, что облегчает проникновение чужеродной ДНК. Этим методом были получены трансгенные растения кукурузы, риса и сахарного тростника. *Микроинъекцию* проводят под микроскопом с помощью стеклянной иглы, вводя через нее ДНК в ядро клетки. Таким способом были получены трансгенные растения ячменя и капусты. *Баллистический метод* является на сегодняшний день одним из самых эффективных для трансформации однодольных растений. В этом методе на частицы вольфрама диаметром 0,6—1,2 мкм напыляется ДНК вектора, содержащая необходимую для трансформации генетическую конструкцию. Затем эти частицы наносят на целлофановую подложку и помещают внутрь баллистической пушки, принцип действия которой состоит в изменении давления. В момент сбрасывания давления вольфрамовые частицы с большой скоростью выбрасываются из баллистической пушки и проникают в цитоплазму и ядро трансформируемых клеток. Этим методом удалось получить растения-трансформанты риса, ячменя, пшеницы и кукурузы.

Прямой перенос ДНК в растительные протопласты может проводиться с использованием липосом — сферических частиц, мембраны которых содержат фосфолипиды. Частицы обволакивают трансформирующую ДНК и тем самым защищают ее от нуклеаз. Липосомы можно создать путем встряхивания или обработки ультразвуком водных эмульсий фосфолипидов: фосфатидилсерина, лецитина и др. С помощью липосом в протопласты растений удается ввести ДНК Ti-плазмид, а также целые метафазные хромосомы.

При повышении эффективности выращивания культурных растений реализуются два взаимодополняющих момента: улучшение пищевых (или технологических) качеств, а также повышение устойчивости к биотическим и абиотическим факторам внешней среды. Так, у злаков наибольшую ценность представляют белки, запасаемые в эндосперме, которые, однако, имеют несбалансированный для питания человека и животных состав: их запасные белки бедны лизином, триптофаном и метионином. Мало метионина и в запасных белках бобовых. Введение дополнительных кодонов лизина в гены запасных белков (проламинов) злаков могло бы скорректировать аминокислотный состав зерна. В настоящее время охарактеризованы десятки генов запасных белков злаков, бобовых и некоторых других растений; клонированы гены запасных белков (гордеинов) ячменя, глиадинов и глютенина пшеницы, зеинов кукурузы и др. Эксперименты по трансформации пшеницы модифицированным геном, кодирующим одну из субъединиц запасного белка глютенина, привели к активному синтезу модифицированного белка, что способствовало улучшению хлебопекарного качества пшеничной муки.

Использование генно-инженерных подходов позволяет повысить резистентность растений к насекомым-вредителям. Известно, что бактерии *Bacillus thurengiensis* выделяют белок-прототоксин, который в кишечном тракте насекомых расщепляется их протеазами до активного токсина, убивающего вредителей. Трансформация генами эндотоксина, выделенными из генома *B. thurengiensis*, растений томата, вызывает у растений защитный эффект, сопоставимый с использованием инсектицидных препаратов.

В ответ на инфицирование вирусами, бактериями и грибами у растений индуцируется синтез белков-ферментов, называемых PRP (pathogen related proteins), которые способны деструктировать оболочки патогенов. С их помощью были получены трансгенные растения табака и турнепса, устойчивые к фитопатогенам. Антивирусный эффект у растений можно вызывать путем трансформирования в их геномы генов человеческих интерферонов. Трансформация турнепса соответствующей векторной конструкцией приводит к продуцированию интерферона в клетках растения и повышению их устойчивости к вирусу желтой мозаики. Применяя эту методику, удается также повысить устойчивость картофеля и табака к вирусным заболеваниям.

Существенное практическое значение имеют работы в области получения трансгенных растений, устойчивых к гербицидам, которые, уничтожая сорняки, угнетающе действуют и на посевы сельскохозяйственных культур. В большинстве случаев гербициды отрицательно воздействуют на фотосинтез и биосинтез аминокислот. Ряд гербицидов (паракват, ацифлуорфен и др.) ингибируют фотосинтез путем повышения содержания активных форм (радикалов) кислорода. Известно, что нейтрализация (дисмутация) супероксиданиона ($O_2^{\cdot}$)

катализируется ферментом супероксиддисмутазой, кодируемой геном SOD. С целью защиты от гербицидов была получена кДНК гена SOD из табака и введена в составе вектора на основе Ti-плазмиды в геном люцерны, что резко повысило устойчивость этого растения к вышеназванным гербицидам.

Как следует из вышеизложенного, генетическая инженерия создает основу для современной биотехнологии, которая в основном базируется на использовании живых клеток микроорганизмов, дрожжей и, реже, на культурах клеток животных и высших растений. В биотехнологических работах принципиально важное значение имеют три момента оперирования с генами: возможность осуществлять рекомбинацию, репликацию и эффективную экспрессию (транскрипцию и трансляцию) генетического материала. С появлением генетической инженерии исследователи научились выделять индивидуальные гены, видоизменять их вне клеток и переносить с помощью векторов в удобные для амплификации клетки, которые способны как «микробиофабрики» синтезировать белки, кодируемые модифицированными генами. Таким образом, вначале удалось сделать полностью контролируемым процесс рекомбинации, вынести его за пределы клетки, т. е. придать ему внеклеточный уровень. Затем, после создания метода полимеразной цепной реакции, процесс репликации (амплификации) генов также в значительной мере стал контролируемым и доступным для выполнения на внеклеточном уровне. Однако последний и решающий этап получения функционально активных белковых продуктов — транскрипция генов и трансляция мРНК — остался связанным с жизнедеятельностью клетки. Это обстоятельство создает большие сложности при получении значительных количеств чистых (гомогенных) белковых продуктов. Если учесть при этом и другие трудности, возникающие в работе традиционными методами генетической инженерии (прежде всего к ним относится трудность выделения достаточных количеств индивидуальных мРНК), то становится ясным, что единственным путем повышения эффективности наработки белковых продуктов может стать переход к полностью внеклеточным технологиям белкового синтеза. Теоретическое обоснование такого перехода и его научно-практическое обеспечение были впервые сделаны в нашей стране в лаборатории акад. А. С. Спирина в начале 90-х годов XX в. Речь идет о том, чтобы отказаться от живых клеток, заменив их биореакторами, способными осуществлять избирательный синтез нужных белков. Сами по себе бесклеточные системы в молекулярной биологии известны достаточно давно (см. гл. 1) и с успехом применяются в научных целях, однако они работают медленно и с очень низким выходом белкового продукта. Центральная идея А. С. Спирина и его соавторов состоит в том, чтобы использовать в качестве исходного материала биотехнологии биосинтеза белков не ДНК, а РНК. Для этого необходимо иметь возможность быстро амплифицировать индивидуальные мРНК в бесклеточной

системе. Сделать это можно с использованием особого фермента — РНК-зависимой РНК-полимеразы — *репликазы*, содержащейся в бактериофаге Q_β. Репликаза Q_β — единственная фаговая репликаза, которую удалось очистить и получить в стабильном состоянии и которая способна быстро наращивать количество цепей РНК (теоретически из 1 молекулы РНК за 100 мин можно получить 10^{12} копий). При исследовании структуры реплицируемых этим ферментом РНК оказалось, что РНК фага и РНК бактерии способны к рекомбинации. Отсюда появляется перспектива целенаправленного получения рекомбинантных РНК на базе репликации с репликазой Q_β.

Таким образом, рассматриваемая система, вероятно, может включать все три необходимых биотехнологических компонента: репликацию, рекомбинацию и интенсивную экспрессию в бесклеточной системе. Именно в нашей стране учеными под руководством А. С. Спирина были созданы первые биореакторы непрерывного действия, в которых осуществляется внеклеточный синтез белков. С использованием только определенных мРНК удается получать высокоочищенные белковый гормон кальцитоцин и интерфероны человека.

СПИСОК ЛИТЕРАТУРЫ

Албертс Б. Молекулярная биология клетки. / [Б. Албертс и др.]. — М.: Мир, 1994.

Алмазов В. П. p53-инструмент для терапии злокачественных заболеваний человека / [В. П. Алмазов и др.] // Молекулярная биология. — 2007. — Т. 41. — Вып. 6. — С. 947—963.

Баранов В. С. Молекулярная медицина: молекулярная диагностика, превентивная медицина и генная терапия // Молекулярная биология. — 2000. — Т. 34. — Вып. 4. — С. 684—695.

Велц Р. Моделирование и синтез рибозимов / [Р. Велц и др.] // Молекулярная биология. — 2000. — Т. 34. — Вып. 6. — С. 1090—1096.

Георгиев П. Г. Различные механизмы регуляции длины теломер / [П. Г. Георгиев и др.] // Молекулярная биология. — 2000. — Т. 34. — № 5. — С. 743—752.

Глик Б. Молекулярная биотехнология. Принципы и применение / Б. Глик, Дж. Пастернак ; под ред. Н. К. Янковского. — М.: Мир, 2002.

Глякина А. В. Поиск структурных факторов, ответственных за стабильность белков из термофильных организмов / [А. В. Глякина и др.] // Молекулярная биология. — 2007. — Т. 41. — Вып. 4. — С. 681—687.

Зверева М. Э. Структура и функция тмРНК (10Sa РНК) / [М. Э. Зверева и др.] // Молекулярная биология. — 2000. — Т. 34. — Вып. 6. — С. 1081—1089.

Зеленин А. В. Введение в геномику растений / [А. В. Зеленин и др.] // Молекулярная биология. — 2000. — Т. 35. — Вып. 3. — С. 339—348.

Зенгер В. Принципы структурной организации нуклеиновых кислот. — М.: Мир, 1987.

Жимулёв И. Ф. Общая и молекулярная генетика. — Новосибирск: Изд-во Сибирского ун-та, 2003. — 479 с.

Киселев Л. Л. Геном человека и биология XXI века // Вестник РАН 2000. — Т. 70. — Вып. 5. — С. 412—424.

Кларк Д. Молекулярная биология: простой и занимательный подход / Д. Кларк, Л. Рассел ; пер. с англ. — М.: ЗАО «Компания КОНД», 2004. — 472с.

Коничев А. С. Основные термины молекулярной биологии / А. С. Коничев, Г. А. Севастьянова. — М.: КолосС, 2006. — 188 с.

Коничев А. С. Биохимия и молекулярная биология: Словарь терминов / А. С. Коничев, Г. А. Севастьянова. — М.: Дрофа, 2008. — 359 с.

Лебедев Ю. Б. Эндогенные ретровирусы: возможная роль в функционировании клеток человека // Молекулярная биология. — 2000. — Т. 34. — Вып. 4. — С. 635—645.

Макарова Ю. А. Некодирующие РНК / Ю. А. Макарова, Д. А. Крамеров // Биохимия. — 2007. — Т. 72. — Вып. 11. — С. 1427—1448.

Молекулярная биология: структура и биосинтез нуклеиновых кислот / под ред. А. С. Спирина. — М.: Высшая школа, 1990.

Мушкамбаров Н. Н. Молекулярная биология / Н. Н. Мушкамбаров, С. Л. Кузнецов. — М.: ООО «Медицинское информационное агентство», 2003.

Мюльберг А. А. Фолдинг белка — СПб.: Изд-во С.-Пб. ун-та, 2004.

Овчинников Ю. А. Биоорганическая химия. — М.: Просвещение, 1987.

Патрушев Л. И. Искусственные генетические системы. — Т. 1. Генная и белковая инженерия. — М.: Наука, 2004. — 526с.

Примроус С. Геномика. Роль в медицине / С. Примроус, Р. Твайман. — М.: Бином. Лаборатория знаний, 2008. — 277с.

Прохорчук А. В. Метилирование генома и его роль в функционировании эукаритического организма / А. В. Прохорчук, А. С. Рузов // Генетика. — 2000. — Т. 36. — Вып. 11. — С. 1475—1486.

Разин С. В. Хроматин: упакованный геном / С. В. Разин, А. А. Быстрицкий. — М.: Бином, 2009.

Ридли М. Геном: автобиография вида в 23 главах. — М.: Эксмо, 2008. — 432с.

Рис Э. От клеток к атомам. Иллюстрированное введение в молекулярную биологию / Э. Рис, М. Стернберг. — М.: Мир, 1988.

Рогаев Е. И. Микро-РНК в норме и патологии / [Е. И. Рогаев и др.] // Молекулярная биология. — М., 2008. — Т. 42. — № 5. — С.751—764.

Роллер Э. Открытие основных законов жизни. — М.: Мир, 1978.

Рысков А. П. Мультилокусный ДНК-фингерпринтинг в генетико-популяционных исследованиях биоразнообразия // Молекулярная биология. — 1999. — Т. 33. — Вып. 6. — С. 997—1011.

Рыхлевская А. И. Моно- и бифункциональные ДНК-гликозилазы, участвующие в репарации окислительных повреждений в ДНК / А. И. Рыхлевская, С. А. Кузнецова // Молекулярная биология. — 2000. — Т. 34. — Вып. 6 — С. 1007—1024.

Самуилов В. Д. Программируемая клеточная смерть / [В. Д. Самуилов и др.] // Биохимия. — 2000. — Т. 65. — Вып. 8. — С. 1029—1046.

Свердлов Е. Д. Взгляд на жизнь через окно генома. — М.: Наука, 2009. — Т. 1: Очерки структурной молекулярной генетики. — 2009. — 525 с.

Сингер М. Гены и геномы / М. Сингер, П. Берг. — М.: Мир, 1998.

Скулачев В. П. Старение организма — особая биологическая функция, а не результат поломки сложной живой системы: биохимическое обоснование гипотезы Вейсмана // Биохимия. — М., 1997. — Т. 62. — Вып. 11. — С. 1394—1399.

Спирин А. С. Молекулярная биология. Рибосомы и биосинтез белка — М.: Издательский центр «Академия», 2011.

Спирин А. С. Рибонуклеиновые кислоты как центральное звено живой материи // Вестник Российской Академии наук. — 2003. — Т. 73. — № 2. — С. 117—127.

Тарантул В. З. Геном человека: энциклопедия, написанная четырьмя буквами. — М.: Языки славянской культуры, 2003. — 392 с.

Франк-Каменецкий М. Д. Век ДНК. — М., 2004.

Чемерис А. В. Секвенирование ДНК / [А. В. Чемерис и др.]. — М.: Наука, 1999.

Чумаков С. П. Организация и регуляция ядерно-цитоплазматического транспорта / С. П. Чумаков, В. С. Прасолов // Молекулярная биология. — 2010. — Т. 44. — Вып. 2. — С. 211—228.

Шульц Г. Принципы структурной организации белков / Г. Шульц, Р. Ширмер. — М.: Мир, 1982.

Sanger F., Nicklen S., Coulson A. R. DNA sequencing with chain — terminating inhibitors. Proc. Natl. Acad. Sci. USA. V. 74: 5463—5467.

Watson J. D. DNA.The secret of life. WilliamHememann. London, 2003.

СОДЕРЖАНИЕ

Предисловие 3
Введение 5

Глава 1. Методы молекулярной биологии 12

Глава 2. Белки 22
Аминокислотный состав белков 23
Пептиды 28
Структурная организация белков 32

Глава 3. Нуклеиновые кислоты 74
Первичная структура нуклеиновых кислот 74
Макромолекулярная структура ДНК 83
Полиморфизм двойной спирали 88
Разнообразие форм ДНК 95
Сверхспирализация ДНК. Топоизомеразы 96
Структура и функции РНК 100
Макромолекулярная структура РНК 102
Транспортные РНК 104
Рибосомные РНК 106
Матричные РНК 108
Гетерогенная ядерная РНК 109
Малые ядерные РНК 110
Малые цитоплазматические РНК 110
Концепция «Мир РНК» 111

Глава 4. Структура генома вирусов и фагов 116
Типы генетического материала и механизм его репликации у различных вирусов 118
Типы взаимодействия вируса с клеткой-хозяином 121
Характеристика некоторых вирусов 124
Происхождение вирусов и их роль в эволюции 134

Глава 5. Геном прокариот 136
Структура бактериальной хромосомы 136
Структура прокариотических генов 141

Бактериальные плазмиды 144
IS-элементы и транспозоны бактерий 147

Глава 6. Структура генома эукариот 151

Кинетика реассоциации денатурированной ДНК и сложность генома эукариот 152
Последовательности нуклеотидов эукариотического генома 155
Структура эукариотических генов 157
Гены, кодирующие белки 158
Регуляторные элементы генов, кодирующих белки 161
Рибосомные гены 163
Гены тРНК 166
Гистоновые гены 167
Тандемные повторы 168
Мини- и микросателлиты. ДНК-фингерпринтинг 170
Онкогены и антионкогены 174
Подвижные генетические элементы эукариот 176
Программа «Геном человека» 186
Картирование генома человека 186
Генетическое картирование 187
Структура генома человека (по данным секвенирования на 2001 — 2003 гг.) 194
Геномы органелл эукариот: ДНК митохондрий и хлоропластов 200
Геномы митохондрий 201
Репликация митохондриальной ДНК 203
Полиморфизм митохондриальной ДНК и эволюция человека 204
ДНК хлоропластов 205
Происхождение ДНК органелл 206

Глава 7. Репликация ДНК 207

Белки и ферменты, участвующие в репликации ДНК 208
Репликация хромосомы *E. coli* 216
Репликация хромосом у эукариот 225
Биосинтез ДНК на РНК-матрице (обратная транскрипция) 236

Глава 8. Генетическая рекомбинация 240

Глава 9. Транскрипция 246

Транскрипция у прокариот 249
Регуляция транскрипции 254
Регуляции транскрипции у прокариот 254
Регуляция транскрипции у бактериофага λ 261
Транскрипция у эукариот 264
Хроматин и общая (тотальная) регуляция транскрипции у эукариот 274

Глава 10. Процессинг РНК 282
Процессинг у прокариот 283
Процессинг тРНК и рРНК у эукариот 286
Процессинг мРНК у эукариот 292
Глава 11. Биосинтез белка 300
Генетический код 300
Активация аминокислот 302
Этапы трансляции 313
Регуляция трансляции 319
Перепрограммирование трансляции 326
Глава 12. Репарация ДНК 333
Глава 13. Программируемая клеточная смерть (апоптоз) 346
Глава 14. Генетическая инженерия 354
Методы генетической инженерии (технология получения рекомбинантных ДНК) 354
Рестрикция ДНК 355
Гибридизация нуклеиновых кислот 363
Полимеразная цепная реакция и другие методы амплификации нуклеиновых кислот 366
Клонирование ДНК 370
Определение нуклеотидных последовательностей 376
Химический синтез гена 380
Достижения и перспективы генетической инженерии 381
Список литературы 395

Учебное издание

Коничев Александр Сергеевич,
Севастьянова Галина Андреевна

Молекулярная биология

Учебник

Технический редактор *Е. Ф. Коржуева*. Компьютерная верстка:
Н. В. Протасова. Корректор *А. Б. Глазкова*

Изд. № 104116090. Подписано в печать 28.06.2012. Формат 60×90/16.
Гарнитура «Таймс». Печать офсетная. Бумага офс. № 1. Усл. печ.л. 25,0.
Тираж 1 000 экз. Заказ № 1279.

ООО «Издательский центр «Академия». www.academia-moscow.ru
125252, Москва, ул. Зорге, д. 15, корп. 1, пом. 26б.
Адрес для корреспонденции: 129085, Москва, пр-т Мира, 101В, стр. 1, а/я 48.
Тел./факс: (495) 648-0507, 616-00-29.
Санитарно-эпидемиологическое заключение № РОСС RU. АЕ51. Н 16067 от 06.03.2012.

Отпечатано с электронных носителей издательства.
ОАО «Тверской полиграфический комбинат», 170024, г. Тверь, пр-т Ленина, 5.
Телефон: (4822) 44-52-03, 44-50-34. Телефон/факс: (4822) 44-42-15.
Home page — www.tverpk.ru Электронная почта (E-mail) — sales@tverpk.ru

МОЛЕКУЛЯРНАЯ БИОЛОГИЯ

ISBN 978-5-7695-9147-1

Издательский центр «Академия»
www.academia-moscow.ru