

НУКЛЕИНОВЫЕ КИСЛОТЫ

Строение нуклеиновых кислот

Синтез нуклеиновых кислот

Химическая модификация
нуклеиновых кислот

Нуклеопротеиды

Процессы
с участием нуклеиновых кислот:
репликация, транскрипция
и трансляция

Генная инженерия



Мишер (Miescher) Иоган Фридрих (1844—1895), швейцарский врач. Окончил Базельский университет. В 1869 г. из ядер лейкоцитов выделил вещество, названное им нуклеином, и установил его кислотные свойства; эта дата считается датой открытия нуклеиновых кислот.

Нуклеиновые кислоты — важнейшие биополимеры, осуществляющие хранение и передачу генетической информации в живой клетке.

Существуют два различных типа нуклеиновых кислот — дезоксирибонуклеиновые кислоты (ДНК) и рибонуклеиновые кислоты (РНК). ДНК представляет собой генетический материал большинства организмов. В прокариотических клетках, кроме основной хромосомной ДНК, часто встречаются внехромосомные ДНК — плазмиды. В эукариотических клетках основная масса ДНК расположена в клеточном ядре, где она связана с белками в хромосомах. Эукариотические клетки содержат ДНК также в различных органеллах (митохондриях, хлоропластах). Что же касается РНК, то в клетках имеются матричные РНК (мРНК), рибосомные РНК (рРНК), транспортные РНК (тРНК) и ряд других; кроме того, РНК входят в состав многих вирусов.

Исторический очерк. К середине прошлого века было установлено, что способность к наследованию признаков определяется материалом клеточного ядра. В 1869 г. Ф. Мишер, исследуя химический состав ядер гнойных клеток, выделил из них вещество кислого характера, названное им нуклеином. Это событие расценивается сейчас как открытие нуклеиновых кислот. Сам термин «нуклеиновые кислоты» был введен в 1889 г., а в 1891 г. немецкий биохимик А. Кёсель описал гидролиз нуклеиновой кислоты, установив, что она состоит из остатков сахара, фосфорной кислоты и четырех гетероциклических оснований, принадлежащих к пуринам и пиримидинам. Он же впервые указал на существование двух типов нуклеиновых кислот.

С начала нашего века началось интенсивное изучение продуктов расщепления нуклеиновых кислот. Э. Фишер внес большой вклад в химию пуринов и пиримидинов, а позднее Ф. Левен, Д. Гулланд и др. определили строение углеводных компонентов и природу нуклеозидных звеньев (названия «нуклеозид» и «нуклеотид» были предложены Ф. Левеном еще в 1908—1909 гг.). Окончательно строение нуклеозидов, нуклеотидов и роль фосфодиэфирной связи были выяснены в 1952 г. в результате работ английской школы под руководством А. Тодда.

Что касается полимерной природы нуклеиновых кислот, то с конца 30-х годов существовало убеждение, что ДНК представляет собой тетра-нуклеотид с четырьмя различными гетероциклическими основаниями; это позволило Ф. Левену сформулировать позднее тетра-нуклеотидную теорию строения ДНК. Теория была опровергнута лишь в 1950 г. благодаря работам Э. Чаргаффа, который при тщательном анализе нашел значительные различия в нуклеотидном составе ДНК из разных источников; он же сформулировал правила о попарной эквивалентности содержания аденина и тимина, гуанина и цитозина в ДНК. Пионерские работы по изучению химического состава ДНК растений были выполнены в этот период А. Н. Белозерским и его сотрудниками.

Накопленные данные позволили сформулировать концепцию: генетическая информация в полимерной цепи ДНК заключена в порядке чередования четырех мономерных звеньев. К этому времени в лаборатории М. Уилкинса (Р. Франклин) были получены первые рентгенограммы ДНК. Венцом исследований по строению нуклеиновых кислот явилась модель двойной спирали ДНК, предложенная в 1953 г. Дж. Уотсоном и Ф. Криком и ознаменовавшая рождение молекулярной биологии.

К началу 40-х годов были получены прямые доказательства участия нуклеиновых кислот в передаче генетической информации. В 1940 г. Дж. Бидл и Э. Тейтем, основываясь на многочисленных экспериментальных данных по изучению мутантных штаммов хлеб-

ной плесени *Neurospora crassa*, сформулировали важный принцип «один ген — один фермент». Несколько позднее О. Эйвери, К. Мак-Леод и М. Мак-Карти однозначно установили, что передача наследственной информации от клетки к клетке осуществляется ДНК.

Идея самовоспроизводства, или репликация, ДНК, следовавшая из модели Уотсона — Крика, была подтверждена открытием фермента ДНК-полимеразы (А. Корнберг, 1957).

Начало работам, приведшим к выяснению биологической роли РНК, было положено Т. Касперсоном и Ж. Браше еще в 30-х годах. Значительно позже (в 50-х годах) было показано, что синтез белка осуществляется рибосомами, но концепция о существовании РНК-посредника (мРНК), переносящего информацию от ДНК к рибосомам, была сформулирована Ф. Жакобом и Ж. Моно только в 1961 г. Главный фермент транскрипции — ДНК-зависимая РНК-полимераза — был открыт в 1960 г. С. Вейссом, Ж. Гурвицем и О. Стивенсом.

Одним из самых важных этапов в изучении функции нуклеиновых кислот явилась расшифровка способа записи информации в ДНК и принцип передачи ее на белковую структуру, т. е. формулирование генетического кода. В 1961 г. Ф. Крик и С. Бреннер показали, что каждой аминокислоте в белке соответствует триплет нуклеотидов. К этому времени были уже обнаружены транспортные РНК (М. Б. Хагланд и К. Огата, 1957) и выяснена их функция, проявляемая во взаимодействии антикодонов и кодонов (П. Замечник, 1957). Сам же генетический код, состоящий из 64 кодонов, был установлен в 1966 г. благодаря работам М. Ниренберга, Г. Кораны и С. Очоа.

С момента открытия генетического кода начинается новый этап в изучении структуры и функции нуклеиновых кислот. Он характеризуется стремительным накоплением огромного количества экспериментальных данных, приведших к новым качественным представлениям.

К 1970 г. исследования так называемой системы рестрикции и модификации, которая существует в прокариотических клетках и предохраняет от попадания внутрь клетки чужеродной генетической информации, привели к выделению первого из ферментов, осуществляющих рестрикцию. Им оказалась эндонуклеаза, расщепляющая ДНК по определенной последовательности оснований. Затем было найдено около 400 подобных ферментов, способных узнавать свыше 90 различных последовательностей в ДНК.

Таким образом, появился метод специфических расщеплений молекул ДНК, который составил основу современной методологии установления структуры нуклеиновых кислот и геной инженерии. Исследования рестрикционных эндонуклеаз были отмечены Нобелевской премией, присужденной Х. Смиуту, В. Арберу и Д. Натансу.

В 1970 г. американские ученые Г. Темин и Д. Балтимор сообщили об открытии в вирионах опухолеродных РНК-содержащих вирусов фермента — обратной транскриптазы, способного синтезировать ДНК, используя РНК в качестве матрицы. Существование такого фермента ранее было предсказано советским генетиком С. М. Гершензоном. Это открытие явилось настоящей революцией в фундаментальных представлениях о путях биосинтеза нуклеиновых кислот и дало возможность получения *in vitro* ДНК, являющихся копиями информационных РНК.

В 1972 г. П. Берг сообщил о соединении *in vitro*, с помощью специально разработанных приемов, ДНК двух различных вирусов: бактериофага λ и SV-40. Таким образом, было положено начало генетической инженерии, которая вскоре стала использовать для получения фрагментов ДНК рестрикционные эндонуклеазы, а для



Кёссель [Kossel] Альбрехт (1853—1927), немецкий биохимик, иностранный почетный член АН СССР (1926). Окончил Страсбургский университет (1877), с 1887 г. — профессор Берлинского, а в 1901—1923 гг. — Гейдельбергского университетов. Основные работы посвящены химии белков и нуклеопротеидов. Впервые описал гистоны и протамины. Автор одной из первых теорий строения белков. Лауреат Нобелевской премии по физиологии и медицине (1910).



Чаргафф (Chargaff) Эрвин (р. 1905), американский биохимик. Окончил Венский университет (1928); с 1935 г. — в Колумбийском университете в Нью-Йорке. Основные работы — по изучению химического состава и структуры нуклеиновых кислот, определил количественное соотношение азотистых оснований, входящих в их состав (правило Чаргаффа). Это открытие было использовано Ф. Криком и Дж. Уотсоном при построении модели структуры ДНК.

их соединения открытые и хорошо охарактеризованные к тому времени ферменты — ДНК-лигазы. Были разработаны (Г. Бойер, С. Коэн, Д. Хелинский) системы введения чужеродных ДНК в различные клетки, которые использовали в качестве носителя ДНК внехромосомные элементы, способные существовать и размножаться в этих клетках, зачастую вместе с ним. Встраивание чужой ДНК в носители (векторы) осуществлялось путем разрезания их рестрикционными эндонуклеазами и последующего сшивания лигазами.

Методы генной инженерии в сочетании с доступностью ДНК, являющихся копиями мРНК, открыли три возможности: а) выделение ранее труднодоступных генов эукариот в индивидуальном виде; б) получение их в больших количествах, необходимых для структурного анализа; в) создание системы их экспрессии в чужеродных организмах, например в бактериальных клетках.

Последнее обстоятельство открывало путь к созданию живых организмов с заданными свойствами. Вскоре эта возможность реализовалась в биотехнологии. Были получены микробы, способные продуцировать самые разнообразные белки, обычно синтезируемые эукариотическими, в том числе человеческим, организмами.

Две первые возможности сделали реальным анализ структуры ранее недоступных генов. Недоставало только методов быстрого определения первичной структуры больших ДНК.

В 1975 г. А. Максам и У. Гилберт в США и Ф. Сенгер в Великобритании разработали такие методы, ставшие последним звеном в цепи методов, необходимых для широкомасштабных исследований структуры и функционирования геномов. В разработку этих методов существенный вклад был внесен А. Д. Мирзабековым и Е. Д. Свердловым (СССР).

Строение нуклеиновых кислот

Первичная структура нуклеиновых кислот

Нуклеиновые кислоты представляют собой биополимеры, построенные из нуклеотидов, соединенных фосфодиэфирной связью. Каждый нуклеотид, в свою очередь, состоит из остатков гетероциклического основания, углевода и фосфорной кислоты.

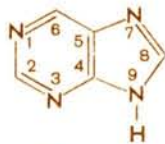
Одним из важнейших компонентов нуклеиновых кислот являются гетероциклические основания. Все они представляют собой производные пиримидина или пурина.

В подавляющем большинстве случаев нуклеиновые кислоты в качестве гетероциклических оснований содержат урацил (только в РНК), тимин (только в ДНК) и цитозин, являющиеся производными пиримидина, а также аденин и гуанин, относящиеся к производным пурина.

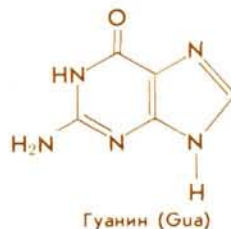
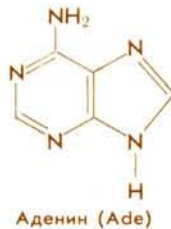
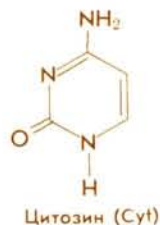
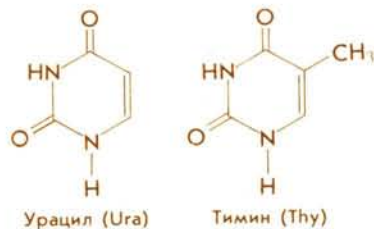
В соответствии с принятой номенклатурой основания могут записываться трехбуквенным кодом, представляющим три первые буквы их латинского названия.



Пиримидин

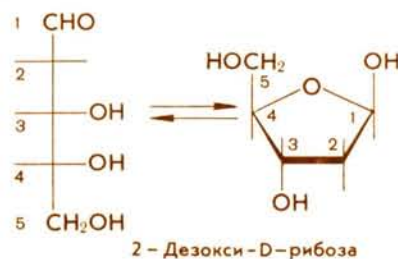
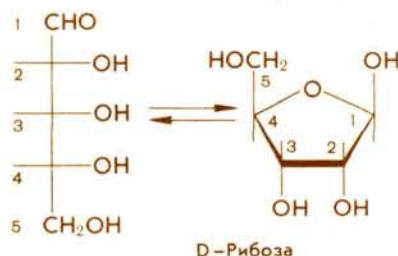


Пурин



Белозерский Андрей Николаевич (1905—1972), советский биохимик, академик АН СССР (1962). Окончил Среднеазиатский университет в Ташкенте (1927), работал в Московском университете и в Институте биохимии им. А. Н. Баха АН СССР, в 1971—1972 гг. — вице-президент АН СССР. Основные работы посвящены химии и биохимии нуклеиновых кислот. Первым установил (1936) наличие в нуклеопротеидах растительной клетки ДНК и нещелочных белков, исследовал их структуру и локализацию. Показал видовую специфичность нуклеиновых кислот, обнаружил связь между составом нуклеиновых кислот и филогенезом организмов. Герой Социалистического Труда (1969).

Нуклеозиды. В составе нуклеиновых кислот гетероциклические основания связаны с D-рибозой в РНК или с 2-дезоксид-рибозой в ДНК, образуя соединения, называемые соответственно рибонуклеозидами или дезоксирибонуклеозидами. Нуклеозиды являются β-N-пентафуранозидами гетероциклических оснований

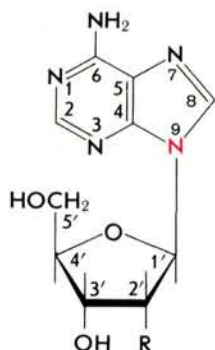


Пиримидиновые основания соединяются с углеводом своими N-1, а пуриновые — N-9 атомами азота. Нуклеозиды, содержащие аденин и гуанин, называются аденозином или дезоксиаденозином и гуанозином или дезоксигуанозином. Нуклеозиды — производные урацила и цитозина называются соответственно уридином или дезоксиуридином и цитидином или дезоксцитидином. Дезоксирибонуклеозид тимина принято называть тимидином, а рибонуклеозид — риботимидином.

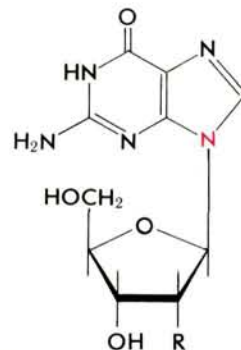


Бидл (Beadle) Джордж Уэлс (р. 1903), американский генетик. Образование получил в университете штата Небраска и Корнеллском университете; в 1937—1968 гг. — профессор Стэнфордского университета, Калифорнийского технологического института, президент Чикагского университета. Основные работы — в области цитологии и генетики. Один из основоположников биохимической генетики, выдвинул (1944, совместно с Э. Тейтемом) концепцию «один ген — один фермент». Лауреат Нобелевской премии по физиологии и медицине (1958, совместно с Э. Тейтемом и Дж. Ледербергом).

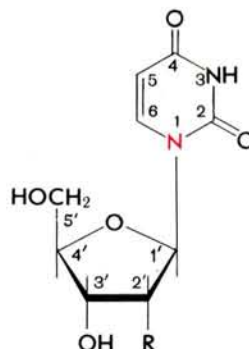
Для сокращения названий нуклеозидов используется трехбуквенный или однобуквенный код. В первом варианте к двум начальным буквам латинского названия нуклеозида добавляется третья так, чтобы отличить название нуклеозида от названия основания (например, Adenine = Ade, Adenozine = Ado). Во втором варианте используются начальные буквы латинских названий. Дезокси-нуклеозиды отличаются от рибонуклеозидов добавлением префикса d; dAdo, dThd или dA, dT и т. д. Для обобщенного обозначения любого нуклеозида применяется символ N, для пиримидинового нуклеозида символ Y, пуринового — символ R



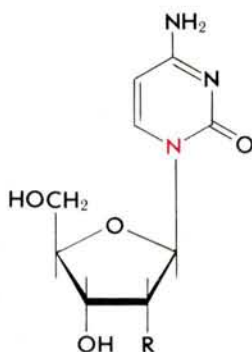
R=OH, аденозин (A, Ado)
R=H, дезоксиаденозин (dA, dAdo)



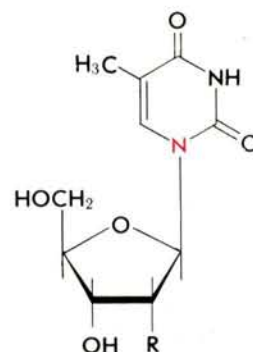
R=OH, гуанозин (G, Guo)
R=H, дезоксигуанозин (dG, dGuo)



R=OH, уридин (U, Urd)
R=H, дезоксиуридин (dU, dUrd)



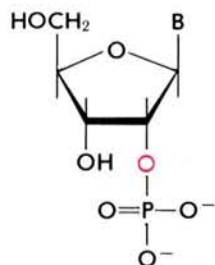
R=OH, цитидин (C, Cyd)
R=H, дезоксицитидин (dC, dCyd)



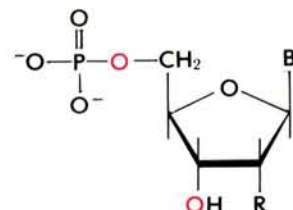
R=OH, риботимидин (T, Thd)
R=H, тимидин (dT, dThd)

Для того чтобы различать номера атомов оснований и атомов сахара, к последним в нуклеозидах добавляется сверху штрих. Например, С-3 углеродный атом рибозы в нуклеозиде называется С-3' атомом, а связанный с ним гидроксил — 3'-гидроксильной группой.

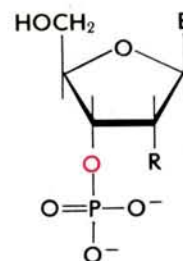
Нуклеотиды. Третий компонент нуклеиновых кислот — ортофосфорная кислота — образует сложноэфирные связи со спиртовыми группами рибозы или дезоксирибозы. Путем расщепления нуклеиновых кислот в контролируемых условиях удается выделить сложные эфиры нуклеозидов и фосфорной кислоты — нуклеотиды. Названия нуклеотидов производятся от названия гетероциклического основания, входящего в их состав, с добавлением слова «кислота»: цитидиловая кислота, адениловая кислота и т. д. В современной номенклатуре указываются также положения фосфатной группы или групп (аденозин-5'-фосфат, аденозин-3'-фосфат, дезоксиаденозин-5'-фосфат); часто используются однобуквенные сокращения: для 5'-фосфатов — рА, рG, рC, рU, рN, рdА, рdG, рdC, рdU, рdN, для 3'-фосфатов — Ар, Gr, Cr, Ur, Nr, dAr, dGr, dCr, dUr, dNr, для 2'-фосфатов — А(2')р, G(2')р, C(2')р, U(2')р, N(2')р



Нуклеозид-2'-фосфат N(2')р



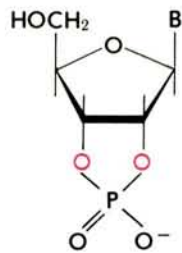
R = OH, нуклеозид-5'-фосфат (рN)
R = H, дезоксинуклеозид-5'-фосфат (рdN)



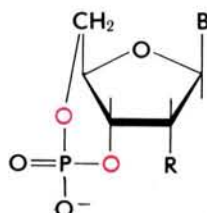
R = OH, нуклеозид-3'-фосфат (рN)
R = H, дезоксинуклеозид-3'-фосфат (рdN)

B — гетероциклическое основание
(от англ. base — основание)

При определенных условиях расщепления рибонуклеиновых кислот образуются циклические фосфаты, представляющие собой диэфиры ортофосфорной кислоты с 2'- и 3'-гидроксильными группами рибозы. Они называются нуклеозид-2', 3'-циклофосфатами и обозначаются N > р (например, аденозин-2', 3'-циклофосфат, или А > р). Циклофосфаты типа (а) N(3', 5') > р не образуются при расщеплении нуклеиновых кислот. Некоторые из них, например аденозин-3', 5'-циклофосфат (см. с. 240), являются продуктами ферментативных реакций, протекающих в клетке, и играют важную биологическую роль.

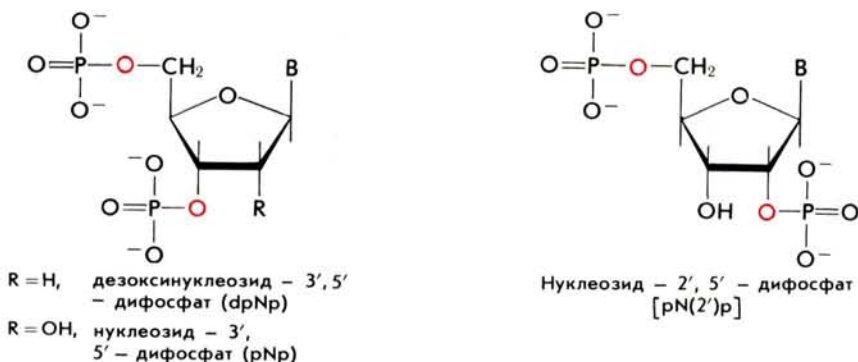


Нуклеозид-2', 3'-циклофосфат (N > р)

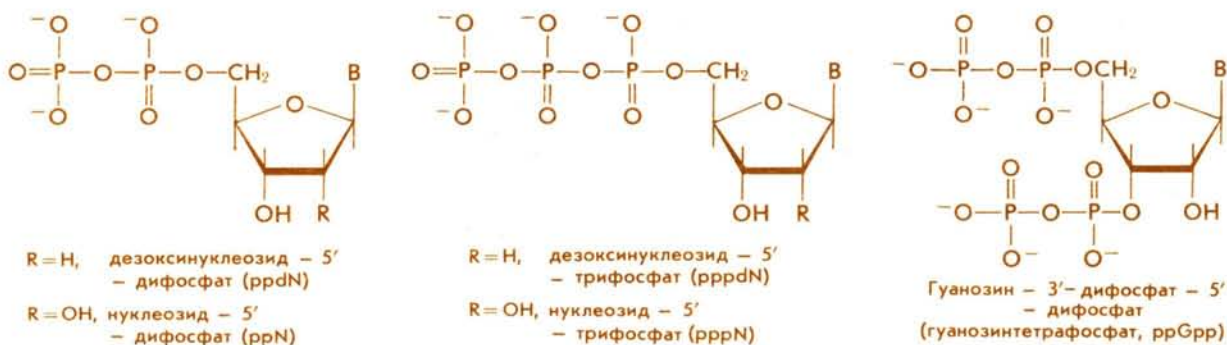


а) R = OH, нуклеозид-3', 5'-циклофосфат N(3', 5') > р
б) R = H, дезоксинуклеозид-3', 5'-циклофосфат N(3', 5') > р

При гидролизе нуклеиновых кислот могут образовываться нуклеотиды, содержащие два остатка фосфорной кислоты, один из которых связан с 5'-гидроксильной, а другой — с 3'- или 2'-гидроксильной группами. Такие производные называют нуклеозид-3', 5'- или нуклеозид-2', 5'-дифосфатами.

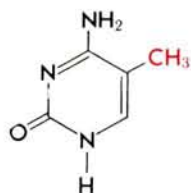
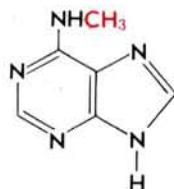


Важную роль в жизнедеятельности клетки играют сложные эфиры полифосфорных кислот и нуклеозидов. Наиболее широко распространены эфиры, образованные ди- или трифосфатными группами и 5'-ОН-группами нуклеозидов (см. с. 348). Они называются нуклеозидди- или нуклеозидтрифосфатами и обозначаются ppN, ppdN, pppN, pppdN. Пирофосфатные группы могут быть соединены с 5'- и 3'-гидроксилами рибозы. Наиболее интересным соединением такого рода является гуанозинтетрафосфат: гуанозин-5'-дифосфат-3'-дифосфат, ppGpp

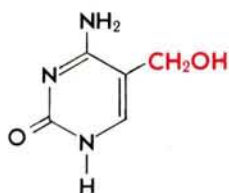
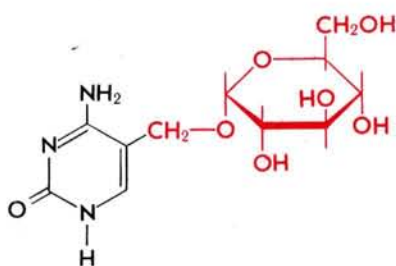
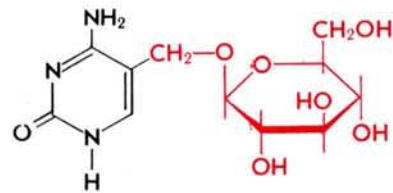


Редкие (минорные) компоненты нуклеиновых кислот. Помимо основных компонентов, в состав нуклеиновых кислот могут входить нуклеозиды с необычными гетероциклическими основаниями или с модифицированным углеводным остатком.

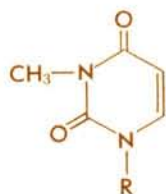
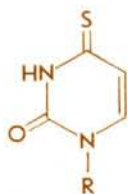
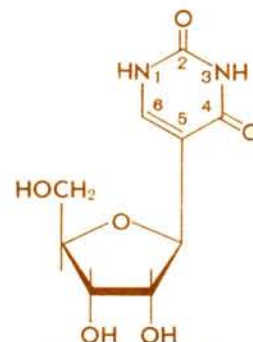
Редкими компонентами ДНК являются метилированные основания: 5-метилцитозин, 6-N-метиладенин и некоторые другие. В ДНК

5-Метилцитозин ($m^5\text{Cyt}$)6-N-Метиладенин ($m^6\text{Ade}$)

некоторых бактериофагов вместо цитозина имеются 5-гидроксиметилцитозин и его гликозилированные производные — α -D-гликопиранозил или β -D-гликопиранозил

5-Гидроксиметилцитозин ($hm^5\text{Cyt}$)5-(α -D-Глюкопиранозилгидроксиметил)цитозин5-(β -D-Глюкопиранозилгидроксиметил)цитозин

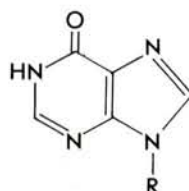
РНК также содержат редкие компоненты. Они в большей степени встречаются в тРНК, в меньшей степени — в рибосомных и ядерных РНК. Наиболее распространены производные обычных оснований (цитозина, урацила, аденина и гуанина) с заместителями в гетероциклическом ядре или в экзоциклических аминогруппах. Известны некоторые производные урацила, такие, как 3-метилуридин, 4-тиоуридин и т. д. Необычная модификация наблюдается в случае псевдоуридина, в котором урацил соединяется с сахаром С-гликозидной, а не N-гликозидной связью за счет атома С-5 гетероциклического ядра

3-Метилуридин ($m^3\text{U}$)4-Тиоуридин ($S^4\text{U}$)Псевдоуридин (ψ)



Тейтем (Tatum) Эдвард Лори (1909—1975), американский генетик и биохимик. Окончил Висконсинский университет в Милуоки (1931), с 1957 г. — профессор Рокфеллеровского института медицинских исследований. Основное направление научных исследований — молекулярная генетика. Сформулировал концепцию «один ген — один фермент», ставшую основой биохимической генетики (1944, совместно с Дж. Бидлом). Открыл явление генетической рекомбинации у бактерий — конъюгацию (1946, совместно с Дж. Ледербергом). Лауреат Нобелевской премии по физиологии и медицине (1958, совместно с Дж. Бидлом и Дж. Ледербергом).

Очень часто в тРНК встречается продукт дезаминирования аденозина — инозин. Известны компоненты, являющиеся продуктами метилирования 2'-гидроксила рибозы (2'-О-метилнуклеозиды)



Инозин



2'-О-Метилнуклеозид (Nm)

Олиго- и полинуклеотиды. *Олигонуклеотидами* называют полимеры, в которых несколько нуклеозидов (до 20) соединены друг с другом фосфодиэфирными связями; более длинные цепи называют *полинуклеотидами*.

В обычных нуклеиновых кислотах 5'-гидроксильная группа одного нуклеозида связана с 3'-гидроксильной группой другого посредством остатка фосфорной кислоты, образующей с этими группами сложноэфирные связи. Простейшими олигонуклеотидами являются динуклеозидмонофосфаты [нуклеотидил-(3' → 5')-нуклеозиды]. В молекуле аденилил-(3' → 5')-цитидина (рис. 173) два углеводных остатка неодинаковы с химической точки зрения. Тот, который принадлежит аденозину, имеет свободную 5'-гидроксильную группу, тогда как его 3'-гидроксильная группа участвует в образовании фосфодиэфирной связи. Другой, принадлежащий цитидину, напротив, содержит свободную 3'-гидроксильную группу, а его 5'-гидроксильная группа вовлечена в образование фосфодиэфирной связи.

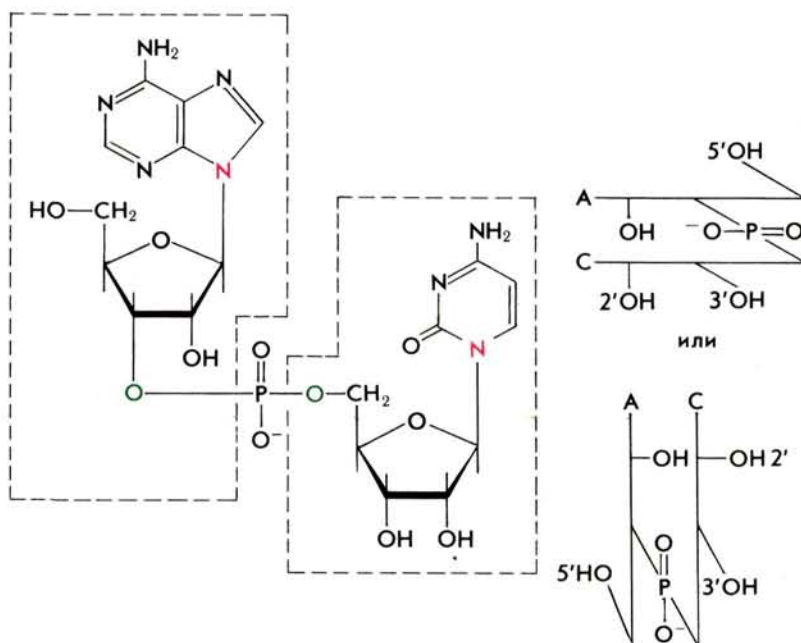
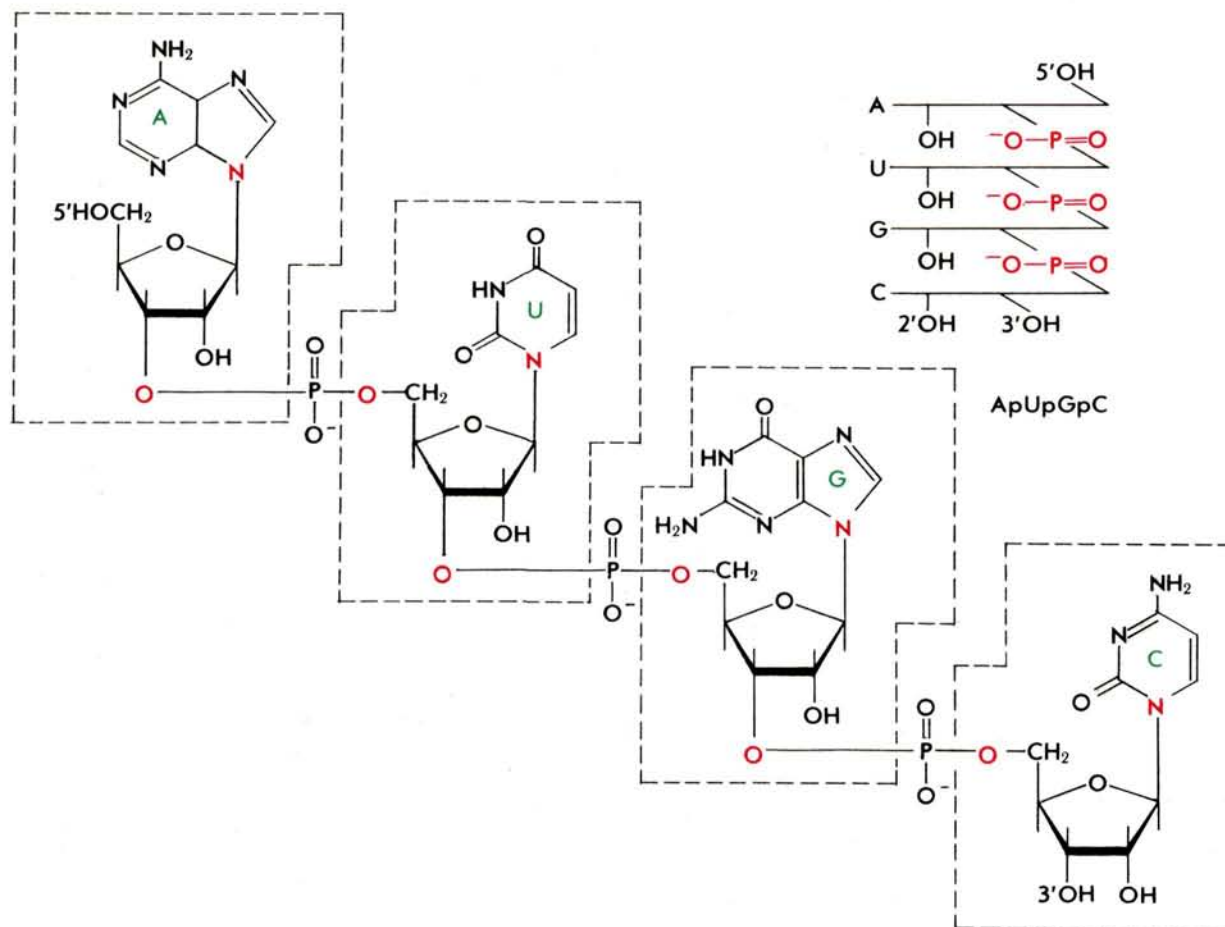


Рис. 173. Аденилил-(3' → 5')-цитидин (ApC).

Нуклеозид со свободной 5'-ОН-группой называется 5'-концевым, нуклеозид со свободной 3'-ОН-группой — 3'-концевым. Это определение относится и к более длинным олигонуклеотидам, например тетра-нуклеотиду, изображенному на рисунке 174. Полное название олигонуклеотида: аденилил-(3' → 5')-уридил-(3' → 5')-гуанилил-(3' → 5')-цитидин. Для простоты написания и произношения нуклеозиды обозначаются однобуквенным кодом и нуклеотидные цепи записываются слева направо, начиная с 5'-концевого нуклеозида в порядке следования мономерных звеньев. Возможны два варианта сокращенного написания: динуклеозидмонофосфат (рис. 173) может быть записан как ApC или A—C, а тетра-нуклеотид (рис. 174) как ApUpGpC или A—U—G—C. В случае дезоксирибоолиго- или полинуклеотидов добавляется символ d: d(A—T—G—C—...).

Для обозначения полинуклеотидов используются способы, которые иллюстрируются примерами: поли (A), поли (A—U); они представляют собой соответственно полимеры: A—A—A—A—... и A—U—A—U—A—U—..., а в случае дезоксирибополинуклеотидов: поли d(A), поли d(A—T). В сокращенной записи фосфорилирование 5'-концевого звена символизирует буква «р» слева от 5'-концевого звена: pApC, pApUpGpC, pA—C, pA—U—G—C. Фосфорилирование 3'-концевого звена обозначается той же буквой, но справа от него: ApCp, ApUpGpCp, A—Cp, A—U—G—Cp.

Рис. 174. Аденилил-(3' → 5')-уридил-(3' → 5')-гуанилил-(3' → 5')-цитидин (ApUpGpC).





Ледерберг (Lederberg) Джошуа (р. 1925), американский генетик. Окончил Колумбийский университет (1944), с 1959 г. — профессор Стэнфордского университета. Заложил основы генетики микроорганизмов. Открыл явление конъюгации у бактерий (1946, совместно с Э. Тейтемом). Выяснял причины возникновения устойчивости бактерий к антибиотикам. Лауреат Нобелевской премии по физиологии и медицине (1958, совместно с Дж. Бидлом и Э. Тейтемом).

При описании химических процессов с участием сахара и фосфорной кислоты применяются также условные обозначения полинуклеотидных цепей, как это показано на рисунках 173 и 174: остатки сахара изображаются прямыми линиями, соединенными между собой; 2'-, 3'- и 5'-гидроксилы символизируются отрезками, исходящими из прямой.

Различные олиго- и полинуклеотиды отличаются друг от друга содержанием нуклеотидов каждого типа. Выраженные в процентах относительные количества мономерных звеньев называют нуклеотидным составом, а их последовательность — первичной структурой нуклеиновой кислоты.

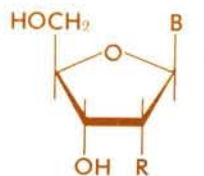
При доказательстве химического строения нуклеозидов, нуклеотидов и полинуклеотидов определяются тип гетероциклического основания, природа углеводного компонента и место присоединения его к основанию, конфигурация гликозидного центра; устанавливаются также число и место присоединения остатков фосфорной кислоты.

Гетероциклические основания в составе нуклеиновых кислот идентифицируются после жесткого кислотного гидролиза сравнением их хроматографических характеристик, электрофоретической подвижности, а также УФ-, ИК- и ЯМР-спектров.

Идентификация углеводных компонентов также проводится после кислотного гидролиза нуклеиновых кислот. При этом из пуриновых или предварительно гидрированных по 5, 6-двойной связи пиримидиновых рибонуклеотидов образуется D-рибоза. Основания дезоксирибонуклеотидов отщепляются в значительно более мягких условиях, чем рибонуклеотиды, но вследствие нестабильности образующаяся дезоксирибоза превращается в левулиновую кислоту. При чрезвычайно мягкой обработке кислотой пуриновых или восстановленных пиримидиновых дезоксирибонуклеотидов удается выделить дезоксирибозу или ее легко идентифицируемые производные



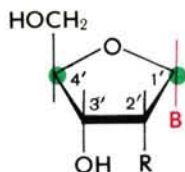
Возможность отщепления гетероциклического основания от углевода в условиях мягкого кислотного гидролиза позволяла предположить, что нуклеозиды представляют собой N-гликозиды, причем в образовании гликозидной связи участвует один из гетероциклических атомов. С помощью спектральных и химических методов анализа было установлено, что основание соединено с углеводом своим N-9 атомом в случае пуринов и N-1 в случае пиримидинов. В состав нуклеотидов входят только два остатка сахара — D-рибоза и 2-дезоксид-рибоза. С помощью периодатного окисления было показано, что оба углевода находятся в форме фуранозы. Наличие в циклической форме углевода асимметрического (C-1') атома углерода обуславливает возможность ее существования в виде двух различных стереоизомеров. В соответствии с принятой номенклатурой стереоизомеры, отличающиеся только конфигурацией гликозидного центра, называются аномерами. Тот из аномеров,



Фуранозная форма

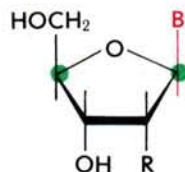
R = H, дезоксирибоза; R = OH, рибоза

который имеет одинаковую конфигурацию у C-1' и у атома, определяющего принадлежность сахара к D- или к L-ряду (см. с. 446), называется α -аномером. Аномер с противоположными конфигурациями у этих атомов называется β -аномером



R=H, α -аномер гликозида
D- дезоксирибозы
R=OH, α -аномер гликозида
D- рибозы

(Конфигурации атомов C-1' и C-4', отмеченных кружками, одинаковы)



R=H, β -аномер гликозида
D- дезоксирибозы
R=OH, β -аномер гликозида
D- рибозы

(Конфигурации атомов C-1' и C-4', отмеченных кружками, противоположны)



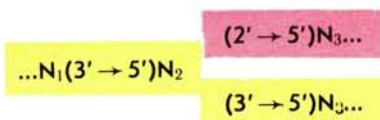
Касперсон (Caspersson) Тобийорт Оскар (р. 1910), шведский биофизик. Окончил Стокгольмский университет (1936), с 1944 г.— руководитель медицинского отдела Нобелевского института медицинских исследований клетки и лаборатории экспериментальных исследований клетки в Стокгольме. Основные работы — в области физики клетки. Исследовал содержание, распределение в клетке и биологическую роль нуклеиновых кислот, в первую очередь РНК.

В нуклеиновых кислотах оба моносахарида имеют β -конфигурацию аномерного центра. Современные доказательства конфигурации аномерного центра базируются на исследовании ЯМР- и ИК-спектров, спектров кругового дихроизма и дисперсии оптического вращения.

При щелочном и кислотном гидролизе ДНК и РНК ведут себя различно. ДНК устойчива к действию щелочи, тогда как РНК очень быстро гидролизует и дает смесь нуклеозид-3'- и нуклеозид-2'-фосфатов. Эти данные свидетельствуют о том, что в случае РНК наиболее вероятен 5' → 3' тип фосфодиэфирной связи.

Лучшие результаты дают ферментативные методы гидролиза, в том числе использование фосфодиэстеразы змеиного яда, рибонуклеазы и других ферментов. Совокупность химических и ферментативных методов однозначно показывает, что и в ДНК, и в РНК полинуклеотидные цепи построены однотипно за счет 5' → 3'-фосфодиэфирных связей между каждыми двумя соседними нуклеозидными звеньями.

До недавнего времени среди нуклеиновых кислот не обнаруживались разветвленные структуры, что позволяло считать их исключительно линейными полимерами. В последние годы показана возможность существования и разветвленных структур РНК. В частности, гетерогенные ядерные РНК в клетках опухоли HeLa содержат до 10% молекул, в которых ответвление обусловлено присоединением одной цепи РНК к другой посредством 2' → 5'-фосфодиэфирных связей. Таким образом, в точке разветвления существует структура:



Определение первичной структуры

Общие принципы расшифровки первичной структуры нуклеиновых кислот те же, что и в случае других биополимеров. Полинуклеотидная цепь расщепляется с помощью ферментов и химических агентов, обладающих повышенной избирательностью, на фрагменты, которые дешифруются специфическими методами, и затем реконструируется вся цепь.

дезоксирибонуклеиновые кислоты. Были разработаны также методы быстрого определения последовательностей длинных участков ДНК. Успехи в анализе последовательности ДНК привели к тому, что структуры РНК стали определяться путем их «переписывания» в соответствующую дезоксирибонуклеотидную последовательность с помощью обратной транскриптазы.

Принцип блочного метода определения последовательности показан на рисунке 175. Олигонуклеотид неизвестной структуры расщепляется двумя способами: X и Y, различающимися по специфичности. При расщеплении по способу X образуются три олигонуклеотида, а по способу Y — два. Структуры всех полученных фрагментов устанавливаются соответствующим методом. Далее проводится сопоставление структур олигонуклеотидов Y со структурами X с целью нахождения частично совпадающих (перекрывающихся) последовательностей и реконструкции таким образом исходной цепи. Поскольку каждый полученный олигонуклеотид представляет собой блок, из суммы которых строится исходная структура, метод называется методом перекрывающихся блоков.

Ферменты, расщепляющие ДНК. Фрагментация нуклеотидной цепи при анализе последовательности осуществляется с помощью ферментов, называемых нуклеазами. Известно значительное число таких ферментов. Одни из них расщепляют фосфодиэфирные связи внутри полинуклеотидной цепи (эндодезоксирибонуклеазы), другие гидролизуют цепь начиная с 5'- или 3'-конца — они называются 5'-или 3'-экзодезоксирибонуклеазами. Эндонуклеазы различаются по специфичности — расщепляют определенные последовательности или не предъявляют к последовательности никаких требований. Существуют ферменты, гидролизующие только одноцепочечные ДНК или двухцепочечные ДНК, или те и другие. Наиболее важным классом эндодезоксирибонуклеаз являются рестрикционные эндонуклеазы.

Рестрикционные эндонуклеазы. Название рестрикционных эндонуклеаз (или, сокращенно, рестриктаз) образуется из первой буквы названия рода и первых двух букв названия вида микроорганизма, из которого получен фермент. Например, если источником фермента является *Escherichia coli*, фермент называется Eco, *H. influenzae* — Hin. Название штамма следует за этими тремя буквами — Hind. Если в микроорганизме содержится несколько ферментов, то они нумеруются римскими цифрами EcoRI, EcoRII и т. д.

Эти ферменты являются компонентами системы рестрикции — модификации прокариотических клеток, которая предназначена для их защиты от чужеродных ДНК. К рестрикционным эндонуклеазам относятся также и все другие ферменты, специфически расщепляющие ДНК по определенной последовательности, хотя часто их роль в процессах рестрикции — модификации не доказана. В зависимости от характера узнаваемой нуклеотидной последовательности («сайта»), мест расщепления и условий реакции рестриктазы делятся на три класса.

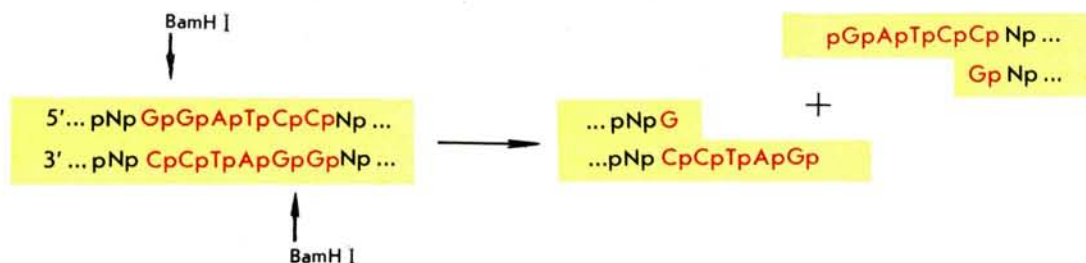
Рестрикционные эндонуклеазы первого класса узнают специфическую последовательность ДНК, но расщепляют нуклеиновую кислоту в различных местах, находящихся на неопределенном расстоянии от участка узнавания. Они содержат в одном ферментном комплексе рестриктирующую и модифицирующую активности и требуют присутствия в среде S-аденозилметионина, АТФ и Mg^{2+} .

Ферменты второго класса расщепляют участок ДНК, имеющий ось симметрии второго порядка. К ним относится также несколько ферментов, узнающих несимметричные последовательности и расщепляющих ДНК на некотором расстоянии от участка узнавания. Общим свойством рестриктаз этого класса является необходимость

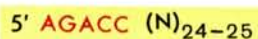


Берг (Berg) Пол (р. 1926), американский биохимик. Окончил Пенсильванский университет (1948), с 1959 г. — профессор Стэнфордского университета. Основные работы посвящены генной инженерии, изучению роли тРНК в биосинтезе белка. Впервые получил рекомбинантные молекулы ДНК бактериального вируса лямбда и вируса обезьяны SV-40. Лауреат Нобелевской премии по химии (1980, совместно с У. Гилбертом и Ф. Сенгером).

присутствия в среде только ионов Mg^{2+} . Некоторые рестриктазы расщепляют полинуклеотидную цепь так, что в результате образуются фрагменты, содержащие на 5'-концах выступающие четырехзвенные фрагменты с фосфорилированными 5'-концевыми звеньями. Вследствие симметрии расщепляемой последовательности эти тетра-нуклеотиды комплементарны друг другу. Они называются липкими, так как могут взаимодействовать друг с другом, образуя совершенные двуспиральные участки. Другие рестриктазы этого класса расщепляют полинуклеотидную цепь с образованием фрагментов, не содержащих однонитевых концевых участков, или структур с выступающими 3'-концевыми однонитевыми участками.



Для рестриктаз третьего класса участок узнавания не обязательно должен быть симметричным, а расщепление происходит на расстоянии 24 — 27 нуклеотидов от участка узнавания. Так же как рестриктазы первого класса, они содержат в одном комплексе и рестриктирующую, и метилирующую активности и требуют для рестрикции присутствия АТФ. Пример последовательности, узнаваемой рестриктазой EcoRI, относящейся к третьему классу, приведен ниже:



Известно уже несколько сотен рестриктаз. Часто выделяемые из разных источников ферменты проявляют одинаковую специфичность. Такие ферменты называют изоизомерами.

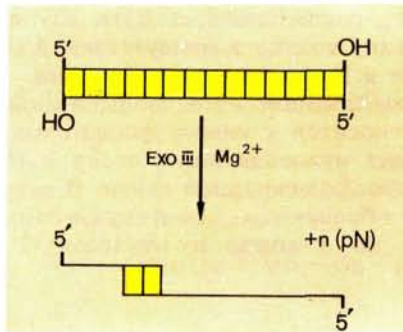
Дезоксирибонуклеазы I и II (ДНазы I и II). Дезоксирибонуклеаза I, или панкреатическую дезоксирибонуклеазу, выделяют из поджелудочной железы крупного рогатого скота. Она расщепляет одно- и двухцепочечные ДНК, давая сложную смесь как моно-, так и олигонуклеотидов с 5'-концами, содержащими фосфатные группы.

Фермент проявляет активность в присутствии ионов Mg^{2+} или Mn^{2+} при оптимальном значении рН среды 7,5. ДНазы I неспецифичны к последовательности субстрата, однако наблюдаются незначительные различия в скоростях расщепления фосфодиэфирных связей, образуемых различными нуклеотидами.

Дезоксирибонуклеаза II гидролизует ДНК до смеси моно- и олигонуклеотидов, содержащих 3'-фосфатные группы, оптимум рН фермента лежит в кислой области.

ДНазы I и II используются для расщепления ДНК с целью получения коротких фрагментов.

Эксонуклеаза III из *E. coli*. Фермент катализирует последовательное удаление 3'-концевых звеньев из двухцепочечных ДНК путем отщепления их в виде 5'-мононуклеотидов

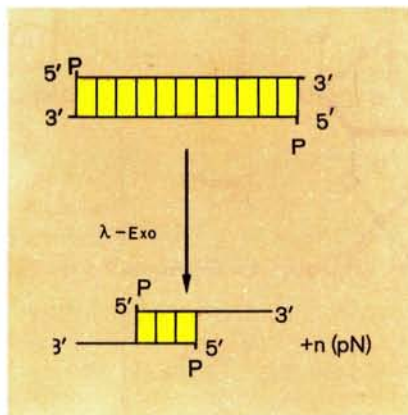


При этом в ДНК остаются одноцепочечные 5'-концевые участки; гидролиз продолжается до тех пор, пока длина двухцепочечного участка в центре гидролизуемой молекулы не уменьшается настолько, что два 5'-концевых фрагмента больше не могут удерживаться вместе. Когда они разделяются, гидролиз останавливается. Эксонуклеаза III используется для укорачивания двухцепочечных фрагментов ДНК.

5'-Эксонуклеаза из *E. coli*, инфицированной бактериофагом λ . Фермент удаляет нуклеозид-5'-фосфаты, последовательно отщепляя их от 5'-конца двухцепочечных ДНК. В ре-



Смит (Smith) Хамилтон (р. 1931), американский микробиолог. Окончил Калифорнийский университет в Беркли (1952), с 1970 г. — профессор университета Дж. Гопкинса в Балтиморе. Основные работы — в области генной инженерии. Впервые выделил рестрикционную эндонуклеазу Hind II (1970), разработал методы определения структуры узнаваемого участка. Лауреат Нобелевской премии по физиологии и медицине (1978, совместно с Д. Натансом и В. Арбером).



зультате образуются ДНК с выступающими 3'-концевыми участками.

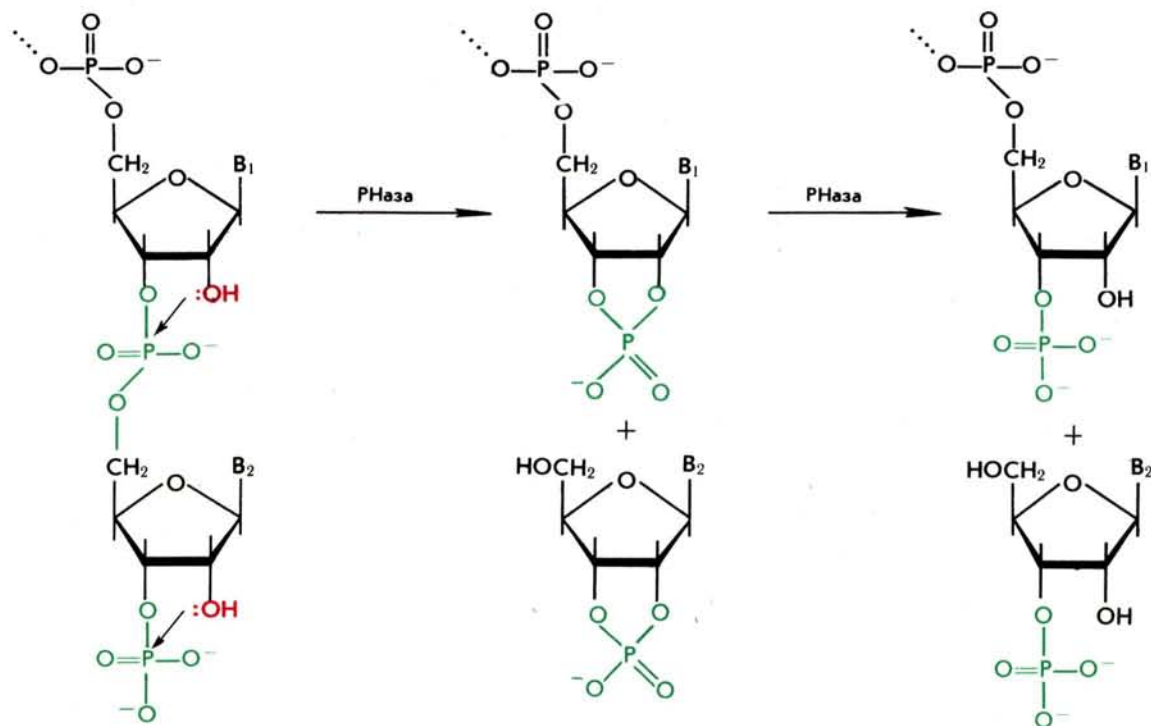
Эксонуклеаза VII из *E. coli* специфична к одноцепочечным ДНК и отщепляет небольшие олигонуклеотиды как с 5'-, так и с 3'-концов ДНК.

Применение находит и эксонуклеаза I из *E. coli*, гидролизующая одноцепочечные ДНК в 40 000 раз быстрее, чем двух-

цепочечные. Гидролизу подвергаются только полимеры со свободной 3'-гидроксильной группой; в результате гидролиза ДНК укорачивается с 3'-конца путем последовательного отщепления нуклеозид-5'-фосфатов. Фермент используется для специфического удаления 3'-концевых одноцепочечных участков ДНК.

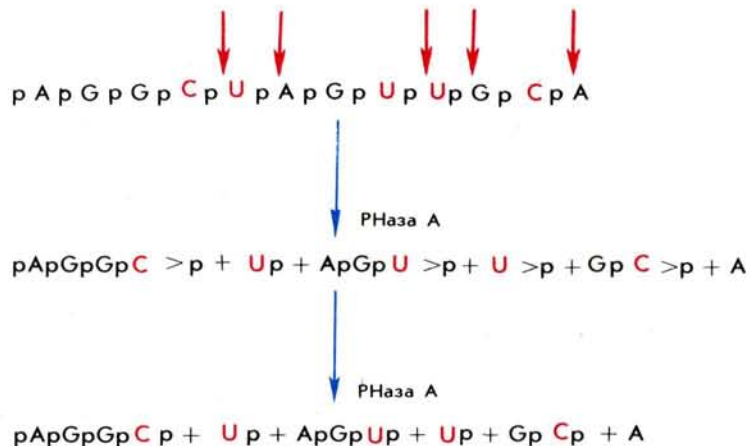
Следует отметить, что в клетках содержится большое количество дезоксирибонуклеаз, расщепляющих ДНК как по экзо-, так и по эндонуклеазному типу только в присутствии АТР. Многие из них принимают участие в процессах рекомбинации.

Ферменты, расщепляющие РНК. Большинство известных рибонуклеаз (РНаз) относится к классу фосфотрансфераз — ферментов, катализирующих нуклеофильную атаку 2'-ОН-группы рибозы на атом фосфора фосфодиэфирной связи. В результате на первом этапе реакции образуются олигонуклеотиды, содержащие 2',3'-циклофосфат на 3'-конце и нуклеозид-2',3'-циклофосфаты



При более высоких концентрациях РНАзы циклофосфатные группы раскрываются, образуя 3'-фосфатные группы. Процесс раскрытия циклофосфатов может быть осуществлен кратковременной обработкой кислотой; при этом образуются смеси 2'- и 3'-монофосфатов. Такие рибонуклеазы относятся к эндонуклеазам.

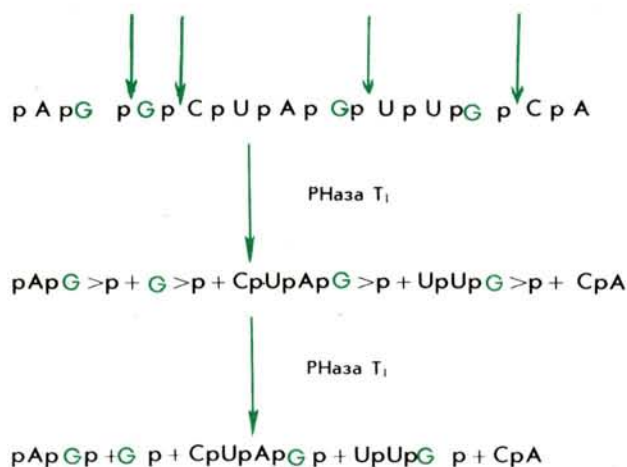
Рибонуклеаза А специфически гидролизует РНК по пиримидиновым основаниям; конечными продуктами (после расщепления циклофосфатов) являются пиримидиновые нуклеозид-3'-фосфаты и олигонуклеотиды, содержащие 5'-ОН- и 3'-фосфатную группы:



Натанс (Nathans) Даниэль (р. 1928), американский вирусолог. Окончил университет штата Делавэр (1950), с 1972 г. — директор отдела микробиологии медицинского факультета университета Дж. Гопкинса в Балтиморе. Основные работы — по изучению структуры и функции генома вируса. Впервые с помощью рестрикционных эндонуклеаз провел картирование ДНК вируса обезьяны SV-40. Лауреат Нобелевской премии по физиологии и медицине (1978, совместно с Г. Смитом и В. Арбером).

Скорость расщепления различных фосфодиэфирных связей убывает в ряду $CpA > UpA \gg CpC \gg UpC > CpU > UpU$; предпочтительно гидролизуются однонитевые РНК.

Рибонуклеаза T_1 из *Aspergillus oryzae* осуществляет расщепление РНК по гуаниновым звеньям, образуя гуанозин-3'-фосфаты и олигонуклеотиды, имеющие на 5'-конце гидроксильную группу, а на 3'-концах — гуанозин-3'-фосфат:



Так же как и рибонуклеаза А, РНаса T_1 со значительно большей скоростью расщепляет одноцепочечные полинуклеотиды; скорость гидролиза убывает в ряду $GpC > GpA > GpG > GpU$.



Арбер (Arber) Вернер (р. 1929), швейцарский генетик. Образование получил в Цюрихе, профессор университета и Биоцентра в Базеле. Основные работы посвящены изучению структуры и функционирования вирусного генома. Открыл (1962) ферменты рестрикции и модификации. Лауреат Нобелевской премии по физиологии и медицине (1978, совместно с Д. Натансом и Г. Смитом).

Рибонуклеаза T_2 из *A. oryzae* при длительном гидролизе РНК расщепляет ее полностью до нуклеозид-3'-фосфатов. При частичном расщеплении РНК образуются преимущественно фрагменты, содержащие на 3'-концах аденозин-3'-фосфат.

Рибонуклеаза U_2 из *Ustilago sphaerogena* преимущественно расщепляет фосфодиэфирные связи, в которых донором 3'-ОН-группы являются пуриновые нуклеотиды, т. е. связи RpN ; скорость расщепления возрастает в ряду $U < C < G < A$.

Среди других ферментов, расщепляющих РНК, следует упомянуть рибонуклеазу III из *E. coli*, проявляющую абсолютную специфичность к двухцепочечным полирибонуклеотидам. Фермент дает олигонуклеотиды, оканчивающиеся 3'-фосфатом.

Ферменты, расщепляющие ДНК и РНК. Известно значительное число ферментов, неспецифичных к типу углеводного остатка в нуклеиновых кислотах. Так, фосфодиэстеразы, выделяемые из яда змей, гидролизуют олиго- и полинуклеотиды начиная с 3'-конца, путем последовательного отщепления нуклеозид-5'-фосфатов. Эти ферменты способны катализировать гидролиз одноцепочечных и со значительно меньшей скоростью двуспиральных полинуклеотидов. Напротив, фосфодиэстеразы из селезенки гидролизуют нуклеиновые кислоты начиная с 5'-конца, последовательно удаляя концевые звенья в виде нуклеозид-3'-фосфатов. Оба фермента гидролизуют рибо-, дезоксирибо-, олиго- и полинуклеотиды и используются для получения мононуклеотидов.

Большая группа ферментов, неспецифичных к типу углеводного остатка, весьма чувствительна к вторичной структуре и расщепляет только одноцепочечные полинуклеотиды. Наиболее широко используемыми ферментами такого типа являются нуклеаза S_1 из *A. oryzae*, нуклеаза из *Mung bean* и нуклеаза из *N. crassa*. Первые две из них расщепляют одноцепочечные ДНК и РНК в слабокислой среде в присутствии ионов цинка до олиго- и мононуклеотидов с фосфатной группой на 5'-конце; третья расщепляет полинуклеотиды в нейтральной и слабоосновной средах, давая 5'-фосфорилированные короткие олигонуклеотиды.

Эти нуклеазы широко используются для исследования вторичной структуры одноцепочечных участков. В генной инженерии они применяются для удаления односторонних концевых участков ДНК.

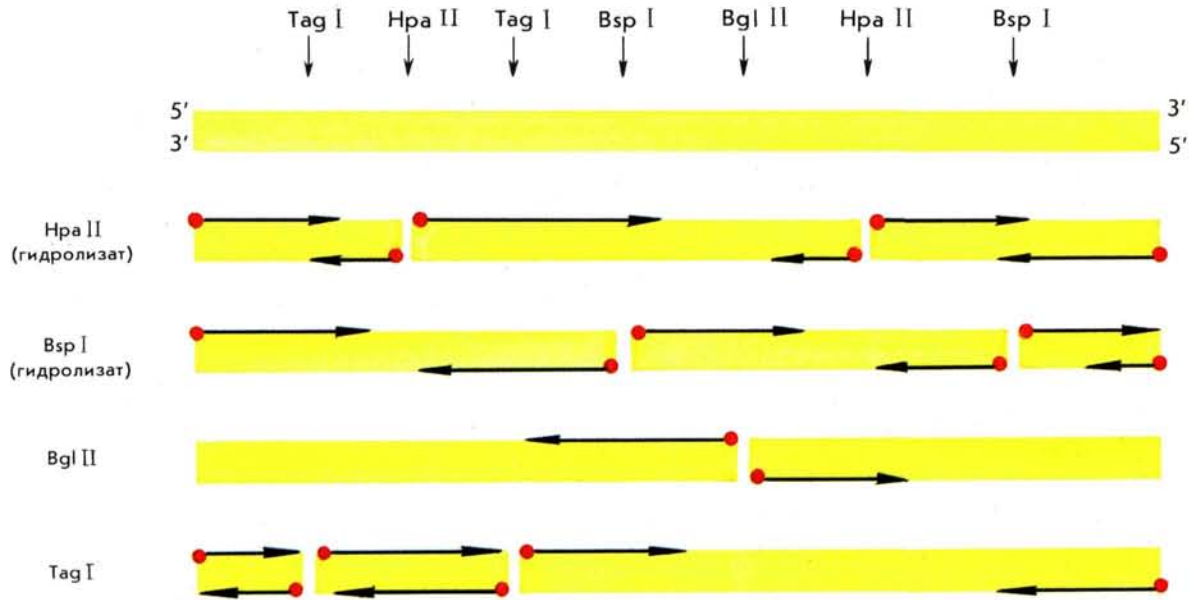
Существует большое число ферментов, способных отщеплять фосфатные группы от моно-, олиго- и полинуклеотидов. Некоторые из них обладают специфичностью в том отношении, что удаляют только 5'- или только 3'-фосфатные группы. Из неспецифичных фосфомоноэстераз, отщепляющих фосфаты вне зависимости от их положения как от моно-, так и от полинуклеотидов, широко используются две фосфомоноэстеразы. Одна из них, так называемая бактериальная щелочная фосфатаза, выделяется из *E. coli*, другая — из тонкого кишечника телят.

В блочном методе определения последовательности нуклеиновых кислот весьма существенна методология структурного анализа. В одном из возможных вариантов используется полное расщепление полинуклеотидной цепи двумя (или более) ферментами с различной специфичностью. В случае РНК это может быть осуществлено различными рибонуклеазами, например гидролиз последовательно рибонуклеазой T_1 , а затем — панкреатической рибонуклеазой. Структуры полученных олигонуклеотидов сравниваются для поиска перекрывающихся последовательностей. Если это оказывается недостаточным, используется третья РН-аза.

Для двухцепочечных ДНК перекрывающиеся блоки получают путем расщепления двумя или более различными рестрикционными эндонуклеазами. При этом часть исследуемой ДНК гидролизует одним ферментом и устанавливается структура полученных фраг-

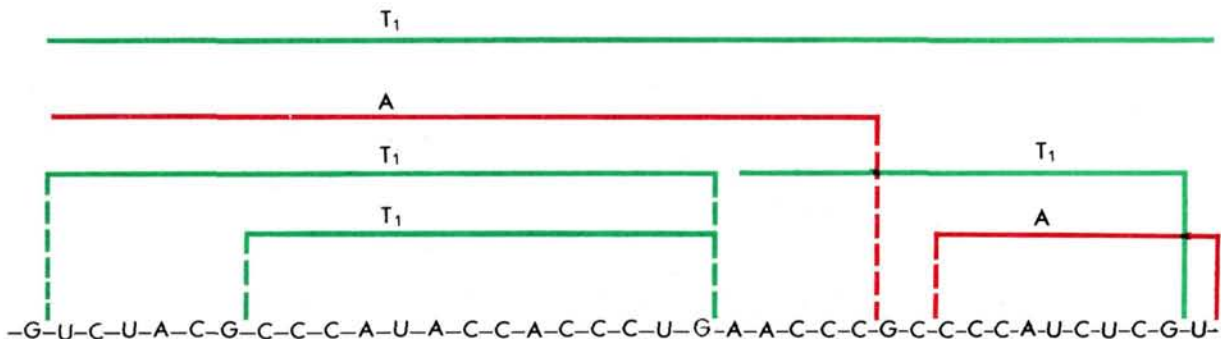
ментов, а затем те же операции повторяются на другой порции ДНК с другим ферментом. Поскольку требование перекрывания приходится сочетать с необходимостью получения сравнительно небольших фрагментов, иногда применяется расщепление несколькими рестрикционными эндонуклеазами одновременно. При этом обязательно устанавливать полную последовательность фрагментов — важно лишь найти «перекрытия». Процесс поиска перекрывающихся блоков изображается специальными схемами, как показано на рисунке 176.

Рис. 176. Схема применения метода перекрывающихся блоков для определения последовательности ДНК с использованием рестриктаз.



Фрагменты ДНК условно обозначены прямоугольниками, верхняя сторона которых представляет собой одну комплементарную цепь, а нижняя сторона — другую. Слева указаны эндонуклеазы, использованные для получения перекрывающихся блоков. Вверху дана рестрикционная карта определяемого фрагмента. Стрелками в прямоугольниках указывается длина расшифрованного участка и направление, в котором определялась последовательность (в данном случае она определялась в направлении от 5'-к 3'-концу).

Рис. 177. Перекрывающиеся блоки, получаемые методом неполных расщеплений РНК РН-азами А и Т₁.



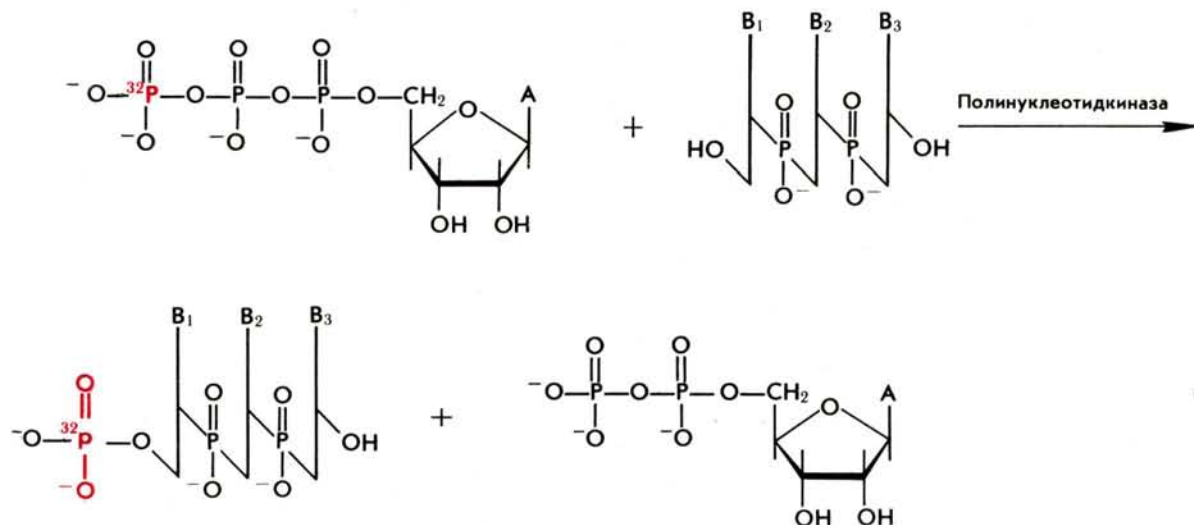
Как видно на схеме, установленные последовательности полностью перекрываются; более того, они определены фактически дважды — для верхней цепи и независимо для нижней: это служит гарантией правильности определения полной структуры и позволяет избежать ошибок.

Другой способ получения перекрывающихся блоков предполагает неполное расщепление исследуемого фрагмента (в частности, при помощи РНаз). Пример определения последовательности рибонуклеотида таким методом изображен на рисунке 177.

Неполные расщепления используются и в случае установления последовательности ДНК. Для этой цели может быть применена обработка как рестрикционными эндонуклеазами, так и неспецифичными эндонуклеазами (например, ДНазой I в присутствии ионов Mn^{2+}).

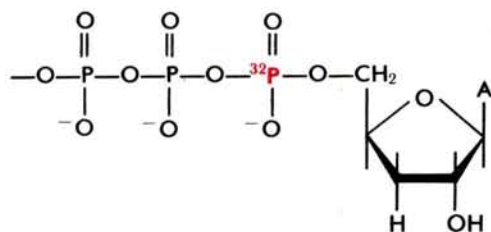
Определение строения олигонуклеотидов. Во всех современных способах определения первичной структуры нуклеиновых кислот первостепенную роль играют методы введения радиоактивных меток в 5'- и 3'-концевые звенья. Чаще всего роль концевой метки играет фосфатная группа, содержащая ^{32}P , но иногда в качестве метки используют также тритий (3H) или иод (^{125}I).

Для введения меченых фосфатных групп в 5'-концевые звенья олиго- и полинуклеотидов используется фермент полинуклеотидкиназа (см. с. 353); радиоактивным реагентом служит АТФ, содержащий меченый фосфор в γ -положении ($^{32}pppA$)

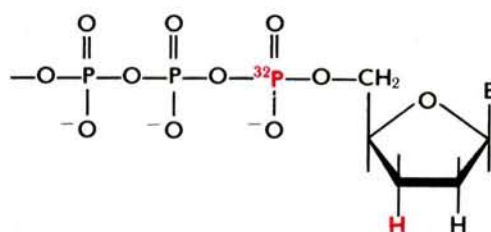


Методы включения метки в 3'-концевые группы более разнообразны. Так, в двухцепочечные ДНК метка вводится в составе $[\alpha\text{-}^{32}P]$ -дезоксирибонуклеозидтрифосфатов с помощью ДНК-полимеразы I E. coli или фага T4 (см. с. 348). Двух- и одноцепочечные ДНК могут быть помечены по 3'-концевому звену с использованием концевой нуклеотидилтрансферазы. В одном из вариантов этого метода к 3'-ОН-группе олигодезоксирибонуклеотида присоединяются $[\alpha\text{-}^{32}P]$ -меченые рибонуклеозидтрифосфаты. После ферментативной реакции продукт подвергается щелочному гидролизу, в результате чего на 3'-конце образуется только одно рибонуклеотидное звено. Во втором способе используются аналоги природных нук-

леозидтрифосфатов, не содержащие 3'-ОН-группы. Такие аналоги могут быть введены в 3'-концевое звено олигонуклеотида, но не способны присоединять к себе последующие звенья, поэтому их называют терминирующими. Обычно для этих целей применяются кордицепинтрифосфат или 2',3'-дидезоксинуклеозидтрифосфаты, содержащие $[\alpha\text{-}^{32}\text{P}]$ -меченую фосфатную группу:



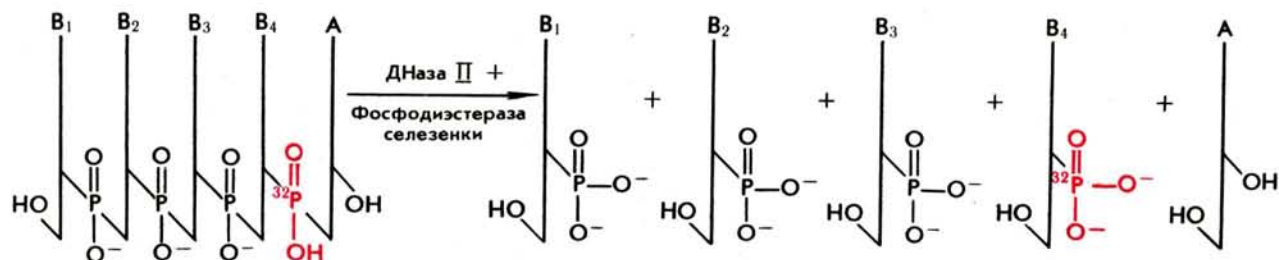
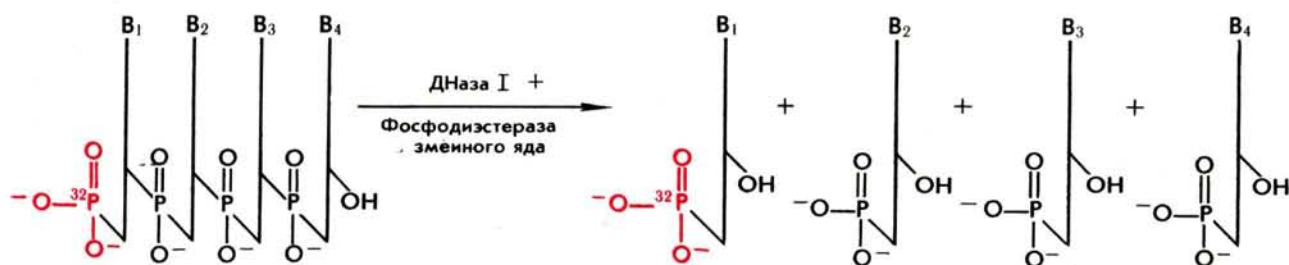
Кордицепин трифосфат



2', 3' – Дидезоксинуклеозидтрифосфат

Введение в 3'-концевое звено РНК 3',5'-дифосфатов, меченных ^{32}P по 5'-фосфатной группе, может быть осуществлено с помощью РНК лигазы, присоединяющей субстрат к 3'-ОН-группе. 3'-Конец РНК может быть маркирован также тритием с помощью химических методов: окислением периодатом и восстановлением бортригидом натрия. При этом используется РНК, не содержащая 3'-фосфатной группы.

Необходимым составным элементом анализа нуклеотидной последовательности во многих случаях является определение кон-



цевых групп. Это достигается введением метки (5'- или 3'-[³²P]-фосфата) в 5'- и 3'-концевые звенья олиго- или полинуклеотидов с последующим ферментативным гидролизом до соответствующих нуклеозидмонофосфатов. После разделения продуктов гидролиза идентифицируются меченые концевые звенья.

Для определения первичной структуры коротких олигонуклеотидов (до 15 — 20 звеньев) обычно используют метод «блуждающего пятна», или «сэнгерпринта». Меченный по одному из концов олигонуклеотид гидролизуется экзонуклеазой с противоположного конца в условиях неполного расщепления, так что образуется набор всех возможных фрагментов. При расположении полученных таким путем олигонуклеотидов в порядке возрастания длины каждые два соседних продукта будут отличаться друг от друга на одно концевое звено. Меченое же звено у всех олигонуклеотидов является общим (рис. 178).

Совершенно очевидно, что если разделить все полученные олигонуклеотиды и для каждого из них определить концевое звено, противоположное меченому, то после расположения всех фрагментов в порядке возрастания длины можно реконструировать последовательность исходного олигонуклеотида. Таким образом, проблема заключается в разделении продуктов и идентификации их концевых групп.

В методе «блуждающего пятна» эти две задачи решаются одновременно. Смесь меченых фрагментов подвергается двумерному разделению. Сначала оно проводится путем электрофореза на полосках ацетата целлюлозы при pH 3,5 (разделение по составу), а затем влажная полоска прикладывается к краю пластинки с ДЭА целлюлозой, которая сорбирует частично разделенные олигонуклеотиды, и разделение ведется в направлении, перпендикулярном первому (разделение по длине). В качестве элюента используется гидролизат РНК, содержащий олигонуклеотиды всех возможных длин и составов (такая процедура называется гомохроматографией).

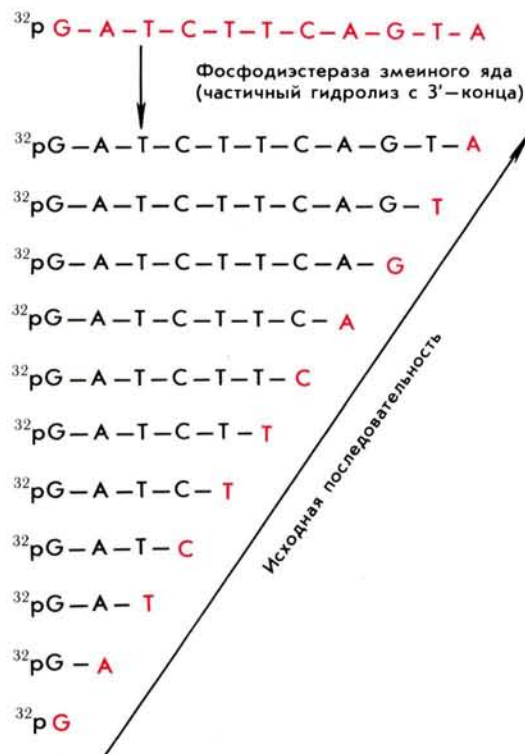


Рис. 178. Фрагменты частичного гидролиза 5'-меченого олигонуклеотида фосфодиэстеразой змеиного яда.

Результат разделения фиксируется радиоавтографией: для этого к пластинке прикладывается рентгеновская пленка, и те участки, которые контактируют с радиоактивными продуктами, засвечиваются. В результате каждому фрагменту соответствует отдельное пятно на пленке и фактически анализируется относительное расположение полученных пятен.

Если сравнивать подвижность каждого олигонуклеотида при электрофорезе с подвижностью получаемого из него продукта, укороченного на одно звено, то наблюдаются следующие закономерности: отщепление рТ приводит к уменьшению подвижности продукта по сравнению с исходным олигонуклеотидом вследствие резкого уменьшения заряда. Отщепление рG вызывает качественно такой же, но количественно меньший эффект. Отщепление рА, уменьшая длину олигонуклеотида и в значительно меньшей степени (по сравнению с рТ и рG) заряд, приводит к некоторому увеличению подвижности. Наконец, отщепление рС очень мало меняет заряд, но резко уменьшает длину, что существенно увеличивает подвижность продукта.

При гомохроматографии разделение происходит в соответствии со следующими закономерностями: чем больше длина олигонуклеотида, тем он менее подвижен, так что каждые два соседних пятна представляют собой олигонуклеотиды, отличающиеся на одно звено; если два соседних пятна получены в результате отщепления пуринового нуклеотида, то расстояние между ними по вертикали больше, чем при отщеплении пиримидинового нуклеотида.

В качестве примера рассмотрим анализ радиоавтограммы, приведенной на рисунке 179. Очевидно, что наименее подвижный при гомохроматографии компонент представляет собой исходный олигонуклеотид. Соседний с ним продукт образуется путем отщепления 3'-концевого звена: при электрофорезе подвижность его слегка увеличивается, а при гомохроматографии — резко возрастает. Отсюда делается вывод, что отщепленным звеном является А. Следующий продукт (3) отличается от продукта (2) также одним звеном. Он значительно менее подвижен при электрофорезе и несколько более подвижен при гомохроматографии, следовательно, (3) образуется из (2) в результате отщепления Т. Продолжая такой анализ, определяют звено, которым отличается друг от друга каждая пара соседних пятен, т. е. концевые звенья каждого олигонуклеотида смеси. «Читая» эти концевые нуклеотиды, можно восстановить исходную последовательность.

Используемые в методе «блуждающего пятна» правила схематически представлены в рамке на рисунке 179.

Отщепление Т резко уменьшает подвижность при электрофорезе и мало увеличивает при гомохроматографии.

Отщепление G незначительно уменьшает подвижность при электрофорезе и сильно увеличивает при гомохроматографии.

Отщепление А незначительно увеличивает подвижность при электрофорезе и сильно — при гомохроматографии.

Отщепление С сильно увеличивает подвижность при электрофорезе и незначительно — при гомохроматографии.

С помощью метода «блуждающего пятна» могут определяться последовательности олигонуклеотидов, содержащих до 20 звеньев. Одним из наиболее существенных его недостатков является недостоверность определения последовательности двух или трех звеньев, соседних с меченым звеном.

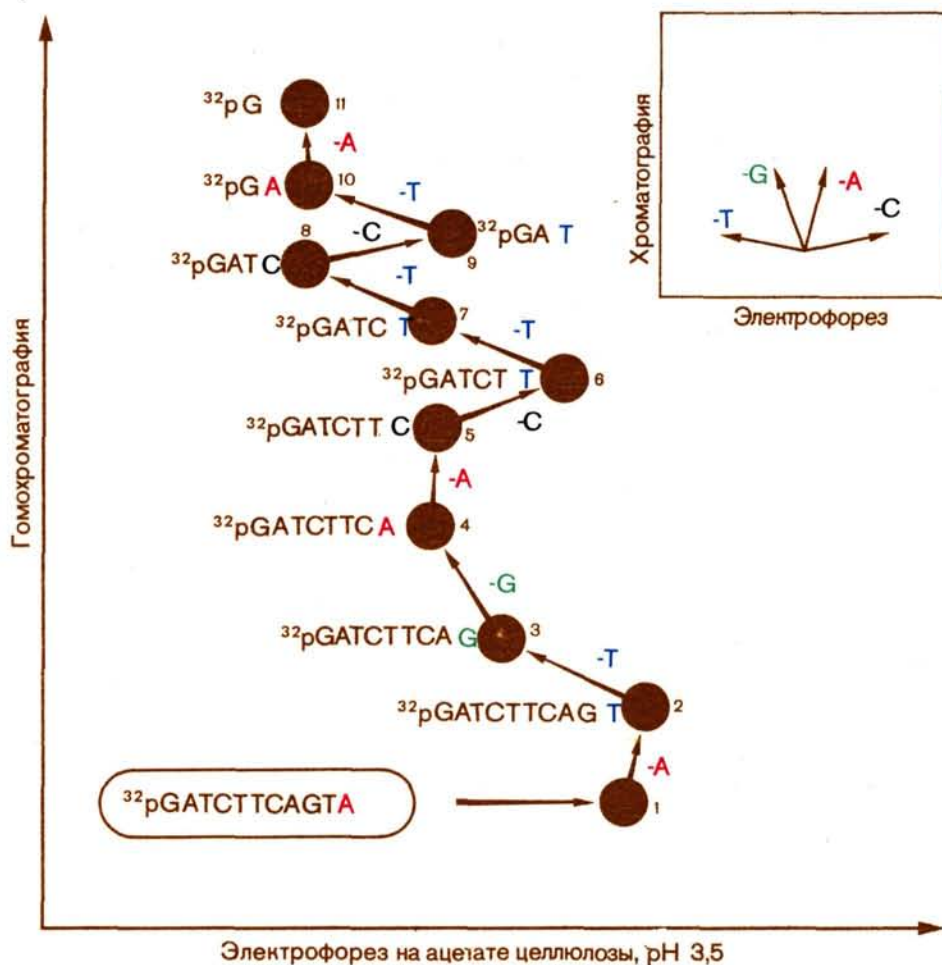
Современные методы определения последовательности фрагментов нуклеиновых кислот основаны на одном общем принципе. Этот принцип заключается в сравнении длин всех возможных концевых продуктов, полученных из исходного полимера таким образом, что все они имеют на одном конце одну и ту же последовательность, а на другом — один и тот же нуклеотид.

В таблице 12 представлены наборы концевых продуктов разной длины, имеющих одно и то же 5'-концевое звено (G), а оканчивающихся аденозином (колонка А), гуанозином (колонка G), цитидином (колонка С) или тимидином (колонка Т). В колонке А находятся пентануклеотид, ундека- и тридекануклеотид, длины которых совпадают с положением А в исходном полинуклеотиде. Аналогично в колонке С длины фрагментов (три-, гекса-, нона-, дека-, тетрадека- и пентадека-) соответствуют положению С в исходном нуклеотиде. То же справедливо для соединений, оканчивающихся G и Т.

Следовательно, если для полинуклеотида получить полный набор концевых продуктов, специфически «оборванных» у определенного типа звеньев, то из анализа их длины легко определить положение данного звена в исходном полинуклеотиде. Проводя подобную операцию для каждого из четырех нуклеотидных звеньев, можно вывести полную последовательность исследуемого полимера.

Однако определение абсолютных длин фрагментов — операция весьма затруднительная, и поэтому современные методы анализа заменяют его сравнением длин продуктов, оканчивающихся разными типами звеньев. Если нуклеотиды (табл. 12) записывать в последовательности увеличения их длины на одно звено начиная с самого короткого продукта (колонка G, затем колонки Т, С, снова Т, А и т. д.), то получается формула GTCTACGGCCATACC, соот-

Рис. 179. Двумерное разделение смеси радиоактивных нуклеотидов.



ветствующая последовательности исходного олигонуклеотида. При таком рассмотрении учитываются не абсолютные значения длин олигонуклеотидов, сравнение ведется по принципу «короче — длиннее».

Таким образом, для полного анализа последовательности любого полинуклеотида необходимо иметь, во-первых, метод получения четырех наборов специфических концевых продуктов, подобных тем, которые приведены в таблице 12, и, во-вторых, метод, позволяющий проводить сравнение длин этих продуктов. Различие современных методов анализа заключается в способах получения наборов специфических фрагментов. Общность — в способе сравнения длин, которое во всех случаях производится путем электрофореза радиоактивных концевых продуктов в пластинках полиакриламидного геля. По окончании электрофореза положение продуктов в геле определяют путем радиоавтографии. Каждый продукт проявляется в виде темной полосы на рентгеновской пленке; сравниваются относительные подвижности полос в разных дорожках, подобно тому как сравнивались длины продуктов в таблице 1. Картина распределения полос на рисунке 180 соответствует последовательности олигонуклеотидов в таблице 12. Это следует из анализа относительных подвижностей продуктов: самый подвижный и, следовательно, самый короткий продукт обнаруживается в дорожке G, следующий по подвижности и длине — в колонке T, далее — в колонке C, следующие два — в колонках T и A соответственно и т. д. Выписывая таким образом названия колонок в порядке, в котором в них обнаруживаются последовательно удлиняющиеся продукты, получают формулу исходного олигонуклеотида.

Метод неполных специфических химических расщеплений (метод Максама—Гилберта) используется для определения последовательности ДНК. Меченые концевые продукты получают путем специфического расщепления полинуклеотида химическими реагентами по одному из звеньев. Обработка проводится в условиях неполной «статистиче-

Таблица 12.

Схема определения последовательности полинуклеотидов

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
5' р GTCTACGGCCATACTACCC 3'

Длина концевого продукта	Концевые продукты, заканчивающиеся звеном				Последовательность, выведенная из анализа
	A	G	C	T	
1	pG				5'
2				pGT	G
3			pGTC		T
4				pGTCT	C
5	pGTCTA				T
6			pGTCTAC		A
7		pGTCTACG			C
8		pGTCTACGG			G
9			pGTCTACGGC		G
10			pGTCTACGGCC		C
11	pGTCTACGGCCA				C
12				pGTCTACGGCCAT	A
13	pGTCTACGGCCATA				T
14			pGTCTACGGCCATAC		A
15			pGTCTACGGCCATACC		C
					3'

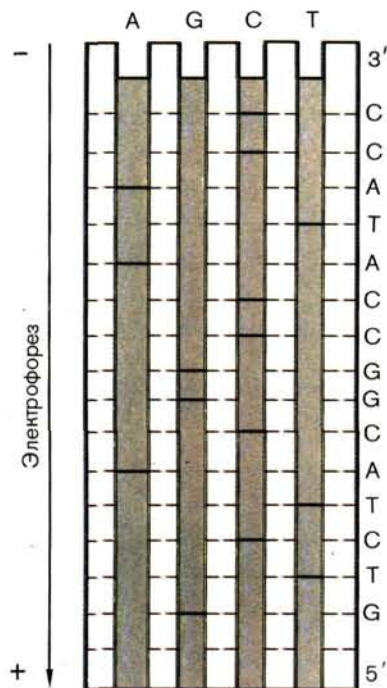


Рис. 180. Схематическое изображение результатов разделений в полиакриламидном геле 4 серий концевых продуктов, представленных в таблице 12.

ской» модификации. В результате получается большой набор различных продуктов, но только концевые олиго- и полинуклеотиды являются мечеными и обнаруживаются на радиоавтограммах. Так, например, при частичном расщеплении по тимидиновым звеньям (в положениях, указанных стрелками) меченого по 5'-концу олигонуклеотида $^{32}\text{pG} - \text{T} - \text{p} - \text{C} - \text{p} - \text{T} - \text{p} - \text{A} - \text{C}$ могут образоваться следующие олигонуклеотиды: $^{32}\text{pGpT}$, $^{32}\text{pGpTpCpT}$, pCpTpApC , pApC , pCpT , из которых мечеными оказываются только два концевых. Для специфических расщеплений используются реакции, приводящие к удалению гетероциклического основания из полинуклеотидной цепи. После этого цепь может быть расщеплена по образующемуся дезоксирибозильному звену или его производному путем β -элиминирования действием щелочи или аминов.

Далее рассматриваются четыре реакции, используемые в одном из вариантов метода. В качестве объекта выбран олигонуклеотид, приведенный в таблице 12.

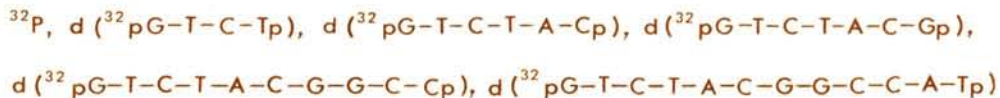
А. Расщепление ДНК по гуанозиновым звеньям осуществляется неполной модификацией ДНК диметилсульфатом. Метилирование затрагивает атом азота в положении (7) в G, а также N(1) и N(3) положения в A. Продукт неполной модификации обрабатывается пиперидином при повышенной температуре. При этом N-гликозидная связь разрывается и полинуклеотидная цепь подвергается фрагментации с образованием продуктов, содержащих на 5'- и 3'-концах фосфатные группы (рис. 181).

В этих условиях не происходит расщепления полинуклеотидной цепи по метилированным производным аденозина. Из модельного нуклеотида в результате реакции должны получиться следующие меченые олигонуклеотиды, образовавшиеся при расщеплении по G:



Эти фрагменты отличаются от продуктов, приведенных в колонке G таблицы 12, лишь тем, что на месте гуаниновых остатков содержатся только фосфатные группы.

Б. Ввиду отсутствия специфического метода расщепления ДНК по остаткам аденина применяется неполное расщепление цепи по пуриновым звеньям. Сравнительный анализ результатов двух типов расщепления (суммарно по пуриновым остаткам и только по гуанинам) дает возможность дифференцировать олигонуклеотиды, являющиеся продуктами расщепления по А. Для фрагментации по остаткам пурина используется лабильность N-гликозидной связи пуриновых дезоксирибозидных производных в кислой среде. ДНК подвергается кратковременной обработке муравьиной кислотой, в результате которой часть пуриновых оснований удаляется, и полинуклеотидная цепь затем расщепляется по образующимся дезоксирибозильным остаткам под действием пиперидина. Из модельного олигонуклеотида при этом должны образоваться следующие меченые соединения:



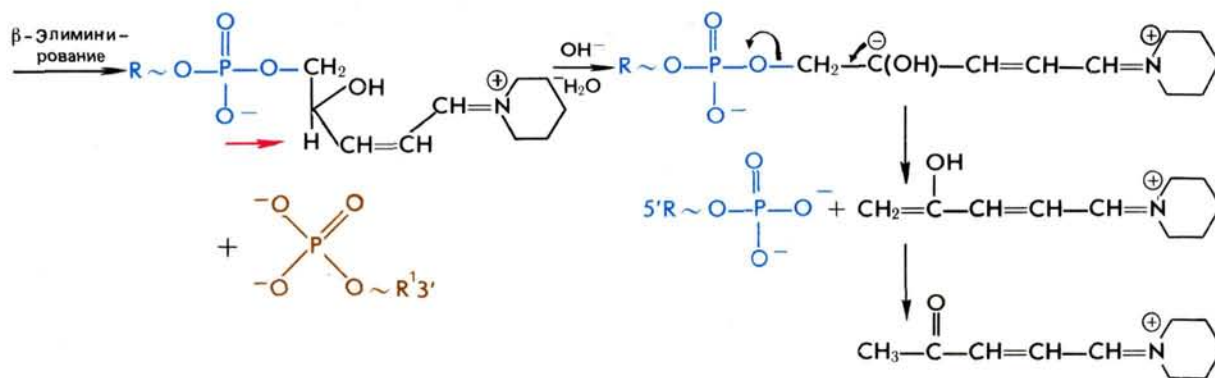
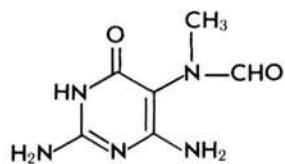
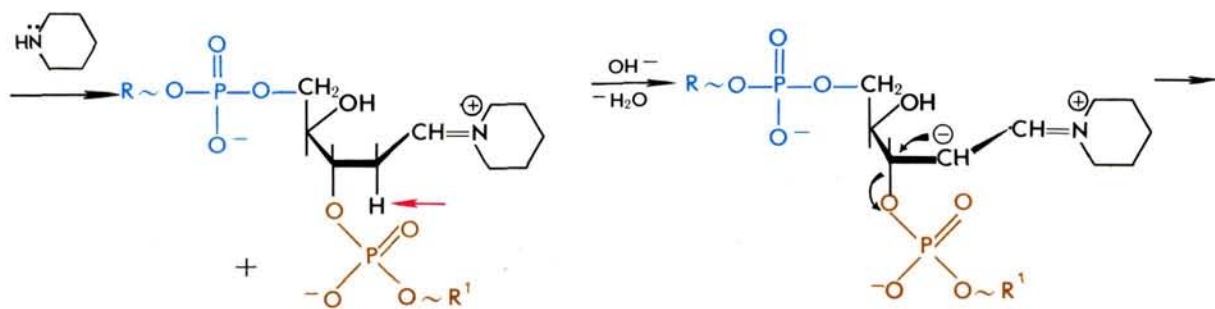
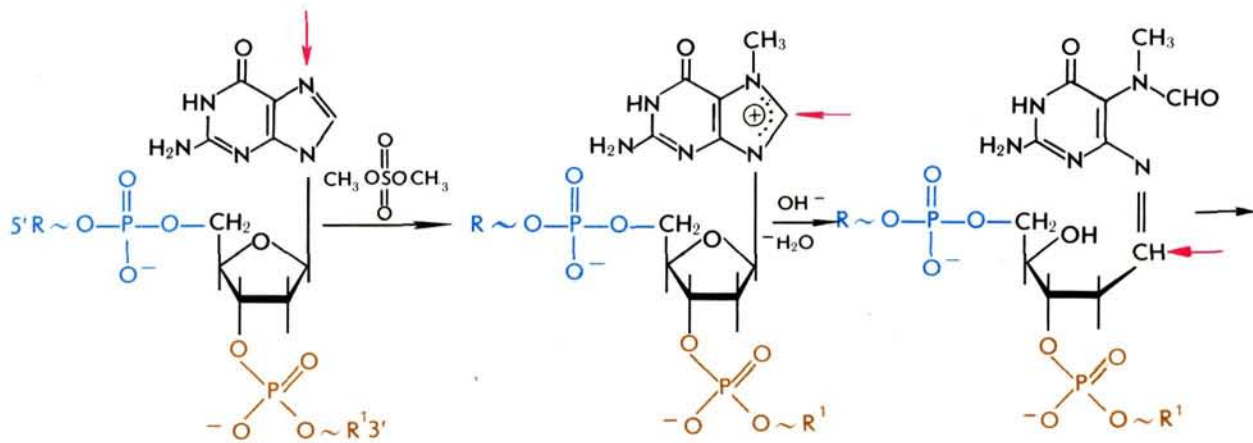


Рис. 181. Химическая модификация по гуанозиновым звеньям.

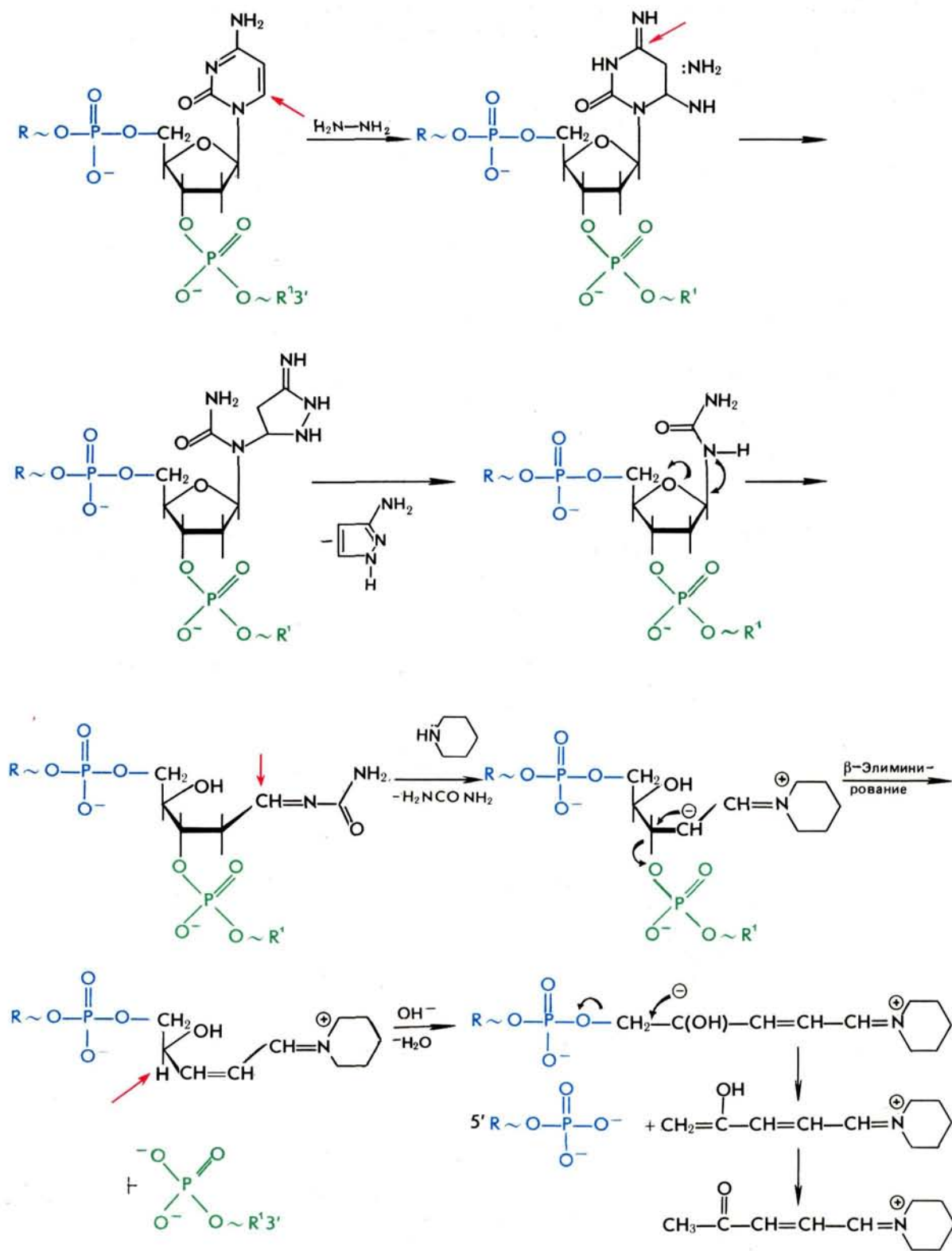
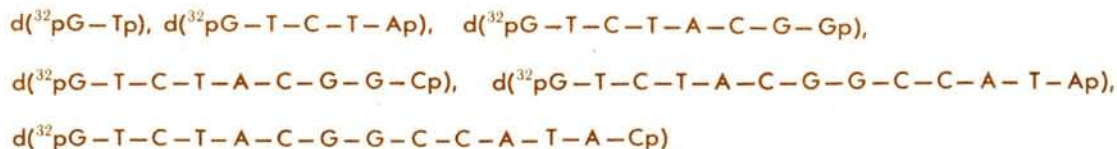


Рис. 182. Химическая модификация по цитозинным звеньям.

В. Расщепление по цитозину осуществляется гидразином при высокой концентрации NaCl с последующей обработкой продуктов модификации пиперидином (рис. 182), что приводит к образованию следующих олигонуклеотидов:



Г. Для определения положения тимидина используется деградация по остаткам пиримидинов, для чего ДНК последовательно обрабатывается гидразином при низкой концентрации NaCl, а затем пиперидином. Из модельного олигонуклеотида при этом, кроме соединений, приведенных выше, должны образоваться:



Продукты всех четырех реакций наносят в соседние лунки полиакриламидного геля, разделяют электрофорезом и радиоавтографируют. Анализ результатов, схематически представленных на рисунке 183, приводит к следующим выводам: самый «короткий» фрагмент (фосфат) находится в колонках G и A + G и, вероятно, образовался в результате расщепления по гуанозину; наличие полосы в колонке T + C и ее отсутствие в колонке C свидетельствует о том, что это продукт расщепления по тимидину; следом идет олигонуклеотид, присутствующий и в колонке C, и в колонке T + C, т. е. получившийся в результате расщепления по цитозину. Затем по подвижности располагаются продукты в колонке T + C (но не в C), A + G (но не в G); продолжая этот анализ, восстанавливаем структуру исходного олигонуклеотида.

В приложении к реальным полинуклеотидам метод дает возможность «прочесть» последовательность свыше 200 нуклеотидных звеньев на одном геле.

Проблемой, которую каждый раз приходится решать при использовании метода Максама — Гилберта, является получение полинуклеотидов, меченных только по одному из концов. Обычно для анализа используются двухцепочечные фрагменты ДНК (например, образовавшиеся после расщепления рестриктазами), а все методы введения концевых меток приводят к образованию фрагментов, меченных по двум 5'- или двум 3'-концам. Поэтому приходится прибегать к дополнительным операциям, таким, как разделение комплементарных цепей ДНК после введения метки.

Метод ДНК-полимеразного копирования в присутствии терминирующих аналогов трифосфатов (метод Сенгера). Этим методом чаще всего анализируются фрагменты ДНК, клонированные в одноцепочечных фагах, хотя в настоящее время разработаны варианты, применимые и к двухцепочечным ДНК. При этом исследуется не сам фрагмент, а комплементарная ему ДНК, которая синтезируется с помощью ДНК-полимеразы в условиях, приводящих к набору продуктов разной длины, содержащих одинаковый 5'-концевой участок, а на

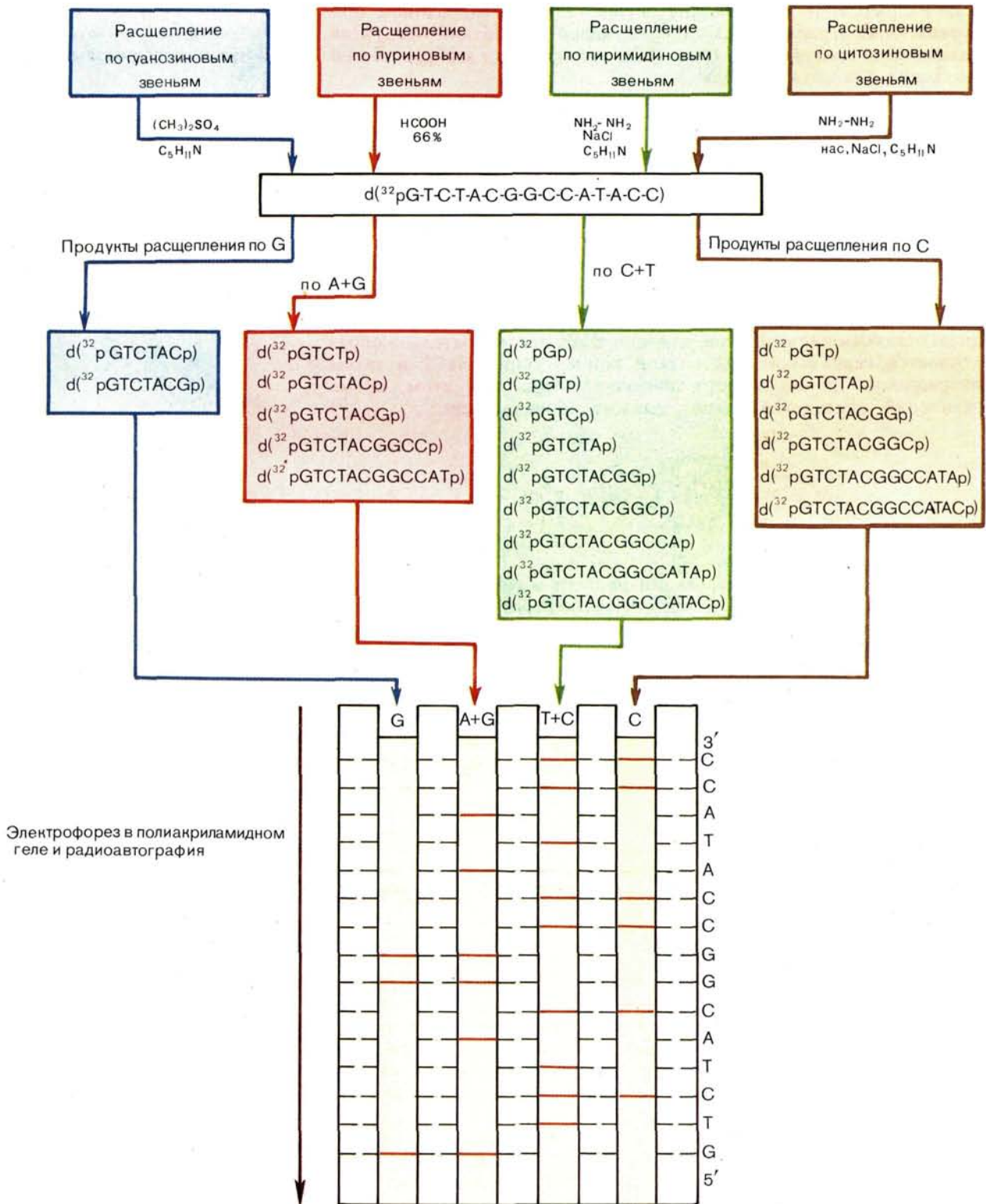


Рис. 183. Схема определения нуклеотидной последовательности (метод Максама — Гилберта).

3'-конце — звенья определенного типа (А, G, С или Т). Для получения такого набора фрагментов к смеси четырех природных дезоксирибонуклеозидтрифосфатов, используемых ДНК-полимеразой для синтеза копии, добавляют трифосфат аналога того нуклеозида, который желательнее иметь в качестве 3'-концевого звена. Этот аналог отличается от соответствующего природного субстрата лишь строением углеводного компонента, что позволяет ДНК-полимеразе включать его в цепь, но не дает возможности присоединять последующие звенья. Чаще всего в качестве таких терминирующих аналогов используют трифосфаты 2',3'-дидезоксинуклеозидов. Так, для получения набора фрагментов, оканчивающихся тимидином, в смесь для полимеризации добавляют 2',3'-дидезокситимидинтрифосфат; для обрыва по аденозину — 2',3'-дидезоксиаденозинтрифосфат и т. д. При введении в реакцию любого аналога ДНК-полимераза с определенной вероятностью включает его в цепь вместо природного субстрата, после чего синтез цепи обрывается. Включение аналога происходит по закону случая в любом месте синтезируемых цепей; в итоге получаются наборы олиго- и полинуклеотидов различной длины, но с одинаковыми звеньями на 3'-концах, как это показано на рисунке 184.

Для получения продуктов с уникальным 5'-концом синтез ДНК начинают с помощью затравки — олиго- или полинуклеотида, комплементарного матрице на участке, расположенном рядом с 3'-концевой областью матрицы. Затравка содержит свободную 3'-ОН-группу, необходимую ДНК-полимеразе для начала синтеза ДНК. Гибридуясь с матрицей в определенном месте и включаясь во вновь синтезируемую ДНК в качестве ее 5'-концевого участка,



Гилберт (Gilbert) Уолтер (р. 1932), американский биохимик. Окончил Гарвардский университет (1953); с 1968 г. — профессор этого университета, затем президент фирмы «Биоген». Выполнял основополагающие исследования по изучению механизма специфического взаимодействия белков и ДНК, установлению первичной структуры ДНК, предложил (1977, совместно с А. Максамом) метод расшифровки первичной структуры ДНК. Лауреат Нобелевской премии по химии (1980, совместно с Ф. Сенгером и П. Бергом).

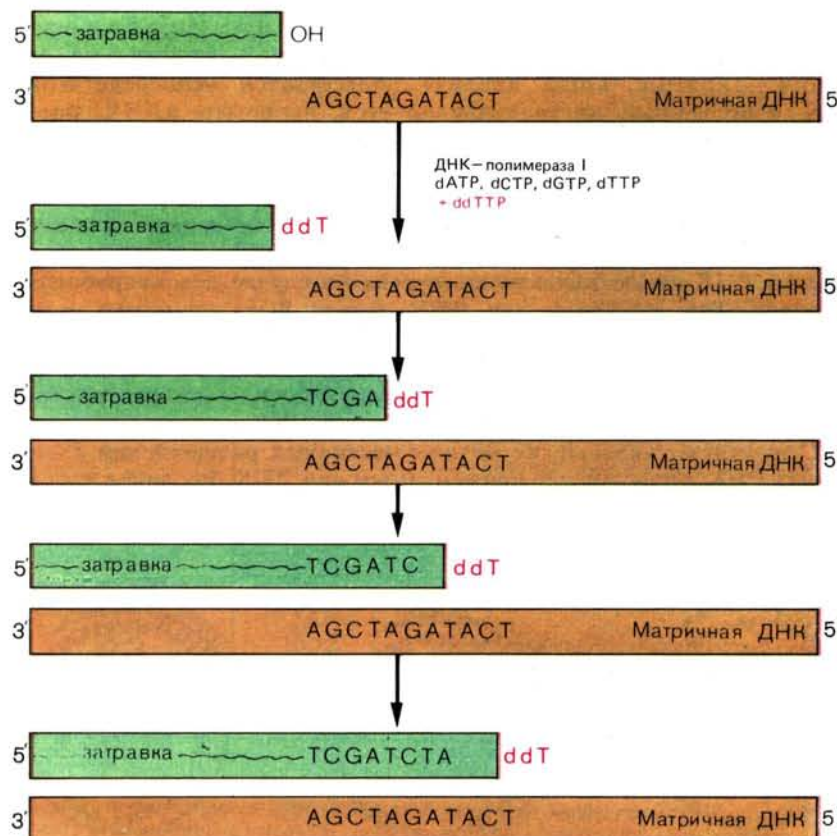


Рис. 184. Пример специфического обрыва синтезируемой полинуклеотидной цепи в присутствии терминирующего аналога 2,3-дидезокситимидинтрифосфата.

затравка обеспечивает уникальность этого звена для всех наборов специфически оборванных копий.

Для того чтобы синтезируемые продукты были радиоактивными, в качестве одного из четырех природных дезоксирибонуклеозидтрифосфатов используется $[\alpha^{32}\text{P}]$ -меченый продукт. С этой целью смесь затравки и матрицы разделяется на четыре части: с одной проводится синтез в присутствии ddTTP с образованием фрагментов, терминируемых тимидином, с другой — в присутствии ddATP, с третьей — в присутствии ddGTP и с последней — в присутствии ddCTP. Все четыре радиоактивных набора продуктов наносятся на параллельные дорожки полиакриламидного геля, разделяются, радиоавтографируются и анализируются так же, как в методе Максама — Гилберта.

Как уже упоминалось, чаще всего в качестве матрицы для анализа используются фрагменты ДНК, клонированные в одноцепочечных фагах. Для унификации анализа последовательности любых фрагментов, клонированных в составе данного фага-вектора, можно использовать одну и ту же затравку. Для этого затравка должна быть комплементарна не клонированному фрагменту, а (+)-цепи фага вектора в том ее месте, которое граничит с 3'-концом клонированной вставки (рис. 185). В качестве затравок чаще всего используются синтетические олиго- и дезоксирибонуклеотиды. Для исключения дегградации затравки с 5'-конца в качестве ДНК-полимеразы применяется обычно «фрагмент Кленова» ДНК-полимеразы I *E. coli*.

Метод Сенгера наиболее часто используется для анализа последовательности ДНК после ее неспецифического расщепления и клонирования суммы получаемых фрагментов. Метод весьма эффективен и экономичен, так же как метод Максама — Гилберта, и позволяет анализировать последовательность свыше 200 звеньев на одном геле.

Методы быстрого определения последовательности РНК. Эти методы могут быть разбиты на две группы — прямые, когда анализу подвергается непосредственно РНК, и косвенные, когда анализируется, например, кДНК, полученная путем копирования РНК с помощью обратной транскриптазы и клонированная в составе какого-либо вектора, или соответствующий ген, кодирующий данную РНК после выделения его из библиотеки генов. Косвенные методы, естественно, являются методами анализа последовательности ДНК; прямые же методы анализа РНК аналогичны используемым в случае дезоксирибонуклеотидов. Так же как и для ДНК, может быть применен метод копирования с терминирующими аналогами трифосфатов, но вместо ДНК-полимеразы используют обратную транскриптазу, а в качестве затравки — синтетические олигонуклеотиды или рестриктивные фрагменты, комплементарные данной РНК.

Существуют химические методы частичных расщеплений 3'-меченных РНК, аналогичные применяемым для ДНК. Различие заклю-

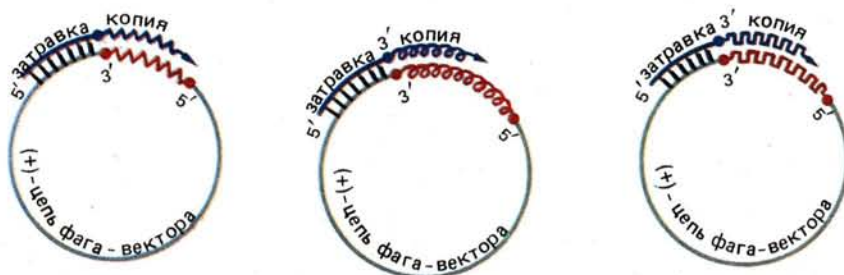


Рис. 185. Иллюстрация использования одной и той же затравки с данным вектором для анализа последовательности различных фрагментов.

чается в том, что при расщеплении модифицированной РНК по образующимся рибозильным звеньям или их производным нельзя использовать щелочные условия, поскольку происходит неспецифический гидролиз РНК. Поэтому расщепление проводят в слабокислой среде в присутствии аминов. Для модификации применяют диметилсульфат (G), гидразин + NaCl (C), гидразингидрат (U) и диэтилпирокарбонат, который модифицирует A и G, но преимущественно A. Однако наиболее широко распространены методы неполных расщеплений РНК специфическими рибонуклеазами. Как и в случае метода Максама — Гилберта, РНК маркируют с помощью полинуклеотидкиназы для 5'-конца или РНК-лигазы для 3'-конца. Затем меченую РНК разделяют на несколько порций и каждую порцию подвергают неполному расщеплению одной из специфических РНаз. Продукты, полученные из каждой порции, наносят в соседние лунки плоского полиакриламидного геля, разделяют и радиоавтографируют. Недостатком метода является малая специфичность РНаз к данному типу оснований, а также зависимость скорости расщепления от типа основания. Вследствие этого приходится использовать набор различных РНаз с перекрывающейся специфичностью.

Некоторые из применяемых для частичных расщеплений РНаз приведены в таблице 13.

Быстрые методы определения последовательности не решают всех задач, стоящих перед исследователем первичной структуры нуклеиновых кислот, поскольку они не дают информации о положении и природе минорных компонентов нуклеиновых кислот. Такие задачи встречаются при изучении структуры тРНК. В этих случаях сохраняют свое значение «старые» методы определения последовательности, заключающиеся в исчерпывающем или частичном гидролизе рибонуклеазами, выделении индивидуальных олигонуклеотидов, определении их нуклеотидного состава и идентификации минорных компонентов.

Таблица 13.

Рибонуклеазы, используемые для анализа последовательности РНК

РНаза	Специфичность	Применение
T2	Неспецифична	Для получения статистического набора олигонуклеотидов всех возможных длин
T1	G	Для анализа распределения гуаниловых остатков
U2	A > G	Для определения положения A остатков (разработаны условия расщепления исключительно по связям Ap↓N ¹)
A	U + C	Для анализа последовательности используется редко ввиду резкого различия в скоростях расщепления связей (Pu↓Pu ≫ Pu↓Pu)
Vc (пиримидил-РНаза из <i>Bacillus cereus</i>)	U + C	Для определения положения пиримидиновых звеньев
PhyM (из <i>Physarum polycerphalum</i>)	A + U	Широко используется как РНаза с перекрывающейся специфичностью для проверки расщеплений по A и для отличия U от C

¹ РНазы, выпускаемые некоторыми фирмами специально для анализа последовательности нуклеотидов в РНК (sequence grade), в определенных условиях проявляют исключительно высокую специфичность по отношению к основаниям.



Баев Александр Александрович (р. 1904), советский биохимик, академик АН СССР (1970). Окончил Казанский университет (1927), с 1959 г.— заведующий лабораторией функциональной энзимологии Института молекулярной биологии АН СССР и отделом молекулярной биологии и молекулярной генетики микроорганизмов Института биохимии и физиологии микроорганизмов АН СССР. Основное направление исследований — строение нуклеиновых кислот и генная инженерия. Установил (1967) первичную структуру валиновой тРНК дрожжей. Разработал метод изучения функциональной топологии биополимеров (метод разрезанных молекул). Герой Социалистического Труда (1981), лауреат Государственной премии СССР (1969).

Для иллюстрации применяемых здесь методов и приемов рассмотрим определение структуры одного из гексануклеотидов, полученных при исчерпывающем гидролизе валиновой тРНК РНазой Т1 (А. А. Баев и сотр.).

А. Определение нуклеотидного состава, проведенное после щелочного гидролиза, дало результат:



Б. Из способа получения (гидролиз гуанилрибонуклеазой) было ясно, что Gr должно находиться на 3'-конце.

В. Исчерпывающий гидролиз фосфодиэстеразой змеиного яда дал единственный нуклеозид — цитидин. Отсюда следует, что олигонуклеотид содержит на 5'-конце цитидин и, таким образом, его структура С(А, U, ψ, С)Gr.

Г. Полный гидролиз пиримидил-РНазой привел к следующим продуктам: 2Cr + Arψr + Ur + Gr и, значит, его структура С(А, ψ, U, С)Gr.

Д. При частичном гидролизе пиримидил-РНазой был выделен пентануклеотид С(U, ψ, А, С). Дефосфорилирование и последующий исчерпывающий гидролиз последнего панкреатической РНазой обнаруживают среди нуклеотидов один нуклеозид — уридин, который, следовательно, находится на 3'-конце пентануклеотида. Следовательно, остаются возможными только две структуры гексануклеотида: С — А — ψ — С — U — Gr или С — С — А — ψ — U — Gr. Выбор между ними был сделан на том основании, что среди продуктов частичного гидролиза гексануклеотида панкреатической РНазой был обнаружен динуклеотид CrUr, а это возможно лишь в том случае, если верна первая из двух структур.

В заключение раздела, посвященного анализу последовательности нуклеиновых кислот, следует отметить, что новые методы обеспечили возможность полностью расшифровать строение ряда простейших геномов, к которым относятся бактериофаги φX174 (5255 звеньев), G-4 (5577 звеньев), T7 (39 936 п. о.), λ (48 592 п. о.), некоторых других фагов и вируса обезьян SV-40 (5226 п. о.), больших участков генома бактерий, животных, растений и т. п. Эти результаты заставили по-новому взглянуть на структуру и функцию генома и на его эволюцию. И тем не менее сегодня в середине 80-х годов расшифрована еще только очень незначительная часть генетической информации. Общая длина расшифрованных последовательностей составляет всего лишь несколько миллионов нуклеотидных звеньев, а это — только 0,001 длины генома человека.

Пространственная структура нуклеиновых кислот

Если мономерной единицей нуклеиновых кислот считать нуклеотидное звено, то для описания его конформации следует учитывать возможность вращений вокруг шести связей (рис. 186). Кроме того, надо принимать во внимание положение основания относительно углеводной части, которое определяется вращением вокруг N-гликозидной связи, и конформационные изменения фуранозного цикла. Все это делает конформационный анализ нуклеиновых кислот весьма сложным.

Конформации нуклеиновых кислот и их компонентов, так же как и белков (см. с. 87), часто описываются с помощью так называемых торсионных (двугранных) углов. При определении торсионного угла рассматриваются три смежные связи, например А — В, В — С и С — D (рис. 187). Торсионным углом вращения связи В — С называют угол между проекциями связей А — В и С — D на плоскость, перпендикулярную связи В — С. Если торсионный угол равен $\sim 180^\circ$, то это соответствует *транс*-конформации (t-конформация). При значениях торсионных углов около $+60^\circ$ или -60° конформации называются *гош*⁺ и *гош*⁻ (*g*⁺ и *g*⁻).

На рисунке 186 приведен фрагмент нуклеотидной цепи: жирной линией выделен ее остов, который включает в себя атомы О(5'), С(5'), С(4'), С(3'), О(3') и Р. Остальные атомы углеводных остатков и гетероциклические основания рассматриваются как боковые группы. В основной цепи можно выделить повторяющуюся единицу. Если выбрать в качестве начала отсчета атом С(3') основной цепи, тогда такой единицей будет последовательность атомов С(3')^α—О(3')^β—Р^γ—О(5')^δ—С(5')^ε—С(4')^ζ—С(3'), включающая в себя шесть различных связей α, β, γ, δ, ε, ζ. Соответствующие торсионные углы обозначают так же. Выбор начала отсчета может быть и другим, но использованный здесь является одним из наиболее простых. Повторяющаяся единица основной полинуклеотидной цепи называется нуклеотидной единицей.

Углеводные компоненты. Проблема пространственного строения углеводов обсуждается в главе «Углеводы», однако здесь целесообразно рассмотреть вопросы, непосредственно касающиеся нуклеиновых кислот. Фуранозное кольцо играет определяющую роль

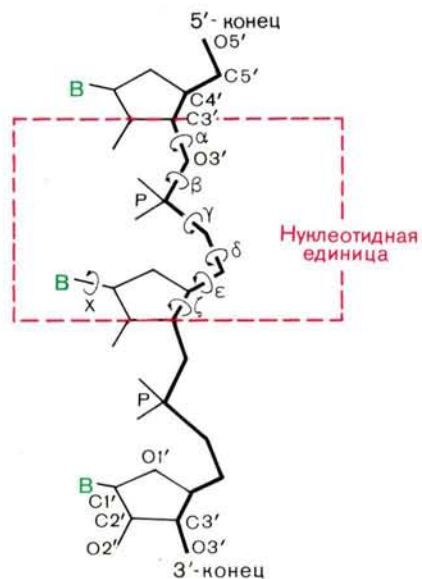


Рис. 186. Фрагмент нуклеотидной цепи. Нуклеотидная единица выделена рамкой.

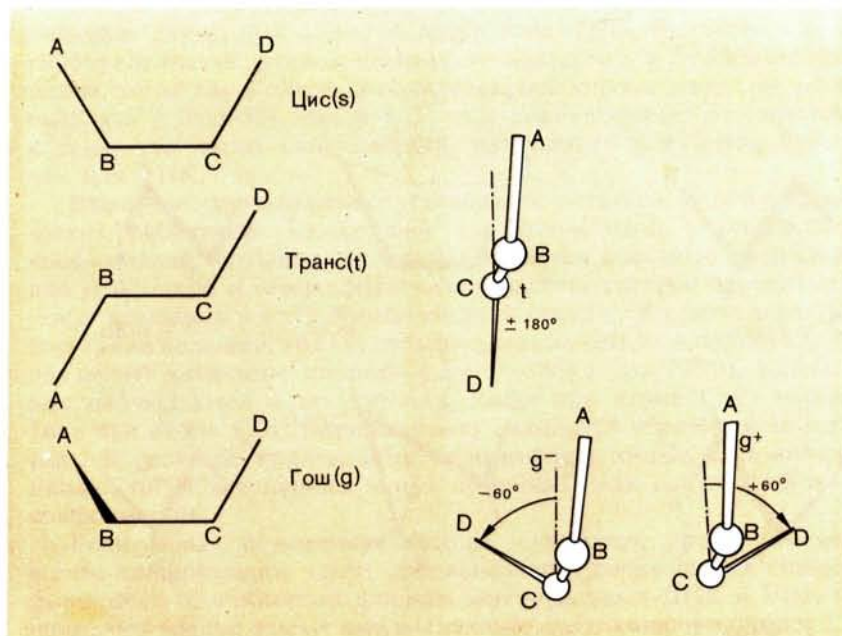


Рис. 187. Цис-, транс- и гош-конформации.

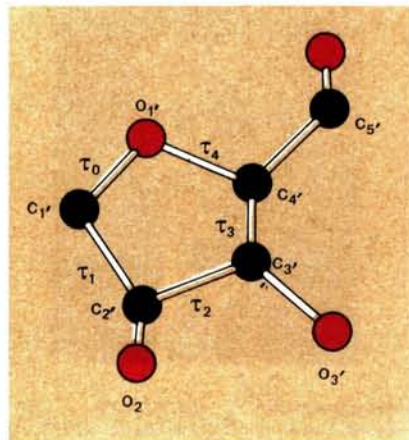


Рис. 188. Конформации циклопентана (а) и углеводного цикла в нуклеозидах (б).

в конформации нуклеотидной единицы. Ниже приведены обозначения связей и торсионных углов в углеводном цикле.

Конформация углеводного цикла может быть определена как функция всех торсионных углов $\tau_0 - \tau_4$. Однако для более простого и наглядного описания конформации рассматривают отклонения определенных атомов от плоскости, образуемой тремя или четырьмя другими атомами. Пятичленный цикл углевода может принимать различные конформации аналогично циклопентану (рис. 188). Если четыре его атома лежат в плоскости, а пятый выведен из нее, то такие конформации называются конформациями типа «конверт» и обозначаются Е или V. Если три атома образуют плоскость, а два соседних выведены из нее, то конформация называется «твист» (Т). Дальнейшее определение конформации основывается на взаимном расположении атома $C(5')$ и атома (или атомов), выведенного из плоскости цикла. Если атом n [$C(2')$ или $C(3')$] выведен из цикла в ту сторону, с которой находится атом $C(5')$, то конформация называется *эндо*, а если отклонен в сторону, противоположную атому $C(5')$, — *экзо*-конформацией.

Варианты конформаций типа «конверт» и «твист» обозначаются следующим образом:

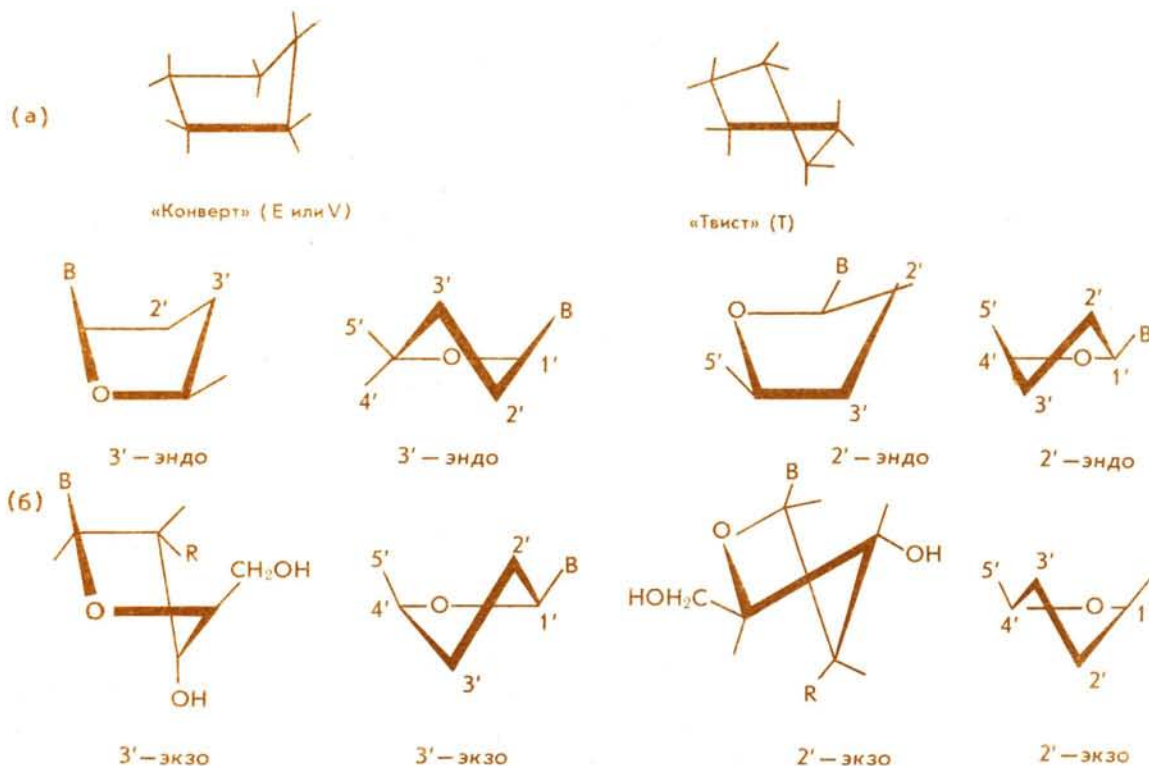
V^n или E^n — конверт с C_n -эндо-конформацией

V_n или E_n — конверт с C_n -экзо-конформацией

T_{n+1}^n — твист с C_n -эндо- и C_{n+1} -экзо-конформациями

T_n^{n+1} — твист с C_{n+1} -эндо- и C_n -экзо-конформациями.

Нуклеозиды и нуклеотиды. Для изучения пространственного строения нуклеиновых кислот широко используется метод рентгеноструктурного анализа моно-, олиго- и полинуклеотидов, который дает точные геометрические параметры молекул биополимеров и их компонентов в кристаллическом или твердом состоянии. Определение конформаций в растворах в настоящее время в основном про-



водится методом ЯМР-спектроскопии, хотя и такие методы, как ДОВ и КД, не потеряли своего значения.

Представляют интерес теоретические подходы, основанные либо на полуэмпирических расчетах потенциальных энергий при различных взаимных расположениях атомов, либо на квантово-химических расчетах (Б. Пюльман).

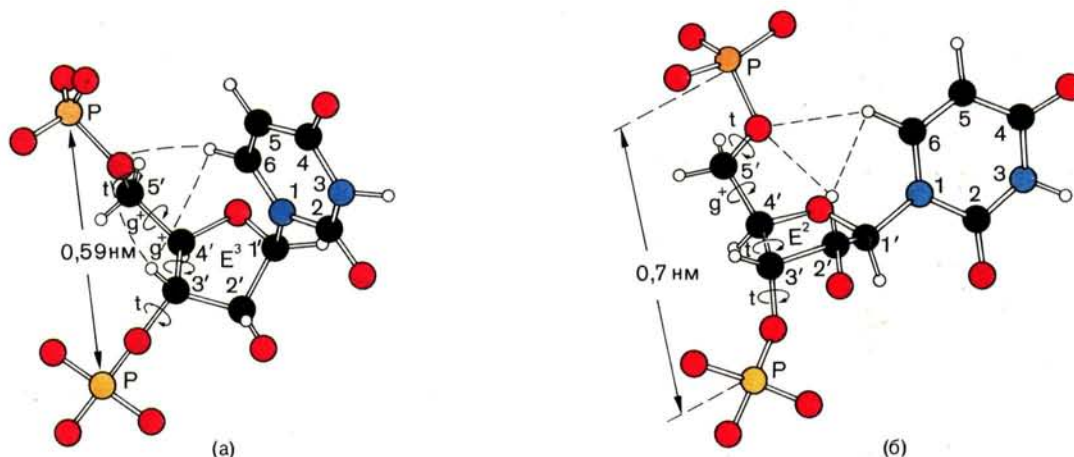


Рис. 189. Возможные конформации нуклеотидных единиц:

(а) — $C(3')$ -эндо-конформация;
(б) — $C(2')$ -эндо-конформация.

Гетероциклические основания нуклеиновых кислот являются практически плоскими структурами. Их атомы выходят из плоскости не более чем на 0,01 нм. С помощью рентгеноструктурного анализа получены геометрические характеристики оснований; оказалось, что связи $C—O$ (0,123 — 0,127 нм) короче связей $C—N$ ($>0,13$ нм) и связей $C—C$ ($>0,13$ нм). Следует отметить, что гликозидная связь незначительно выведена из плоскости оснований.

Фуранозный цикл углеводного компонента в нуклеозидах, нуклеотидах, олиго- и полинуклеотидах в подавляющем большинстве случаев, по данным рентгеноструктурного анализа и ЯМР-спектроскопии, имеет предпочтительную $3'$ -эндо- или $2'$ -эндо-конформации. Нуклеотидная единица с $3'$ -эндо-конформацией имеет меньшую длину, чем $2'$ -эндо-изомер (рис. 189). В олиго- и полирибонуклеотидах углевод чаще всего находится в $3'$ -эндо-конформации, тогда как в олиго- и полидезоксирибонуклеотидах он может быть как в $3'$ -эндо-, так и в $2'$ -эндо-конформации. Это приводит к тому, что число конформаций, возможных для ДНК, больше, чем для РНК.

Взаимное расположение углеводных остатков и гетероциклических оснований. Важнейшей характеристикой в определении конформаций нуклеиновых кислот является взаимное расположение углеводной и гетероциклической частей, которое определяется углом вращения вокруг N-гликозидной связи χ . В кристаллическом состоянии большинство нуклеозидов, моно-, олиго- и полинуклеотидов имеют *анти*-конформацию ($\chi = 0 - 90^\circ$) (рис. 190). Исключения наблюдаются в тех случаях, когда при атоме $C(8)$ пуринов [или при атоме $C(6)$ пиримидинов] находится объемный заместитель. В растворе производные пиримидинов имеют *анти*-конформацию, тогда как пурины могут принимать как *син*-, так и *анти*-конформации.

Если моно- и олигонуклеотиды, существуя предпочтительно в *анти*-конформации, могут довольно легко переходить в *син*-конформацию, то в обычных формах двуспиральных ДНК и РНК мономерные звенья имеют исключительно *анти*-конформацию.



Пюльман (Pullman) Бернар (р. 1919), французский химик, ведущий ученый в области квантовой химии, иностранный член АН СССР (1982). Окончил Парижский университет (1946), с 1963 г. — директор Института физико-химической биологии в Париже. Основные работы посвящены квантовой химии органических молекул, биохимии и биофизике. Выяснил электронную структуру молекул ряда биологически важных соединений, в том числе ДНК.

В конформационном анализе производных нуклеиновых кислот иногда используется концепция жесткой нуклеотидной единицы. Согласно этой концепции нуклеотиды имеют более жестко фиксированную конформацию, чем нуклеозиды. Конформационная гибкость полинуклеотидов в основном обусловлена легкостью враще-

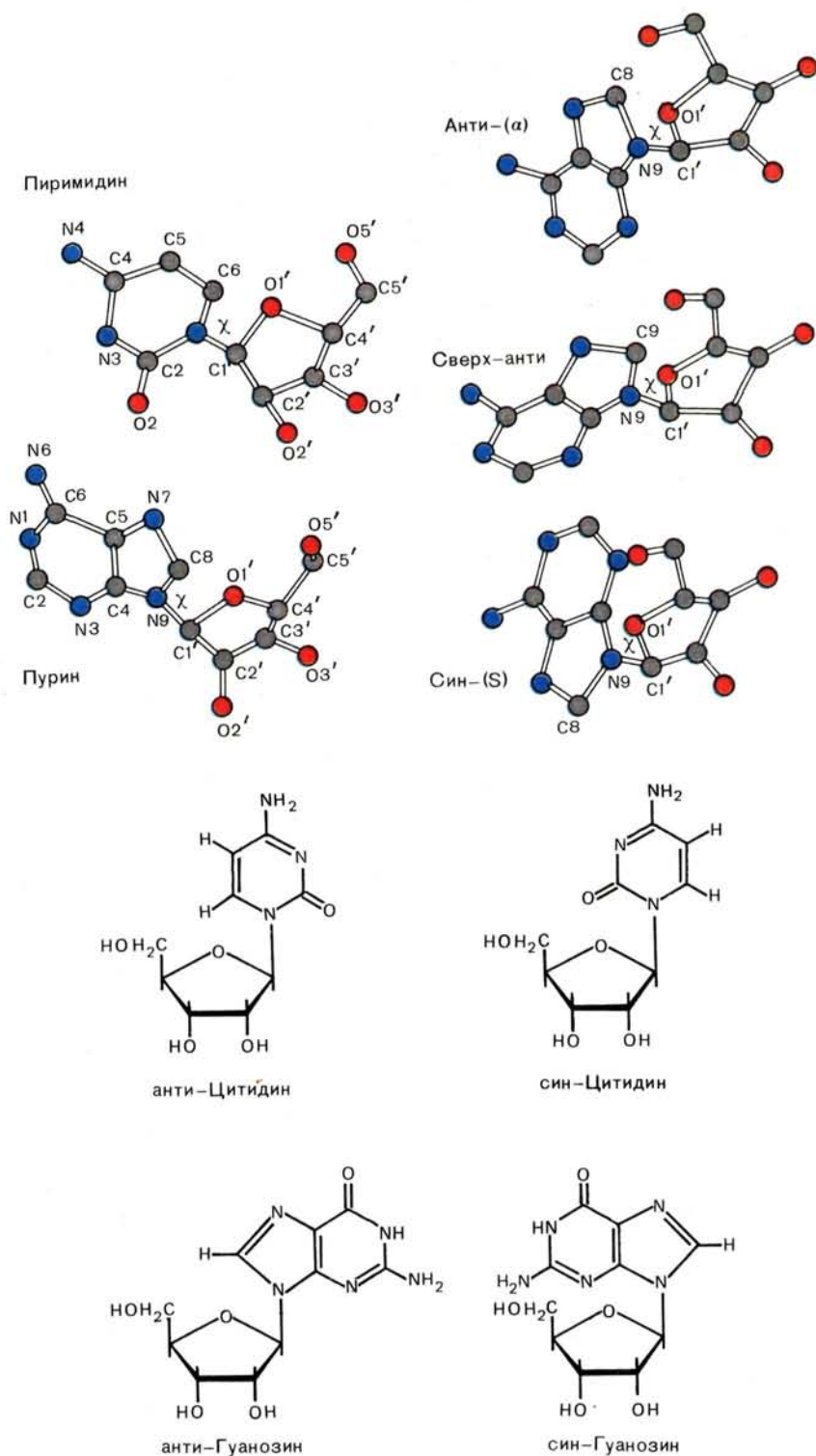


Рис. 190. Анти- и син-конформация в нуклеотидах.

ния вокруг фосфодиэфирных связей при сохранении конфигурации остальных. Правда, в тРНК и некоторых других случаях анализ детальной структуры обнаружил целый ряд отклонений от ожидаемой жесткой конформации.

При сопоставлении свойств моно- и полинуклеотидов следует иметь в виду, что принципиальным отличием полинуклеотидов от соответствующих мономеров является сильно выраженное взаимодействие между соседними основаниями, которые в водных растворах и в кристаллах стремятся расположиться таким образом, чтобы их плоскости были приблизительно параллельны. «Вертикальные» взаимодействия между основаниями, расположенными «стопкой», называются «стэкинг»-взаимодействиями; они обусловлены в основном ван-дер-ваальсовыми силами.

На существование таких взаимодействий в растворах олигонуклеотидов указывают отклонения их физико-химических характеристик от аналогичных характеристик, наблюдаемых для смеси мономерных компонентов, состав которых идентичен нуклеотидному составу олигонуклеотидов. Так, коэффициент молярной экстинкции олиго- и полинуклеотидов меньше, чем сумма коэффициентов молярных экстинкций входящих в него мономеров (этот эффект называется гипохромным эффектом).



Уилкинс [Wilkins] Морис Хью Фредерик (р. 1916), английский биофизик. Окончил Кембриджский университет, с 1946 г.— в Королевском колледже в Лондоне. Получил высококачественные рентгенограммы ДНК, на основе которых была установлена ее пространственная структура. Лауреат Нобелевской премии по физиологии и медицине (1962, совместно с Дж. Уотсоном и Ф. Криком).

Конформации нуклеиновых кислот

Двуспиральные полинуклеотиды. В большинстве случаев ДНК существует в виде «двойной спирали» Уотсона — Крика (рис. 191). Ее основные характеристики сводятся к следующему. Две полидезоксирибонуклеотидные цепи соединены друг с другом с помощью водородных связей и образуют правовинтовую спираль вокруг общей оси. Цепи двойной спирали антипараллельны и комплементарны, т. е. образование водородных связей (поперечных) всегда происходит между основаниями С и G или А и Т.

Гетероциклические основания обращены внутрь двойной спирали, и их плоскости приблизительно перпендикулярны ее оси. Двойная спираль в первом приближении регулярна, т. е. все ее витки имеют практически одинаковые размеры, на каждый виток приходится одинаковое число пар оснований и каждая пара оснований повернута относительно другой на один и тот же угол. С наружной стороны спирали находится гидрофильный углеводно-фосфатный остов.

Комплементарность двух цепей приводит к очень простому принципу удвоения генов, или репликации (рис. 192). Для этого достаточно, чтобы цепи ДНК разделились и на каждой из них была синтезирована новая комплементарная цепь. В результате образуются две дочерние молекулы ДНК, идентичные исходной.

Качественного представления о структуре ДНК недостаточно для понимания многих аспектов ее функционирования, в частности механизмов взаимодействия с белками. Для этого необходима информация о деталях структуры и возможностях ее изменения под действием различных факторов. В настоящее время накоплен очень большой материал по дифракции рентгеновских лучей на ориентированных волокнах ДНК и по строению моно- и олигонуклеотидов в кристаллах, позволяющий дать сравнительно точное описание возможных структур ДНК. Известно, что существует большой набор различных конформаций ДНК, которые меняются и переходят друг в друга в зависимости от внешних условий.



Уотсон [Watson] Джеймс Дьюи (р. 1928), американский биохимик, один из основоположников молекулярной биологии. Образование получил в Чикагском университете и университете штата Индиана. Основные работы посвящены биосинтезу белка и изучению структуры ДНК. Совместно с Ф. Криком расшифровал структуру ДНК и предложил ее модель в виде двойной спирали (1953). Лауреат Нобелевской премии по физиологии и медицине (1962, совместно с Ф. Криком и М. Уилкинсом).

Рассмотрим более детально конформационные характеристики двойной спирали ДНК (рис. 193). Шагом спирали (H) называют длину одного витка; расстояние между соседними звеньями обозначается h , а число звеньев на один шаг n . Легко видеть, что $h = H/n$. Угол поворота на один остаток $t = h/H \cdot 360$.

Конформации двойной спирали ДНК могут существенно различаться между собой, что прежде всего определяется изменениями

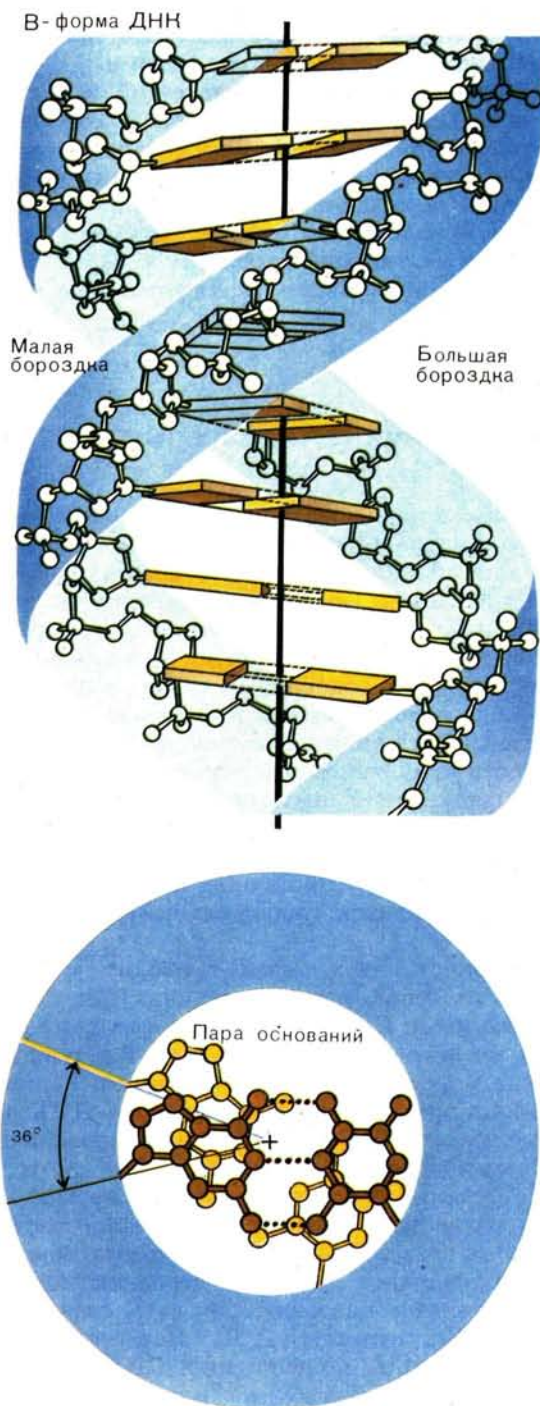
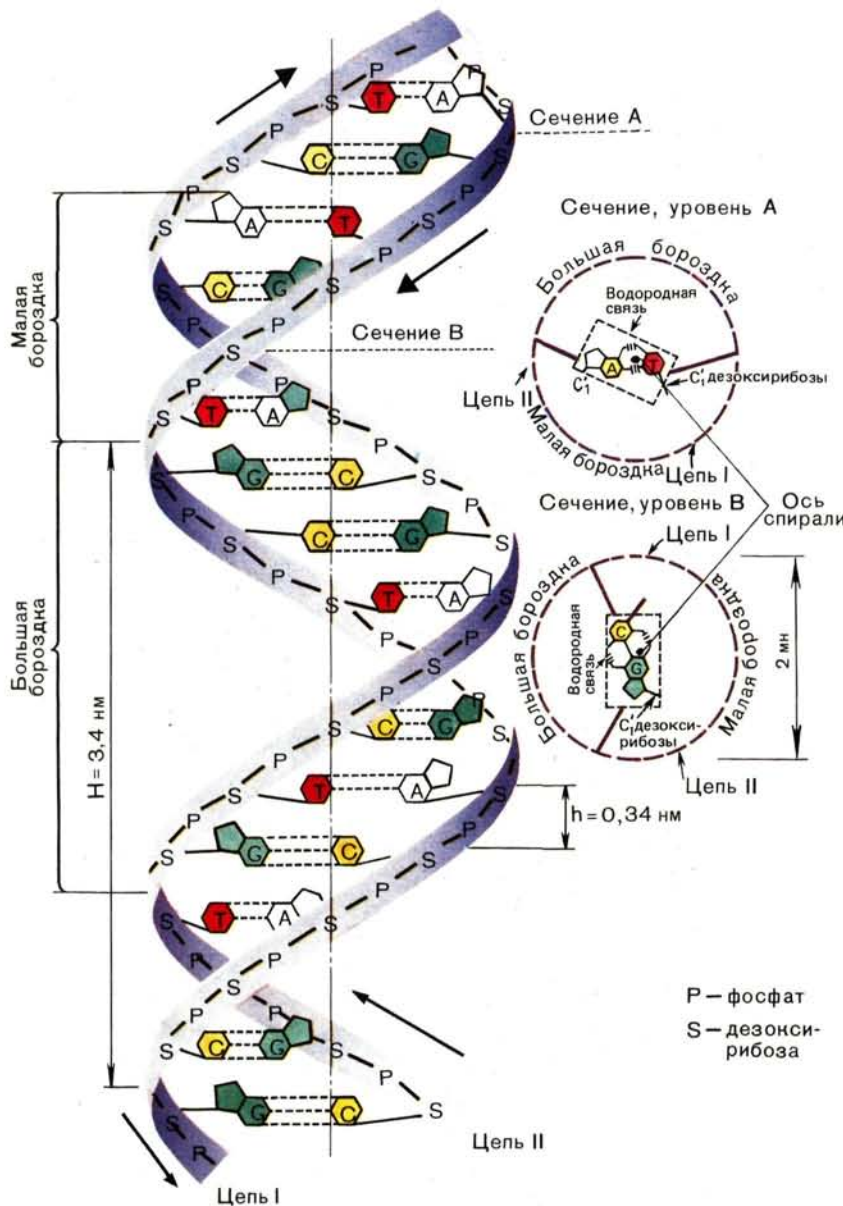


Рис. 191. ДНК в В-форме.

торсионных углов. Каждая конформация ДНК, согласно принятой номенклатуре, может быть описана семибуквенным обозначением торсионных углов $\alpha, \beta, \gamma, \delta, \epsilon, \zeta$ и χ в порядке следования связей в нуклеотидной единице; например, $t\text{tg}^- \text{tg}^+ \text{ta}$ конформация соответствует В-форме ДНК, в которой она находится обычно в кристаллическом состоянии и в водных растворах (рис. 191 и 193). Как видно из рисунков, плоскости пар оснований в В-форме ДНК практически



Крик [Crick] Фрэнсис Харри Комптон (р. 1916), английский физик, один из основоположников молекулярной биологии. Окончил Лондонский университет (1937). Основные работы посвящены изучению структуры нуклеиновых кислот. В 1953 г. совместно с Дж. Уотсоном предложил модель ДНК в виде двойной спирали (модель Уотсона — Крика) и объяснил процесс репликации ее молекулы при делении клеток. В опытах на фагах T4 впервые установил основные принципы генетического кода. Лауреат Нобелевской премии по физиологии и медицине (1962, совместно с Дж. Уотсоном и М. Уилкинсом).

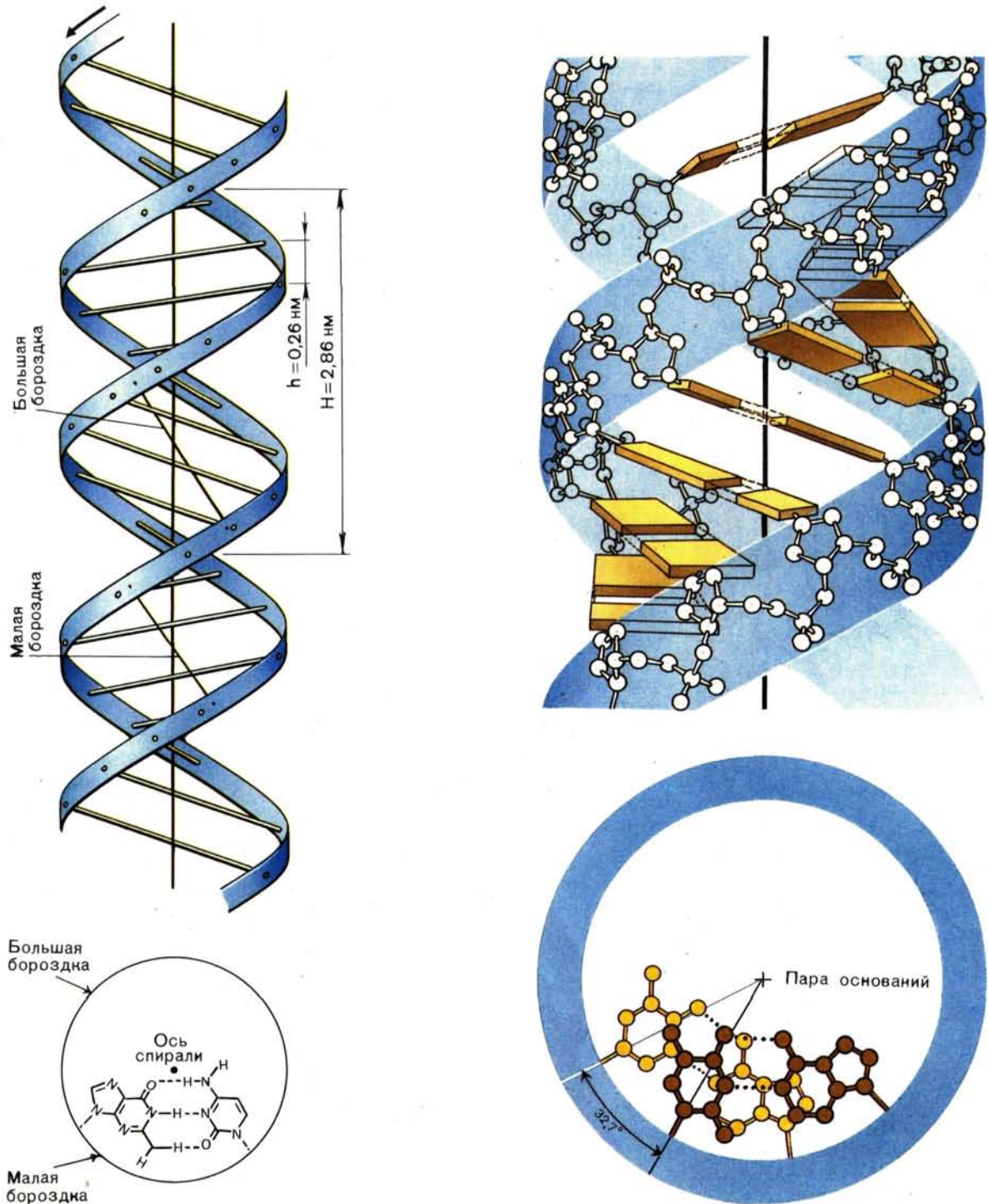


перпендикулярны оси спирали, проходящей, по существу, через центр каждой пары. Геометрические характеристики В-формы: $h = 0,34 \text{ нм}$, один виток спирали содержит 10 п. о. и его длина (H) равна $3,4 \text{ нм}$.

Рис. 192. Конформационные характеристики двойной спирали ДНК в В-форме.

При уменьшении степени гидратации ДНК переходит в А-форму (рис. 194), конформацию которой можно описать как $tg^-g^-tg^+g^+a$. Параметры А-формы следующие: $h = 0,26$ нм, один виток содержит 11 пар оснований и его длина равна $2,86$ нм. Нетрудно видеть, что в А-форме плоскости оснований параллельны друг другу, но наклонены к оси спирали, образуя с ней угол $\approx 20^\circ$. В то же время

Рис. 194. ДНК в А-форме.



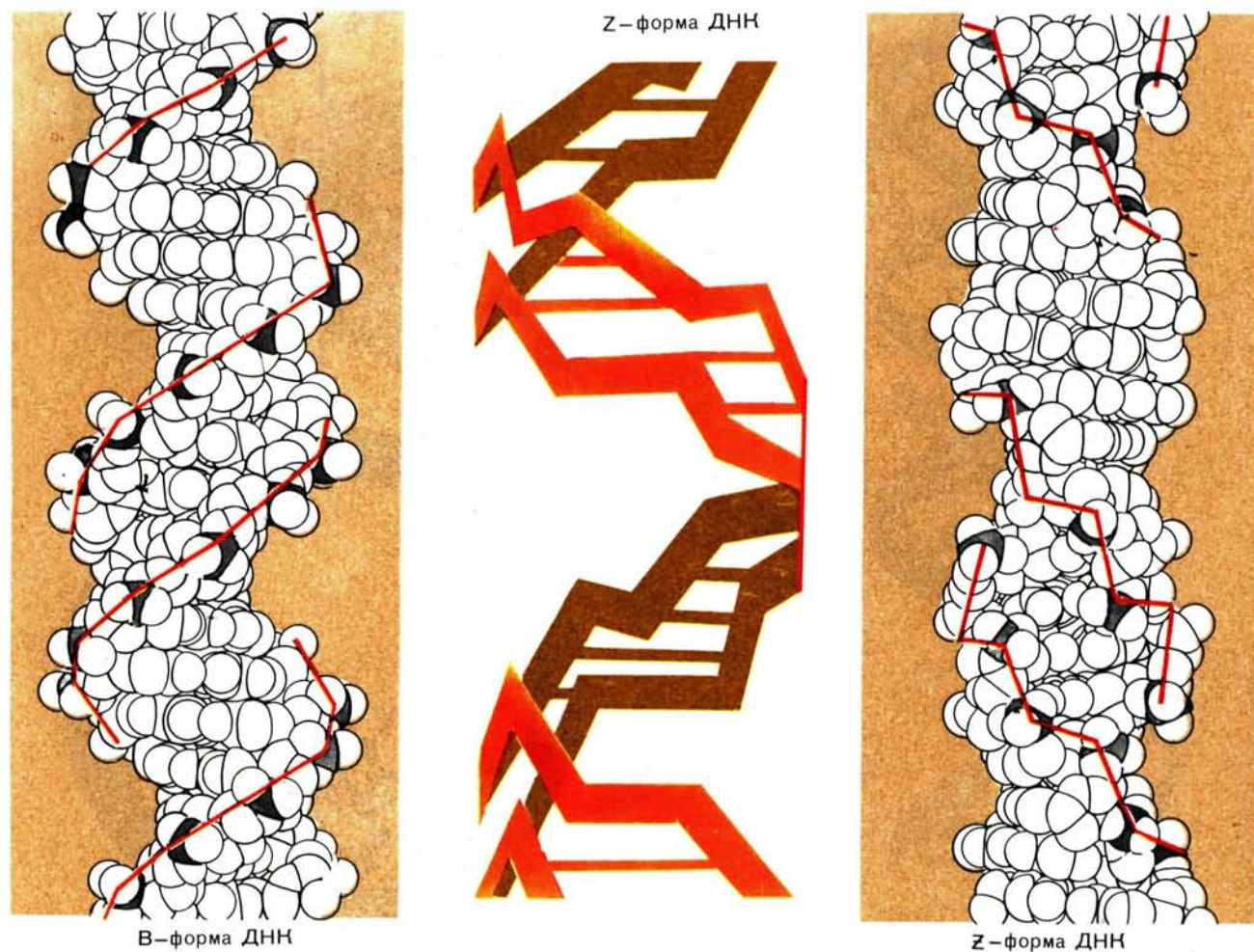
ось спирали проходит в стороне от пар оснований, в результате чего внутри спирали создается пустое пространство; при этом широкая бороздка становится уже, а узкая — шире (их часто называют большая и малая бороздки соответственно).

Уменьшение расстояния между основаниями и большая компактность спирали обусловлены тем, что угол, соответствующий связи $C(3')-C(4')$ в В-форме, имеет конформацию $t[C(2')\text{-эндо}]$, тогда как в А — $g^+[C(3')\text{-эндо}]$. Транс-конформация этой связи приводит к увеличению длины нуклеотидной единицы по сравнению с g^+ -конформацией.

ДНК способна переходить из А-формы в В-форму и обратно в зависимости от природы растворителя. Не исключено, что такие переходы имеют место и при взаимодействии с белками.

Долгое время считалось, что ДНК может быть только правовращающей и левовращающая спираль стереохимически невозможна. Однако недавно было доказано, что левовращающая спираль существует. Она была найдена в случае двуспиральных альтернирующих сополимеров: поли $[d(G-C)] \cdot$ поли $[d(G-C)]$ и поли $[d(G-T)] \cdot$ поли $[d(C-A)]$ (рис. 195). Необычное свойство левых спиралей состоит в том, что повторяющейся единицей в них

Рис. 195. ДНК в Z-форме.



является не мононуклеотид, а динуклеотид, поскольку каждые два соседних нуклеотида находятся в различных конформациях. Например, в одной из левовращающих форм первое нуклеотидное звено имеет конформацию $g^-g^+g^+tg^+s$, характеризующуюся *син*-положением пурина и $C(3')$ -*эндо*-конформацией углевода; второе звено имеет *анти*-положение пиримидина и $C(2')$ -*эндо*-конформацию углевода и характеризуется формулой $g^-g^-g^-tg^+ta$. Таким образом, общая формула повторяющейся динуклеотидной единицы ($g^-g^+g^+tg^+stg^-g^-g^-tg^+ta$). Такая спираль с характерным зигзагообразным строением получила название Z-формы. Получены данные, что Z-форма может существовать и в растворе.

Какие же силы способствуют образованию и сохранению двухспиральных структур? Долгое время считалось, что это — возникающие между цепями водородные связи. Однако сейчас стало ясно, что основной вклад в стабильность двухспиральных молекул вносят другие силы, прежде всего взаимодействия между плоскостями оснований в стопках, т. е. «стэкинг»-взаимодействия; существенную роль играют также взаимодействия между ДНК и водой, в результате которых ДНК стремится принять максимально компактную структуру для уменьшения поверхности соприкосновения с водой. При этом гидрофобные основания локализуются внутри спирали, а на ее поверхности образуется гидратируемая углеводно-фосфатная оболочка.

Дестабилизирующее воздействие оказывают силы электростатического отталкивания между отрицательно заряженными фосфатами комплементарных цепей.

Циклические ДНК и суперспирализация. Многие двухцепочечные ДНК в природе являются циклическими: плазмиды, ДНК митохондрий и хлоропластов, ДНК многих вирусов и бактерий. Такие ДНК, как правило, существуют в суперспиральном состоянии. При этом двойная спираль закручивается сама на себя, как показано на рисунке 196, количество витков образующейся суперспирали зависит от внешних условий. Суперспирализация циклических ДНК приводит к сильному изменению физических свойств молекулы, в особенности гидродинамических и электрофоретических. В клетках суперспирализация осуществляется особыми ферментами, которые для бактерий сравнительно хорошо изучены и называются ДНК-гиразами (или топоизомеразами II). Другие ферменты — топоизомеразы I — могут уменьшать число супервитков в кольцевых молекулах, давая набор «изомеров» с различным числом витков.

Знак суперспирали определяется по взаимодействию ДНК с определенными химическими веществами, способными связываться с ней и раскручивать двойную спираль. Такими веществами являются, например, интеркалирующие красители. Наиболее часто для этих целей используется этидийбромид; степень раскручивания увеличивается с концентрацией красителя

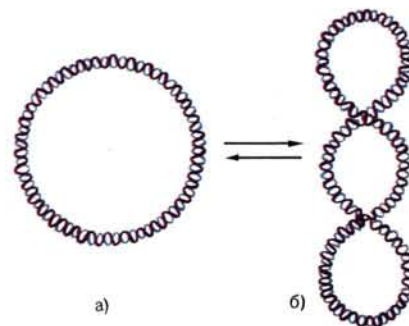
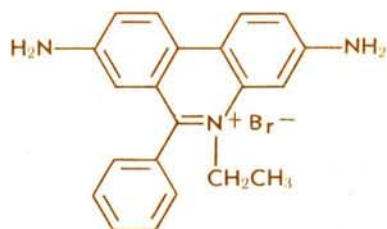


Рис. 196. Суперспирализация двухцепочечной ДНК под действием ДНК-гиразы: (а) — релаксированная форма; (б) — суперспирализованная форма.



Этидийбромид



Рич [Rich] Александр (р. 1924), американский ученый, работающий в области молекулярной биологии и биофизики. Образование получил в Гарвардском колледже и Гарвардской медицинской школе, с 1974 г. — профессор Массачусетского технологического института в Кембридже. Один из ведущих специалистов в области нуклеиновых кислот.

Суперспиральная ДНК при центрифугировании осаждается быстрее, чем релаксированная форма, т. е. лишенная суперспиральных витков. Если проводить центрифугирование в присутствии этидийбромиды и определять зависимость скорости седиментации циклической ковалентно замкнутой ДНК от ее концентрации, то можно определить знак суперспирали.

При суперспирализации ДНК приобретает «скрытую» конформационную энергию, влияющую на любой процесс, происходящий с изменением числа витков двойной спирали. Благодаря этому некоторые ферменты, вызывающие локальное расплетание ДНК (например, РНК-полимеразы, см. ниже), связываются с отрицательно суперспирализованной ДНК более эффективно. Другим следствием суперспирализации является реальная возможность образования участков вторичной структуры, которые термодинамически невыгодны и практически не существуют в линейной или релаксированной циклической ДНК.

Разрушение и восстановление двуспиральных структур. Денатурация, ренатурация, и гибридизация. Двуспиральные структуры очень стабильны в физиологических условиях при значениях рН, близких к нейтральным, температурах около 30 — 40 °С и в присутствии 0,15 моля соли. Однако увеличение или уменьшение рН, повышение температуры и добавление в среду некоторых веществ, таких, например, как мочевины, приводит к разрушению двухцепочечной структуры, предельным случаем которого является *денатурация*, т. е. полное разделение отдельных нитей.

Температура, при которой доли спирального и неспирального состояний равны, называется *температурой плавления* и обозначается T_m . Для каждой ДНК в постоянных условиях T_m постоянна и характеризует стабильность двуспиральной структуры. Другой характеристикой ДНК является ширина температурного перехода ΔT_m , которая отражает кооперативность перехода. Если бы все звенья разрушались одновременно, то ΔT_m была бы равна нулю. Однако процесс плавления идет через ряд промежуточных состояний, и поэтому ΔT_m имеет конечное значение.

Ширина интервала плавления для двухцепочечных синтетических полинуклеотидов, таких, как поли (dA) · поли (dT), значительно уже, чем у природных ДНК, что является следствием гетерогенности последних. Температура плавления ДНК линейно возрастает с увеличением доли G · C пар. На этом основан один из методов определения нуклеотидного состава ДНК.

Разрушенная двуспиральная структура в определенных условиях может быть восстановлена, по крайней мере частично. Этот процесс называют *ренатурацией*. Ренатурация происходит, если раствор денатурированной ДНК выдерживать определенное время в условиях, когда двуспиральная структура стабильна. Скорость ренатурации зависит от многих факторов, и в первую очередь от так называемой сложности ДНК, которая представляет собой число нуклеотидных пар в неповторяющихся последовательностях ДНК. Для ДНК вирусов и бактерий сложность ДНК равна числу нуклеотидных пар в геноме, поскольку в этом случае ДНК не содержит повторяющихся последовательностей достаточно большой длины. Чем выше сложность ДНК, тем медленнее идет ренатурация. Если ренатурации подвергают не целые, а фрагментированные молекулы, то скорость ренатурации возрастает с увеличением длины фрагментов.

Ренатурация широко применяется для исследования структуры ДНК. В частности, она используется для изучения сходства и различий разнородных ДНК. Две ДНК из разных источников денатурируют и затем смесь их ренатурируют. При этом, наряду с исходными нативными молекулами, образуются гибридные молекулы,

содержащие цепи из различных ДНК. Исследование таких гибридов под электронным микроскопом позволяет определить в них положение одно- и двухцепочечных участков. Двухцепочечные участки образуются в областях гомологии исследуемых ДНК. Ренатурация в этом случае называется *гибридизацией*.

Конформации одноцепочечных нуклеиновых кислот. Многие нуклеиновые кислоты — большинство РНК и ряд ДНК — существуют в одноцепочечной форме. Тем не менее они принимают в растворах и кристаллическом состоянии (в тех случаях, когда удается их закристаллизовать) конформации, в которых двухцепочечные участки чередуются с одноцепочечными. Наиболее хорошо изученными одноцепочечными полинуклеотидами являются тРНК.

Конформация тРНК. Транспортные РНК выполняют в клетках разнообразные функции. Однако основная их задача заключается в осуществлении трансляции. В 1965 г. Р. Холли установил первичную структуру тРНК^{Ala} из дрожжей. Тогда же, исходя из представления, что наиболее стабильное состояние тРНК соответствует образованию максимально возможного количества водородно-связанных пар оснований, а также основываясь на экспериментальных данных по неравномерному гидролизу молекулы рибонуклеазами, Р. Холли предложил модель вторичной структуры тРНК необычной формы — с чередующимися одно- и двухцепочечными участками — и назвал эту структуру «клеверным листом» (рис. 197).

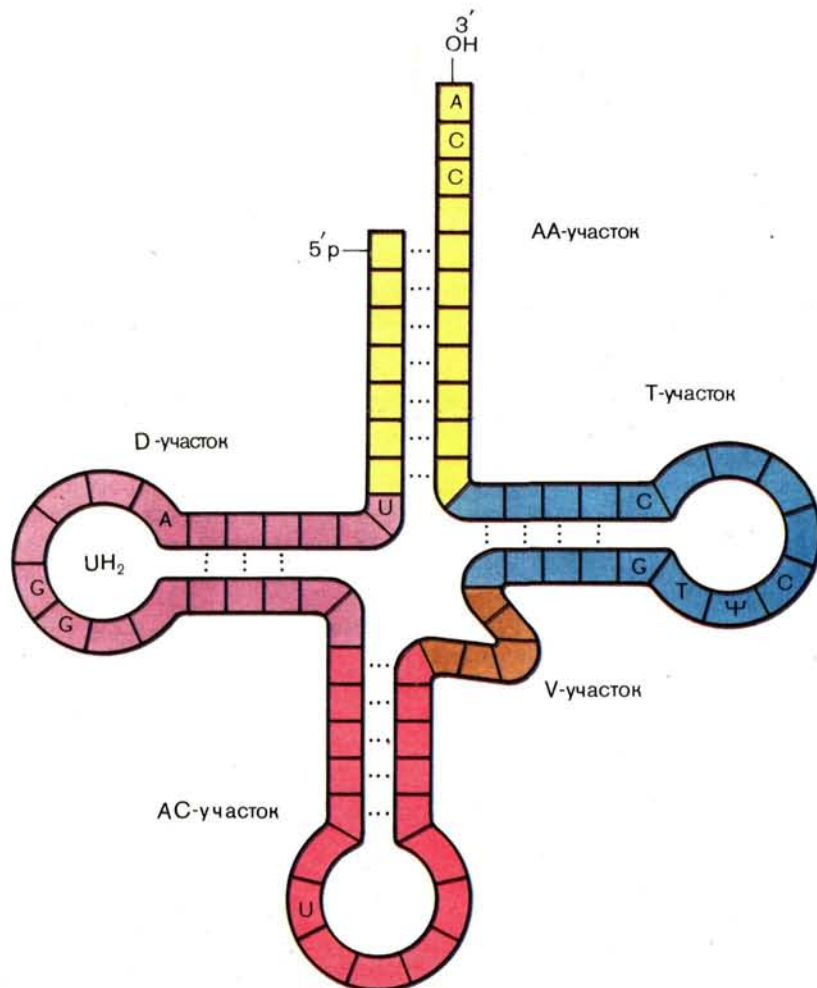


Рис. 197. Структура «клеверного листа» тРНК.



Клуг (Klug) Аарон (р. 1926), английский кристаллограф. Образование получил в университете Витватерсранда в Йоханнесбурге и Кейптаунском университете (ЮАР), с 1962 г. работает в лаборатории молекулярной биологии Медицинского исследовательского совета в Кембридже. Известен работами по изучению нуклеопротеидов методом рентгеноструктурного анализа. Лауреат Нобелевской премии по химии (1982).

С тех пор были определены последовательности более 100 тРНК из разных источников, и оказалось, что, несмотря на значительные различия первичных структур, все тРНК имеют сходную вторичную структуру типа «клеверного листа». Эта структура характеризуется пятью ответвлениями («лепестками»), в составе которых имеется одноцепочечный участок и двухцепочечный стебель. Каждому ответвлению присвоено свое название либо по структурному, либо по функциональному признаку: участок, включающий 3'-концевое звено тРНК, к которому присоединяется аминокислота, получил название аминоацильного (AA); содержащий в петле антикодон — антикодонового (AC); участок, в петле которого содержится общая для всех тРНК последовательность ТψС — Т-участка; участок, содержащий дигидроуридин, — D-участка. Наконец, фрагмент структуры, расположенный между Т- и AC-ответвлениями, имеет для всех тРНК различные размеры и называется переменным V-участком.

Структура типа «клеверный лист» объясняла характерную реакционную способность нуклеотидных звеньев в разных участках тРНК по отношению к химическим агентам и к действию рибонуклеаз. Однако гидродинамические характеристики молекулы свидетельствовали о ее более компактной упаковке, которая могла бы осуществляться за счет третичной структуры. Способ образования третичной структуры стал ясен после рентгеноструктурного анализа первой тРНК, которую удалось получить в кристаллическом состоянии (тРНК^{Phc} из дрожжей). Рентгеноструктурные исследования были выполнены одновременно двумя группами — в лабораториях А. Рича (США) и А. Клуга (Великобритания). Впоследствии были установлены третичные структуры еще нескольких тРНК. Третичная структура молекулы тРНК^{Phc} изображена на рисунке 198, она напоминает латинскую букву L.

Общие закономерности, найденные для структуры тРНК, по-видимому, реализуются и в других одноцепочечных полинуклеотидах. Те из них, для которых уже определены первичные структуры, могут быть представлены как образования с чередующимися двух- и одноцепочечными участками. Например, молекулы рибосомных 5S РНК имеют вторичные структуры, сходные с «клеверным листом» тРНК. Значительно более сложно выглядят структуры высокомолекулярных рибосомных или вирусных РНК (рис. 199). Несомненно, что такие РНК находятся в компактной форме, как это следует из их гидродинамических свойств, однако детали пространственной организации пока неизвестны.

Так же как и двухцепочечные полинуклеотиды, одноцепочечные молекулы могут денатурировать с разрушением вторичной и третичной структур. Денатурация наблюдается при повышении температуры, понижении ионной силы или добавлении в среду денатурирующих агентов — мочевины, органических растворителей и т. д.

Изучение плавления РНК различными методами (УФ-, КД- и ЯМР-спектроскопия) позволяет сделать вывод о многостадийном характере этого процесса. Кооперативность процесса плавления одноцепочечных молекул значительно меньше, чем в случае двуспиральных.

До сих пор рассматривались статические структуры ДНК и РНК, которые представляют собой среднее и наиболее вероятное состояние молекулы в данных конкретных условиях. Важно тем не менее понимать, что даже при низких температурах, далеких от температуры плавления, полинуклеотидная цепь находится в постоянном движении, определяемом тепловыми колебаниями. Нуклеотидные звенья сохраняют достаточную степень свободы и совершают локальные движения, значительные по амплитуде и частоте.

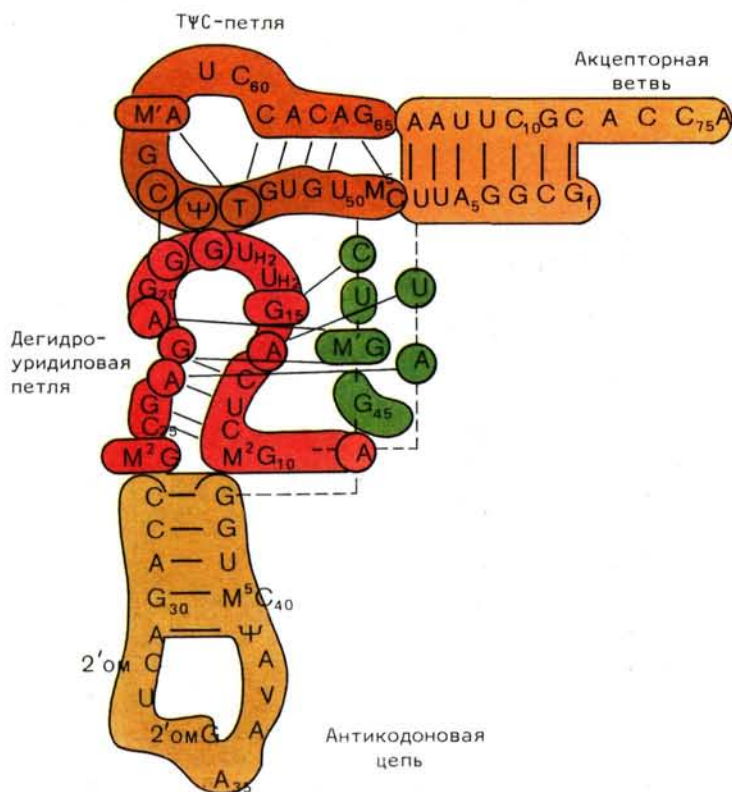
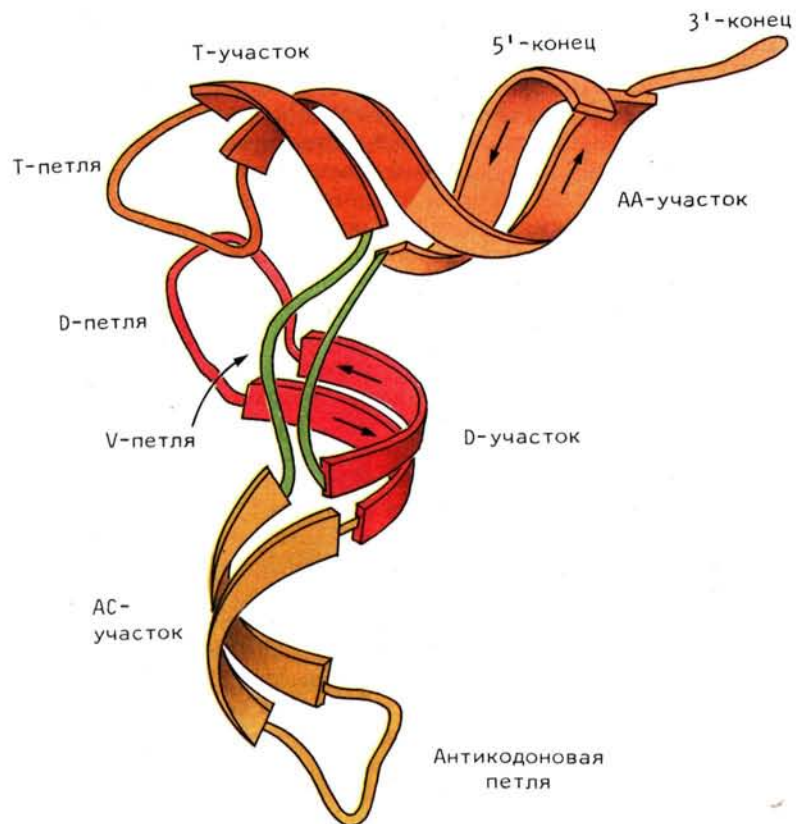


Рис. 198. Пространственная структура тРНК^{Phe} из дрожжей.

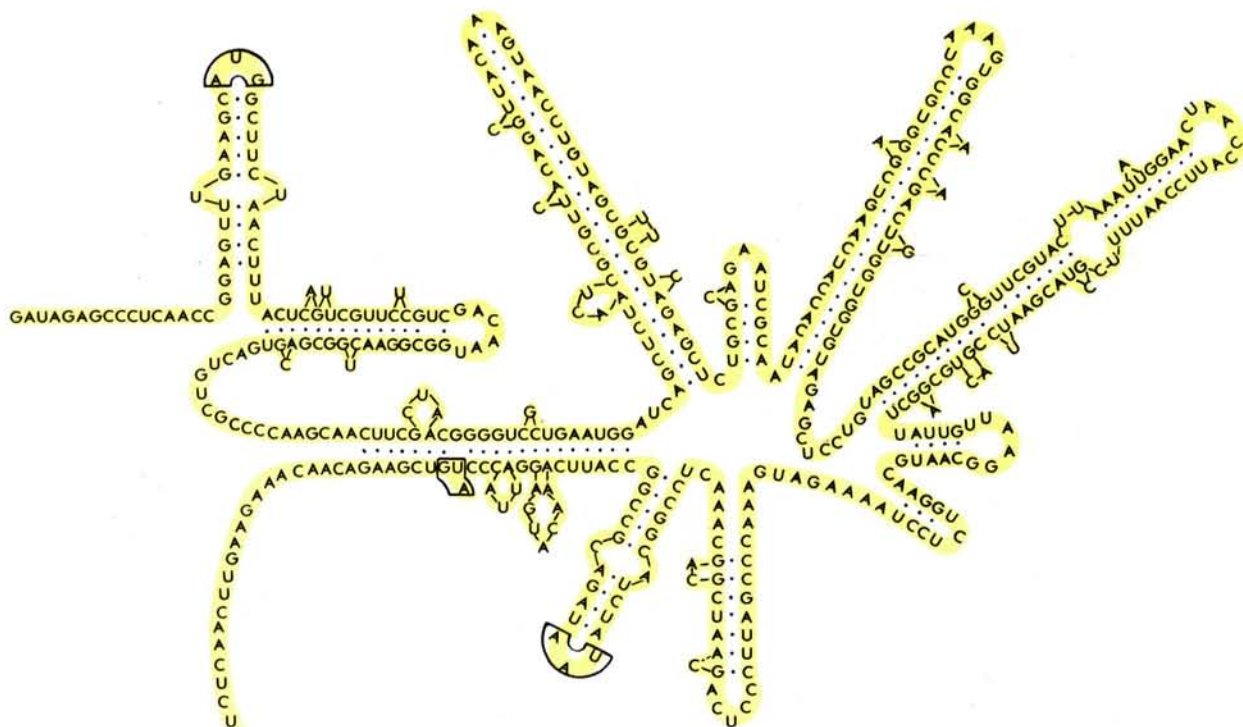
Вероятно, в основе их лежит конформационная гибкость фуранозного цикла нуклеотидной единицы, способной к переходу от $C(2')$ -эндо- к $C(3')$ -эндо-конформации, и наоборот. Подобные переходы, в свою очередь, сказываются на торсионных углах фосфодиэфирных и N-гликозидных связей и приводят к таким эффектам, как изменение перед плавлением и «дыхание» полинуклеотида.

О динамичности структуры полинуклеотидов свидетельствует ряд данных, полученных химическими методами. В частности, существование равновесия водородно-связанной и свободной форм оснований следует из результатов экспериментов по тритиевому или дейтерообмену. Другая группа данных, которая также позволяет предположить динамические нарушения вторичной структуры ДНК, основана на изучении ее взаимодействия с так называемыми интеркалирующими веществами. Эти соединения имеют плоские ароматические хромофоры, как, например, этидийбромид (см. с. 341).

Рентгеноструктурный анализ комплексов таких соединений с синтетическими двухцепочечными олигонуклеотидами показывает, что плоские ароматические кольца красителей внедряются между парами оснований двойных спиралей. Механизм внедрения предполагает проникновение молекулы красителя между парами оснований в момент возникновения локального нарушения структуры, при этом водородные связи между парами оснований сохраняются, тогда как «стэкинг»-взаимодействия нарушаются.

Одним из вариантов такого нарушения структуры является образование так называемых «кинков», представляющих собой из-

Рис. 199. Участок вторичной структуры РНК фара MS2.



ломы двойной спирали, схематически изображенные на рисунке 200. Впервые гипотеза о возможности «кинков» была высказана Ф. Криком и А. Клугом для объяснения способа укладки ДНК в нуклеосомах хроматина. Хотя гипотеза не была прямо подтверждена экспериментально, она породила целый ряд гипотез о способах образования «кинков» в двуспиральных ДНК и их роли во взаимо-

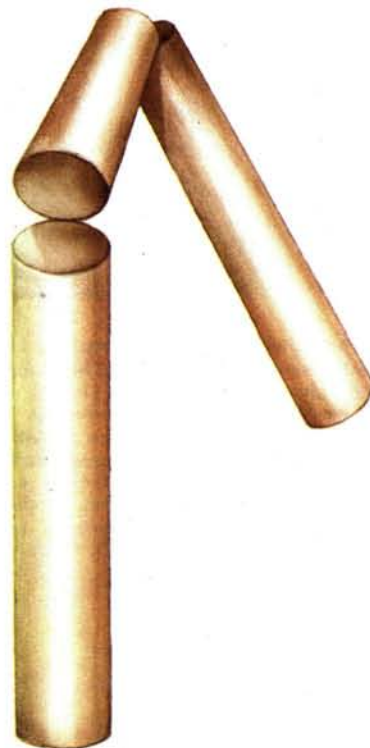
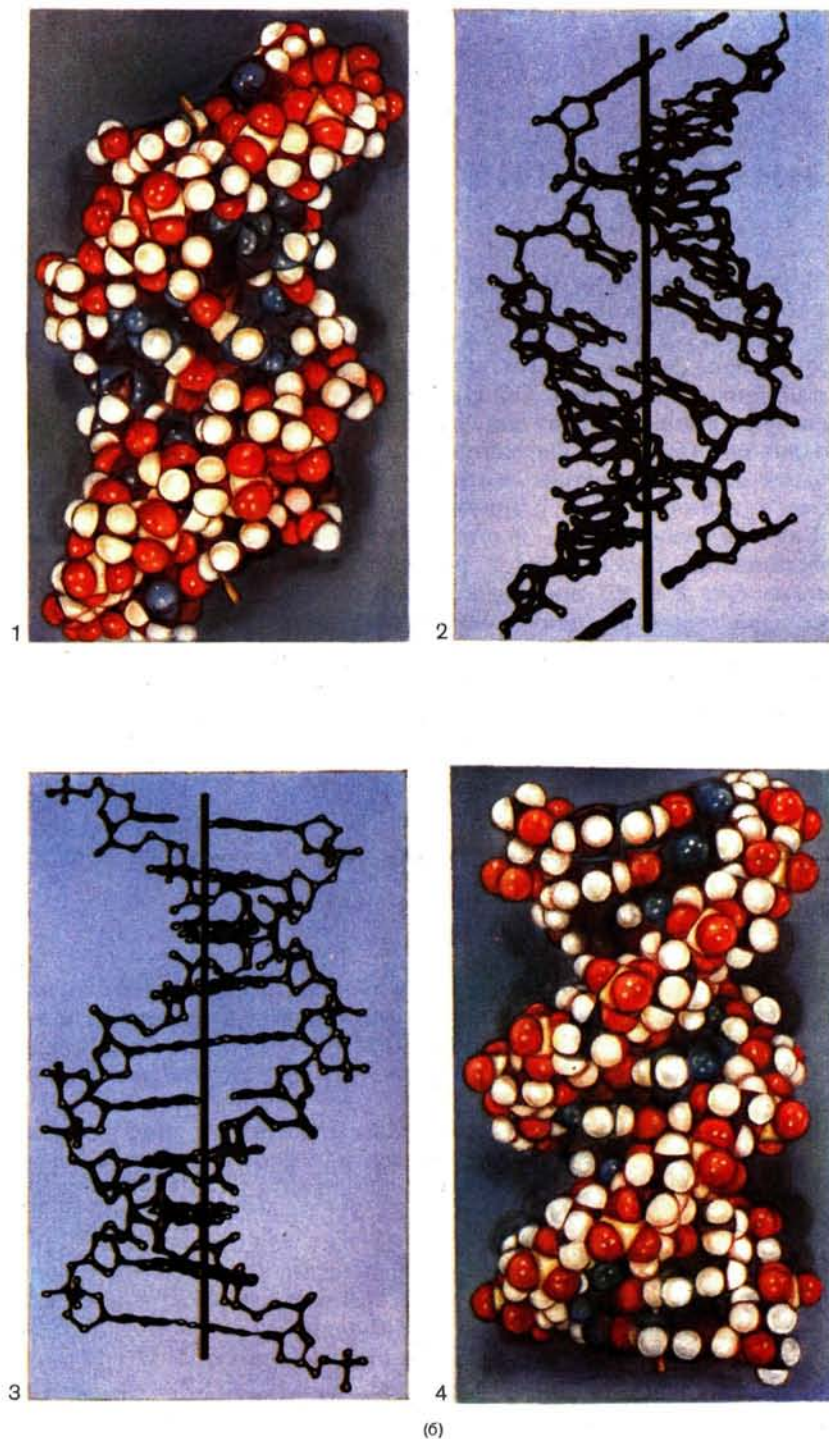


Рис. 200. Схема образования «кинков» в ДНК (а). Образование «кинков» в ДНК (б).

1 — β-«кинковая» структура ДНК;
2 — β-«кинковая» компьютерная структура ДНК;
3 — компьютерная структура ДНК в В-форме;
4 — молекулярная модель ДНК в В-форме.



Корнберг [Kornberg] Артур (р. 1918), американский биохимик. Окончил Рочестерский университет (1941). Внес значительный вклад в изучение биосинтеза белков. Открыл фермент ДНК-полимеразу, с помощью которого синтезировал биологически активную молекулу ДНК. Лауреат Нобелевской премии по физиологии и медицине (1959, совместно с С. Очоа).

действию с интеркалирующими красителями и белками, содержащими способные к интеркаляции ароматические аминокислоты. Эти гипотезы сводятся к тому, что в момент образования «кинка» интеркалирующее вещество способно внедряться между парами оснований, которые удаляются друг от друга в месте излома. Иногда предполагают, что возникновение «кинков» предшествует плавлению двухцепочечных полинуклеотидов.

Синтез нуклеиновых кислот

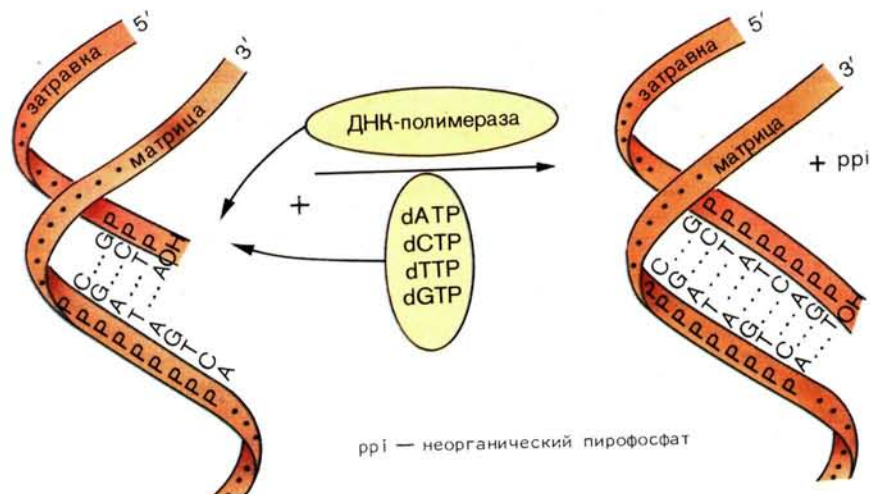
В настоящее время методы синтеза олиго- и полинуклеотидов разработаны в такой степени, что в короткие сроки могут быть получены искусственные фрагменты ДНК большой длины и практически любого состава. Достигнут существенный прогресс и в синтезе полирибонуклеотидов. Общая стратегия синтеза полинуклеотидов и нуклеиновых кислот заключается в комбинированном использовании химических и ферментативных методов. Относительно небольшие олигонуклеотиды синтезируют химически, а затем «соединяют» в длинные цепи с помощью соответствующих ферментов.

Синтетические олиго- и полинуклеотиды, а также полученные синтетическим путем гены и регуляторные области (промоторы, терминаторы и т. д.) широко используются в исследовании структуры и функции нуклеиновых кислот, генетической и белковой инженерии, биотехнологии. Синтез олиго- и полинуклеотидов, представляющий собой важный раздел биоорганической химии, имеет сегодня большое теоретическое и прикладное значение.

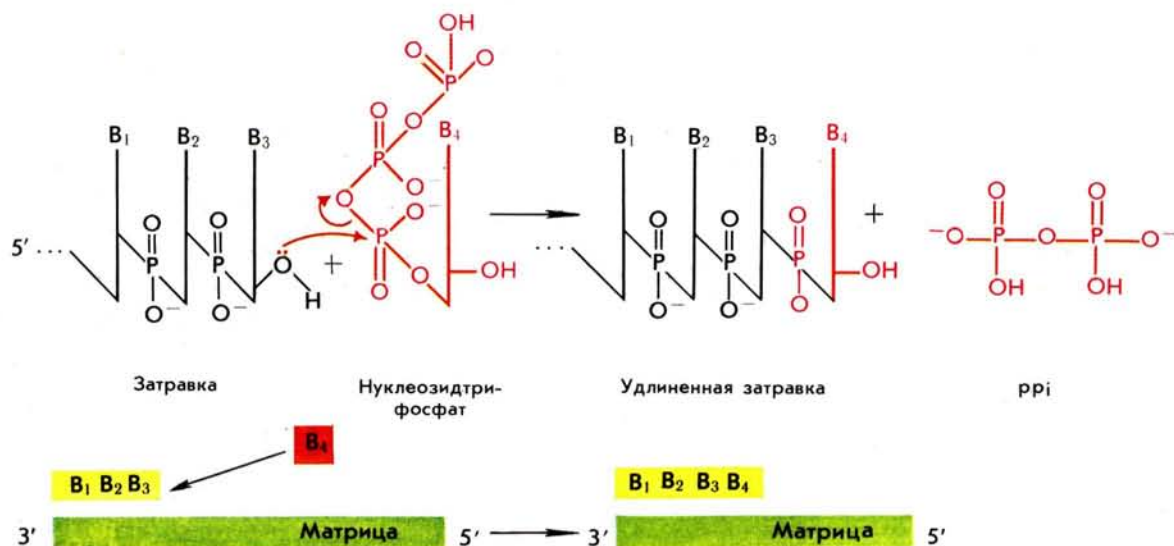
Ферменты биосинтеза нуклеиновых кислот

В этом разделе рассматриваются основные ферменты биосинтеза нуклеиновых кислот, поскольку многие из них широко используются в химико-ферментативном синтезе полинуклеотидов. Среди этих ферментов особый интерес представляют полимеразы и лигазы.

ДНК-полимеразы. ДНК-полимеразы относятся к типу ферментов, образующих фосфодиэфирные связи и использующих в качестве субстрата дезоксинуклеозидтрифосфаты (dNTP). Необходимым условием работы таких ферментов, как правило, является наличие матрицы и затравки. Матрицей служит одноцепочечная ДНК, и ДНК-полимеразы ее копируют, синтезируя комплементарную полидезоксирибонуклеотидную цепь. Синтез не может начаться без затравки, представляющей собой олиго- или полинуклеотид, комплементарный матрице и имеющий свободную 3'-ОН-группу, с которой начинается присоединение первого нуклеотида вновь синтезируемой цепи



Таким образом, фермент выполняет функцию отбора нуклеозидтрифосфатов, комплементарных каждому следующему звену матрицы, и катализирует нуклеофильную атаку 3'-ОН-группы затравки на α -фосфатную группу трифосфата:



После присоединения каждого нового звена удлиненная цепь содержит на 3'-конце свободную гидроксильную группу, к которой может присоединяться новое звено. Процесс продолжается, пока ДНК не станет полностью двухцепочечной. Новая цепь растет в направлении от 5'-к 3'-концу, а матрица «считывается» в направлении от 3'-к 5'-концу.

Вышесказанное относится к общим свойствам всех матрично-зависимых ДНК-полимеризующих ферментов. Однако ферменты из разных источников обладают и индивидуальными особенностями.

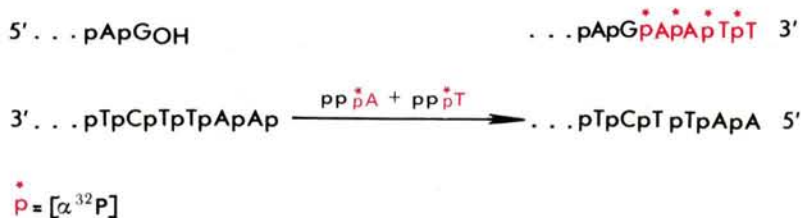


Темин (Temin) Говард (р. 1934), американский вирусолог. Окончил Калифорнийский технологический институт в Пасадене (1959), с 1971 г. — заведующий лабораторией в Висконсинском университете в Милуоки. Основные работы — по изучению РНК-содержащих вирусов. Обнаружил в вирусах саркомы Рауса ферментную систему, способную синтезировать ДНК на матрице РНК (1970, независимо от Д. Балтимора). Лауреат Нобелевской премии по физиологии и медицине (1975, совместно с Р. Дульбекко и Д. Балтимором).

Так, ДНК-полимераза I из *E. coli* (открыта А. Корнбергом), помимо полимеразной активности, обладает еще двумя типами экзонуклеазной активности: как $5' \rightarrow 3'$ -экзонуклеаза она осуществляет с $5'$ -конца двухцепочечных ДНК деградацию, в результате которой происходит отщепление $5'$ -моно- и $5'$ -фосфорилированных коротких олигонуклеотидов, а как $3' \rightarrow 5'$ -экзонуклеаза последовательно отщепляет $5'$ -мононуклеотиды с $3'$ -концов двухцепочечных и одноцепочечных ДНК. Фермент работает в присутствии ионов Mg^{2+} .

При мягком протеолитическом расщеплении субтилизином ДНК-полимераза I дает два полипептида, один из которых обладает только $5' \rightarrow 3'$ -экзонуклеазной активностью, тогда как другой сохраняет полимеразную и $3' \rightarrow 5'$ -экзонуклеазную активности. Последний фрагмент носит название большого фрагмента ДНК-полимеразы I или фрагмента Кленова.

ДНК-полимеразу I и фрагмент Кленова широко используют в генной инженерии и при анализе структуры ДНК для «заполнения» $5'$ -выступающих одноцепочечных концов или для удаления $3'$ -выступающих концов. В первом случае эти ферменты проявляют ДНК-полимеразную активность, а во втором — $3' \rightarrow 5'$ -экзонуклеазную. При заполнении выступающих концов содержащими радиоактивную метку нуклеозидтрифосфатами ДНК оказывается меченой по концевым звеньям:



ДНК-полимераза I (но не фрагмент Кленова) широко используется для получения меченых ДНК методом так называемой *ник-трансляции* (от англ. nick, обозначающего «одноцепочечный разрыв»). Если одна из цепей ДНК содержит одноцепочечный разрыв так, что $3'$ -концевой нуклеотид в месте разрыва имеет $3'$ -ОН-группу, то ДНК-полимераза I в присутствии четырех нуклеозидтрифосфатов осуществляет матричный синтез. Одновременно она гидролизует существовавшую цепь за счет $5' \rightarrow 3'$ -экзонуклеазной активности, как показано на схеме:



По мере синтеза удлиняется 5'-концевой фрагмент и укорачивается 3'-концевой фрагмент. В результате 3'-концевой фрагмент может быть полностью замещен вновь синтезированной ДНК. Если для *ник*-трансляции использовать меченые трифосфаты, то синтезированная ДНК получается меченой.

ДНК-полимераза бактериофага Т4 получается из клеток *E. coli*, инфицированных бактериофагом Т4. Фермент обладает ДНК-полимеразной и значительно более высокой по сравнению с ДНК-полимеразой I $3' \rightarrow 5'$ -экзонуклеазной активностью, но, в отличие от последней, не проявляет $5' \rightarrow 3'$ -экзонуклеазную активность.

ДНК-полимераза Т4 может быть использована в основном так же, как ДНК-полимераза I. Сильная $3' \rightarrow 5'$ -экзонуклеазная активность позволяет применять фермент для введения метки в ДНК без выступающих концов и с $3'$ -выступающими концами.

К двухцепочечному полинуклеотиду добавляют ДНК-полимеразу и один меченый dNTP. Фермент за счет $3' \rightarrow 5'$ -экзонуклеазной активности удаляет нуклеотиды с $3'$ -конца каждой из цепей вплоть до нуклеотида, который содержит то же гетероциклическое основание, что и добавленный в смесь dNTP. За удалением этого звена следует его восстановление за счет полимеразной активности фермента. Таким образом, устанавливается псевдоравновесие: происходит постоянное отщепление и включение $3'$ -концевого звена. При каждом акте включения используется молекула dNTP, при отщеплении удаляется $5'$ -нуклеозидмонофосфат. Результатом реакции является превращение трифосфата в монофосфат. Реакция идет до тех пор, пока не израсходуется весь нуклеозидтрифосфат. После этого начинается экзонуклеазное расщепление, которое может привести к полному гидролизу ДНК.

Особое значение имеет фермент **РНК-зависимая ДНК-полимераза (обратная транскриптаза)** (открыта Г. Теминым и независимо Д. Балтимором), источником которой обычно служит вирус миелобластома птиц. Фермент, подобно ДНК-зависимым ДНК-полимеразам, осуществляет зависимость от затравки и матрицы полимеризацию дезоксирибонуклеозидтрифосфатов, синтезируя ДНК, комплементарную матрице. Затравкой служат дезоксирибо- или рибонуклеотиды, однако в качестве матрицы, наряду с одноцепочечной ДНК, фермент может использовать и одноцепочечную РНК. Фермент обладает также активностью рибонуклеазы и способен деградировать РНК в гибридном комплексе ДНК·РНК. Деградация происходит по экзонуклеазному типу как с $3'$ -, так и с $5'$ -концов РНК.

Обратные транскриптазы наиболее широко используются в генной инженерии для получения ДНК, комплементарных матричным РНК.

Важнейшим классом являются также **ДНК-зависимые РНК-полимеразы**. Эти ферменты катализируют синтез РНК, комплементарных ДНК-матрицам (транскрипцию), используя в качестве субстратов рибонуклеозидтрифосфаты. Общая схема синтеза такая же, как и у ДНК-полимераз, однако РНК-полимеразы способны начинать синтез без затравки, т. е. инициировать синтез РНК. Сигнал инициации синтеза заключен в специальных регуляторных последовательностях ДНК, называемых промоторами. (Более подробно о транскрипции см. с. 411).

Существует также большое число ферментов, аналогичных полимеразам, но способных к безматричному синтезу полинуклеотидов. Из них наибольшее применение нашла **концевая дезоксирибонуклеотидилтрансфераза (терминальная трансфераза)**. Концевая дезоксирибонуклеотидилтрансфераза (из тимуса телят) осуществляет последовательное присоединение нуклеотидных звеньев к $3'$ -ОН-концевым группам ДНК. Субстратами служат дезоксирибонуклео-



Балтимор (Baltimore) Дэвид (р. 1938), американский вирусолог. Образование получил в Массачусетском технологическом институте в Кембридже и Рокфеллеровском институте; с 1968 г.— профессор Центра онкологических исследований Массачусетского онкологического института. Основные работы — по расшифровке механизма биосинтеза белка. Изучал влияние онкогенных вирусов на генетический аппарат клетки. Открыл (1970, независимо от Г. Темина) явление обратной транскрипции, обнаружив в вирусе фермент-обратную транскриптазу. Лауреат Нобелевской премии по физиологии и медицине (1975, совместно с Р. Дульбекко и Г. Теминым).

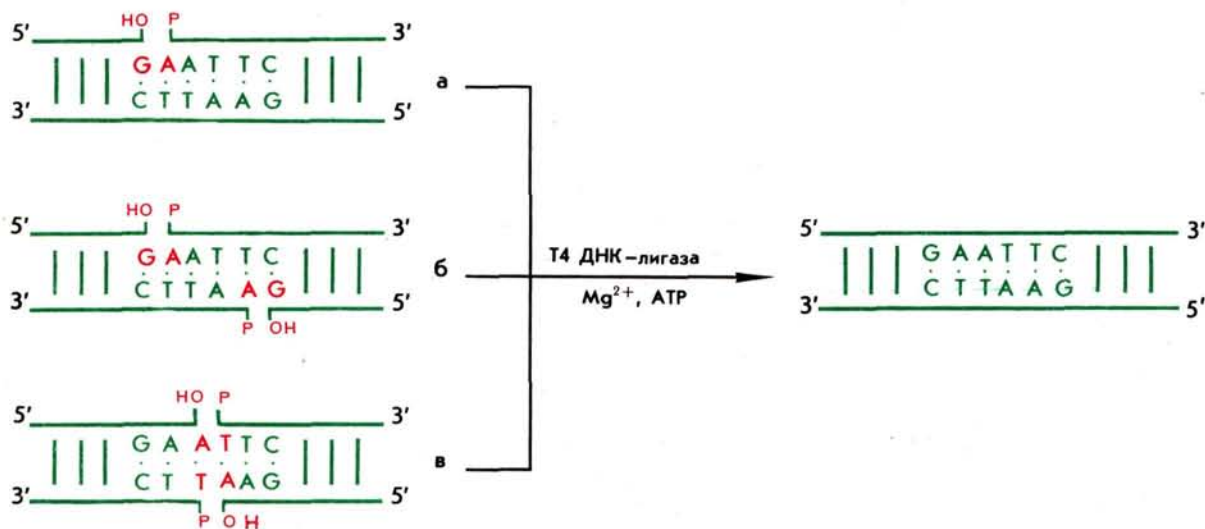
зидтрифосфаты. В присутствии ионов Mg^{2+} фермент преимущественно использует в качестве затравки одноцепочечные или двухцепочечные ДНК с 3'-выступающими концами (последние могут быть получены из полностью двухцепочечных ДНК с помощью экзонуклеазы фага λ). Однако в присутствии ионов Co^{2+} присоединение идет по любым 3'-ОН-концевым группам ДНК:



В результате к 3'-ОН-группам ДНК присоединяется полидезоксирибонуклеотидная одноцепочечная последовательность, состав которой зависит от состава смеси дезоксирибонуклеозидтрифосфатов, использованных в реакции. Если имеется один единственный трифосфат, то синтезируется гомополимер — поли·(dA), поли·(dT), поли·(dG) или поли·(dC).

Фермент используется в генной инженерии при клонировании фрагментов ДНК (см. с. 433) и в анализе последовательности ДНК для введения метки по 3'-концевым звеньям.

ДНК- и РНК-лигазы. Существуют ферменты, способные соединять между собой фосфодиэфирной связью целые фрагменты ДНК или РНК. Такие ферменты называются лигазами. Из многочисленного семейства лигаз наибольшее применение получили ДНК- и РНК-лигазы, синтезирующиеся в клетках, инфицированных бактериофагом Т4. Так, *Т4 ДНК-лигаза* катализирует образование фосфодиэфирной связи между 3'-ОН- и 5'- PO_4 -группами в одноцепочечных разрывах двухцепочечных ДНК (схема а) или между такими же группами двухцепочечных молекул ДНК, содержащих на концах комплементарные одноцепочечные последовательности (схема б).



Такими комплементарными последовательностями могут быть, в частности, одноцепочечные участки, образующиеся при расщеплении ДНК какой-либо рестриктазой. На схеме б представлен пример образования «липких» концов при расщеплении ДНК рестриктазой EcoRI. Со значительно меньшей скоростью, но все же достаточно эффективно фермент соединяет между собой полностью двуспиральные фрагменты ДНК с 3'-ОН- и 5'-PO₄-группами (схема в).

Фермент использует в качестве кофактора АТФ и требует для работы присутствия в среде ионов Mg²⁺.

ДНК-лигаза широко используется в генной инженерии для соединения фрагментов ДНК и при химико-ферментативном синтезе ДНК.

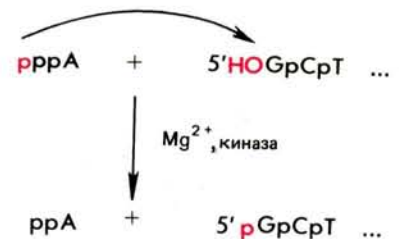
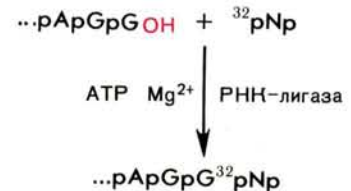
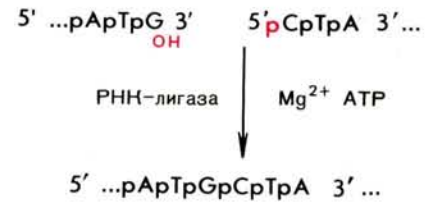
РНК-лигаза катализирует ковалентное связывание одноцепочечных ДНК или РНК, при условии, что они содержат концевые 3'-ОН- и 5'-PO₄-группы. Молекулы, содержащие 3'-ОН-группы, принято называть акцепторными, а 5'-PO₄ — донорными. В роли доноров выступают олигонуклеотиды и двухцепочечные ДНК. Минимальными донорами могут быть нуклеозид 3',5'-дифосфаты pNp и pdNp. Полирибонуклеотиды более эффективны в качестве акцепторов, чем полидезоксирибонуклеотиды. Тринуклеозиддифосфаты NpNpN — ОН являются минимальными акцепторами. Эта реакция лигирования иллюстрируется схемой.

РНК-лигазу широко используют для введения метки в 3'-концевые звенья РНК. В качестве доноров служат 5'-³²P-меченые 3',5'-нуклеозиддифосфаты. Фермент находит применение в химико-ферментативном синтезе полирибонуклеотидов. Добавление его к ДНК-лигазе значительно ускоряет реакцию «сшивания» двухцепочечных молекул ДНК.

Следует упомянуть также ферменты, которые присоединяют фосфатные группы к ДНК и РНК, — так называемые *полинуклеотидкиназы*. Для включения метки в 5'-концы ДНК (или в синтетические олигонуклеотиды) широко применяется полинуклеотидкиназа бактериофага Т4. Основная реакция, катализируемая ферментом, заключается в переносе γ-фосфатной группы АТФ на 5'-ОН-группу олиго- и полинуклеотидов. Фосфорилируются как рибо-, так и дезоксирибопроизводные. Скорость фосфорилирования двухцепочечных молекул несколько ниже, чем одноцепочечных.

В результате реакции АТФ превращается в АДФ. Киназа осуществляет также обмен 5'-концевых фосфатных групп олиго- или полинуклеотидов с γ-фосфатной группой АТФ в присутствии избытка АДФ. Фермент используют для введения меченых фосфатных групп в 5'-концевые звенья ДНК. Для этого применяют АТФ, содержащий радиоактивный атом фосфора (³²P) в γ-положении: ³²pppA.

Т4-полинуклеотидкиназа широко используется для фосфорилирования олиго- и полинуклеотидов при работе с рекомбинантными ДНК, при химико-ферментативном синтезе нуклеиновых кислот и определении их последовательности.



Химический синтез олигонуклеотидов

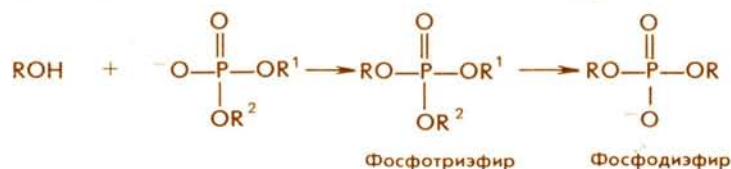
Ключевой стадией химического синтеза олигонуклеотидов является создание фосфодиэфирных связей между нуклеотидами. Можно выделить два основных подхода к решению этой задачи. Один из них использует в качестве источника фосфатной группы моноэтерифицированные фосфаты (фосфомоноэфиры), а другой — диэтерифицированные фосфаты (фосфодиэфиры).

В первом случае при химическом синтезе непосредственно образуется фосфодиэфирная связь, а во втором — получается триэфир фосфорной кислоты, который затем при полном деблокировании превращают в диэфир. Поэтому первый метод носит название *фосфодиэфирного*, а второй — *фосфотриэфирного*. Одним из распространенных вариантов фосфотриэфирного метода является *фосфитный метод*, в котором на первом этапе образуется фосфитный триэфир, окисляющийся затем в триэфир фосфорной кислоты

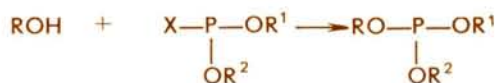
Фосфодиэфирный метод



Фосфотриэфирный метод



Фосфитный метод



R, R¹ — остатки нуклеозидов Фосфитный триэфир
R² — защитные группы

N-,O-Защитные группы. Поскольку нуклеозиды и нуклеотиды — полифункциональные соединения, все реакционноспособные группы, кроме тех, которые должны вступать в реакцию конденсации, необходимо блокировать, чтобы избежать нежелательных побочных реакций. защите подвергаются экзоциклические аминогруппы гетероциклических оснований, гидроксильные группы дезоксирибозы и,

в случае фосфотриэфирных методов, один из двух фосфатных гидроксидов. Основными требованиями к защитным группам являются стабильность в условиях конденсации, возможность введения и удаления без деструкции и модификации исходных и конечных продуктов и, наконец, возможность селективного удаления одних защитных групп при сохранении других.

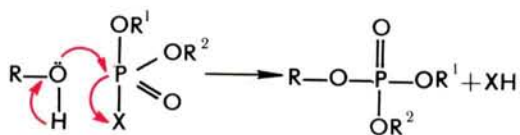
Основные защитные группы, применяемые в синтезе олигонуклеотидов, приведены в таблице 14. Аминогруппы оснований и 3'-гидроксильные группы дезоксирибозы защищаются путем ацилирования. Наиболее распространенными защитными группами при этом являются ацетильная — для 3'-гидроксила углевода, бензоильная — для аминогрупп цитидина и аденозина, анизоильная — для аминогруппы цитидина и изобутирильная — для аминогруппы гуанина. Все ацильные защитные группы могут быть удалены действием концентрированного водного аммиака.

Для блокирования 5'-гидроксильных групп обычно используют производные трифенилметана, которые легко удаляются действием кислот, в том числе и апротонных кислот Льюиса.

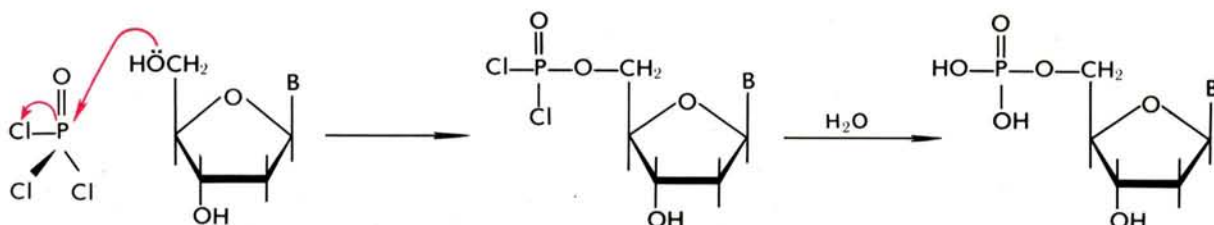
Фосфатная группировка защищается β-цианэтильной группой, которая может быть снята обработкой щелочью. В фосфотриэфирном методе в качестве второй Р-защитной группировки используют *p*- или *o*-хлорфенильные группы. Для их элиминирования применяются аммиак или ароматические альдоксины.

Две другие задачи, возникающие при создании фосфодиэфирных связей, состоят в выборе рациональных методов введения фосфатных групп в нуклеозиды и способов конденсации нуклеозидного и нуклеотидного компонентов.

Фосфорилирование нуклеозидов. В синтезе олигонуклеотидов разработаны два пути фосфорилирования нуклеозидов. В одном из них используются стабильные производные фосфорной кислоты, содержащие высокорекреационноспособную группировку. В другом такая группировка создается прямо в процессе реакции за счет взаимодействия производного фосфорной кислоты с активирующим реагентом. В обоих случаях фосфорилирование гидроксильных групп сахара проходит в результате нуклеофильного замещения у атома фосфора фосфорилирующего компонента.

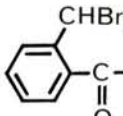

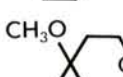
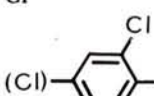
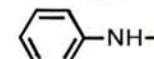
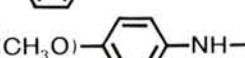


Первый путь реализуется, например, при взаимодействии оксихлорида фосфора с нуклеозидами, при этом с максимальной скоростью фосфорилируется 5'-ОН-группа. Первичный продукт реакции, содержащий связанные с фосфором атомы хлора, гидролизуетея водой:



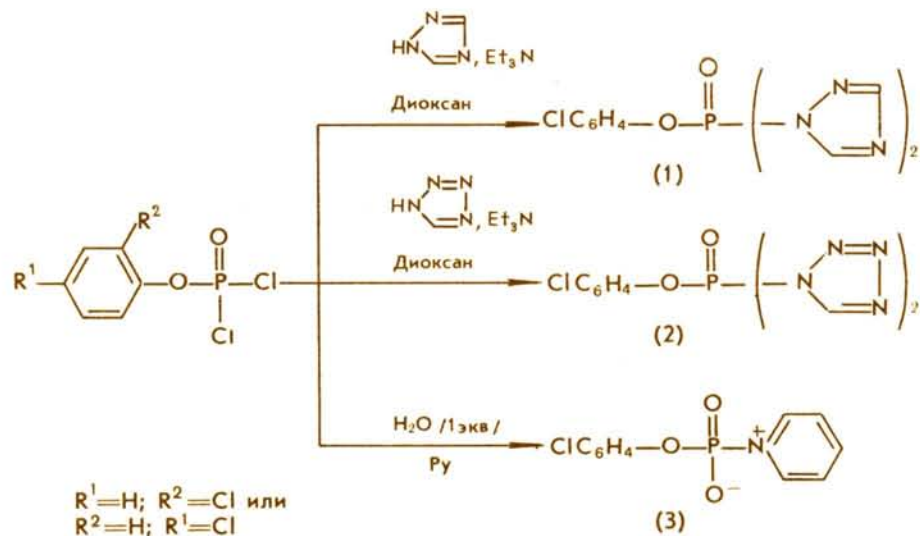
N-, O- и P-Защитные группы, используемые в синтезе олигонуклеотидов

Нуклеиновые кислоты

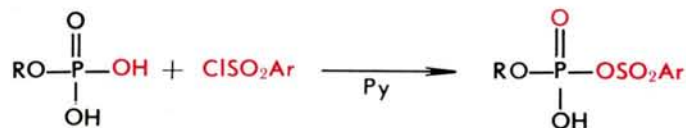
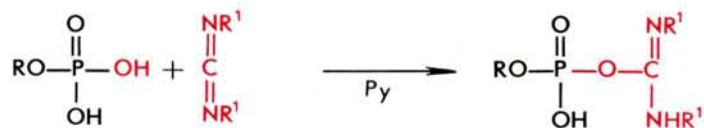
№ п/п	Защищаемая группа	Защитная группа	Формула	Сокращения	Условия отщепления
1	-NH ₂	Ацетильная	$\text{CH}_3-\text{C}(=\text{O})-$	ac	Конц. NH ₃
2	"	Бензоильная	$\text{C}_6\text{H}_5-\text{C}(=\text{O})-$	bz	"
3	"	Анизоильная	$\text{CH}_3\text{O}-\text{C}_6\text{H}_4-\text{C}(=\text{O})-$	an	"
4	"	Изобутирильная	$\begin{array}{c} \text{CH}_3 \\ \\ \text{CH}-\text{C}(=\text{O})- \\ \\ \text{CH}_3 \end{array}$	ib	"
5	5'-OH	Тритильная	$(\text{C}_6\text{H}_5)_3\text{C}-$	Tr	Бензолсульфо- кислота, три- хлоруксусная кислота, три- фторуксусная кислота, ZnBr ₂
6	"	Монометокситритильная	$\text{CH}_3\text{O}-\text{C}_6\text{H}_4(\text{C}_6\text{H}_5)_2\text{C}-$	MeOTr	
7	"	Диметокситритильная	$(\text{CH}_3\text{O}-\text{C}_6\text{H}_4)_2(\text{C}_6\text{H}_5)\text{C}-$	(MeO) ₂ Tr	
8	"	Тритилоксиацетильная	$(\text{C}_6\text{H}_5)_3\text{COCH}_2\text{C}(=\text{O})-$	Trac	Разб. NH ₃
9	"	Левулинильная	$\text{CH}_3\text{C}(=\text{O})-\text{CH}_2\text{CH}_2-\text{C}(=\text{O})-$	Lev	NH ₂ NH ₂
10	"	о-Дибромметилбензоильная		DBMB	AgClO ₄
11	3'-OH	Бензоильная	$\text{C}_6\text{H}_5-\text{C}(=\text{O})-$	Bz	Конц. NH ₃ , NaOH
12	"	Ацетильная	$\text{CH}_3-\text{C}(=\text{O})-$	Ac	"
13	2'-OH	Тetraгидропиранильная		Thp	Разб. HCl
14	"	Метокситетрагидропиранильная		Mthp	"
15	Фосфатная	β-Цианэтильная	$\text{N}\equiv\text{C}-\text{CH}_2-\text{CH}_2-$	CNEt	NH ₃ , NaOH, триэтиламин
16	"	β, β, β-Трихлорэтильная	$\begin{array}{c} \text{Cl} \\ \\ \text{C}-\text{CH}_2- \\ \\ \text{Cl} \end{array}$	TCEt	H ₂ /Zn-Cu
17	"	(п-) или о-Хлорфенильная		ClPh	NH ₃ , разб. NaOH, син-2-нитро- бензальдегид или пиридин-2- карбальдегид
18	"	Анилидная и замещенные анилидные	 		Изоамиленитрит

В более жестких условиях при замещенной 5'-ОН-группе аналогичным образом можно получить 3'-фосфаты.

Для фосфорилирования гидроксильных групп остатка рибозы или дезоксирибозы используют также активированные фосфамиды, например фосфотри-(1) и тетразолиды (2), которые получают взаимодействием хлорфосфатов с триазолом или тетразолом. Эффективные фосфорилирующие реагенты (3) образуются также при действии 1 экв. H_2O в пиридине на арилдихлорфосфаты.

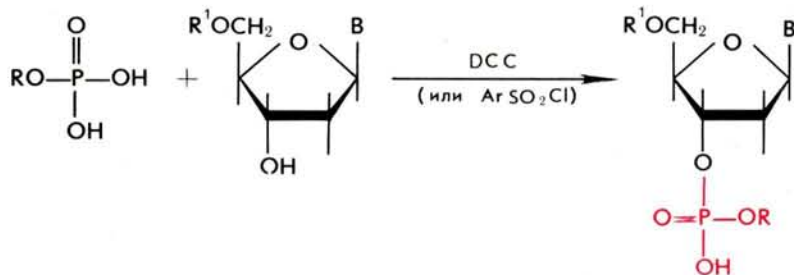


Как уже было отмечено, фосфорилирование нуклеозидов осуществляется также производными фосфорной кислоты в присутствии активирующих агентов, в роли которых могут выступать дициклогексилкарбодиимид (DCC) или арилсульфохлориды ArSO_2Cl . И в том и в другом варианте на первом этапе реакции происходит активация фосфатной группы с образованием промежуточных соединений, приведенных ниже

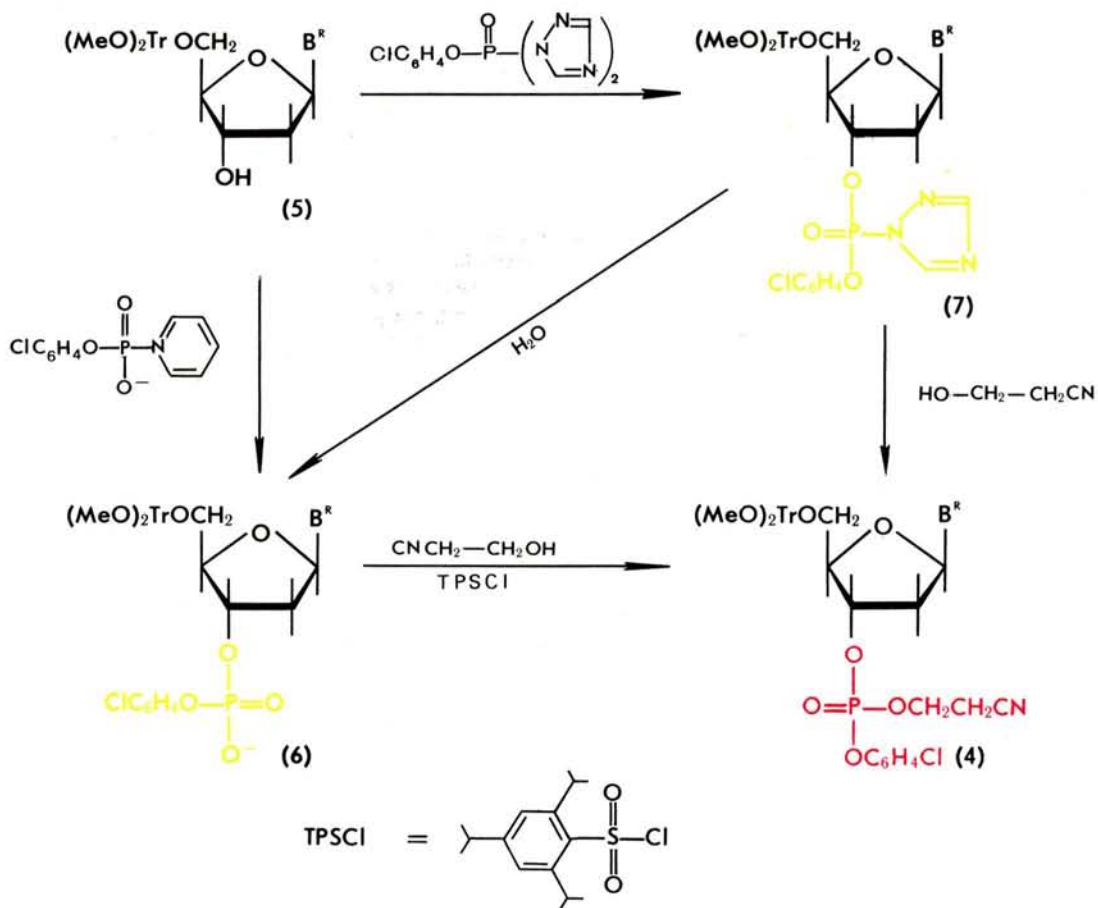


R — остаток нуклеозида

$R^1 = \text{C}_6\text{H}_{11}$



Механизмы дальнейших превращений этих соединений сложны, но в конечном счете происходит нуклеофильное замещение в фосфорилирующем агенте, причем роль нуклеофила играет защищенный нуклеозид со свободной гидроксильной группой. В понимание механизмов активации существенный вклад был внесен Д. Г. Кнорре.



Триизопропилбензолсульфохлорид

R—N-защитные группы

B^R—N-защищенное гетероциклическое основание

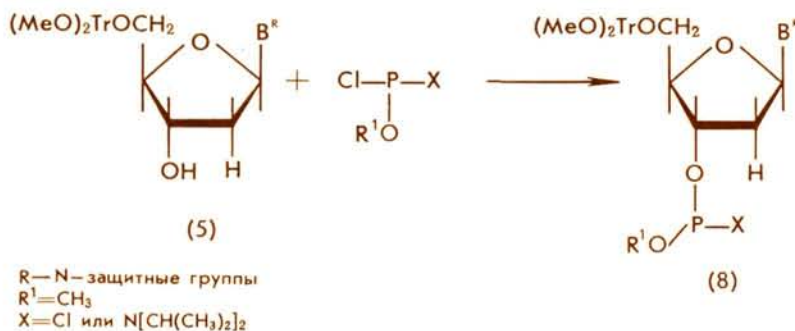
Недостатком карбодиимидного метода является малая скорость реакции и образование большого числа побочных продуктов. Использование арилсульфохлоридов позволяет избежать побочных реакций. В качестве активирующих агентов могут выступать арилсульфотри- или тетразолиды. Вероятно, в этом случае активными фосфорилирующими соединениями являются образующиеся в результате реакции фосфотри- и тетразолиды.

В качестве примера приведены два пути получения фосфорилированного производного (4) — ключевого соединения в фосфотриэфирном методе. Первый из них включает обработку исходного нуклеозида (5) бис-(триазолидом) *n*-хлорфенилфосфорной кислоты с последующим превращением промежуточного продукта (7) в триэфир (4) действием этиленциангидрина. Второй — получение диэфира нуклеотида (6), а затем превращение его в триэфир (4) с помощью конденсирующего агента.

В последние годы в связи с разработкой фосфитного варианта фосфотриэфирного метода синтеза олигонуклеотидов все большее значение приобретают методы введения в нуклеотиды производных фосфористой кислоты с образованием соединений типа (8)



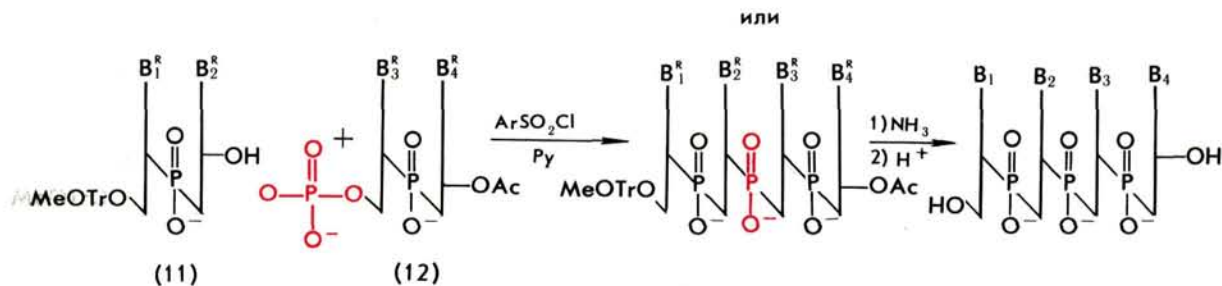
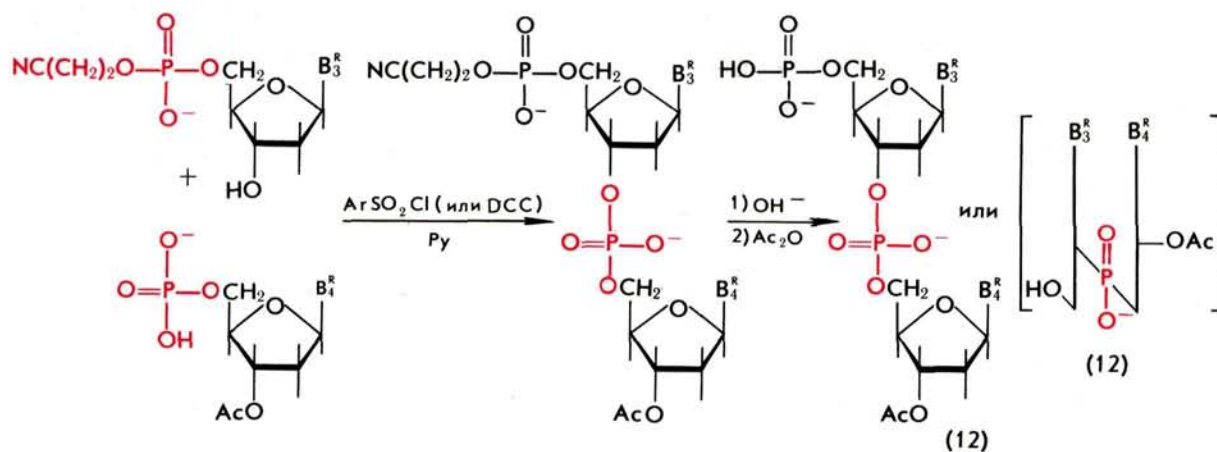
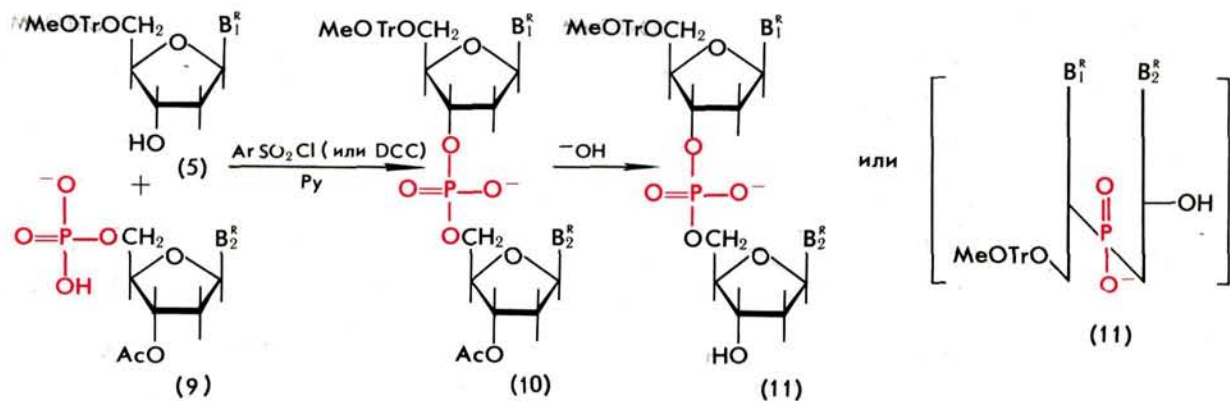
Кнорре Дмитрий Георгиевич (р. 1926), советский химик и биохимик, академик АН СССР (1981). Окончил Московский химико-технологический институт им. Д. И. Менделеева (1947). Директор Института биоорганической химии СО АН СССР (Новосибирск) и профессор Новосибирского университета. Основные работы посвящены изучению направленной химической модификации нуклеиновых кислот и белков, механизма пептидного синтеза.



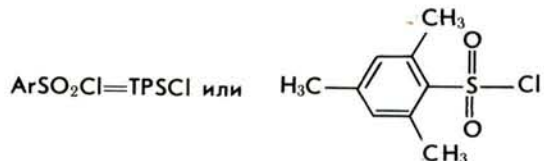
Фосфодиэфирный метод синтеза олигонуклеотидов. Согласно использовавшейся ранее методологии, разработанной в лаборатории Г. Кораны, фосфодиэфирный метод включает в себя конденсацию нуклеозидного компонента с незащищенной 3'-ОН-группой (5) с нуклеотидным компонентом, имеющим свободную 5'-фосфомоноэфирную группу (9) (рис. 201).

В качестве конденсирующих агентов применяются арилсульфохлориды или DCC. Образующийся защищенный динуклеозидмонофосфат (10) либо полностью деблокируется, либо в нем удаляется только ацильная группа, блокирующая 3'-гидроксил. В последнем случае образуется динуклеозидмонофосфат (11) со свободной 3'-ОН-группой, используемый в качестве нуклеозидного компонента для присоединения следующего звена.

Метод синтеза, при котором на каждом последующем шаге присоединяется мономерное звено, называется ступенчатым. Обычно таким образом синтезируются не более чем 3—4-звенные олигонуклеотиды, что связано с трудностями отделения продукта конденсации от исходного олигонуклеотида по мере увеличения длины цепи, поскольку они различаются всего на одно звено. Для синтеза более длинных олигонуклеотидов используют блочный метод. К динуклеотиду со свободной 3'-гидроксильной группой (11)



$\text{B}^R = \text{bzA; anC; ibG}$ или T



Мезитилсульфохлорид

Рис. 201. Общая схема синтеза тетра-
нуклеотида фосфодиэфирным методом.

(нуклеозидный компонент) присоединяют динуклеотид с незащищенной 5'-фосфатной группой, но защищенным 3'-гидроксилом (12) (нуклеотидный компонент).

Однако и при использовании блочного метода синтез сравнительно длинных (более 10—12 звеньев) олигонуклеотидов фосфодиэфирным методом оказывается затруднительным в связи с многочисленными побочными реакциями с участием незащищенной фосфатной группы. Существенный прогресс в синтезе олигонуклеотидов, позволивший получать длинные последовательности, связан с применением усовершенствованного фосфотриэфирного метода.

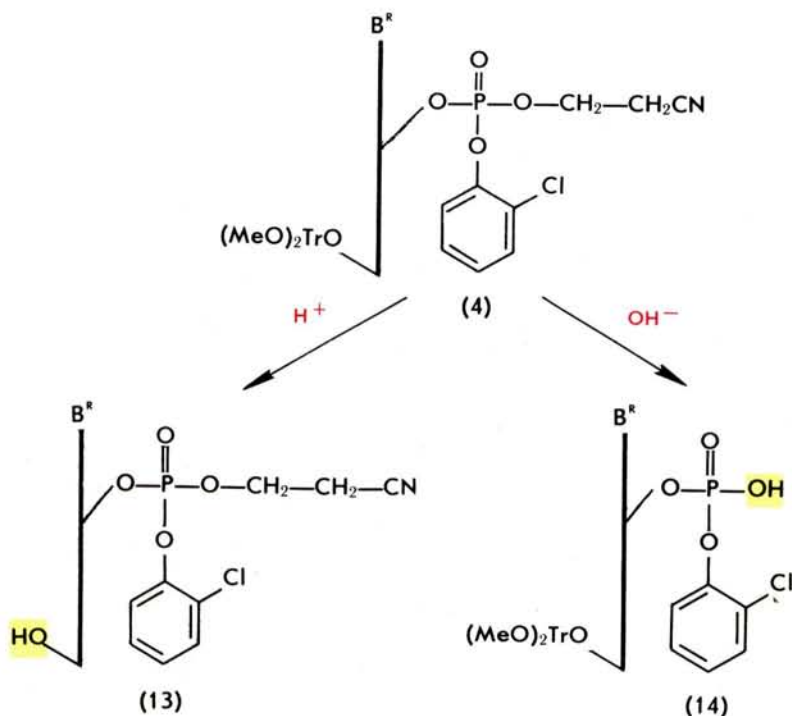
Фосфотриэфирный метод синтеза олигонуклеотидов. Фосфотриэфирный метод является исторически более ранним методом синтеза олигонуклеотидов. Именно таким путем А. Тоддом и М. Микельсоном был получен первый синтетический олигонуклеотид. Однако широкого распространения метод не имел ввиду отсутствия достаточно эффективных защитных групп и конденсирующих агентов. С появлением соответствующих реагентов фосфотриэфирный метод и его варианты практически заменили фосфодиэфирный метод.

В современном фосфотриэфирном методе используется ключевое соединение (4), которое легко может быть превращено как в нуклеозидный (13), так и нуклеотидный (14) компоненты.

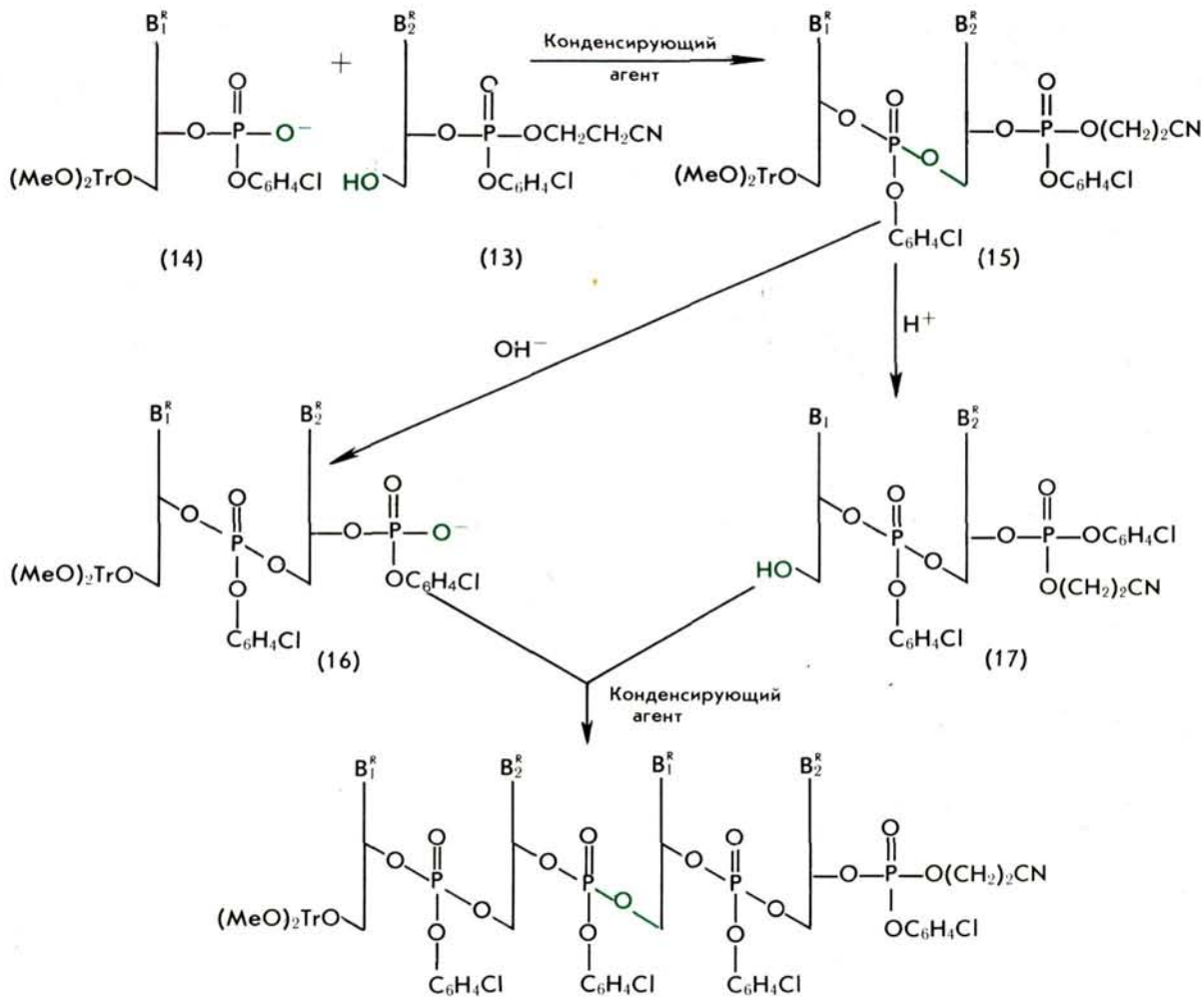
5'-Диметокситригильная группа избирательно удаляется действием слабой кислоты (например, бензолсульфокислоты в хлороформе или с помощью бромистого цинка), β-цианэтильная группировка — действием щелочи, при этом о- или п-хлорфенильная группа остается в нуклеозидном и нуклеотидном компонентах в качестве защиты для межнуклеотидного фосфата.



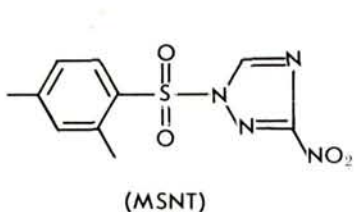
Тодд [Todd] Александер Робертус (р. 1907), английский химик-органик и биоорганик. Образование получил в университетах Глазго и Франкфурта-на-Майне (1931), с 1944 г. — профессор Кембриджского университета. Автор фундаментальных трудов по изучению структуры нуклеиновых кислот, синтезу витаминов. Синтезировал (1936) витамин В₁. Совместно с Д. Брауном предложил общую схему строения РНК. Лауреат Нобелевской премии по химии (1957).



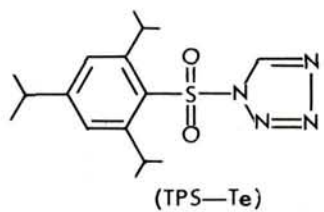
R — N-защитная группа



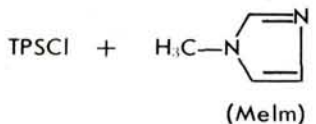
Конденсирующие агенты, используемые в триэфирном методе:



Мезитилсульфо - 3-нитро -
1, 2, 4 - триазаolid



Триизопропилбензолсульфо-
тетразаolid



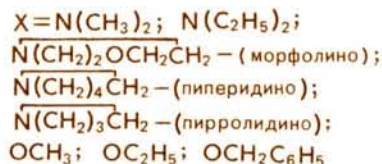
N - Метилимидазол

Рис. 202. Фосфотриэфирный метод синтеза олигонуклеотидов.

При конденсации нуклеозидного (13) и нуклеотидного компонентов (14) образуется полностью защищенный динуклеозид-монофосфат (15). Дальнейшее удлинение цепи (так же как и в диэфирном методе) может идти ступенчатым или блочным методом (рис. 203). В первом случае в качестве нуклеозидных компонентов выступают олигонуклеотиды с 5'-концевой незащищенной гидроксильной группой, а в качестве нуклеотидных — мононуклеотиды с незащищенной 3'-фосфодиэфирной группой. Цепь обычно наращивается от 3'- к 5'-концу. Во втором случае и нуклеозидным и нуклеотидным компонентами служат олигонуклеотиды, подобные (16) и (17). Фосфотриэфирным способом удается стабильно получать 15 — 20-звенные олигонуклеотиды, а также более длинные полинуклеотиды, вплоть до 50 — 60-звенных.

В качестве конденсирующих реагентов в фосфотриэфирном методе ранее чаще всего применялись арилсульфоазолиды: 4-нитроимидазолиды, 1,2,4-триазолиды, 3-нитро-1,2,4-триазолиды и тетразолиды (рис. 202). В то же время хорошо зарекомендовавшие себя в диэфирном методе синтеза арилсульфохлориды оказались неэффективными из-за медленного протекания реакции и низкого выхода целевого продукта. Однако в ходе недавних исследований (В. А. Ефимов, О. Г. Чахмахчева) по синтезу олиго- и полинуклеотидов было обнаружено, что в присутствии некоторых нуклеофильных катализаторов арилсульфохлориды оказываются значительно более эффективными конденсирующими реагентами для создания фосфотриэфирной связи, чем самые активные из арилсульфоазолидов. Одним из таких нуклеофильных катализаторов является N-метилимидазол (MeIm), в присутствии которого реакция межнуклеотидной конденсации протекает за несколько минут. Использование MeIm дало возможность проведения реакций не только в пиридине, но и в других органических растворителях (диоксан, ацетонитрил, хлористый метилен, хлороформ, нитрометан).

Дальнейшее усовершенствование фосфотриэфирного метода, позволившее проводить межнуклеотидную конденсацию за 1 — 2 мин, основано на использовании в качестве конденсирующих реагентов арилсульфохлоридов в присутствии 4-замещенных производных N-окисей пиридина и хинолина



Фосфитный метод. Широкое применение в последнее время получил фосфитный способ синтеза олигонуклеотидов (с. 366). Фосфитное производное (8) может содержать в качестве замести-

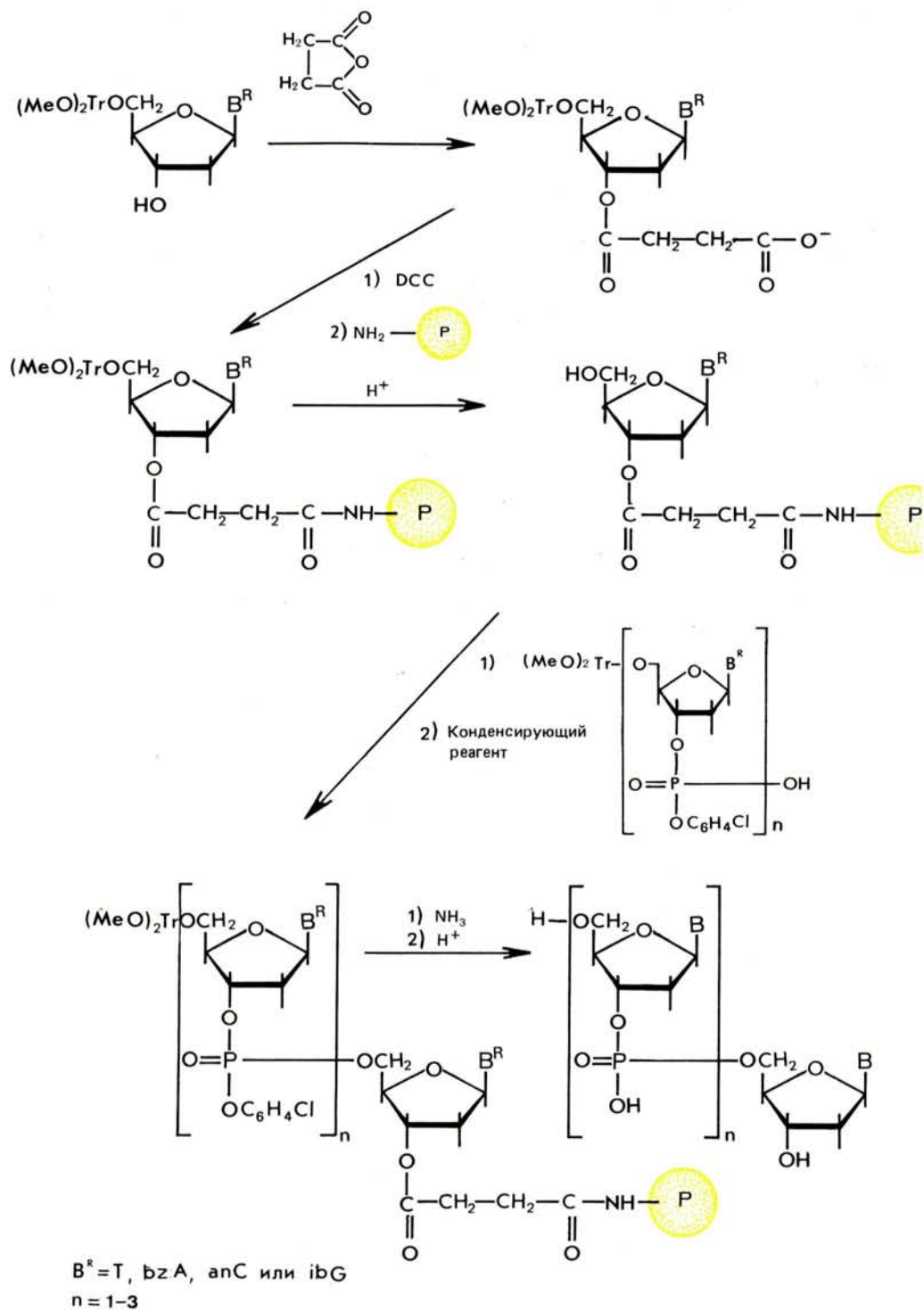
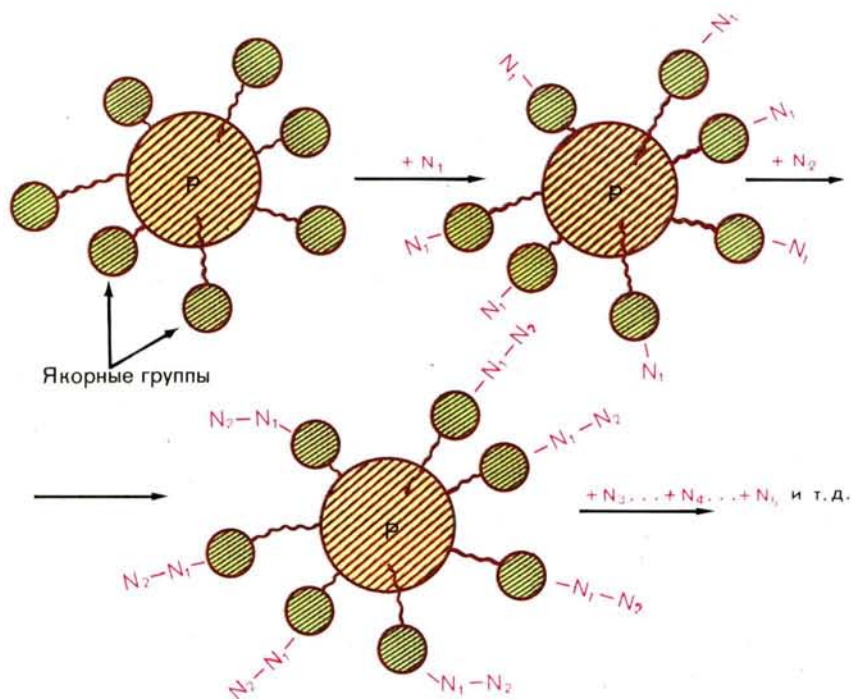


Рис. 203. Твердофазный фосфотриэфирный метод синтеза олигонуклеотидов (фосфатный вариант).

теля галоген, тетразол или вторичный амин (в последнем случае метод называют фосфамидитным). В результате конденсации образуется динуклеотид метоксифосфин (18), который может быть далее окислен в мягких условиях с образованием фосфотриэфира (19). Наиболее эффективен фосфитный метод в твердофазном синтезе олигонуклеотидов (рис. 204).

Синтез на полимерных носителях. Так же как и в случае полипептидов, синтез олигонуклеотидов может быть осуществлен ступенчатым удлинением цепи, ковалентно привязанной одним из концов (5' или 3') к полимерному носителю. Синтез олигонуклеотидов на полимере начал развиваться в 60-х годах усилиями ряда групп: Ф. Крамера (ФРГ) и Г. Кораны с сотр. (США), З. А. Шабаровой, М. А. Прокофьева с сотр. (СССР). Такой подход принципиально более эффективен, чем синтез в растворе, так как существенно облегчается отделение продукта от исходных веществ, что приводит к сокращению времени синтеза. Кроме того, все операции легко могут быть автоматизированы (первые автоматические системы были предложены М. Гайтом и М. Карузерсом). Принципиальная схема синтеза на полимерном носителе выглядит следующим образом:



Для осуществления таким способом синтеза достаточно длинных полимеров необходимы высокие выходы на каждой стадии конденсации. Существенным фактором успеха при твердофазном синтезе является также качество полимерного носителя, который должен быть инертным, содержать стерически доступные функциональные группы для присоединения первого звена, иметь достаточный размер пор, чтобы обеспечить свободную диффузию реагентов, и, наконец, неспецифически не связывать реагенты и продукты.

Найден целый ряд полимеров, удовлетворяющих этим требованиям. К их числу относятся полиакриламид, полиакрилморфолид, полидиметилакриламид, полистирол, силикагель, пористое стекло, бумага и т. д.

Синтезы олигонуклеотидов с использованием фосфатного и фосфитного фосфотриэфирных методов приведены на рисунках 203 и 204. На первой стадии нуклеозид присоединяют с помощью якорной группы к полимерному носителю. Затем его 5'-гидроксильную группу деблокируют кислотной обработкой и конденсируют с нуклеотидным компонентом. У образующегося полностью защищенного динуклеозидмонофосфата удаляют диметокситритильную группу и присоединяют следующий нуклеотид и т. д. На последней стадии продукт отщепляют от полимерного носителя, снимают защитные группы и очищают хроматографией или электрофорезом в полиакриламидном геле.

Фосфотриэфирный и фосфитный методы дают высокие выходы на каждой стадии синтеза и с успехом используются для синтеза олигонуклеотидов на полимерных носителях.

В настоящее время существуют синтезаторы олигонуклеотидов, работающие по заданной программе.

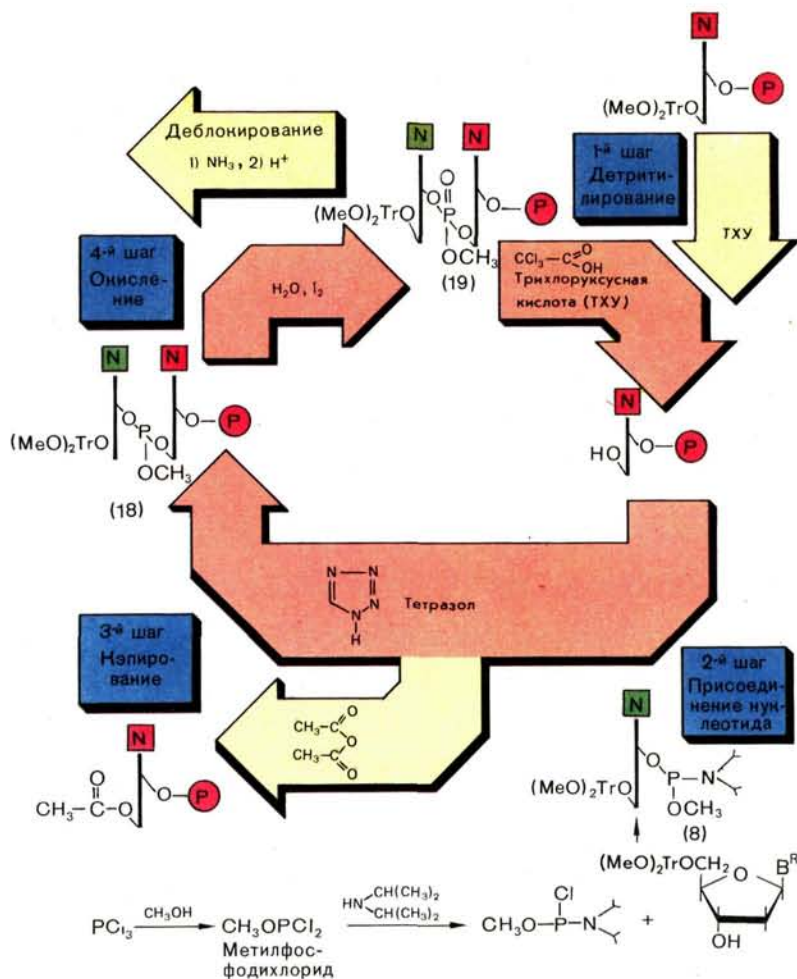
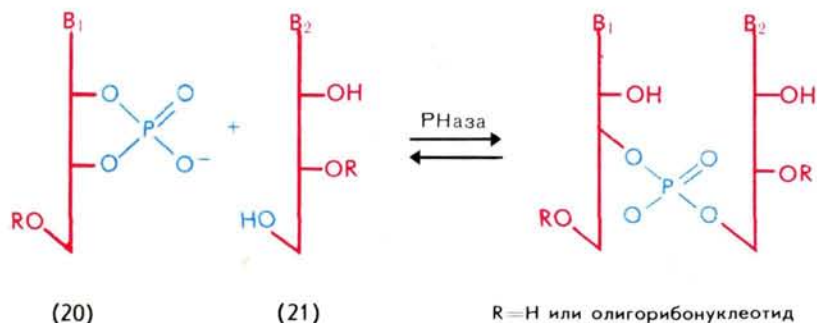


Рис. 204. Твердофазный фосфотриэфирный метод синтеза олигонуклеотидов (фосфитный вариант).

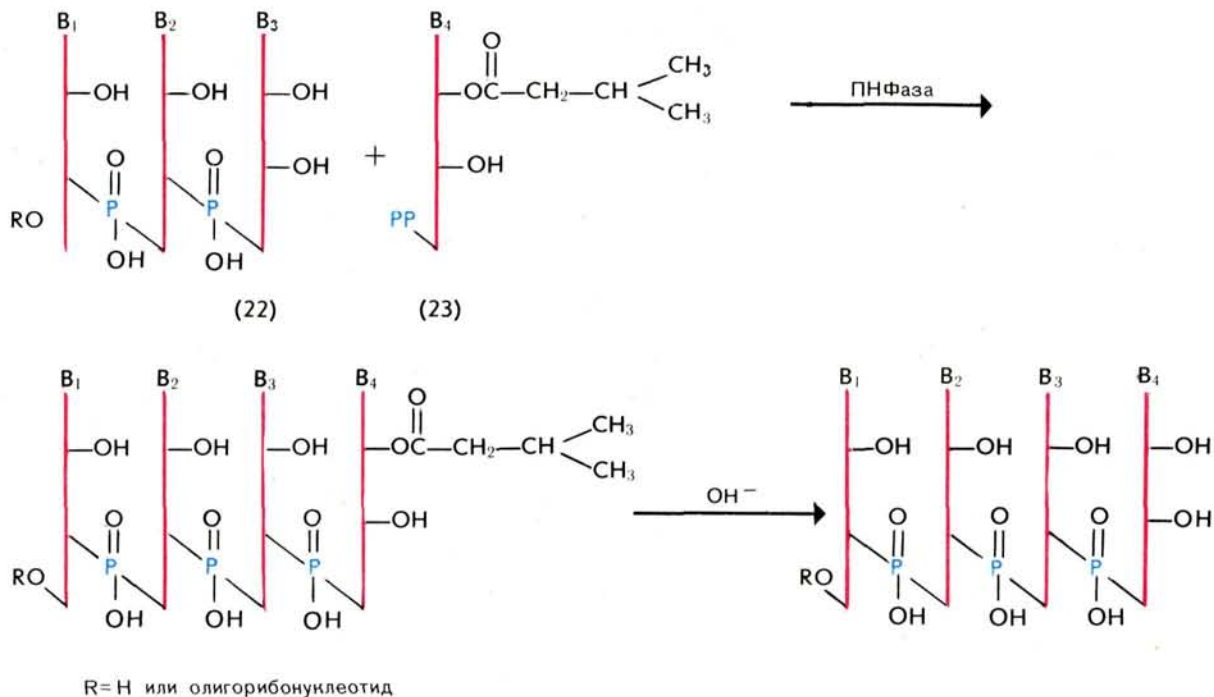
Синтез полирибонуклеотидов

Синтез сравнительно коротких олигорибонуклеотидов может быть осуществлен химическим или ферментативным методом. Более длинные последовательности получают путем соединения коротких фрагментов с помощью РНК-лигазы.

Ферментативный синтез олигорибонуклеотидов. Как правило, для синтеза сравнительно коротких фрагментов — ди- и тетра-нуклеотидов используют реакцию, катализируемую рибонуклеазами:



Эта реакция обратна реакции гидролиза и заключается во взаимодействии 2',3'-рибонуклеозидциклофосфата (20) (нуклеотидный компонент) с 5'-ОН-группой рибонуклеозиды (21) (нуклеозидный



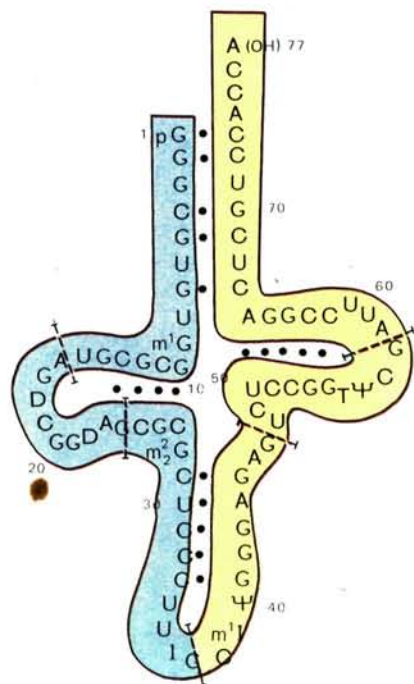


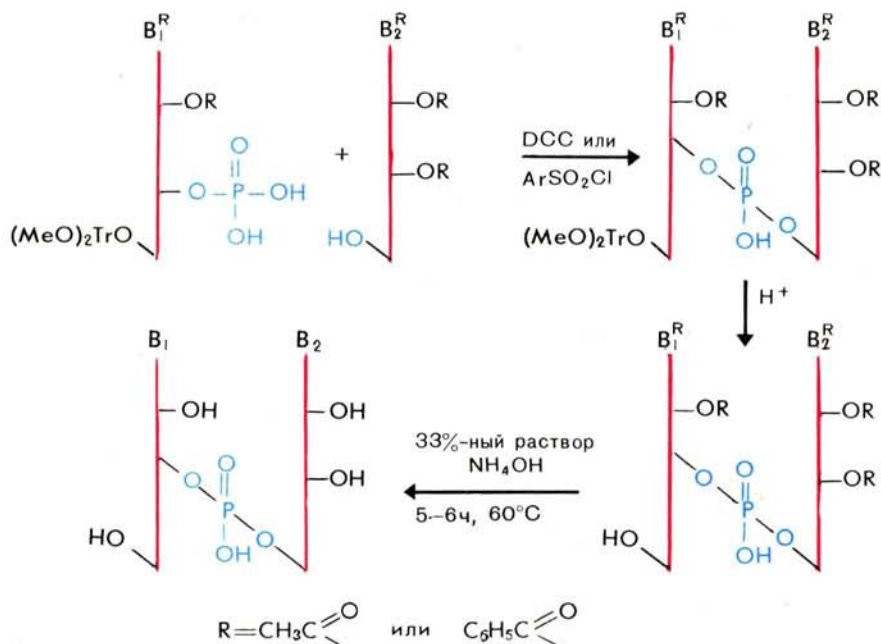
Рис. 205. Структура дрожжевой аланиновой тРНК, полученной химико-ферментативным синтезом.

компонент). В качестве нуклеотидного и нуклеозидного компонентов выступают как мономеры, так и олигонуклеотиды. При использовании неспецифических рибонуклеаз (таких, как РНазы из *Bacillus subtilis* или *Penicillium brevicompactum*) природа оснований B_1 и B_2 не очень существенна. Могут быть применены и специфические РНазы, например гуанилрибонуклеаза (РНаза T_1), если B_1 представляет собой гуанин, или панкреатическая РНаза для пиримидиновых оснований. При увеличении длины синтезируемого продукта все большую роль играет его гидролиз РНазами. В таком случае для наращивания цепи выгодно использовать специфические РНазы, неспособные расщеплять уже имеющуюся цепь. Помимо рибонуклеаз, для ферментативного синтеза олигонуклеотидов находят применение полинуклеотидфосфорилаза (ПНФаза). Для того чтобы фермент присоединял только одно звено к 3'-концу олигонуклеотида-затравки (22), одну из гидроксильных групп дифосфата (23) блокируют объемистым ацильным заместителем, который после реакции удаляют мягкой щелочной обработкой.

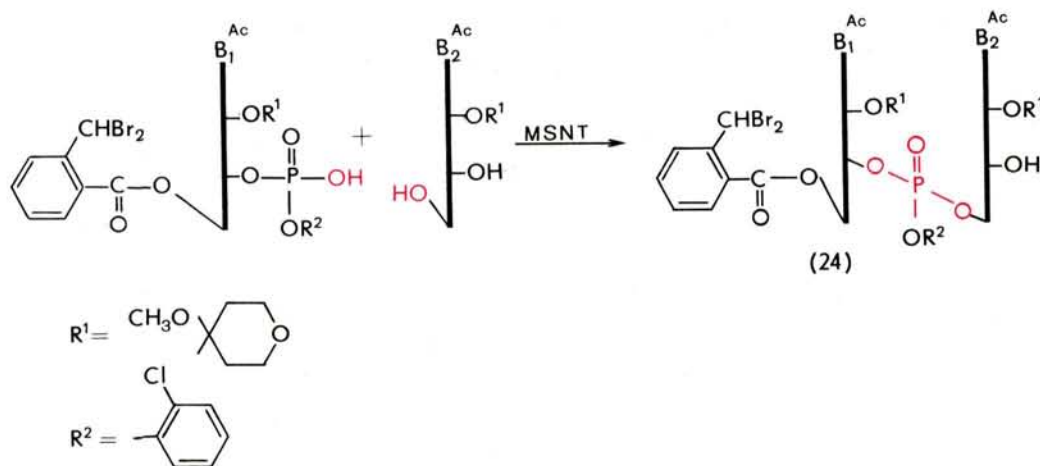
Синтез олигорибонуклеотидов может быть осуществлен комбинацией РНаз и ПНФазы.

Химический синтез олигорибонуклеотидов. Олигорибонуклеотиды синтезируют фосфодиэфирным, фосфотриэфирным и фосфитным методами с использованием в основном аналогичных приемов и реагентов, описанных для ДНК-ряда. Дополнительной проблемой является необходимость избирательной защиты 2'-ОН-группы рибозы, а также лабильность фосфодиэфирной связи олиго- и полирибонуклеотидов в щелочной среде.

Фосфодиэфирный метод был использован Г. Кораной для синтеза всех ди- и тририбонуклеотидов при установлении генетического кода. Пример получения рибодинуклеозидмонофосфата с применением фосфодиэфирного способа дан на схеме:



Рассмотрим один из примеров синтеза защищенного рибонуклеозидмонофосфата (24) с помощью фосфотриэфирного метода. В этом случае с целью защиты 5'-гидроксильной группы использована *o*-дибромметилбензоильная защита, которая легко снимается действием ионов серебра, а 2'-гидроксильная группа рибозы блокируется неустойчивой в кислых условиях метоксипиранильной группировкой. Остальные защитные группы и конденсирующие агенты не отличаются от используемых в синтезе олигодезоксирибонуклеотидов



Одним из примеров синтеза природных рибополинуклеотидов является полный химико-ферментативный синтез дрожжевой аланиновой тРНК (рис. 205), осуществленный большой группой

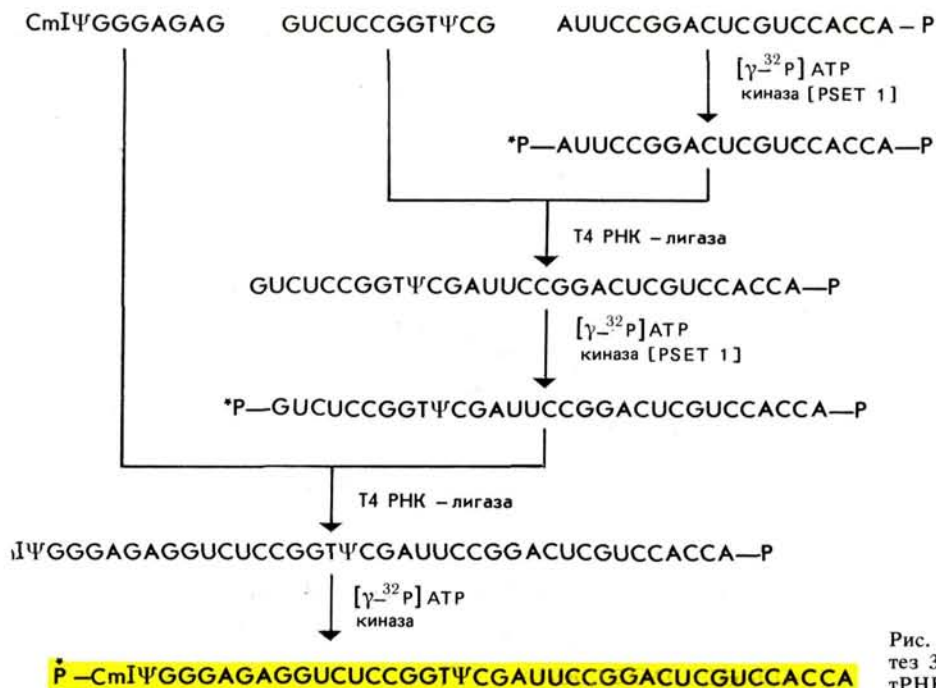


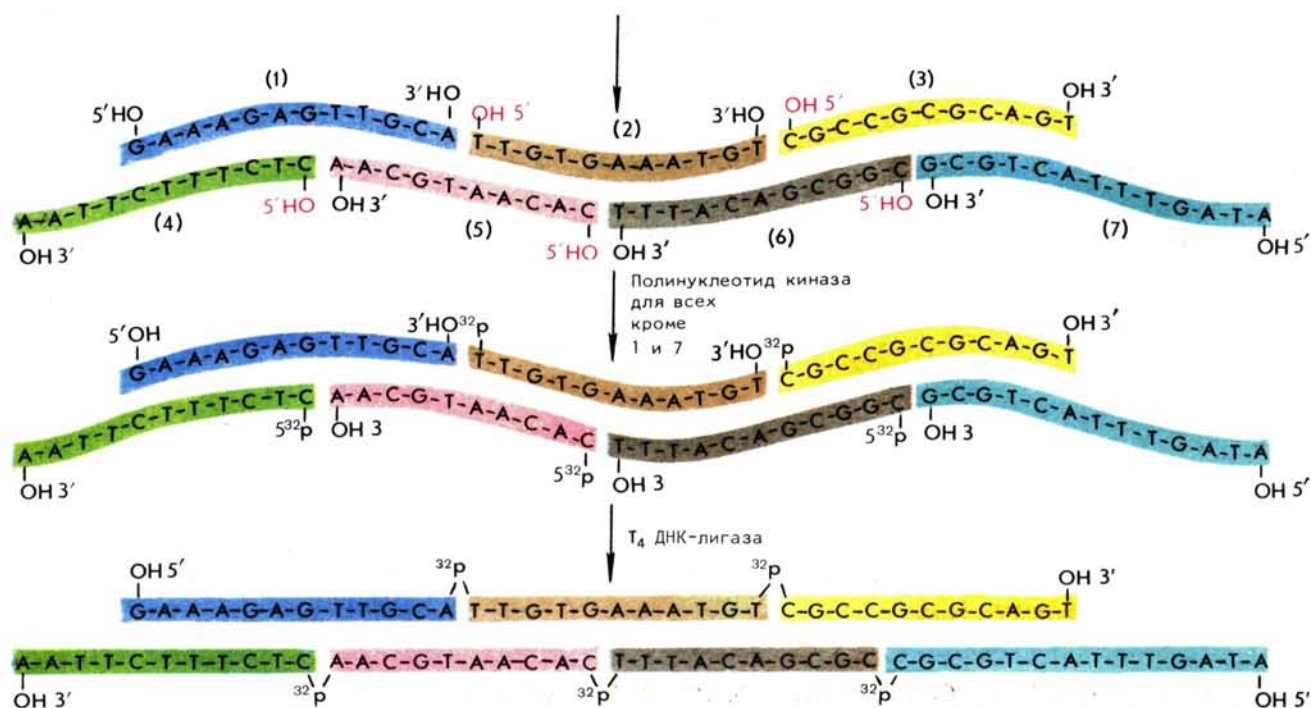
Рис. 206. Химико-ферментативный синтез 3'-половины дрожжевой аланиновой тРНК (нуклеотиды 36—76).

китайских химиков под руководством профессора Ван Иньяля. Отличительной особенностью синтеза явилось то, что синтетическая тРНК содержала минорные нуклеотиды. Для синтеза 76-нуклеотидная последовательность тРНК была разбита на 6 олигонуклеотидных сегментов величиной от 9 до 19 звеньев. Синтез этих сегментов был проведен фосфотриэфирным методом, полученные олигонуклеотиды далее «соединялись» друг с другом при помощи Т4 РНК-лигазы, как показано на рисунке 206. Изучение биологической активности синтетической тРНК показало, что она обладает полной активностью природной дрожжевой аланиновой тРНК в реакциях аминокислотирования и образования пептидной связи на рибосоме.

Химико-ферментативный синтез фрагментов ДНК

Комплекс современных методов синтеза нуклеиновых кислот позволяет исходя из мононуклеотидов получать гены, кодирующие белки длиной более 100 аминокислотных остатков. Первым этапом работы является химический синтез олигодезоксирибонуклеотидов, которые затем с помощью ферментов нуклеинового обмена, таких, как Т4 полинуклеотидкиназа, Т4 ДНК-лигаза и ДНК-полимеразы, превращаются в двухцепочечные фрагменты ДНК (рис. 207). Мето-

Рис. 207. Общая стратегия синтеза двухцепочечной ДНК.



дология сборки синтетических олигомеров в протяженные дуплексы впервые была предложена в конце 70-х годов Г. Кораной и сотр. и долгое время оставалась практически неизменной. Разработанный ими подход (рис. 208, *a*) состоит в многократном последовательном соединении при помощи Т4 ДНК-лигазы 10 — 20-звенных олигонуклеотидов с комплементарными, взаимно перекрывающимися последовательностями оснований сначала в небольшие дуплексы с выступающими одноцепочечными последовательностями на концах, а затем — в целый ген, который клонируют в векторной молекуле (см. с. 431). Этим способом был получен целый ряд полинуклеотидов, в том числе гены некоторых интерферонов человека, с длиной более 500 п. о.

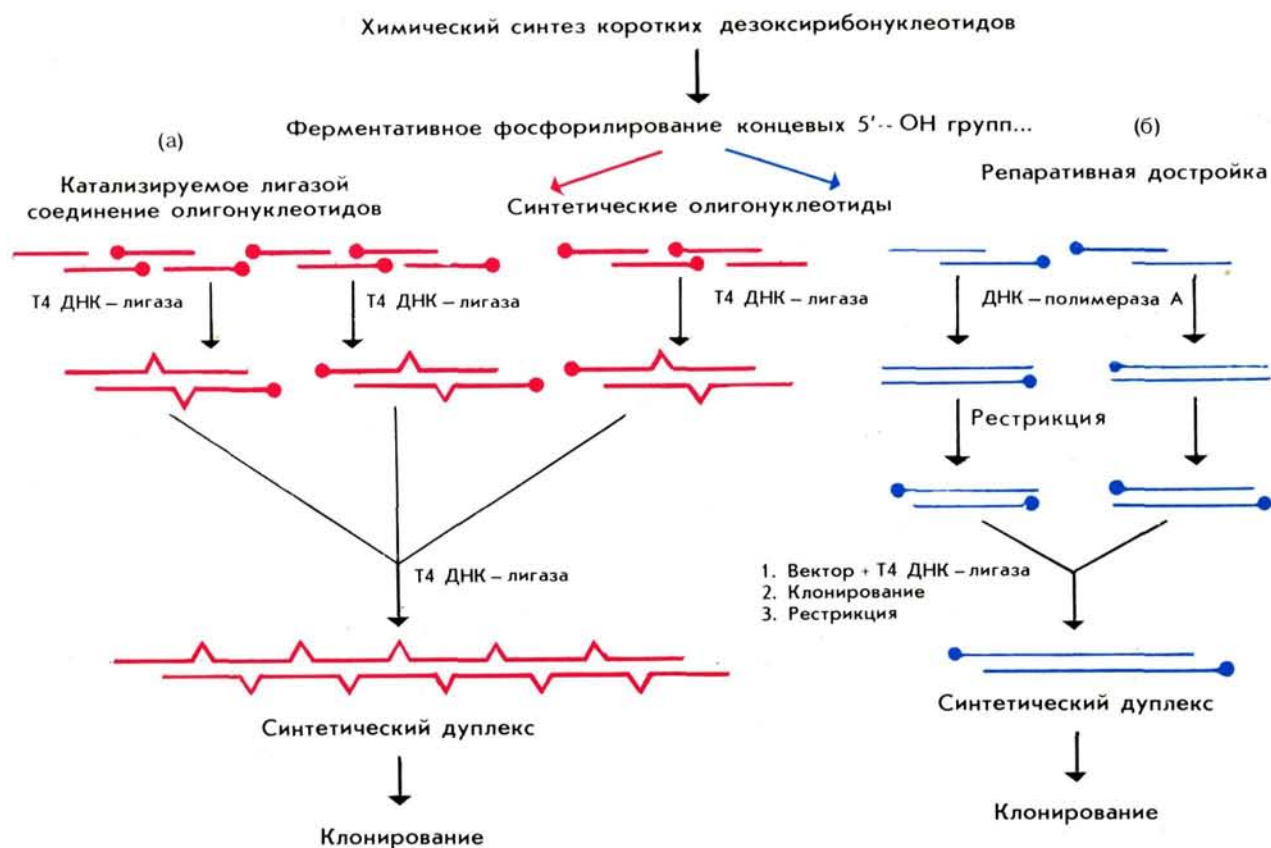


Рис. 208. Два подхода к получению двухцепочечных полинуклеотидов методами химико-ферментативного синтеза: *a* — использующий для соединения отдельных фрагментов реакцию, катализируемую Т4 ДНК-лигазой; *b* — с применением репаративной достройки частичного дуплекса с помощью ДНК-полимеразы (* — 5'-фосфорилированный олигонуклеотид; — олигонуклеотид, не имеющий фосфатной группы на 5'-конце).

С развитием методов синтеза олигонуклеотидов стало возможным получать чисто химическими способами 30 — 60- и даже 100-звенные олигомеры. С одной стороны, это привело к упрощению описанной методологии (за счет существенного сокращения числа лигазных реакций при сборке гена), а с другой — создало основу для разработки принципиально иного подхода к получению синтетических дуплексов. Такой подход, предложенный К. Итакура и сотр., заключается в репаративной достройке с помощью ДНК-полиме-



Корана [Khorana] Хар Гобинд (р. 1922), американский биохимик, иностранный член АН СССР (1971). Образование получил в Пенджабском и Ливерпульском университетах, с 1970 г. — в Массачусетском технологическом институте. Основные работы — в области синтеза нуклеиновых кислот. Внес большой вклад в расшифровку генетического кода. Впервые синтезировал ген аланиновой тРНК (1970). Выполнил ряд важных работ по структуре мембранных белков. Лауреат Нобелевской премии по физиологии и медицине (1968, совместно с Р. Холли и М. Ниренбергом).

разы *E. coli* частичного дуплекса, образованного 3'-концевыми последовательностями двух длинных олигонуклеотидов, в полностью двухцепочечный фрагмент ДНК. Последний после образования на одном из его концов «липкого» сайта (действием соответствующей рестриктазы) соединяется в составе векторной молекулы с другим, аналогично синтезированным дуплексом (рис. 208, б). Одним из примеров применения этого способа явился синтез фрагмента гена человеческого лейкоцитарного интерферона α_2 длиной 132 п. о.

Хотя первый из описанных выше подходов теоретически пригоден для построения генов любой длины, на практике по мере увеличения числа последовательных лигазных «сшивков» становится все труднее очищать продукты реакции и получать их в достаточном для дальнейшей работы количестве. Кроме того, увеличивается расход исходных синтетических олигонуклеотидов. В то же время второй подход пригоден только для получения фрагментов длиной не более 200 — 250 п. о. В связи с этим для облегчения сборки гена (очистки фрагментов и экономии нуклеотидного материала) более целесообразно составлять его из отдельных модулей, предварительно очищенных промежуточным клонированием.

Клонирование синтетических полидезоксирибонуклеотидов

Одним из этапов синтеза искусственных двухцепочечных фрагментов ДНК является их клонирование в составе многокопийного вектора с тем, чтобы синтезированная последовательность сохранялась и при необходимости могла быть наработана в достаточных количествах. С этой целью синтез двухцепочечных фрагментов осуществляют таким образом, чтобы они содержали на концах участки, идентичные получающимся при расщеплении ДНК определенной рестрикционной эндонуклеазой. В качестве примера на рисунке 209 приведена последовательность фрагмента ДНК, кодирующего короткий нейропептид — лейцин-энкефалин. Наряду с собственно кодирующей последовательностью фрагмент содержит иницирующий АТГ (присоединенный к ней для обеспечения инициации трансляции) и терминирующий ТАА кодоны, а также выступающие одноцепочечные последовательности, соответствующие тем, которые образуются на концах ДНК при действии эндонуклеаз рестрикции *EcoRI* и *BamHI*. Таким образом, этот фрагмент может быть вставлен в векторную плазмиду *pBR322*, расщепленную нуклеазами *EcoRI* и *BamHI*, и клонирован в *E. coli*.



Рис. 209. Кодирующая последовательность лейцин-энкефалина с иницирующим и терминирующим кодонами.

При необходимости рекомбинантную плазмиду выделяют из клетки, а требуемый фрагмент вырезают из нее действием двух рестриктаз, как показано на рисунке 210. Если синтезируется очень длинная последовательность, то синтез планируют таким образом, чтобы можно было клонировать составные части последовательности (модули) по отдельности.

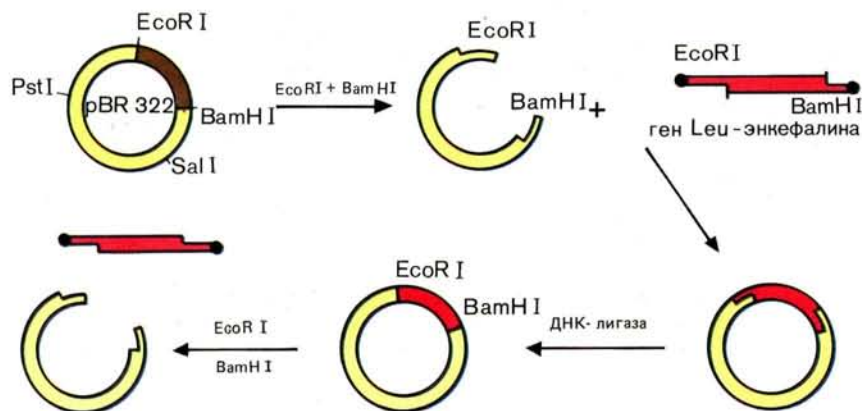


Рис. 210. Клонирование и «вырезание» синтетического фрагмента из вектора pBR322.

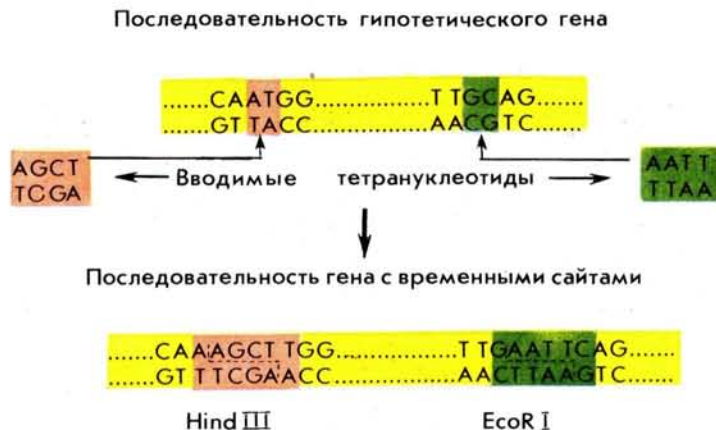
Как показано выше, для успешного клонирования модулей синтетического гена на определенных участках его последовательности должны содержаться подходящие сайты (места), узнаваемые эндонуклеазами рестрикции. Эти сайты должны быть уникальны как для самого гена, так и для вектора, в котором он клонируется. Рестриктные сайты необходимы для введения каждого модуля в плазмиду или фаговую ДНК, его регенерации из векторной молекулы после клонирования и соединения с другими частями целевого гена. Обычно для этого используются сайты, имеющиеся в последовательности синтезируемого гена. Так, последовательность, которую планируется синтезировать (рис. 211), содержит участок расщепления BamHI во внутренней области и участки расщепления рестриктазами EcoRI и SalI на концах. Части I и II этой последовательности синтезируются отдельно и клонируются в плазмиде pBR322. После «размножения» клонированных фрагментов в составе рекомбинантных плазмид они могут быть выделены и соединены по липким концам с образованием нужной последовательности.

Однако чаще всего необходимых внутренних сайтов в гене либо не имеется, либо их мало [не более 1—2 на несколько сотен пар оснований (п.о.)]. Это сильно ограничивает возможности химико-ферментативного синтеза двухцепочечных ДНК и создает необходимость в разработке и использовании для получения полинуклеотидов специальных приемов.

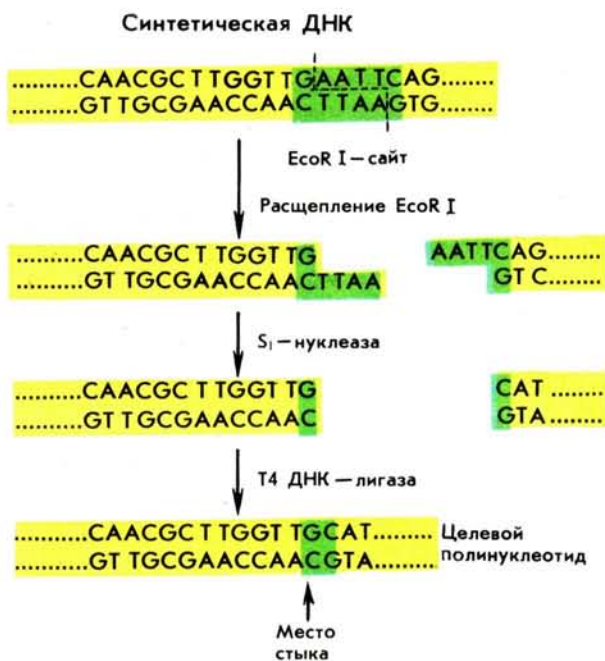
Было предложено несколько таких подходов С. Нарангом (Канада), Р. Ву (США) и В. А. Ефимовым (СССР).

В качестве примера рассмотрим один из них. В его основу были положены модульный принцип конструирования ДНК и использование временных сайтов уникальных эндонуклеаз рестрикции для последовательного клонирования отдельных фрагментов и сборки целевого полинуклеотида. Дополнительные рестриктные сайты вводятся в последовательность гена в произвольно выбранные места,

содержащие какой-либо самокомплементарный динуклеотид АТ, ТА, СG или GС. Для этого последовательность гена как бы раздвигается и между ними вводится тетра- или пентануклеотид, образуя соответствующие рестриктные сайты.



Для удаления созданного дополнительного рестриктного сайта синтезированная ДНК расщепляется соответствующей рестриктазой, образующиеся выступающие одноцепочечные последовательности удаляются действием S_1 -нуклеазы Asp. oгузае и «тупые» концы молекулы соединяются с помощью Т4 ДНК-лигазы. Разбивка последовательности ДНК на модули проводится произвольно.



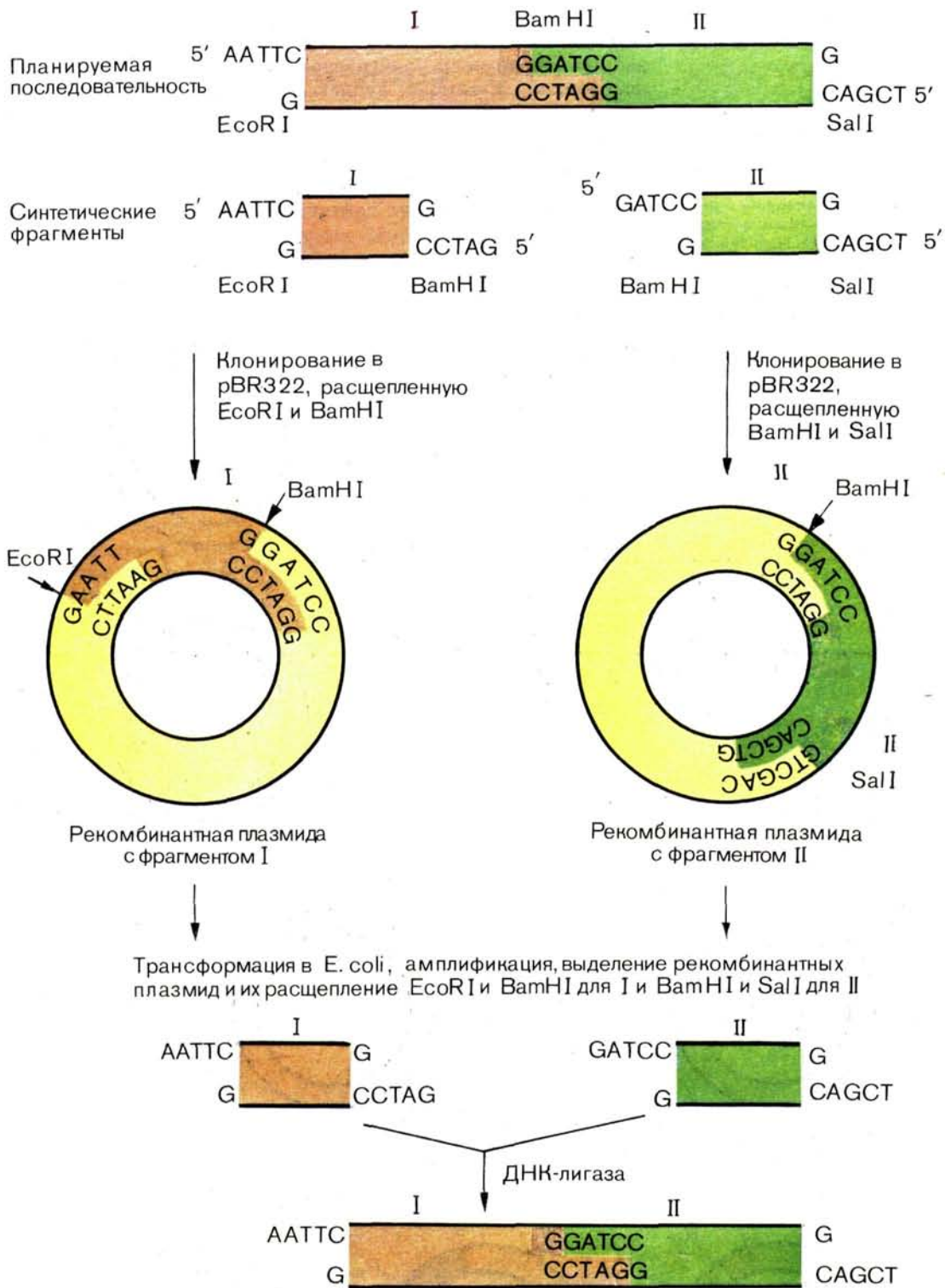
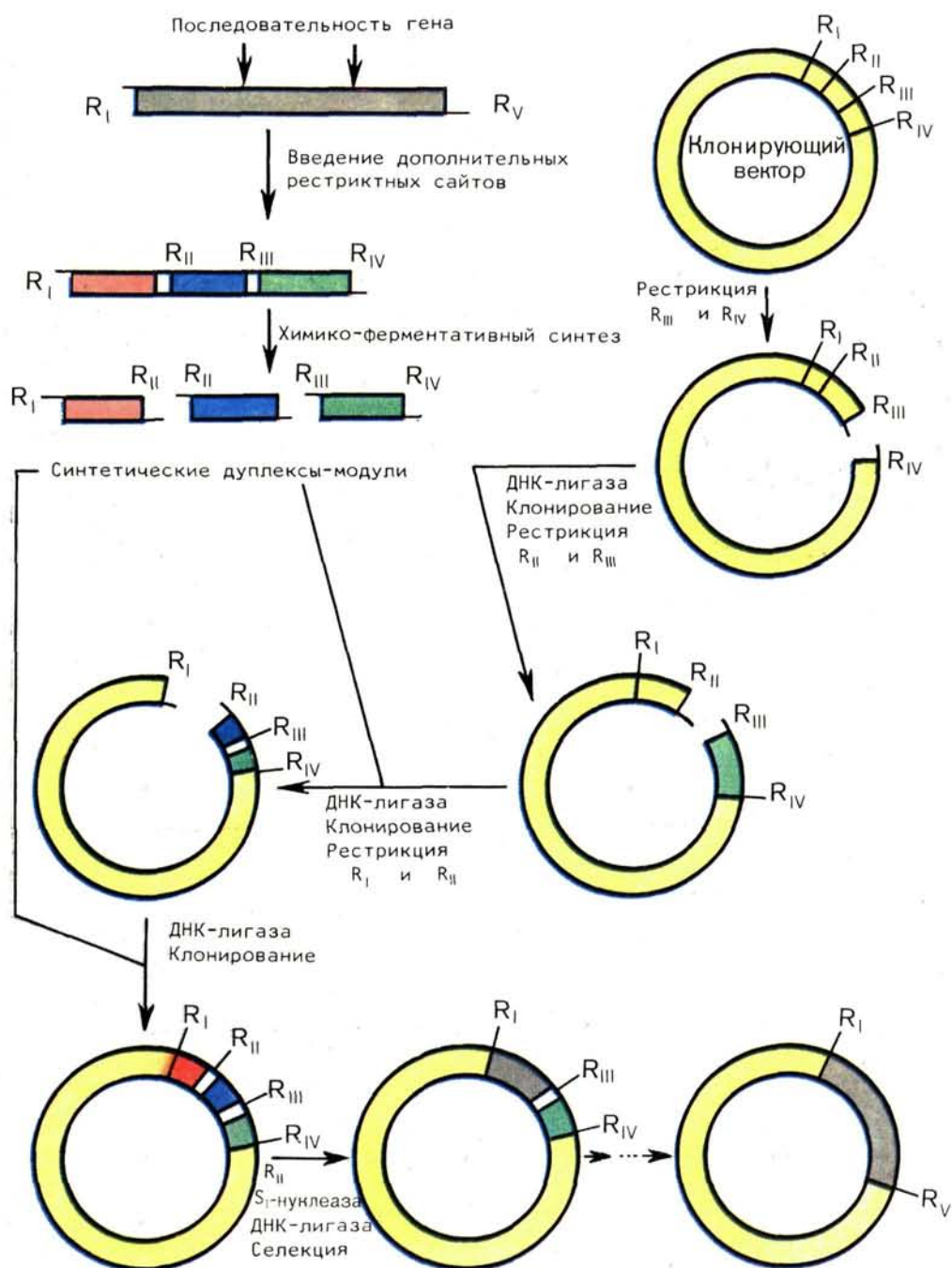


Рис. 211. Создание синтетических генов с промежуточным клонированием.

Рис. 212. Способ сборки искусственного гена с использованием временных рестриктных «сайтов», состоящий в последовательном клонировании модулей синтезируемой ДНК.

Отдельные модули-фрагменты целевой ДНК могут быть получены любым из обсуждавшихся методов: лигированием синтезированных химически олигонуклеотидов, репаративной достройкой 3'-концов дуплекса, образованного двумя частично комплементарными олигомерами, или комбинацией этих методов. Перед соединением в одной плазмиде индивидуальные дуплексы-модули предварительно очищают клонированием. Схема сборки синтетического гена из готовых субфрагментов с использованием временных рестриктных сайтов приведена на рисунке 212.

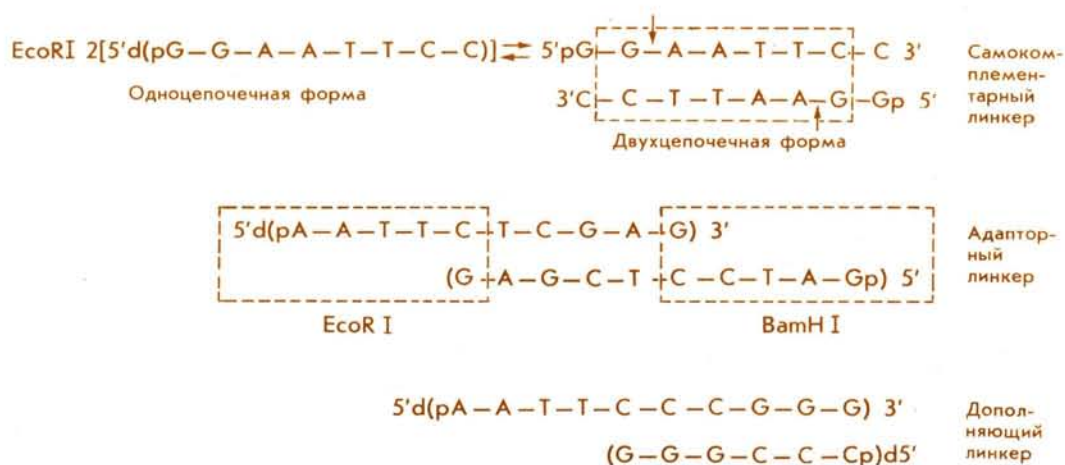


Использование синтетических олиго- и полинуклеотидов в биоорганической химии и биотехнологии

Диапазон применения синтетических олиго- и полинуклеотидов необычайно широк. Прежде всего при изучении белков нередко осуществляется синтез соответствующих генов и их фрагментов (если число аминокислотных остатков не превышает 200—250). Путем экспрессии генов с помощью генноинженерных приемов получают значительные количества труднодоступных белков и пептидов, а также их аналогов.

Однако наиболее широкое применение находят сравнительно короткие олигонуклеотиды. Для того чтобы придать фрагменту нуклеиновой кислоты необходимые для встраивания в определенный участок векторной молекулы липкие концы, синтезируются так называемые линкеры, т. е. двухцепочечные олигонуклеотиды, содержащие последовательность, расщепляемую той или иной рестрикционной эндонуклеазой. Обычно в качестве линкеров применяются самокомплементарные олигонуклеотиды длиной 8—10 нуклеотидных звеньев. На рисунке 213 демонстрируются некоторые типы линкеров.

Рис. 213. Типы линкеров, используемые в генноинженерных исследованиях.



При работе с рекомбинантными молекулами ДНК возникает необходимость замены одного типа липких концов на другой. Для этой цели используются адапторные линкеры — частично двухцепочечные нуклеотиды, содержащие липкие концы, соответствующие двум разным рестриктазам. Третий тип линкеров позволяет перейти от тупого конца ДНК к липкому, и их можно назвать дополняющими.



Фёршт (Fersht) Алан Рой (р. 1943), английский химик-биоорганик. Окончил Кембриджский университет (1968), с 1978 г.— профессор Лондонского университета. Один из основоположников белковой инженерии.

С помощью синтетических фрагментов ДНК можно модифицировать природные гены, вводя в них определенные изменения. Так, синтетические олигонуклеотиды были использованы для превращения гена интерферона человека в форму, пригодную для прямой экспрессии в клетках *E. coli* (рис. 214). Интерферон синтезируется в виде предшественника, и его превращение в истинный белок происходит в результате отщепления сигнального пептида. Клетки *E. coli* не могут осуществлять данный процесс, и поэтому от гена интерферона человека была *in vitro* с помощью рестрикционной эндонуклеазы *Sau3AI* отщеплена часть, кодирующая сигнальный пептид. Однако фермент вместе с частью, кодирующей сигнальный пептид, отщепляет и кодон первой аминокислоты зрелого белка. Для восстановления полного гена к его остатку присоединяли синтетический фрагмент, содержащий кодон первой аминокислоты, иницирующий кодон, а также липкий конец *EcoRI* для введения промотора. Таким образом был восстановлен целый ген, пригодный для прямой экспрессии интерферона в клетках *E. coli*.

Очень важным является использование синтетических олигонуклеотидов в качестве гибридизационных проб (зондов) для поиска нужных рекомбинантных колоний в геноинженерных экспериментах. Олигонуклеотид синтезируется в соответствии с данными, полученными из известной структуры белка или ДНК, и после введения концевой метки ^{32}P используется для гибридизации *in situ* с рекомбинантной ДНК бактериальных клонов, выращенных на нитроцеллюлозных фильтрах и иммобилизованных на них. Для того чтобы поиск нужного гена был эффективным, олигонуклеотид должен иметь последовательность, комплементарную фрагменту искомого гена. Естественно, чем больше длина нуклеотида, тем меньше вероятность встретить соответствующую ему последовательность одновременно в нескольких участках генома. Необхо-

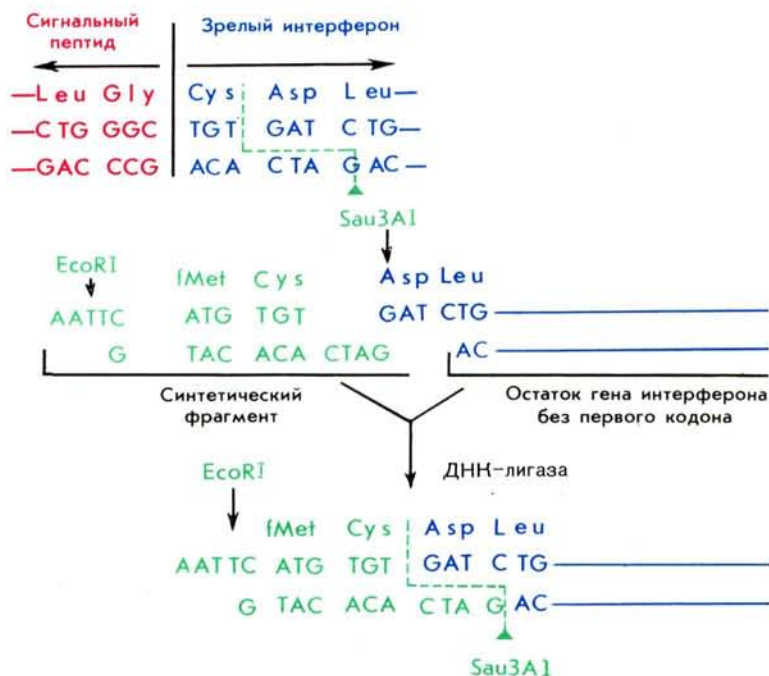


Рис. 214. Превращение гена интерферона человека в форму, пригодную для экспрессии в *E. coli*.

длина олигонуклеотида, в свою очередь, зависит от сложности генома: чем он сложнее, тем она должна быть больше для обеспечения уникального узнавания.

Расчеты показывают, что для специфичного связывания с геном в геноме, например для бактериофага λ , достаточна длина 12 нуклеотидных звеньев, последовательность 15 звеньев идентифицирует уникальный ген в дрожжевом геноме, а 18—20-членные олигонуклеотиды используют для поиска генов в банке генов млекопитающих.

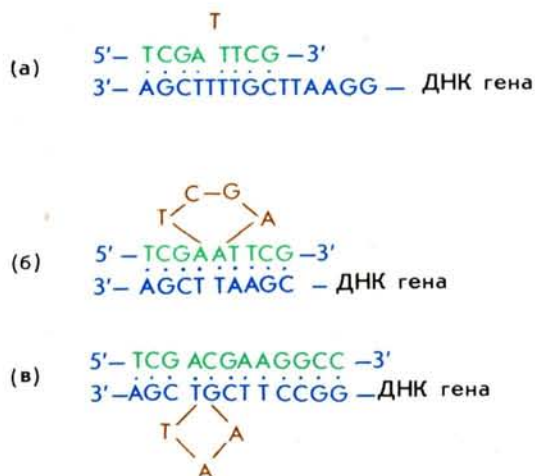


Рис. 215. Структура синтетических олигонуклеотидов, предназначенных для замены одного нуклеотида на другой (а), введение вставки (б) и делеции (в) в ген, клонированный в однонитевом фаге.

Особое место занимает область применения синтетических олигонуклеотидов для введения направленных изменений в структуру гена и затем — в структуру соответствующего белка, называемая белковой инженерией. Чаще всего для этой цели гены клонируют в ДНК нитчатых фагов, хотя возможно применение и других векторов, например плазмид. Для того чтобы ввести мутацию в определенный участок гена с известной структурой, клонированного в фаге (M13 или fd), синтезируется олигонуклеотид, частично комплементарный последовательности, которая содержится в (+)-цепи ДНК однонитевого фага. Отступления от комплементарности в синтезированном олигонуклеотиде определяются задаваемой мутацией. Это может быть простая замена комплементарного нуклеотида на другой, некомплементарный, делеция или вставка нескольких нуклеотидов, как показано на рисунке 216. Частично комплементарный синтетический олигонуклеотид гибридизуется с одной из цепей гена в одноцепочечной фаговой ДНК (рис. 217), и затем с помощью ДНК-полимеразы (не содержащей 5'→3'-экзонуклеазной активности) в присутствии четырех дезоксирибонуклеозидтрифосфатов синтезируется вторая цепь, причем синтетический олигонуклеотид играет роль затравки.

Вновь образованная двухцепочечная ДНК ковалентно замыкается с помощью ДНК-лигазы. Эта ДНК представляет собой гетеродуплекс, в котором комплементарность нарушена в месте введения мутации. Одна цепь гетеродуплекса представляет собой диккий тип, а другая — мутантный. Ковалентно замкнутым гетеродуплексом трансформируются клетки *E. coli*. В результате репликации эти

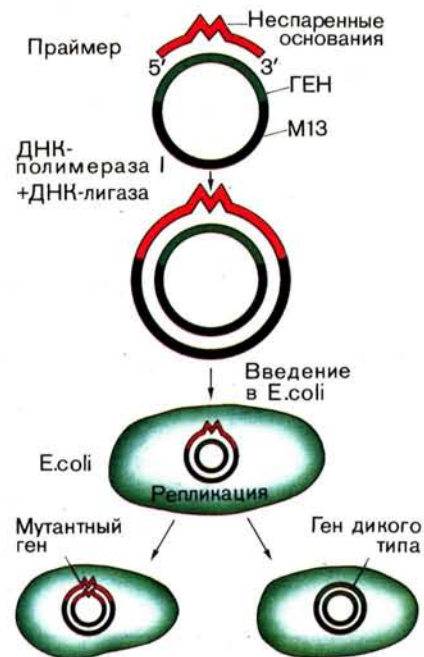
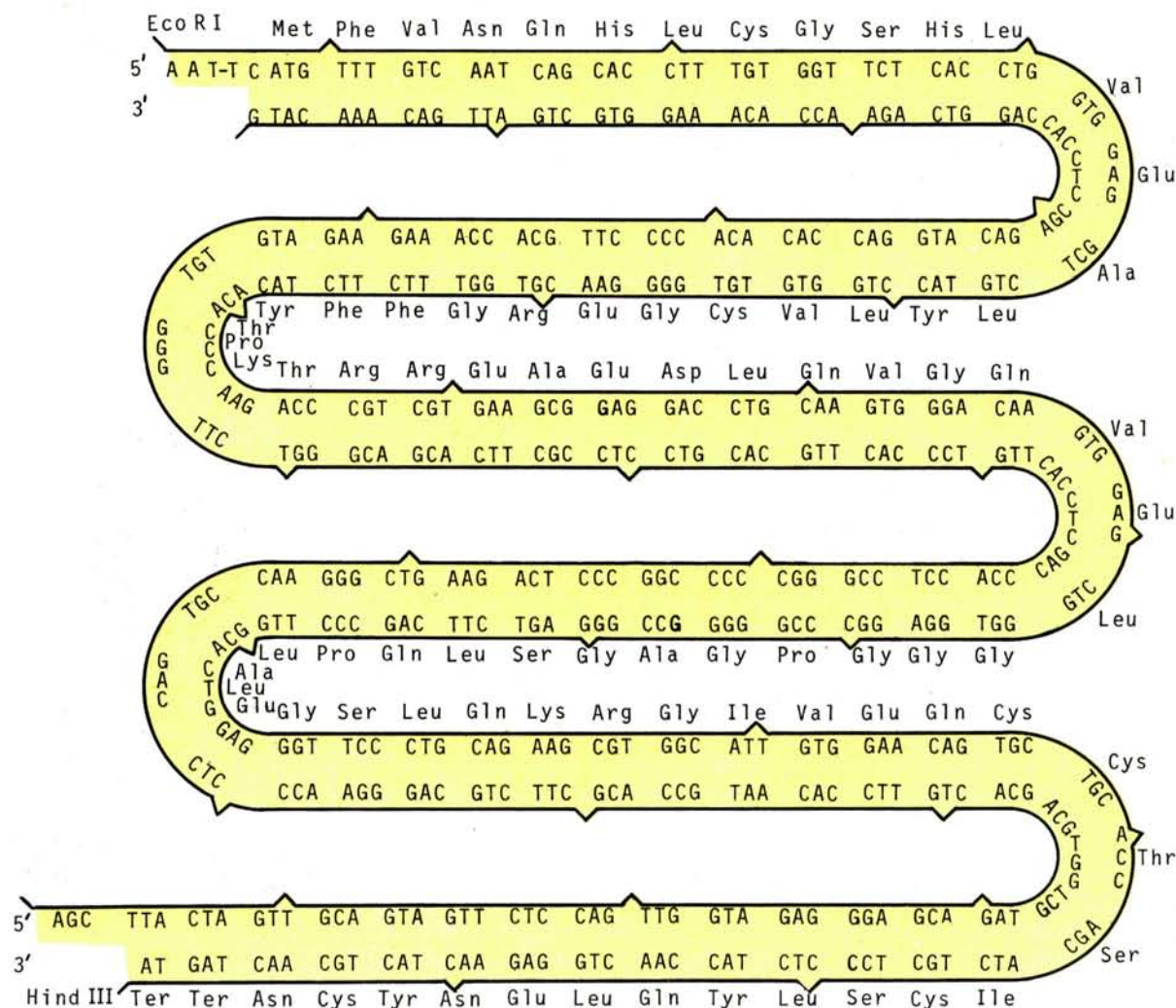


Рис. 216. Принцип олигонуклеотид-направленного мутагенеза.

клетки содержат как мутантные, так и дикие фаги. Смесь фагов выделяется и высевается на чашки Петри таким образом, чтобы каждый фаг дал отдельную колонию. Для отличия диких колоний от мутантных используются разнообразные методы — биологические, если мутация меняет биологическую функцию, рестриктный анализ при нарушении участков расщепления рестриктазами или гибридизация с олигонуклеотидами, использовавшимися для введения мутаций. Последний метод основан на том, что неполностью комплементарные олигонуклеотиды образуют с ДНК значительно менее стабильные комплексы, чем комплементарные олигонуклеотиды и гибридизующиеся при более высоких температурах. Поэтому гибридизация с ДНК фагов, фиксированных на нитроцеллюлозных фильтрах, при низких температурах приводит к образованию комплексов как с мутантными, так и с дикими ДНК, а при высоких — только с мутантными, которым синтетический олигонуклеотид полностью комплементарен.

Направленный мутагенез и синтез генов сделали реальным создание белков, обладающих измененной биологической активностью, например получение ферментов с улучшенными свойствами.

Рис. 217. Синтетический ген проинсулина человека.



ми. Если известна структура активного центра фермента и его комплекса с субстратом, то с помощью направленного мутагенеза и генноинженерных приемов можно заменить определенные аминокислоты активного центра на другие, изменив тем самым каталитические свойства фермента. Так, А. Фёршт, Г. Винтер и сотр. (Великобритания) путем замены треонина на аланин в активном центре тирозил-тРНК-синтетазы улучшили константу связывания фермента с АТР в 100 раз.

Такие пути улучшения старых и создания новых белков открывают большие возможности для медицины и биотехнологии.

В последние годы в связи с интенсивным развитием физико-химической биологии, генной инженерии и биотехнологии все возрастающее практическое значение приобретает химико-ферментативный синтез фрагментов нуклеиновых кислот, в том числе искусственных генов биологически активных пептидов и белков.

В частности, в СССР получены бактериальные штаммы, способные продуцировать проинсулин человека — биосинтетический предшественник инсулина. Проинсулин состоит из одной полипептидной цепи длиной 86 аминокислотных остатков и может быть превращен

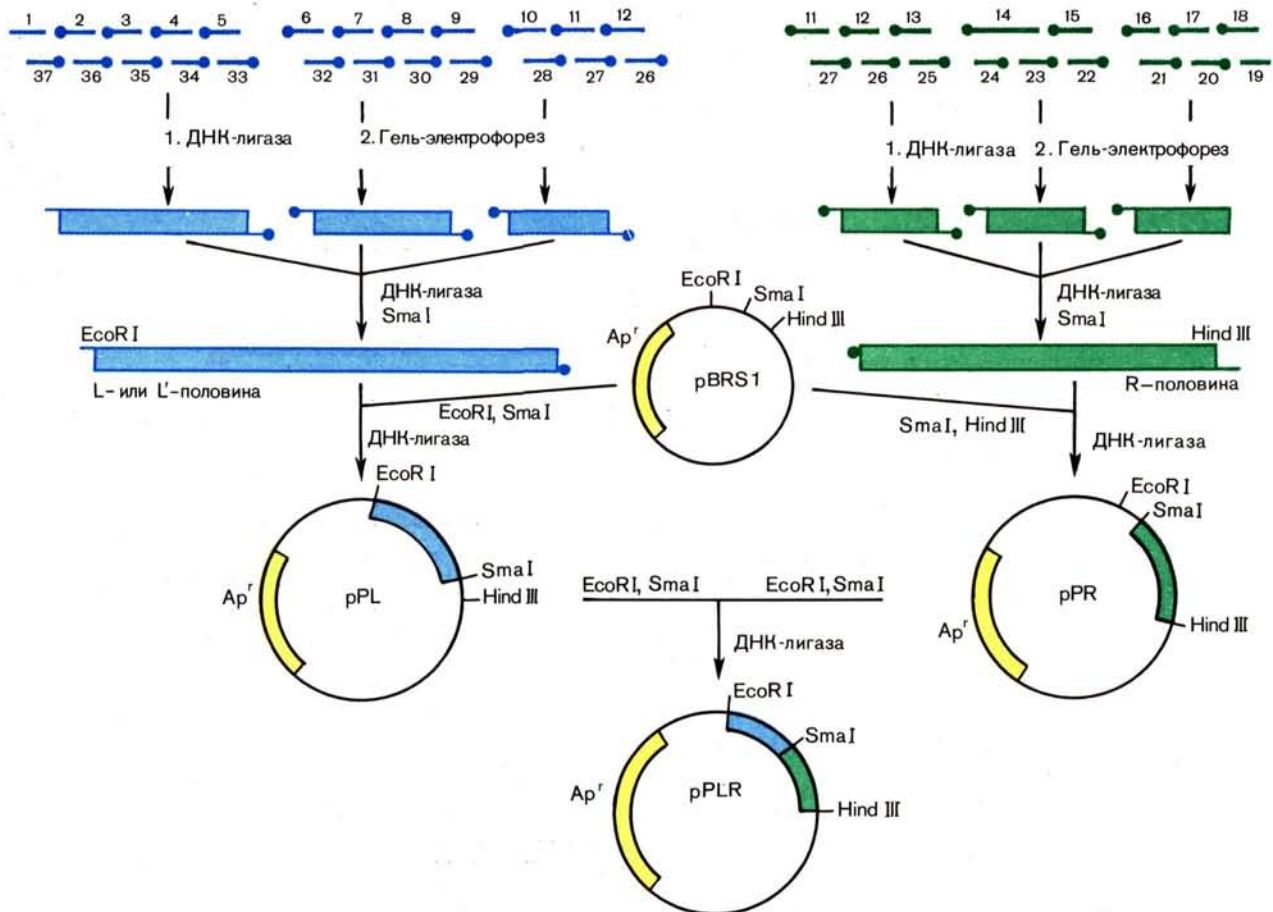


Рис. 218. Схема сборки синтетического гена проинсулина человека из отдельных олигонуклеотидов и его клонирование в плазмидном векторе.



Колосов Михаил Николаевич (1927—1985), советский химик-органик, академик АН СССР (1974). Окончил Московский институт тонкой химической технологии им. М. В. Ломоносова (1948), с 1959 г. работал в Институте биоорганической химии им. М. М. Шемякина АН СССР. Основные работы — в области химии нуклеиновых кислот, антибиотиков и других природных соединений. Синтезировал тетрациклин и изучил механизм его действия. Осуществил синтезы структурных генов валиновой тРНК, α -интерферона человека и др. Лауреат Ленинской премии (1985).

в инсулин путем направленного протеолиза. В основу бактериальных штаммов — продуцентов проинсулина человека положены рекомбинантные плазмиды, несущие синтетический ген проинсулина. Эта последовательность фрагмента ДНК, длиной около 300 нуклеотидных пар, получена химико-ферментативным синтезом (рис. 217, 218). С использованием синтетического гена были получены бактериальные штаммы — продуценты проинсулина человека в составе химерных белков, в частности гибрида с β -галактозидазой *E.coli* (рис. 219) и гибрида с α -амилазой *V.subtilis*. Во всех случаях достигнут высокий уровень экспрессии проинсулина, который превращался в инсулин путем ферментативного гидролиза.

В 1982 г. группами М. Н. Колосова и Л. С. Сандахчиева был завершен полный химико-ферментативный синтез структурного лейкоцитарного интерферона человека α_2 . Искусственный ген интерферона (рис. 220) кодирует полипептид длиной 165 аминокислотных остатков. В процессе синтеза кодирующие триплеты были

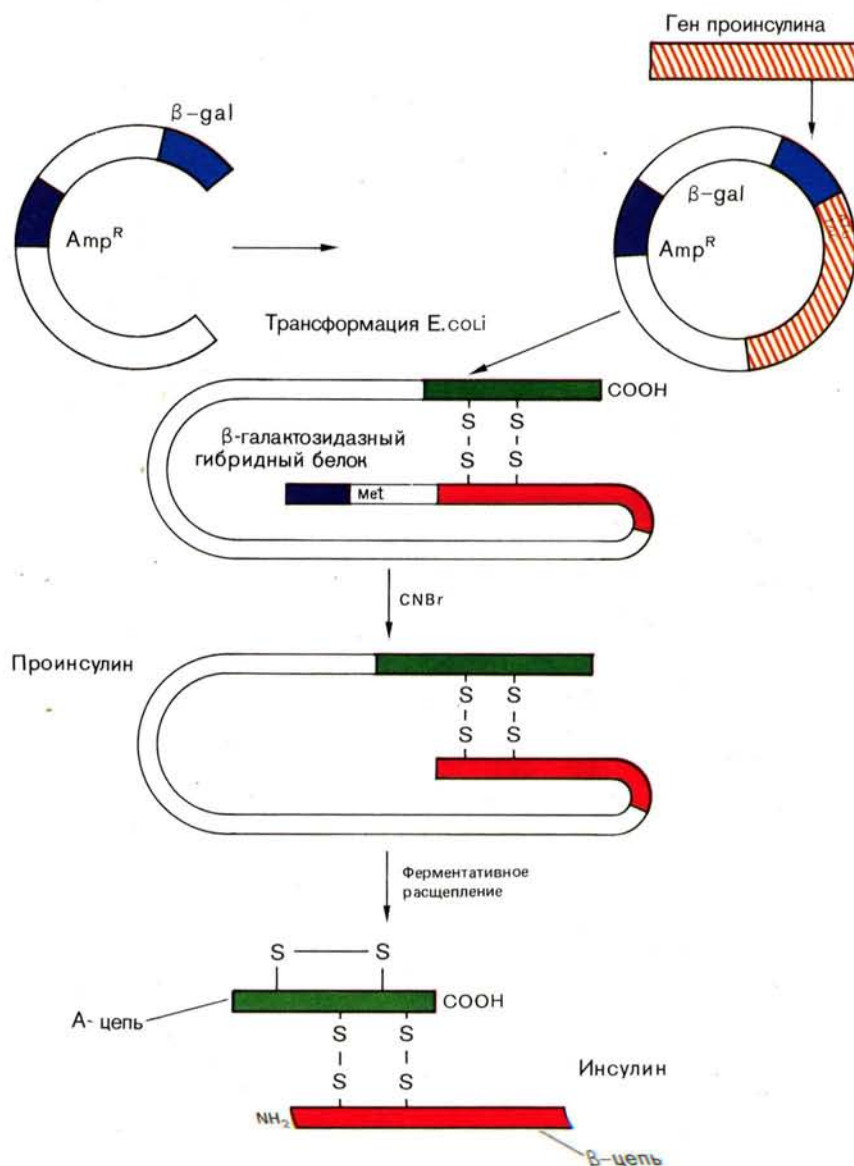


Рис. 219. Экспрессия гена проинсулина человека в составе гибридного белка с β -галактозидазой.

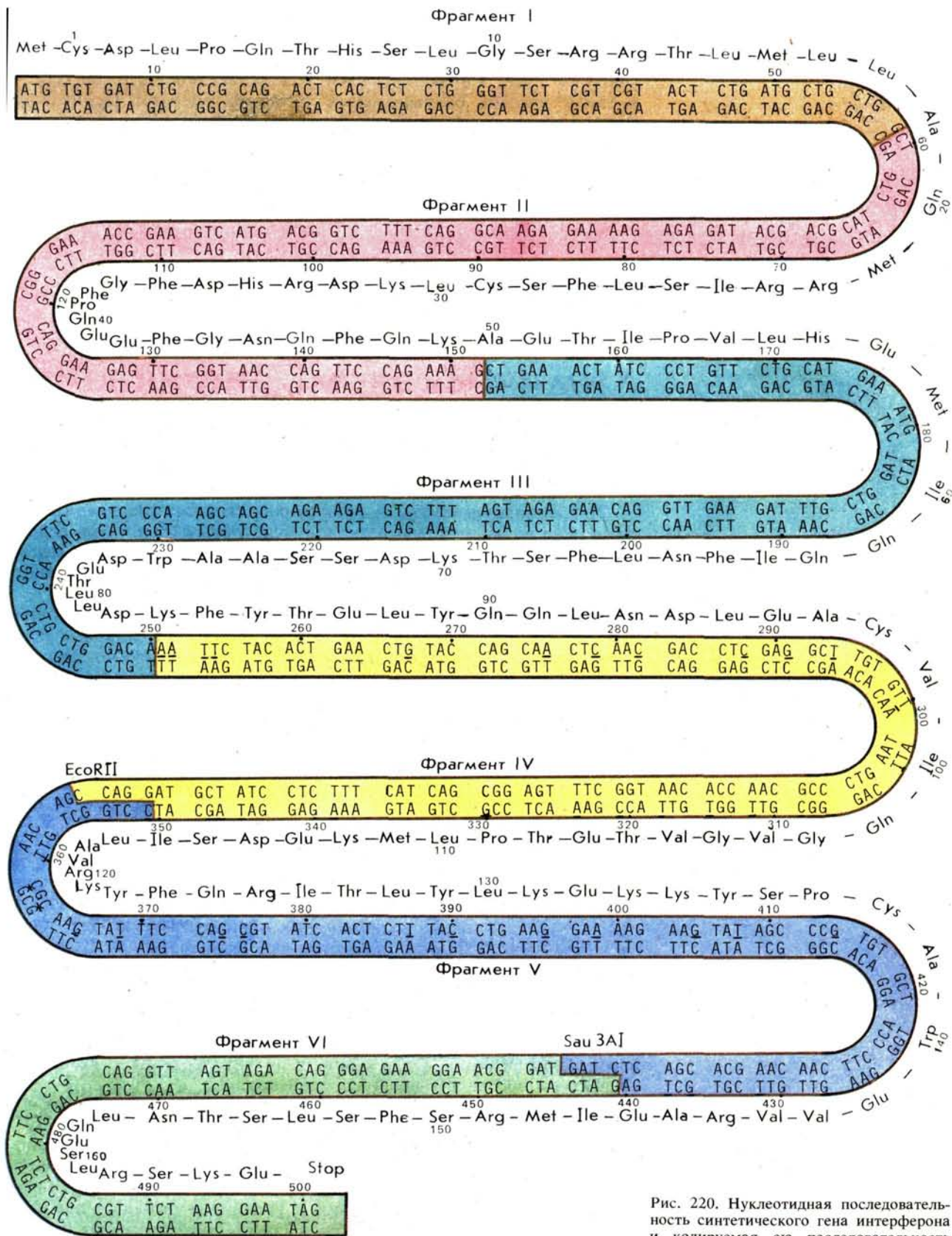


Рис. 220. Нуклеотидная последовательность синтетического гена интерферона и кодируемая ею последовательность аминокислот

выбраны в соответствии с частотой встречаемости в *E.coli* для достижения высокого уровня экспрессии гена. Далее структурный ген интерферона человека α_2 был присоединен к синтетическим участкам ДНК, контролирующим экспрессию генов в *E.coli* (среднестатистический бактериальный промотор, оператор *lac*-оперона *E.coli* и участок связывания рибосомы). Все это вместе составляет фрагмент ДНК величиной свыше 600 п.о., что является к настоящему времени одной из самых длинных синтетических последовательностей ДНК.

Химическая модификация нуклеиновых кислот

Нуклеиновые кислоты — полифункциональные соединения, содержащие большое число реакционноспособных групп. При действии различных химических реагентов в них могут модифицироваться все компоненты: гетероциклические основания, углеводные остатки и остатки фосфорной кислоты.

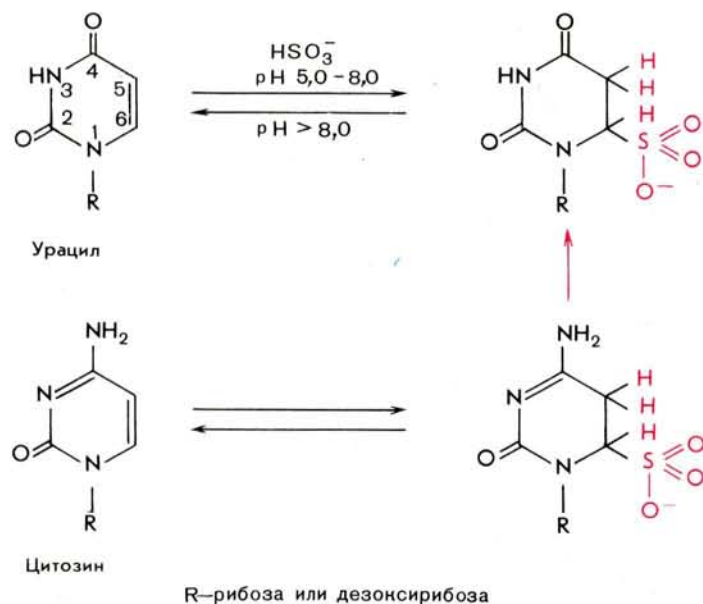
Реакционная способность мономерных структур в составе полимера в значительной степени изменяется под влиянием пространственных факторов. Наибольшее значение имеют методы селективной модификации, широко применяемые при исследовании первичной и пространственной структуры нуклеиновых кислот, введении искусственных мутаций и изучении связи между строением и биологической функцией.

Модификация гетероциклических оснований

Гетероциклические основания определяют специфические свойства нуклеиновых кислот. Их реакционноспособность по отношению к различным агентам зависит от природы основания, условий проведения реакции, а для олиго- и полинуклеотидов — вторичной структуры соответствующего соединения.

Реакции пиримидиновых оснований с бисульфитом натрия. Бисульфит-ион способен обратимо присоединяться к С(5)—С(6)-двойным связям цитозина, урацила и тимина в мягких условиях, образуя соответствующий аддукт, неустойчивый в случае тимина; практическое значение имеют только реакции с урацилом и цитозином. Продукты присоединения к урацилу и цитозину довольно стабильны в нейтральной и кислой средах, но отщепляют бисульфит-ион в щелочной среде.

Важным свойством 5,6-дигидропроизводных цитозина указанного типа является повышенная реакционность аминогруппы, которая легко замещается под действием различных нуклеофильных агентов. Замещение аминогруппы на гидроксигруппу приводит к производному урацила, которое легко превращается в урацил при слабощелочных значениях pH. Таким образом, эта реакция дает возможность специфического преобразования цитозиновых колец в урацильные.

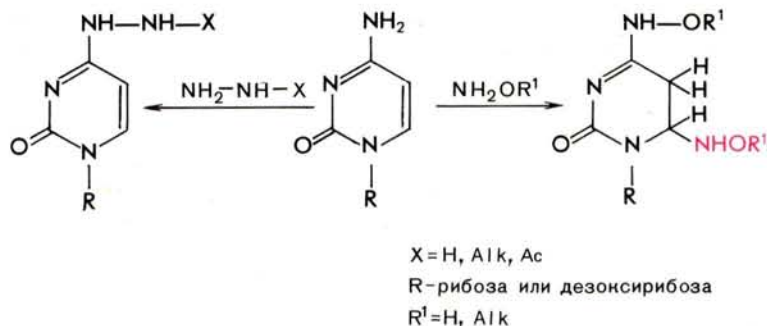


Реакция бисульфита с одноцепочечными нуклеиновыми кислотами протекает значительно медленнее, чем с мономерами, и практически не идет с двухцепочечными молекулами.

Модификация нуклеиновых кислот бисульфитом часто используется как способ введения мутационных замен. С этой целью участок ДНК, выбранный для введения мутаций, превращают в одноцепочечный и затем обрабатывают бисульфитом в условиях неполного дезаминирования. Таким образом, например, пары G·C могут быть заменены на пары A·T (такие замены называются транзициями).

Специфичность к вторичной структуре используется также для анализа пространственного строения полинуклеотидов. Так, обработка тРНК бисульфитом приводит к модификации только цитозиновых оснований, находящихся в петлях.

Реакции с гидразином и гидроксилламин. Гидразин, его алкильные и ацильные производные, а также гидроксилламин и его O-алкильные производные широко применяются в исследованиях нуклеиновых кислот. В водных растворах при кислых значениях pH эти реагенты специфически взаимодействуют с производными цитозина.



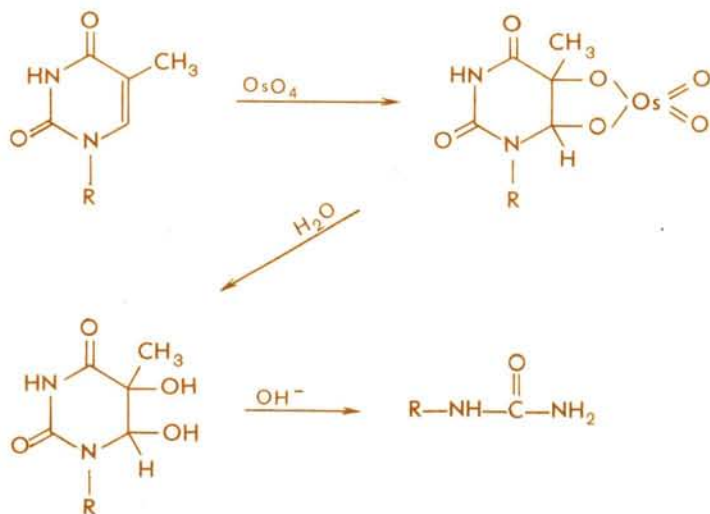
В случае гидразина и его производных наблюдается лишь замещение аминогруппы, тогда как производные гидросиламина одновременно присоединяются по двойной связи, образуя аддукт.

В щелочных условиях реакции с гидразином и гидросиламином приводят к расщеплению гетероциклических колец пиримидинов (см. с. 325), практически не затрагивая пуриновых оснований. Реакция ДНК с гидразином, сопровождающаяся расщеплением по пиримидиновым звеньям, используется при определении нуклеотидной последовательности ДНК в методе Максама — Гилберта.

При анализе первичной структуры нуклеиновых кислот для специфического расщепления по урацильным или цитозиновым звеньям может быть использована также и реакция с гидросиламином в щелочной среде.

Окисление. Реакция с OsO_4 , широко применяемая в органической химии для гидроксирования двойных связей, гладко протекает и в случае пиримидиновых оснований. Скорость реакции возрастает в ряду $\text{T} > \text{U} > \text{C}$, так, тимидин реагирует почти на два порядка быстрее цитидина и в 10 раз быстрее уридина. Поэтому ее можно рассматривать как специфический метод модификации тимидина.

При взаимодействии тимидина с OsO_4 образуется циклический эфир осмиевой кислоты, который легко гидролизует до диола, а



R — дезоксирибоза

последний под действием щелочи превращается в замещенную мочевины. Если модификации подвергаются денатурированные полидезоксинуклеотиды, то обработка после реакции щелочью создает возможность их специфического расщепления по тимидиновым звеньям.

Диольные соединения образуются при взаимодействии пиримидиновых оснований и с такими окислителями, как KMnO_4 , H_2O_2 . Однако в этом случае быстро происходит разрушение цикла.

Реакция с OsO_4 может использоваться для установления положения тимидиновых звеньев при определении первичной структуры ДНК методом Максама — Гилберта.

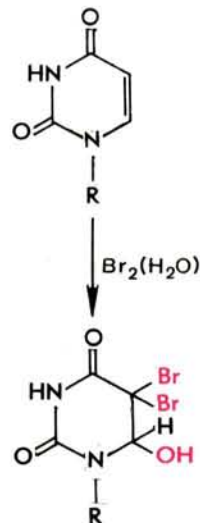
Галогенирование. Одним из распространенных методов модификации оснований в нуклеиновых кислотах является галогенирование. Реакция протекает различно в водных и безводных средах.

Наиболее активные галогенирующие агенты — хлор и бром. При бромировании (или хлорировании) производных нуклеиновых кислот в водных нещелочных средах происходит присоединение галогена по С(5)—С(6) двойной связи колец пиримидина и разрушение гуанина. Аденозин в этих условиях с галогенами не реагирует.

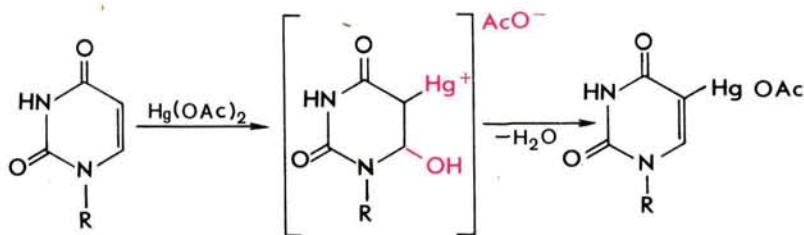
С практической точки зрения важна реакция иодирования нуклеиновых кислот в водных растворах. При действии I_2 или I^- -ионов в присутствии треххлористого таллия (TlCl_3) на одноцепочечные ДНК преимущественно иодируется цитозин, образуя 5-иодопроизводные. При реакции с РНК наряду с 5-иодцитозином образуется 5-иодо-6-гидрокси-5,6-дигидроурацил. При использовании радиоактивного ^{125}I удается получить меченые полинуклеотиды с высокой удельной активностью, которые применяются в опытах по гибридизации.

В безводных растворителях под действием Cl_2 , Br_2 или соответствующих N-галогенимидов происходит электрофильное замещение водорода у атомов С(5) пиримидиновых оснований и С(8) гуанина с образованием соответствующих галогенпроизводных. Эти соединения используются в качестве субстратов РНК- и ДНК-полимераз для синтеза специфически меченных полинуклеотидов.

Меркурирование. Широко распространенная в органической химии реакция непредельных соединений с ацетатом ртути в случае нуклеиновых кислот проходит в мягких условиях (вода, рН 6,0—7,0,



R—рибоза или дезоксирибоза



R—полинуклеотидная цепь

40—50 °С) и специфична по отношению к пиримидиновым основаниям. Результатом реакции является образование производных 5-ацетоксимеркуриоурацила и 5-ацетоксимеркуриоцитозина.

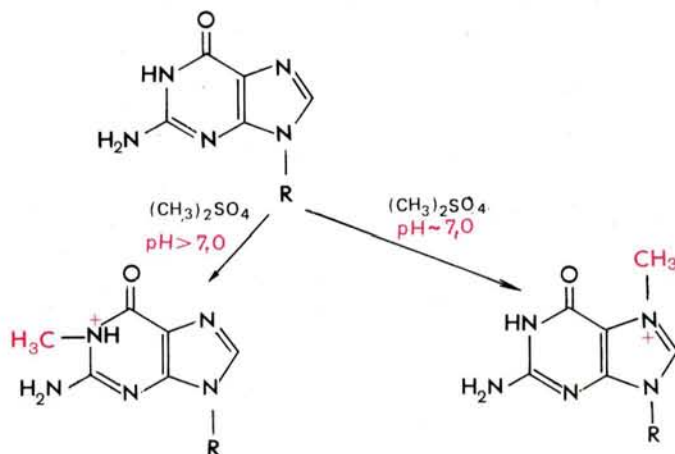
5-Меркурипроизводные вследствие легкого обмена атома ртути на атомы галогена или трития используются для получения специфически меченных пиримидиновых оснований нуклеиновых кислот.

Алкилирование. Реакции алкилирования имеют большое значение для изучения нуклеиновых кислот. С одной стороны, многие алкилирующие реагенты являются мутагенами и канцерогенами, а с другой стороны, реакции алкилирования, осуществляемые в мягких условиях, широко используются для исследования структуры и функции полинуклеотидов.

Наиболее хорошо изучено взаимодействие нуклеиновых кислот и их компонентов с диметилсульфатом $(\text{CH}_3)_2\text{SO}_4$. Эта реакция протекает в мягких условиях в водных растворах.

Направление реакции алкилирования определяется природой гетероциклического основания и условиями эксперимента. Гуаниновое кольцо в нейтральной среде алкилируется преимущественно по атому N(7), а в щелочной — по атому N(1). Исчерпывающее метилирование приводит к 1,7-диметилгуанину.

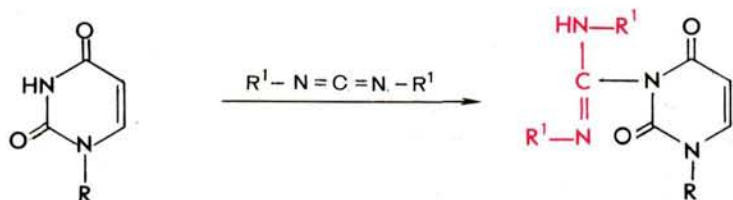
В аденине и его производных все три атома азота могут подвергаться метилированию, однако скорость реакции убывает в ряду $\text{N}(1) > \text{N}(7) > \text{N}(3)$. Единственным продуктом метилирования цитозина является 3-метилцитозин.



В ДНК метилированию подвергаются атомы N(3) аденина, выходящие в малую бороздку В-формы ДНК, и в большей степени атомы N(7) гуанина, выходящие в большую бороздку. Основания двухцепочечных ДНК, метилированные по атомам N(3) и N(7), продолжают образовывать комплементарные пары, но эффективность их взаимодействия понижается.

7-Метилгуанин и 3-метиладенин легко отщепляются от полидезоксирибонуклеотидной цепи в результате разрыва N-гликозидных связей. Далее полинуклеотидная цепь может быть легко расщеплена по модифицированному звену действием щелочи или аминов (см. с. 322). Такой метод деградации используется для определения последовательности ДНК (локализация гуаниновых и адениновых звеньев по методу Максама — Гилберта).

Реакции с карбодиимидом. Для исследования вторичной структуры нуклеиновых кислот применяется реакция оснований с производными карбодиимида, в результате которой образуются соответствующие аддукты.



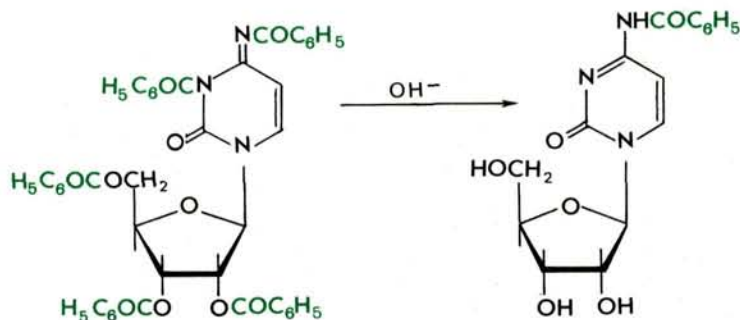
R – полинуклеотидная цепь

R¹-N=C=N-R¹ – водорастворимый карбодиимид

С наибольшей скоростью модифицируются редкие компоненты нуклеиновых кислот — инозин и псевдоуридин. Реакция с карбодиимидом практически не идет с двухцепочечными полинуклеотидами. В нуклеиновых кислотах, содержащих как двухцепочечные, так и одноцепочечные участки, модифицируются только последние, что используется при исследованиях вторичной структуры тРНК.

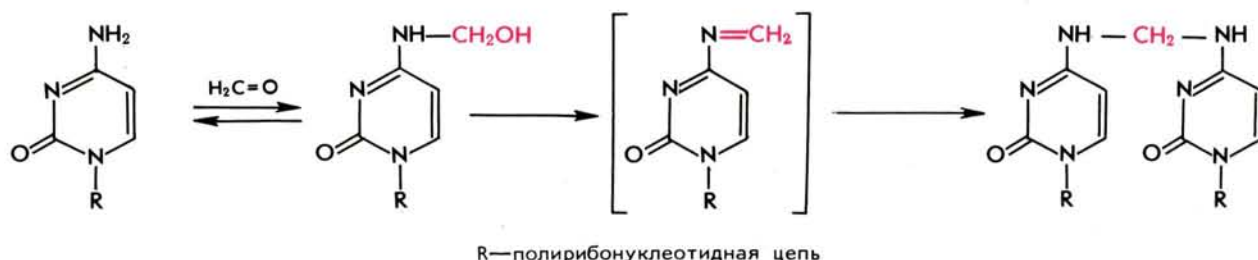
Ацилирование. В реакцию ацилирования под действием ангидридов и галогенангидридов органических кислот вступают как амино- и иминогруппы, так и гидроксильные группы углеводных и фосфатных остатков.

Например, при действии на цитидин бензоилхлорида в безводном пиридине при 100 °С образуется пентаацилированное производное. В мягких щелочных условиях избирательно снимаются бензоильные группы, связанные с гидроксильными группами углевода и атомом N(3), но сохраняется бензоильная группа у экзоциклического атома азота

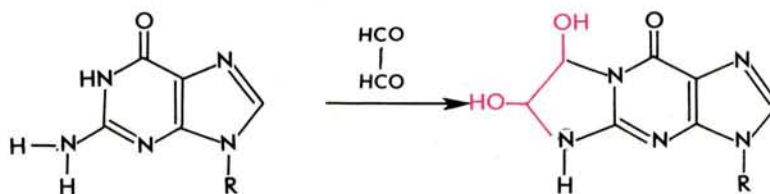


Реакции с альдегидами. Реакция аминогрупп гетероциклических оснований нуклеиновых кислот с формальдегидом приводит к образованию аминотилольных соединений. Следует подчеркнуть,

что реакция обратима: при удалении формальдегида из реакционной смеси и даже при ее разбавлении наблюдается быстрая регенерация исходных продуктов. В щелочной среде метилольное соединение через промежуточно образующееся основание Шиффа может реагировать с экзоциклической аминогруппой другого нуклеозида, давая стабильное бис-(нуклеозидил) метиленовое производное



Широко применяется реакция оснований нуклеиновых кислот с α -дикарбонильными соединениями, такими, как глиоксаль. Аденин и цитозин при этом дают аминометилольные производные, а гуанин образует трициклический продукт за счет конденсации amino- и иминогрупп с обеими карбонильными группами глиоксаля:



Реакция нуклеиновых кислот с глиоксалем может рассматриваться как специфическая по отношению к гуаниновым звеньям; скорость ее взаимодействия с одноцепочечными полинуклеотидами значительно выше, чем с двухцепочечными. Она применялась при изучении первичной структуры РНК как метод ограничения действия гуанилрибонуклеазы. Реакции с альдегидами широко используются для фиксации одноцепочечных участков нуклеиновых кислот при изучении их с помощью электронной микроскопии.

Взаимодействие с азотистой кислотой. Аминосодержащие основания нуклеиновых кислот при обработке HNO_2 претерпевают

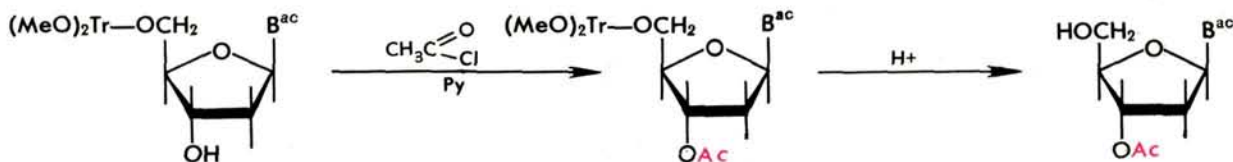
дезаминирование, в результате этого цитозин превращается в урацил, аденин — в гипоксантин, а гуанин — в ксантин. Реакция с HNO_2 применяется для исследования вторичной структуры тРНК. Необходимо также отметить, что азотистая кислота является сильным мутагеном.

Модификация углеводных остатков

Углеводно-фосфатный остов во многом определяет конформацию и физико-химические свойства нуклеиновых кислот. Расщепление нуклеиновых кислот различными ферментами связано со спецификой строения углеводно-фосфатной цепи: в частности, многие ферменты отличают дезоксирибонуклеиновые кислоты от рибонуклеиновых, концевую фосфатную группу от группы, участвующей в образовании фосфодиэфирной связи, 5'-фосфат от 3'-фосфата и т. п.

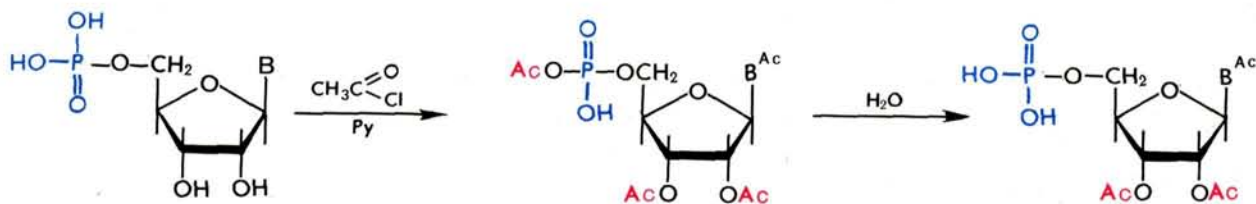
Ацилирование. При действии на нуклеозиды ангидридов или хлорангидридов уксусной или бензойной кислот в безводном пиридине в мягких условиях гладко ацилируются гидроксильные группы остатка углевода. В случае цитидина, аминогруппа которого более реакционноспособна, чем у других оснований, ацилируется также и аминогруппа. Избирательное ацилирование гидроксильных групп сахара достигается проведением реакции в ледяной уксусной кислоте. При действии уксусного ангидрида в нейтральных водных растворах также образуются только O-ацильные производные нуклеозидов.

В реакциях ацилирования первичные гидроксильные группы более реакционноспособны, чем вторичные. Для избирательного ацилирования определенных гидроксильных групп углеводного остатка все другие необходимо защищать. Например, 3'-ацилированные дезоксинуклеозиды получают действием галогенангидридов в пиридине на 5'-O-защищенные нуклеозиды с последующим удалением защитной группы

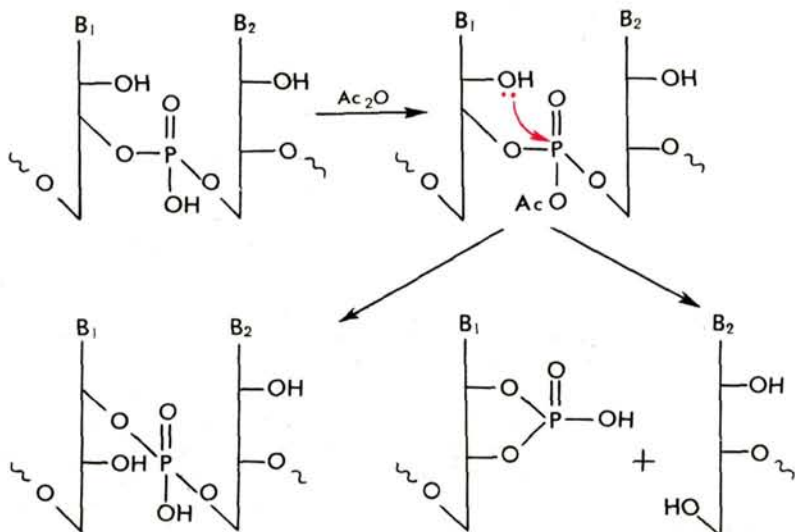


При ацилировании нуклеотидов в безводной среде в реакцию вступает также и фосфатная группа, образуя смешанные ангидриды, которые легко гидролизуются в водных растворах.

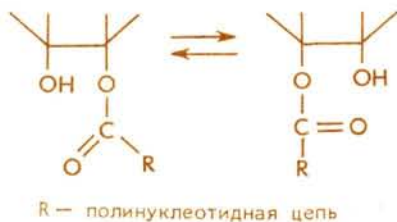
Ацилирование олиго- и полинуклеотидов обычно осуществляется действием уксусного ангидрида в водных растворах при pH 7,0. Олигодезоксинуклеотиды в таких условиях превращаются в 5'-O-ацетаты, в олиго- и полирибонуклеотидах ацетируются также 2'-ОН-группы.



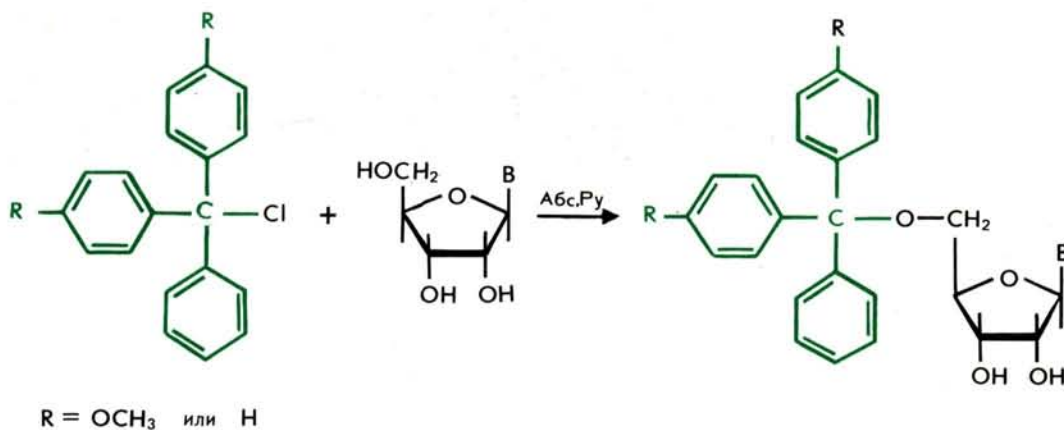
Смешанные ангидриды в случае олигорибонуклеотидов могут образовываться в результате взаимодействия с гидроксильной группой межнуклеотидного атома фосфора, хотя его реакционная способность ниже, чем у концевых монофосфатов. В определенных условиях эта реакция вызывает расщепление олигонуклеотидов и миграцию фосфатной группы от 3'- к 2'-гидроксильной группе рибозы. Обе реакции объясняются атакой 2'-гидроксильных групп рибозы на атом фосфора в смешанном ангидриде, в результате которой разрывается связь фосфата с 3'-, либо с 5'-гидроксильной группой



Рибонуклеотиды и рибонуклеозиды с ацилированными 3'- или 2'-гидроксигруппами легко претерпевают перегруппировки с миграцией ацильной группы на соседний гидроксил, в результате чего из индивидуальных соединений образуются смеси 2'- и 3'-изомеров. Скорость перегруппировки и состав смеси зависят от природы ацильного радикала:

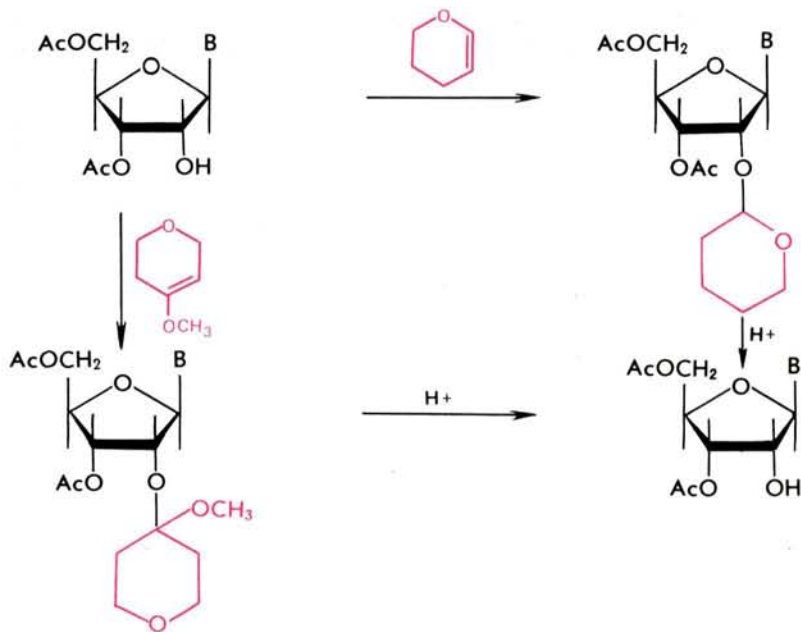


Алкилирование. Из алкилирующих реагентов наиболее широкое применение нашли трифенилхлорметан и его производные, селективно взаимодействующие с первичными гидроксильными группами углеводных остатков:



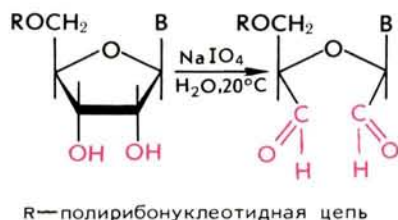
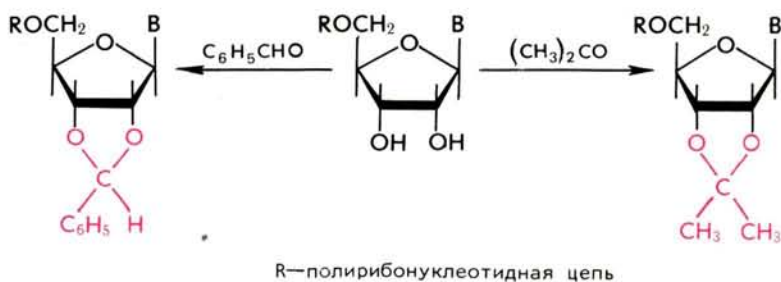
Чаще всего в синтезе олигонуклеотидов для блокирования 5'-гидроксильных групп используются монометокси- и диметокситрифенилметильные производные. Это связано с тем, что образуемые ими эфиры с нуклеозидами гидролизуются в значительно более мягких условиях, чем трифенилметильные. Деблокирование проходит под действием кислот. Введение каждой метоксигруппы ускоряет процесс примерно в 10 раз, что существенно снижает побочные эффекты при синтезе олигонуклеотидов.

Для защиты гидроксильных групп углеводов используется также реакция с виниловыми эфирами, приводящая к образованию ацеталей. Обычно в качестве алкилирующего реагента в этом случае применяют дигидропиран или 4-метокси-3,4-дигидропиран. Реакция проводится в присутствии кислотных катализаторов в органических растворителях



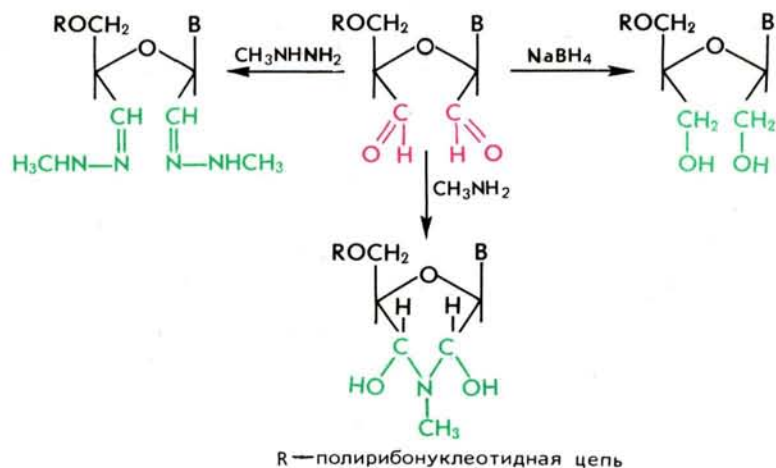
Тетрагидропиранильные защитные группы снимаются в слабкокислой среде.

Реакции с карбонильными соединениями. Рибонуклеозиды, содержащие незамещенные 2'- и 3'-гидроксильные группы, в присутствии кислотных катализаторов легко вступают в реакцию с альдегидами и кетонами, образуя циклические ацетали или кетали. В частности, для защиты α -диольных групп рибонуклеозидов нередко используется взаимодействие с бензальдегидом или ацетоном. В первом случае образуются бензильденные, во втором — изопропилиденные производные рибонуклеозидов. Циклические производные легко расщепляются в слабкокислой среде



Окисление. Широкое применение в химии нуклеиновых кислот находит окисление *цис*-гликольной группы в рибонуклеозидах, олиго- и полирибонуклеотидах. Эта реакция проходит в мягких условиях под действием солей иодной кислоты с образованием диальдегида. Такого рода диальдегиды весьма реакционноспособны.

Они восстанавливаются боргидридом натрия до диолов, вступают в реакцию с гидразинами, давая гидразоны, и реагируют с аминами с образованием циклических производных:



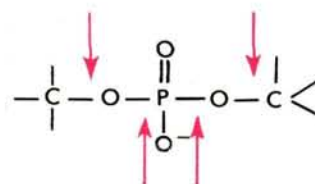
В том случае, если α -диольная группа принадлежит нуклеозид-5'-фосфату, олиго- или полинуклеотиду, окисление с последующей обработкой щелочью или аминами приводит к расщеплению фосфоэфирной связи с одновременным отщеплением основания. Реакция дает возможность избирательно отщеплять 3'-концевые звенья у полирибонуклеотидов и используется при анализе последовательности РНК.

Расщепление N-гликозидных связей. N-Гликозидные связи нуклеозидов и нуклеотидов весьма устойчивы в нейтральной и щелочной средах и расщепляются при этих значениях pH только в жестких условиях. Однако они относительно лабильны в кислой среде, что послужило основой первых методов определения нуклеотидного состава нуклеиновых кислот (гидролиз хлорной кислотой в течение 1 ч при 100 °С). Для нуклеозидов, нуклеотидов и полинуклеотидов соблюдается закономерность: скорость гидролиза соединений дезоксирибы на 2—3 порядка выше скорости гидролиза соединений рибозы, а в каждом ряду N-гликозидные связи пуриновых оснований значительно более лабильны, чем пиримидиновых. На этом различии основан метод селективного расщепления молекул ДНК по пуриновым звеньям (апуринизация), широко применяющийся в структурных исследованиях.

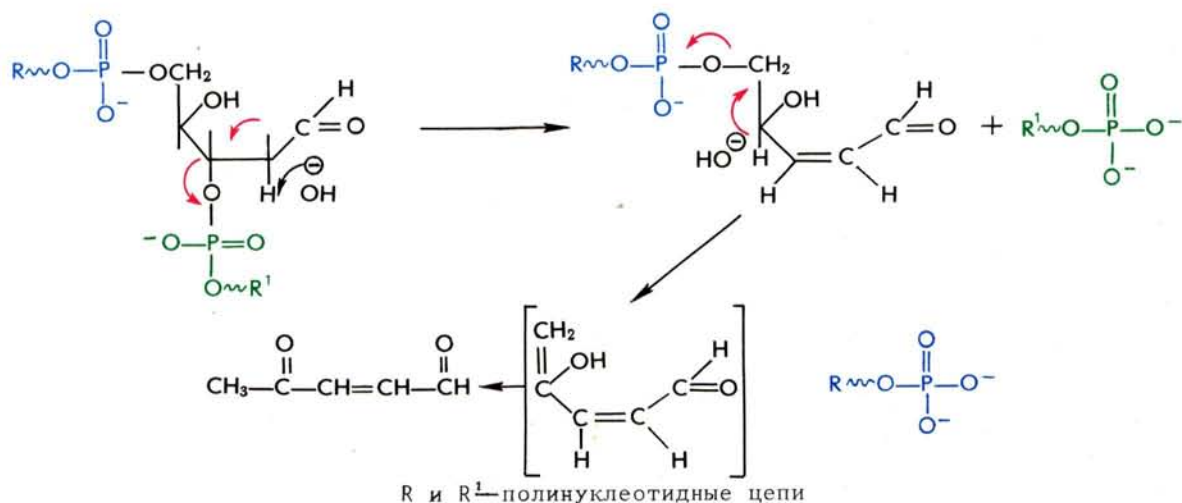
Расщепление фосфоэфирных связей

Расщепление фосфоэфирных связей может проходить с разрывом C—O или P—O связей.

Разрыв по связи C—O (β -элиминирование) становится возможным после удаления соответствующего гетероциклического основания и образования карбонильной группы в β -положении к



фосфодиэфирной группировке. Если β -элиминирование происходит в углеводных остатках, содержащих 3'- и 5'-фосфаты, то на первой стадии образуется неопределенный альдегид с сопряженной системой двойных связей, имеющий в β -положении 5'-фосфатную группу. Эта группа далее отщепляется по механизму β -элиминирования с образованием гипотетического продукта, который, как любой виниловый спирт, претерпевает дальнейшие превращения.



Реакции β -элиминирования широко используются при определении первичной структуры нуклеиновых кислот. На них, в частности, основаны методы специфичных расщеплений ДНК и РНК после деструкции различного типа оснований химическими агентами.

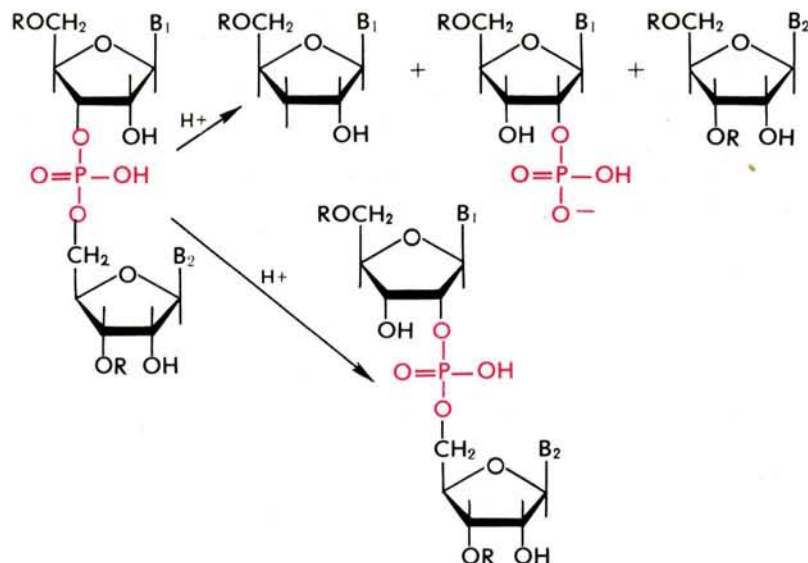
Среди реакций, протекающих с разрывом связей Р—О, практически важны кислый и щелочной гидролиз РНК или олигорибонуклеотидов. При щелочном гидролизе РНК конечным продуктом являются нуклеозид-2'(3')-фосфаты; ДНК в тех же условиях не расщепляется.

Щелочной гидролиз до мононуклеотидов применяется для определения нуклеотидного состава РНК.

Характерной особенностью поведения РНК и олигорибонуклеотидов в кислой среде является изомеризация природной 3' → 5'-фосфодиэфирной связи в 2' → 5'-фосфодиэфирную. Процесс идет конкурентно с расщеплением фосфодиэфирных связей. При неполном гидролизе РНК в образовавшихся олигонуклеотидах присутствует значительный процент олигомеров с изомерными связями. Продолжительный гидролиз в кислой среде, так же как и в щелочной, приводит к образованию мононуклеотидов.

С возможностью изомеризации фосфодиэфирной связи в кислой среде следует считаться во всех случаях, где РНК подвергается кислотной обработке, например в процессе снятия защитных групп при синтезе олигорибонуклеотидов.

Нуклеозид-2',3'-циклофосфаты, в которых фосфодиэфирная связь образована соседними гидроксильными группами рибозы, в



кислой и в щелочной среде гидролизуются быстрее, чем 3' → 5'-фосфодиэфиры, что связано с напряженностью пятичленного цикла. Особенно быстро гидролиз происходит в кислой среде. Поэтому кислотную обработку применяют для раскрытия циклофосфатных групп олигонуклеотидов, образующихся при гидролизе РНК рибонуклеазами.

Основные химические реакции нуклеиновых кислот и их компонентов позволяют получить представление о многообразии химических превращений, в которых участвуют нуклеиновые кислоты *in vivo* и *in vitro*.

Нуклеопротеиды

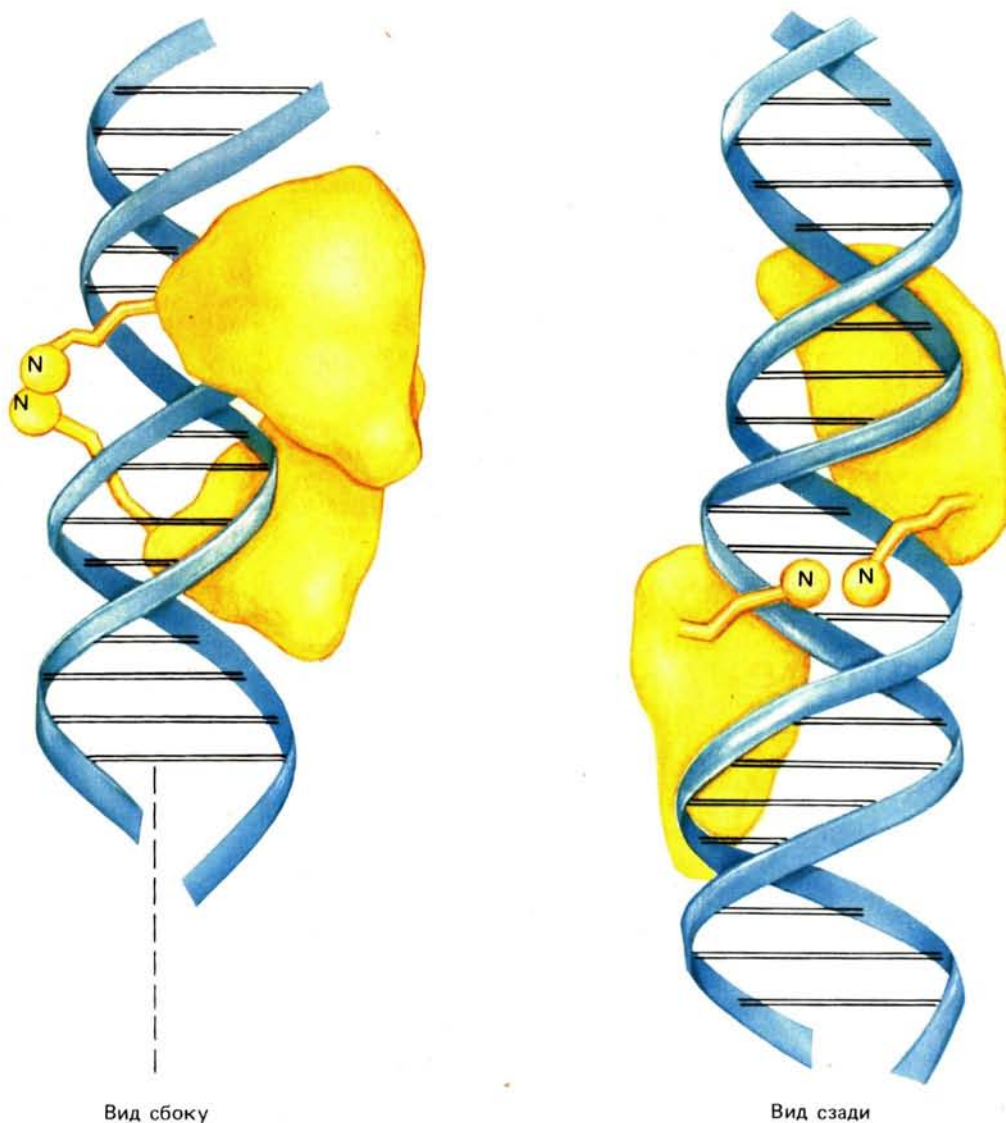
Все функции нуклеиновых кислот в организме осуществляются в комплексах с белками. В то же время лишь некоторые белки выполняют свои функции в комплексе с нуклеиновыми кислотами. Такие комплексы называются нуклеопротеидами. Одни нуклеопротеиды существуют в течение длительного времени, например хроматин, рибосомы, вирусные частицы. Другие возникают на короткое время и, выполнив свою функцию, диссоциируют—к ним относятся комплексы, образуемые ДНК- и РНК-полимеразами, регуляторными белками, репрессорами или активаторами и т. п. Нуклеопротеиды осуществляют такие важные процессы в клетке, как репликация, транскрипция и трансляция, транспорт нуклеиновых кислот из ядра в клетку, секреция белков в эукариотических клетках и т. п.

Структура нуклеопротеидных комплексов

Нуклеопротеиды образуются, как правило, в результате нековалентных взаимодействий белков и нуклеиновых кислот. В связывании принимают участие электростатические и гидрофобные взаимодействия, водородные связи, а также уже упоминавшиеся «стэкинг»-взаимодействия; стабилизирующую роль в комплексах часто играют ионы металлов и другие кофакторы.

Для ряда белков, образующих комплексы с полинуклеотидами, а в некоторых случаях и для самих комплексов по данным рентгеноструктурного анализа построены достоверные трехмерные модели. Однако для сложных многокомпонентных нуклеопротеидов, таких, например, как рибосомы, этот метод пока неприменим. Информацию

Рис. 221. Модель взаимодействия оператора с cI -репрессором бактериофага λ .



об их строении можно получить с помощью электронной микроскопии в сочетании с химическими, иммунохимическими и физическими методами, позволяющими оценить пространственную конфигурацию комплексов и взаимное расположение отдельных компонентов. Широко используются модификация бифункциональными реагентами, моноклональные и поликлональные антитела, спектральные и расчетные методы, нейтронография.

Комплексы репрессоров с операторами. В качестве примера нуклеопротеидов, выполняющих в клетке регуляторную функцию, рассмотрим комплексы белков — репрессоров с ДНК. Репрессоры являются негативными регуляторами синтеза РНК на ДНК: связываясь со специальным участком ДНК, называемым оператором, они запрещают транскрипцию генов, находящихся под их контролем (см. с. 413). В результате этого взаимодействия образуются прочные комплексы, характеризующиеся константами связывания порядка $10^{10}—10^{14} \text{ M}^{-1}$. Как правило, репрессоры представляют собой небольшие белки, состоящие из нескольких идентичных субъединиц. Так, репрессоры умеренных бактериофагов (сI-репрессоры) состоят из двух субъединиц, репрессор лактозного оперона *E. coli* — из четырех. Самым маленьким из известных репрессоров является так называемый *cro*-белок бактериофага λ — каждая из субъединиц димера содержит 66 аминокислотных остатков (субъединица сI-репрессора фага λ состоит из 236, а lac-репрессора *E. coli* — из 347 аминокислотных остатков).

В настоящее время установлены первичные структуры многих репрессоров и операторов. Характерной особенностью нуклеотидных последовательностей операторов является их симметричный характер.

Симметричная структура операторов и субъединичное строение репрессоров позволили предположить, что с каждым из симметричных участков оператора взаимодействует одна из субъединиц соответствующего репрессора. Исследование доступности для различных химических реагентов участков оператора в комплексе с репрессором и в свободном состоянии показало, что основное взаимодействие происходит только с одной стороны двойной спирали ДНК и осуществляется через группы оснований, выходящих в большую бороздку. Эти представления в совокупности с данными рентгеноструктурного анализа ДНК-связывающего N-концевого домена сI-репрессора фага λ и *cro*-белка, которые удалось получить в кристаллическом состоянии, легли в основу построения моделей комплексов репрессоров с операторами.

Полученная для сI-репрессора модель представлена на рисунках 221 и 222. Согласно компьютерному анализу рентгеноструктурных данных, существует единственное удовлетворительное сочетание репрессора с оператором. Две субъединицы белка симметрично расположены с одной стороны двойной спирали ДНК, и лишь два контакта наблюдаются с противоположной стороны. N-Концевые гибкие участки репрессора как бы «обнимают» спираль ДНК, а две α -спирали (2 и 3) оказываются в большой бороздке и взаимодействуют с ДНК своими боковыми группами (рис. 222).

Аналогичное расположение двух α -спиралей в большой бороздке ДНК реализуется, по-видимому, и в случае *cro*-белка.

Анализ первичных структур многих репрессоров показывает, что во всех известных случаях определенные аминокислотные остатки консервативны и соответствующие участки полипептидных цепей могут образовывать две α -спирали, подобно спиралям 2 и 3 сI-репрессора. Возможно, это общий способ специфического взаимодействия регуляторных белков с ДНК. Окончательные выводы о структуре комплексов могут быть сделаны только после рентгеноструктурного анализа.

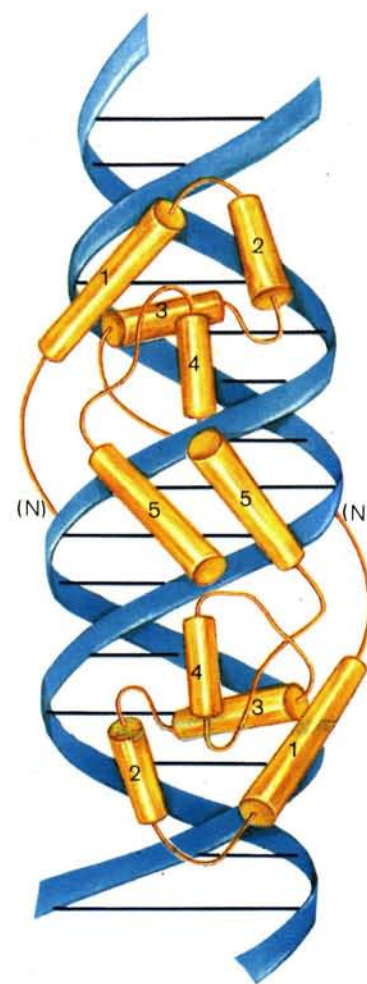


Рис. 222. Предполагаемая структура комплекса оператора с сI-репрессором. α -Спиральные участки репрессора изображены цилиндрами.



Мирзабеков Андрей Дарьевич (р. 1937), советский биохимик, академик АН СССР (1987). Окончил Московский институт тонкой химической технологии им. М. В. Ломоносова (1962); с 1962 г. работает в Институте молекулярной биологии АН СССР, с 1984 г. — директор института. Основные работы — по изучению структур нуклеиновых кислот и нуклеопротеидов и механизма их функционирования. Установил последовательность расположения гистонов вдоль ДНК в нуклеосомах. Лауреат Государственной премии СССР (1969).

Предварительные данные имеются и о структуре комплексов РНК-полимеразы с промоторами — участками ДНК, в которых начинается транскрипция (см. с. 413). Конечным результатом взаимодействия полного фермента с промотором является образование очень прочного комплекса, который и осуществляет инициацию синтеза РНК.

Хроматин. Нуклеосомы. Основной генетический материал в клетках эукариот сосредоточен в хромосомах клеточного ядра, где происходят процессы репликации и транскрипции. В промежутках между делениями клеток хромосомы вытянуты в длинные тонкие нити, в процессе деления они становятся короче и толще. В компактном состоянии ДНК длиной до нескольких сантиметров упаковывается в хромосомы, длина которых измеряется микрометрами. Хромосомы состоят из нуклеопротеида, называемого хроматином. Основными компонентами хроматина являются ДНК, гистоны и негистоновые белки. Гистоны представляют собой небольшие белки основного характера, отличающиеся высокой степенью эволюционной консервативности. По соотношению остатков основных аминокислот (Lys/Arg) гистоны разделяют на 5 классов: Н1, Н2А, Н2В, Н3, Н4. Наиболее консервативны аргининбогатые гистоны Н3 и Н4, в остальных вариации в последовательности аминокислот в зависимости от источника более значительны. Негистоновые белки — это

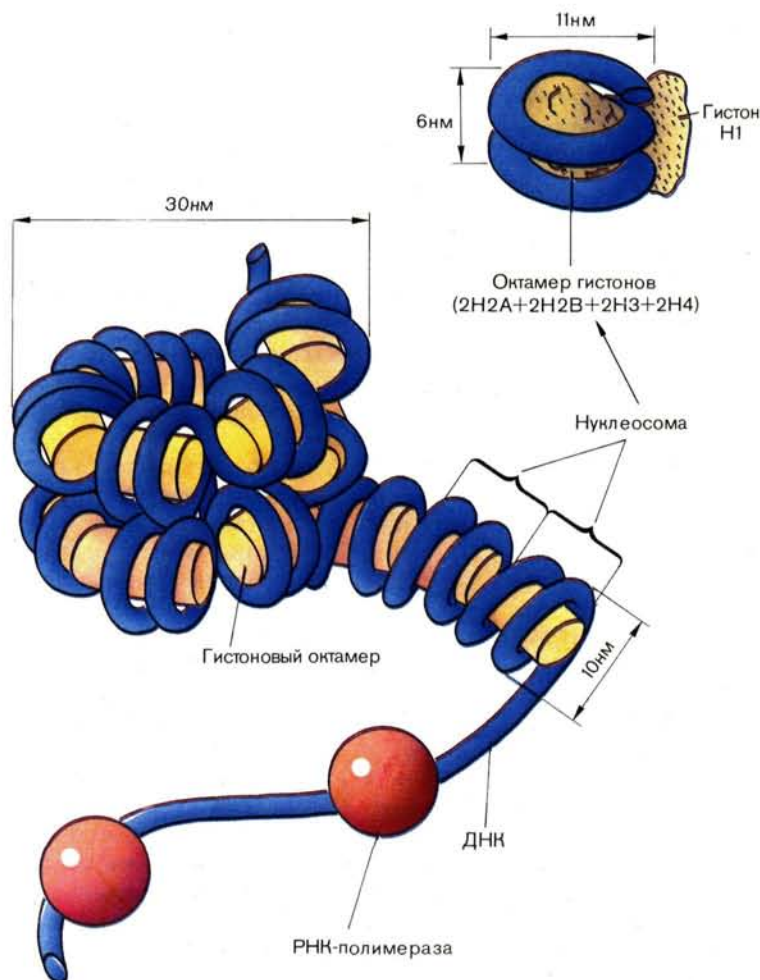


Рис. 223. Структура хроматина и строение нуклеосомы.

общее название всех других белков хроматина разной природы, различных свойств и функций. ДНК в составе хроматина предположительно находится в В-форме, хотя имеются данные, что в отдельных участках хроматина существует Z-форма ДНК. На ДНК приходится около половины массы хромосом, вторую половину их составляют белки.

Знание статических и динамических свойств хроматина необходимо для понимания механизмов репликации и транскрипции. В изучении структуры хроматина за последнее время был достигнут значительный прогресс. В 1974 г. было показано (Р. Корнберг), что хроматин состоит из дискретных частиц — нуклеосом, соединенных «перемычками» — участками свободной ДНК. Непрерывный дуплекс ДНК переходит из одной частицы в другую, образуя структуру, напоминающую нитку бус. Мономерные нуклеосомы можно получить мягкой обработкой хроматина микрококковой нуклеазой. Нуклеосома представляет собой компактную нуклеопроteidную частицу, состоящую из нуклеосомного ядра, линкерной ДНК и связанной с ней одной молекулы гистона Н1 (рис. 223). Размер участка ДНК, приходящегося на одну нуклеосому, — около 200 п. о. — в пределах одного источника является величиной постоянной. При более глубоком нуклеазном гидролизе линкерная ДНК отщепляется, что приводит к потере гистона Н1,

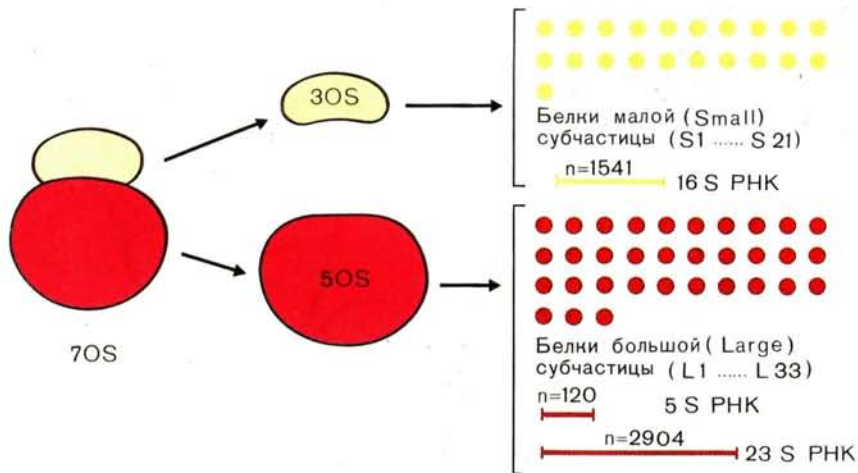


Рис. 224. Схематическое строение рибосом *E. coli* (n — число нуклеотидных остатков в молекуле РНК).

при этом остается нуклеосомное ядро. В нуклеосомном ядре, имеющем форму диска диаметром 11 нм и высотой около 6 нм (А. Круг, Дж. Финч), двойная спираль ДНК (140—150 п. о.) лежит на поверхности и накручена на октамер гистонов Н2А, Н2В, Н3 и Н4, каждый из которых представлен двумя копиями.

В исследованиях взаимодействия гистонов с ДНК в составе хроматина значительная роль принадлежит А. Д. Мирзабекову, Г. П. Георгиеву и др.

В клеточном ядре хроматин уложен очень плотно. Наивысшая степень компактизации хроматина наблюдается в хромосомах на стадии митоза.

Рибосомы. Во всех организмах реализация генетической информации, заключенной в нуклеиновых кислотах, в конкретные

белковые структуры происходят на клеточных органеллах — рибосомах. Прокариотические клетки содержат рибосомы с константой седиментации 70S (S — единица Сведберга) — 70S рибосомы, цитоплазма клеток эукариот — 80S рибосомы. Рибосомы хлоропластов и митохондрий близки по размерам прокариотическим. Все рибосомы состоят из двух неравных субъединиц: 80S рибосомы из 60S и 40S, 70S — из 50S и 30S субъединиц. Рибосома является рибонуклеопротеидом и состоит из РНК и белков. Наиболее полно изучены 70S рибосомы *E. coli*. Их схематическое строение представлено на рисунке 224.

Малая 30S субчастица рибосом *E. coli* состоит из одной молекулы РНК (16S РНК) и 21 молекулы рибосомных белков. В состав 50S субчастицы входят 2 молекулы РНК — 23S и 5S РНК и 33 белка.

В настоящее время установлены первичные структуры всех компонентов рибосом *E. coli*. Оказалось, что белки L7 и L12 имеют идентичные аминокислотные последовательности и отличаются

Рис. 225. Локализация на поверхности 30S субчастиц рибосом *E. coli* 3'-конца 16S РНК с помощью иммунной электронной микроскопии: *a* — фотография димера 30S субчастиц, полученная с помощью электронной микроскопии. Видна бивалентная молекула иммуноглобулина G (IgG), связывающая две 30S субчастицы в точках расположения 3'-концов их 16S РНК, модифицированных гаптемом, к которому и были получены специфические антитела; *b* — перенесение информации на пространственную модель 30S субчастицы.

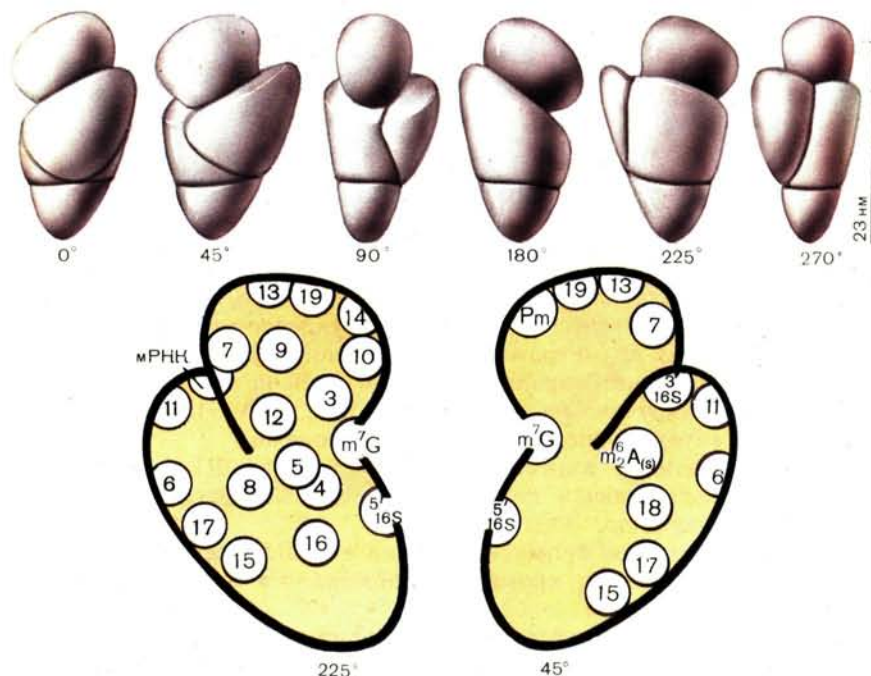
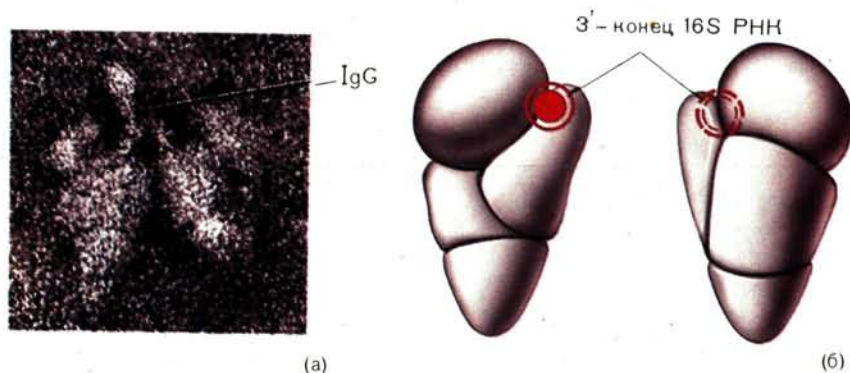


Рис. 226. Пространственная модель 30S субчастицы рибосом в различных проекциях, построенная по данным электронной микроскопии. Цифры в кружках — положение соответствующих белков, 3'-16S и 5'-16S — положение 3'- и 5'-концов 16S РНК; m^7G и $m^5A_{(3)}$ — локализация остатков 7-метилгуанина и двух б-диметиладенинов, Pm — локализация места связывания антибиотика пурамицина (в н и з у).

только ацелированием N-концевой аминокислоты у L7. В отличие от других рибосомных белков, эти белки представлены каждый двумя копиями и в структуре 50S субчастицы образуют тетрамер L7/L12. Кроме того, идентичными оказались белки S20 и L26, так что число индивидуальных белковых компонентов в рибосоме меньше, чем сумма белков обеих субчастиц. Таким образом, 70S рибосома *E. coli* состоит из 3 молекул РНК (16S, 23S и 5S) и 53 различных белков. За исключением некоторых (S1, S2, S6, L7/L12), рибосомные белки имеют основной характер.

Рибосомы эукариот 80S устроены аналогичным образом, однако количество белков в субъединицах больше, РНК длиннее (18S и 28S), а большая субчастица содержит не одну (5S), а две молекулы малых рибосомных РНК — 5S и 5,8S.

Рибосома является нуклеопротеидом, способным к самосборке *in vitro* (М. Номура). Схемы (карты) последовательного присоединения рибосомных компонентов, приводящего к реконструкции

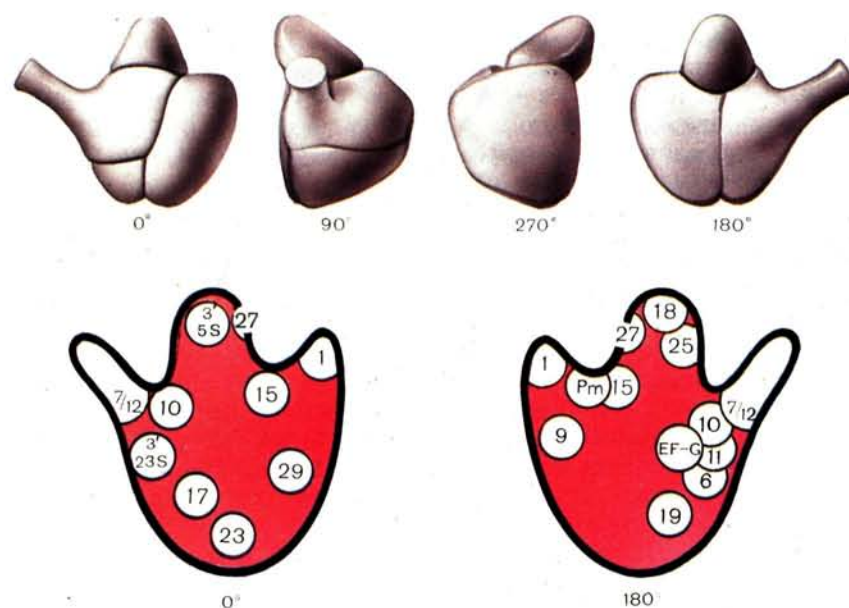


Рис. 227. Пространственная модель 50S субчастицы рибосом *E. coli*, представленная в различных проекциях (вверху). Расположение на поверхности 50S субчастиц структурных и функциональных центров (внизу). Обозначения: цифры в кружках — расположение соответствующих L-белков, 3'-5S и 3'-23S — локализация 3'-концов рибосомных 5S и 23S РНК, Pm и EF-G — локализация мест связывания пурамицина и фактора элонгации G соответственно.

активных субчастиц, хорошо известны. Возможность осуществления сборки с неполным набором компонентов, с заменой одного или нескольких из них на мутантные, химически модифицированные или гомологичные компоненты из другого организма, дает удобный метод структурно-функционального изучения рибосом.

В исследовании пространственной организации рибосом наибольшую роль сыграли методы электронной микроскопии, рассеяния нейтронов и химические методы ковалентного связывания компонентов друг с другом.

Особенно важная роль принадлежит электронной микроскопии в исследовании комплексов рибосом с антителами к индивидуальным рибосомным белкам или отдельным участкам рРНК, модифицированным низкомолекулярными химическими агентами (гаптенами) или несущим природную модификацию (6,6-диметиладенин, 7-метилгуанин). Этот подход называется иммунной электронной

микроскопией. Его возможности проиллюстрированы на рисунке 225, где представлена фотография димера 30S субчастицы рибосом, в котором 3'-концы 16S РНК, модифицированные гаптенем, соединены молекулой специфичного к гаптену иммуноглобулина G. Совокупность данных, полученных с помощью такого подхода, привела к созданию моделей субчастиц рибосом *E. coli* с относительным расположением важных структурных и функциональных центров на их поверхности (рис. 226—228). В этой области существенный вклад внесены тремя группами исследователей: М. Номура (США), Х. Г. Витман (ФРГ), А. С. Спири́н (СССР).

Вирусы и другие нуклеопротеиды. Приведенные примеры далеко не исчерпывают список известных нуклеопротеидных структур. Существует целый мир бактериальных, растительных и животных вирусов, в котором обнаружено поразительное многообразие вирусных частиц (вирионов) как по строению и составу, так и по способам хранения и воспроизведения генетической информации. В отличие от клеток, где хранителем наследственности всегда является двуспиральная ДНК, а РНК служит только для переноса и реализации генетической информации, вирусы в качестве генетического материала используют как ДНК (ДНК-содержащие вирусы), так и РНК (РНК-содержащие вирусы). Геномная ДНК может быть одноцепочечной или двуспиральной, кольцевой или линейной. РНК-содержащие вирусы также чрезвычайно разнообразны: они могут содержать одноцепочечную или двуспиральную РНК, их геном может быть представлен одной или сразу несколькими молекулами РНК, упакованными в одну капсиду.

В состав вирионов входят самые разнообразные белки — от белков, выполняющих чисто структурные функции, до сложных ферментов, необходимых вирусу для взаимодействия с клеткой-хозяином или для воспроизведения.

Из других нуклеопротеидных комплексов следует упомянуть информосомы (комплексы мРНК со специфическими белками в клетках эукариот), в изучение которых основной вклад внесен советскими исследователями (А. С. Спири́н, Г. П. Георгиев, Л. П. Овчинников), и комплексы аминоксил-тРНК-синтетаз с тРНК, исследованиям структуры которых уделяется большое внимание

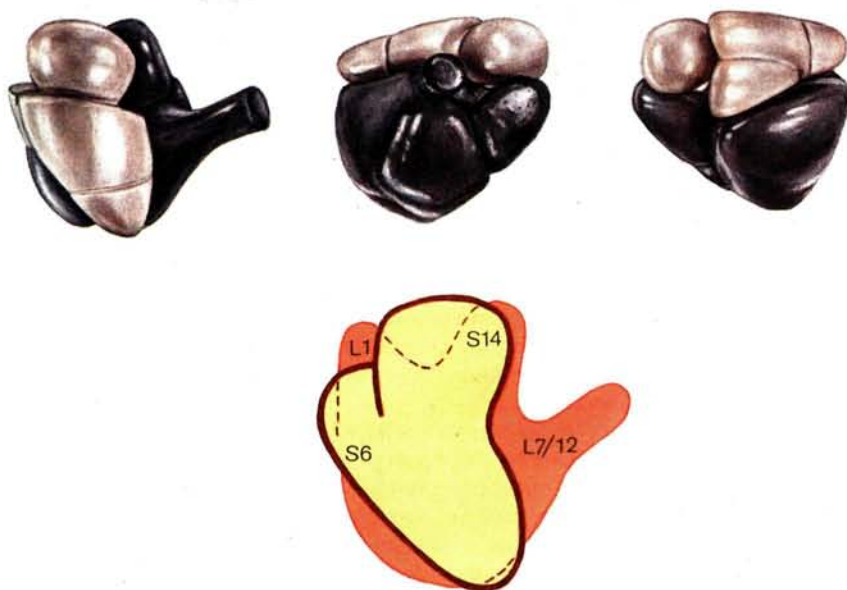


Рис. 228. Относительное расположение большой и малой субчастиц рибосом *E. coli* по данным электронной и иммунной электронной микроскопии.

как в зарубежных лабораториях (Ж.-П. Эбель), так и в СССР (Д. Г. Кнорре, Л. Л. Киселев).

Изучение структуры этих и других нуклеопротеидов является необходимым звеном в решении общей проблемы взаимодействия нуклеиновых кислот и белков в живой клетке — взаимодействия, обеспечивающего этапы воспроизведения и реализации ее наследственной информации.



Спирин Александр Сергеевич (р. 1931), советский биохимик, академик АН СССР (1970). Окончил Московский университет (1954), с 1967 г. — директор Института белка АН СССР. Основные работы посвящены биохимии нуклеиновых кислот и биосинтезу белка. Установил структурные превращения рибосом и сформулировал один из основных принципов их строения и самосборки, открыл информсомы. Предложил модель молекулярного механизма работы рибосомы. Лауреат Ленинской премии (1976) и Государственной премии СССР (1986).

Проблемы нуклеиново-белкового узнавания

Все компоненты нуклеопротеидных комплексов синтезируются отдельно и затем собираются в функционирующую структуру. Для того чтобы сборка произошла и прошла правильно, они должны «узнать» друг друга и получить программу сборки. Узнавание нуклеиновой кислоты белком представляет собой процесс, каждая стадия которого осуществляется за счет нуклеиново-белковых взаимодействий.

Взаимодействия в процессе узнавания могут быть специфическими и неспецифическими. Под специфическим нуклеиново-белковым взаимодействием подразумевается кооперативное взаимодействие определенных групп белка и нуклеиновой кислоты, возникающее за счет характерного для данного белка и данной нуклеиновой кислоты пространственного расположения этих групп. Примеры специфических взаимодействий: репрессоры и операторы, РНК-полимераза и промоторы.

Неспецифические взаимодействия значительно более однородны и обуславливаются только электростатическими или гидрофобными связями. Например, РНК-полимераза образует неспецифические комплексы с любыми полианионами — гепарином, полиинозиновой кислотой, одноцепочечной и двухцепочечной ДНК.

Проблема узнавания значительно более сложна, чем определение структуры уже существующего нуклеопротеида. Например, имеются последовательности ДНК, которые посредством специфических взаимодействий с белками осуществляют регуляцию функционирования ДНК. К таким функционально значимым участкам ДНК относятся промоторы, операторы, терминаторы, участки начала репликации и многие другие последовательности. Эти последовательности должны быть узнаны белком, выполняющим регуляторную функцию, т. е. они должны отличаться от всех других последовательностей генома. Если однотипных участков в данном геноме не один, а несколько, как в случае промоторов, то они должны нести и одинаковые элементы последовательности. Задача поисков регуляторных последовательностей очень сложна даже в случае сравнительно простых геномов, таких, как геном *E. coli*. Так, репрессор лактозного оперона узнает в ДНК *E. coli* только одну последовательность *lac*-оператора и тем самым отличает ее от 10^7 других последовательностей.

Накопление экспериментальных данных по нуклеиново-белковому узнаванию позволяет постепенно подходить к выяснению существующих закономерностей. В частности, выдвинута достаточно обоснованная концепция, что любой полипептид способен специфически связываться с нуклеотидной цепью, комплементарной его матрице, или с кодируемым комплементарной цепью полипептидом (называемым антипептидом).

В качестве подтверждения этого предположения недавно были синтезированы антипептиды для АКТГ и γ -эндорфина. Структуры антипептидов были выведены из последовательностей РНК, комплементарных мРНК для данных пептидов. Оказалось, что соответствующие пептид и антипептид (рис. 229) образуют в растворе

(1-24)

АКТГ H-Ser Tyr Ser Met Glu His Phe Arg Trp Gly Lys Pro Val Gly Lys Lys Arg Arg Pro Val Lys Val Tyr Pro-OH

H-Gly Val His Leu His Arg Ala Pro Leu Leu Ala His Arg Leu Ala Pro Ala Glu Val Phe His Gly Val Arg-OH

Рис. 229. Структуры фрагментов 1—24 АКТГ и антипептида к нему.

прочные комплексы (константа связывания $\sim 10^8 \text{ M}^{-1}$), причем антипептид способен конкурировать при связывании даже со специфическими к пептиду антителами. Эти результаты объясняются гидрофобно-гидрофильными взаимодействиями, специфичность которых вытекает из природы самого генетического кода.

Процессы с участием нуклеиновых кислот: репликация, транскрипция и трансляция

Основная функция нуклеиновых кислот заключается в хранении, воспроизведении и передаче генетической информации. Являясь системами динамическими, нуклеиновые кислоты осуществляют все процессы с высокой скоростью и эффективностью, постоянно взаимодействуя с соответствующими белками, прежде всего с ферментами. Главными процессами с их участием являются репликация, транскрипция и трансляция.

Репликация

Репликацией называется процесс удвоения ДНК. Принципиальный механизм его вытекает из строения двуспиральной молекулы ДНК (см. с. 335). В результате репликации образуются две молекулы ДНК, представляющие собой точные копии исходной молекулы. Каждая из вновь образовавшихся молекул содержит одну цепь исходной ДНК и одну вновь синтезированную цепь. Иными словами, репликация полуконсервативна—половина родительской молекулы сохраняется в дочерней молекуле.

Однако этот принципиальный механизм даже в простейших случаях реализуется путем сочетания многих сложных процессов, в которые вовлечены многочисленные ферменты и регуляторные белки.

Лучше всего процессы репликации изучены для наиболее простых систем — бактерий, бактериофагов и внехромосомных генетических элементов бактерий — плазмид.

Принято использовать понятие «репликон», предложенное в 1963 г. Ф. Жакобом, С. Бреннером и Ф. Кьюзенем для обозначения генетической единицы репликации, т. е. сегмента ДНК, который автономно воспроизводится (реплицируется) в процессе клеточного роста и деления. Каждый репликон должен иметь систему «управления» собственной репликацией. Хромосома *E. coli*, плазмиды, ДНК бактериофагов представляют собой репликоны разной сложности, способные к автономной репликации в клетке и имеющие систему инициации. Репликон может содержать в себе гены, кодирующие синтез всех белков, необходимых для репликации (хромосома *E. coli*), части таких белков (некоторые сравнительно крупные бактериофаги) или использовать для своей репликации практически только чужие белки (мелкие фаги M13 или G-4, содержащие однонитевые циклические ДНК).

Ключевую роль в процессе репликации играют реплицирующие ДНК-полимеразы, которые осуществляют матричный синтез ДНК из дезоксирибонуклеозидтрифосфатов. Фермент синтезирует нить ДНК, комплементарную родительской нити (называемой матрицей), последовательно присоединяя к 3'-концу растущей цепи мононуклеотидные звенья, комплементарные звеньям матрицы (рис. 230). При этом ДНК-полимераза катализирует нуклеофильную атаку 3'-ОН-группы концевой нуклеотида растущей цепи на α -фосфатную группу дезоксирибонуклеозидтрифосфата, отбираемого ферментом на основе его комплементарности соответствующему звену матрицы. В результате отщепляется пирофосфат и образуется фосфодиэфирная связь. Растущая цепь удлиняется на одно звено, и процесс повторяется с новым дезоксирибонуклеозидтрифосфатом. Для того чтобы ДНК-полимераза могла начать синтез, необходимо существование уже готового фрагмента ДНК или РНК, комплементарного матрице и содержащего свободную 3'-ОН-группу. Этот фрагмент называют затравкой. В процессе синтеза дочерних цепей родительская двухцепочечная ДНК расплетается, образуя структуру, по форме напоминающую латинскую букву Y. Такая структура называется репликативной вилкой.

Процессы с участием нуклеиновых кислот: репликация, транскрипция и трансляция



Бреннер [Brenner] Сидней (р. 1927), английский ученый, работающий в области молекулярной биологии и молекулярной генетики. Образование получил в университете Витватерсранда (ЮАР) и Оксфордском университете, с 1957 г. работает в лаборатории молекулярной биологии Медицинского исследовательского совета в Кембридже (Великобритания). Вместе с Ф. Жакобом и М. Меселсоном экспериментально подтвердил идею о существовании РНК-посредника.

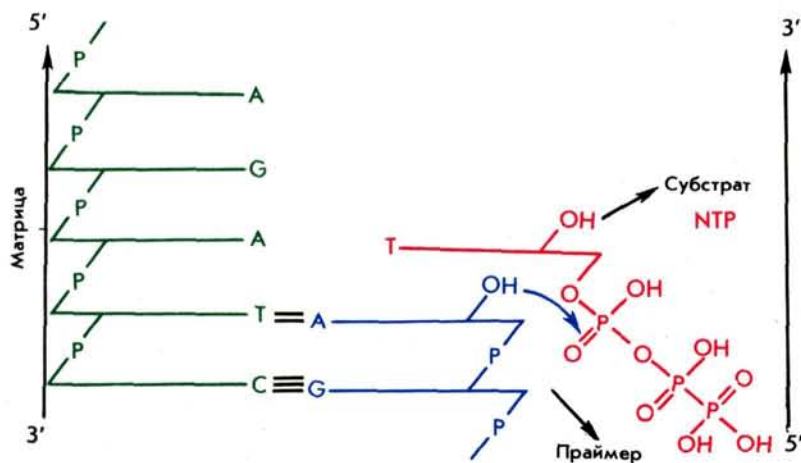


Рис. 230. Синтез комплементарной цепи ДНК, катализируемый ДНК-полимеразой.

Репликация хромосомы *E. coli*

Хромосома *E. coli* представляет собой циклическую двухцепочечную ДНК, длина которой равна приблизительно $3,8 \cdot 10^6$ звеньев. Хромосома существует в комплексе с белками и клеточной мембраной и довольно плотно упакована (такой упакованный комплекс называется нуклеоидом). Схема репликации хромосомы *E. coli* изображена на рисунке 231. Каждый цикл репликации начинается в определенном месте хромосомы, называемом *ori*. Термин *ori* обозначает специфические последовательности ДНК-репликонов, служащие сигналом инициации репликации (от английского *origin of replication*). Начавшись, репликация протекает одновременно в двух направлениях, так что образуются и движутся две репликативные вилки. Скорость движения каждой вилки составляет около 800 нуклеотидов в секунду и не зависит от условий роста клеток; таким образом, бактериальная хромосома всегда реплицируется примерно за 40 мин. Образовавшиеся две циклические ДНК разделяются (рис. 232).

Репликация хромосомы *E. coli* осуществляется с помощью фермента ДНК-полимеразы III, действующего в ансамбле с множеством других белков, выполняющих функции расплетения цепей ДНК (хеликазы), связывания одноцепочечной ДНК (SSb белок), синтеза праймера (праймаза), лигирования ДНК-фрагментов (ДНК-лигаза) и т. д. Сам фермент ДНК-полимераза III не способен начать репликацию. Инициация репликации начинается с расплетения нитей родительской ДНК в области *ori C* (рис. 231) и синтеза короткого РНК-транскрипта, который служит праймером (затравкой) для синтеза ДНК. В некоторых случаях (ColE1 плаزمид, ДНК нитевидных фагов типа M13) праймер синтезируется клеточ-

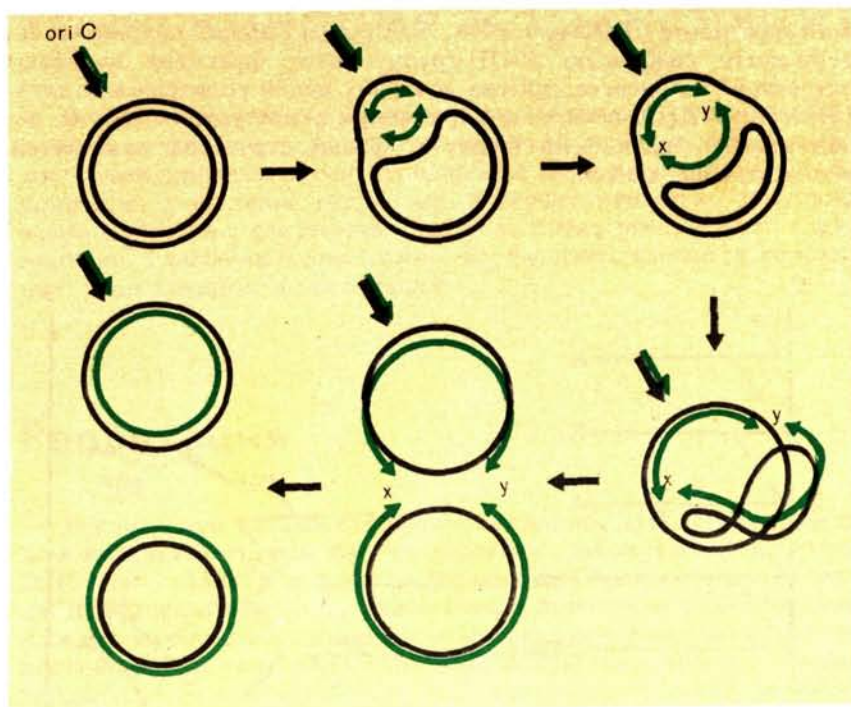


Рис. 231. Схема репликации хромосомы *E. coli* (X и Y — две репликативные вилки).

ной ДНК-зависимой РНК-полимеразой, в остальных — специальными ферментами-праймазами. В *E. coli* праймаза активируется примосомой — мультимерным белковым комплексом, содержащим около 20 полипептидных цепей (7 различных субъединиц), в функцию которого входит изменение конформации одной из цепей ДНК, необходимое для взаимодействия с праймазой (рис. 232).

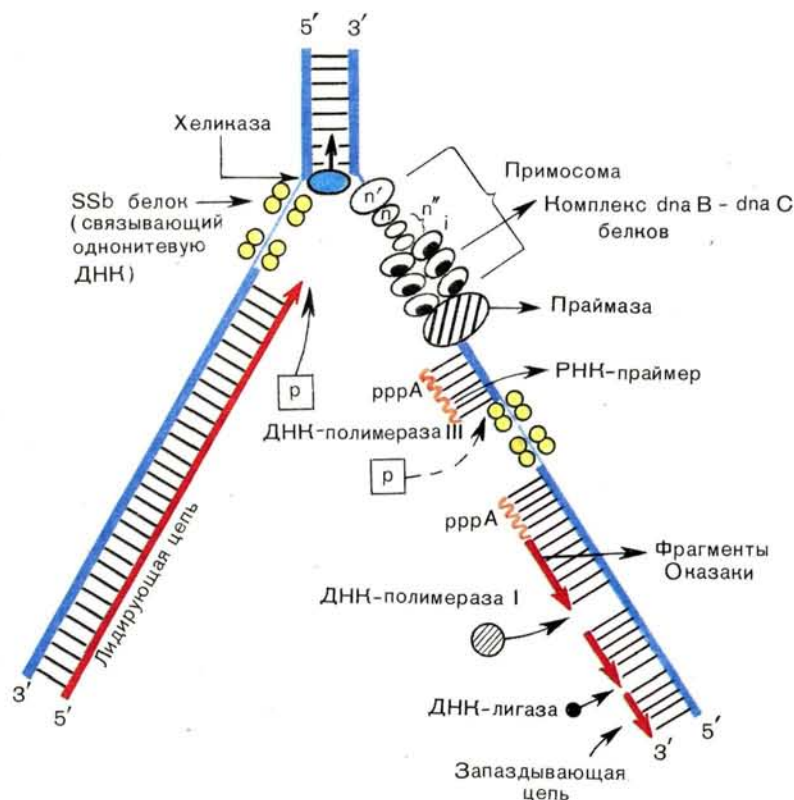


Рис. 232. Репликация ДНК в *E. coli*. Схема событий в репликативной вилке: п', п, п'', i, dna B, dna C, SSb — белковые продукты соответствующих генов хромосомы *E. coli*, вовлеченные в репликацию.

После инициации начинается продвижение репликативных вилок — элонгация. Одна из цепей вновь синтезируемой ДНК удлиняется в том же направлении, в котором движется репликативная вилка, причем синтез осуществляется непрерывно. Эту цепь ДНК называют лидирующей. Другая цепь — запаздывающая, синтезируется короткими фрагментами Оказани (рис. 232). Синтез каждого фрагмента иницируется вблизи начала репликационной вилки и продолжается в противоположную от нее сторону до тех пор, пока 3'-конец вновь синтезируемой ДНК не достигнет 5'-конца предыдущего фрагмента Оказани. Инициация синтеза этих фрагментов осуществляется праймазой, которая в качестве затравки синтезирует короткие фрагменты РНК, далее они удлиняются ДНК-полимеразой III, являющейся основным реплицирующим ферментом *E. coli*. Длины фрагментов Оказани равны примерно 1000 нуклеотидов. В репликативной вилке «работают» также еще два белка. Один из них (SSb белок) специфически связывает одноцепочечные ДНК, облегчая расплетение двойной спирали и одновременно защищая одноцепочечные участки от действия нуклеаз. Другой — ДНК-расплетающий белок (хеликазы) — движется вдоль двойной цепи,

расплетая ее при одновременном гидролизе АТФ (см. с. 408). Наконец, еще один фермент — ДНК-гираза, или ДНК-топоизомераза II, предотвращает накопление положительных супервитков, вызываемое расплетением двойной спирали, образуя в циклической ковалентнозамкнутой ДНК одноцепочечные разрывы и затем снова «сшивая» ее.

Фермент ДНК-полимераза I удаляет РНК-затравку и достраивает фрагменты, а ДНК-лигаза соединяет между собой соседние фрагменты Оказаки фосфодиэфирной связью, которая образуется между 3'-гидроксильной группой одного и 5'-фосфатом другого фрагментов. Терминация синтеза ДНК определяется специфической последовательностью. Детали этого процесса пока неизвестны. Наконец, две дочерние хромосомы, образовавшиеся после репликации, прикрепляются к мембране. Рост участка мембраны между точками прикрепления раздвигает их, и разделение клеток завершает процесс.

Репликация ДНК бактериофагов и плазмид

Наиболее хорошо изучена репликация бактериофагов, содержащих в своем вирионе небольшие (около 5000 нуклеотидных звеньев) циклические одноцепочечные ДНК. К ним относятся, например, фаг ψ X-174 и широко используемые в генной инженерии фаги M13 и fd. Их репликация состоит из трех разных по механизму этапов.

На первом этапе при попадании в клетки *E. coli* одноцепочечная циклическая ДНК фага, или (+)-цепь, превращается в двухцепочечную ковалентно замкнутую сверхспиральную ДНК, так называемую RF-I форму (от английского *replicative form*). В качестве промежуточного продукта образуется циклическая двухцепочечная ДНК с одним разрывом, или RF-II форма. Формы RF-I и RF-II образуются в результате синтеза цепи ДНК, которая комплементарна (+)-цепи; новую цепь называют (-)-цепью. Второй этап заключается в репликации RF-I формы для увеличения ее содержания в клетке. Наконец, на третьем этапе синтезируется в больших количествах вирусная (+)-цепь, которая затем упаковывается в вирусную оболочку.

Репликация плазмид, используемых в качестве векторов в генной инженерии, протекает значительно медленнее, чем репликация хромосомы *E. coli*, и, по-видимому, не требует прикрепления ДНК к клеточной мембране. В частности, репликация плазмиды ColEI протекает только в одном направлении, т. е. осуществляется с образованием одной репликативной вилки. Для репликации ColEI нужны только белки, кодируемые хромосомой, и участок *ori* плазмиды, где происходит инициация репликации. Инициация синтеза ДНК в этом случае осуществляется с участием РНК-полимеразы *E. coli*.

Репликация в эукариотических клетках

ДНК эукариотических хромосом находится в комплексе с равным по весу количеством гистонов. Как уже было отмечено, примерно каждые 200 п. о. ДНК образуют суперспираль, накрученную на октамер гистонов H2A, H2B, H3 и H4. Такая структурная единица

называется нуклеосомой. Следуя друг за другом, нуклеосомы образуют хроматин (см. с. 400).

Отличительным свойством репликации у эукариот является то, что реплицируются нуклеосомы: при синтезе дочерних цепей они на какое-то время разрушаются, но позади репликативной вилки вновь собираются, причем в сборке участвуют как старые, так и вновь синтезированные гистоны (рис. 233). Таким образом, синтез гистонов должен быть скоординирован с репликацией. Действительно, ингибирование синтеза гистонов влечет за собой ингибирование репликации, и наоборот.

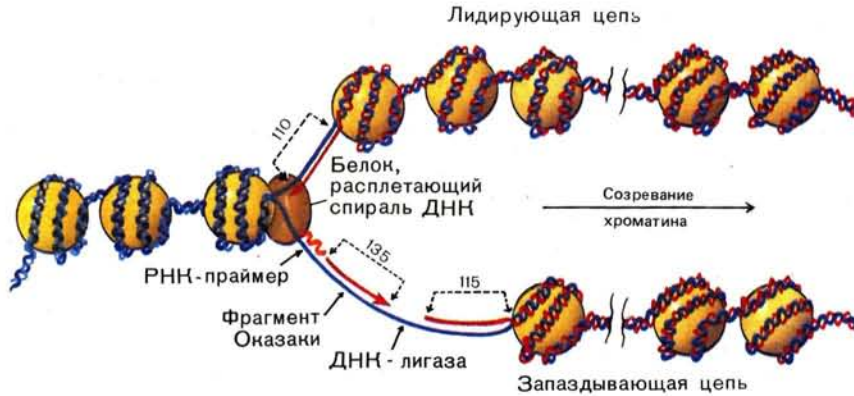


Рис. 233. Схема репликации хроматина.

Так же как у прокариот, репликация состоит из трех основных стадий: инициации, элонгации и терминации. Репликация эукариотической ДНК происходит одновременно во многих областях хромосомы и, по-видимому, иницируется на определенных последовательностях ДНК, которые хорошо идентифицированы у ряда вирусов. Инициация, как и у прокариот, требует участия специфических белков.

Элонгация синтеза осуществляется ДНК-полимеразами. В клетках эукариот известно три типа ДНК-полимераз: α , β и γ . Предполагается, что репликацию основной клеточной ДНК осуществляет полимераз α , репарацию повреждений — полимераз β , а репликацию ДНК митохондрий — полимераз γ . Так же как и у прокариот, в репликативной вилке одна из цепей является ведущей (лидирующей), а другая — отстающей (запаздывающей) (рис. 234). Лидирующая цепь синтезируется непрерывно, тогда как запаздывающая — фрагментами Оказаки. Эти фрагменты также иницируются короткими РНК, которые синтезируются, по-видимому, РНК-полимеразой I. В распространении репликативной вилки принимают участие дестабилизирующие двойную спираль ДНК-связывающие белки.

Транскрипция

Процесс биосинтеза РНК в клетке называется транскрипцией. Его результатом является образование РНК, комплементарных отдельным участкам ДНК. При этом в каждом участке РНК комплементарна только одной определенной нити ДНК.

Процесс биосинтеза РНК осуществляется ферментами — РНК-полимеразами, которые используют ДНК в качестве матрицы. Как и в случае репликации, механизм образования фосфодиэфирных связей включает в себя катализируемую ферментом нуклеофильную атаку 3'-гидроксильной группы растущей цепи на α -фосфатную группу присоединяемого субстрата (рибонуклеозидтрифосфата). При образовании фосфодиэфирной связи от трифосфата отщепляется неорганический пирофосфат. Каждый вновь присоединяемый нуклеозид комплементарен тому звену матрицы, которое является ближайшим 5'-соседом только что скопированного звена. Цепь РНК растет в направлении $5' \rightarrow 3'$ по мере движения РНК-полимеразы по копируемой цепи ДНК в направлении от 3'-конца к 5'-концу. Схема процесса транскрипции показана на рисунке 234.

Как и репликация, транскрипция состоит из трех основных этапов: инициации, элонгации и терминации. В отличие от ДНК-полимераз, РНК-полимеразы способны к самостоятельной инициации синтеза РНК, которая осуществляется в определенных точках ДНК. Место инициации синтеза РНК определяется специальными регуляторными участками ДНК — промоторами. Терминация синтеза также происходит на специфических участках ДНК — терминаторах. Процесс транскрипции регулируется разнообразными способами, что позволяет клетке приспосабливаться к изменениям условий существования. Наиболее хорошо изучены транскрипция и способы ее регуляции у бактерий и бактериофагов.

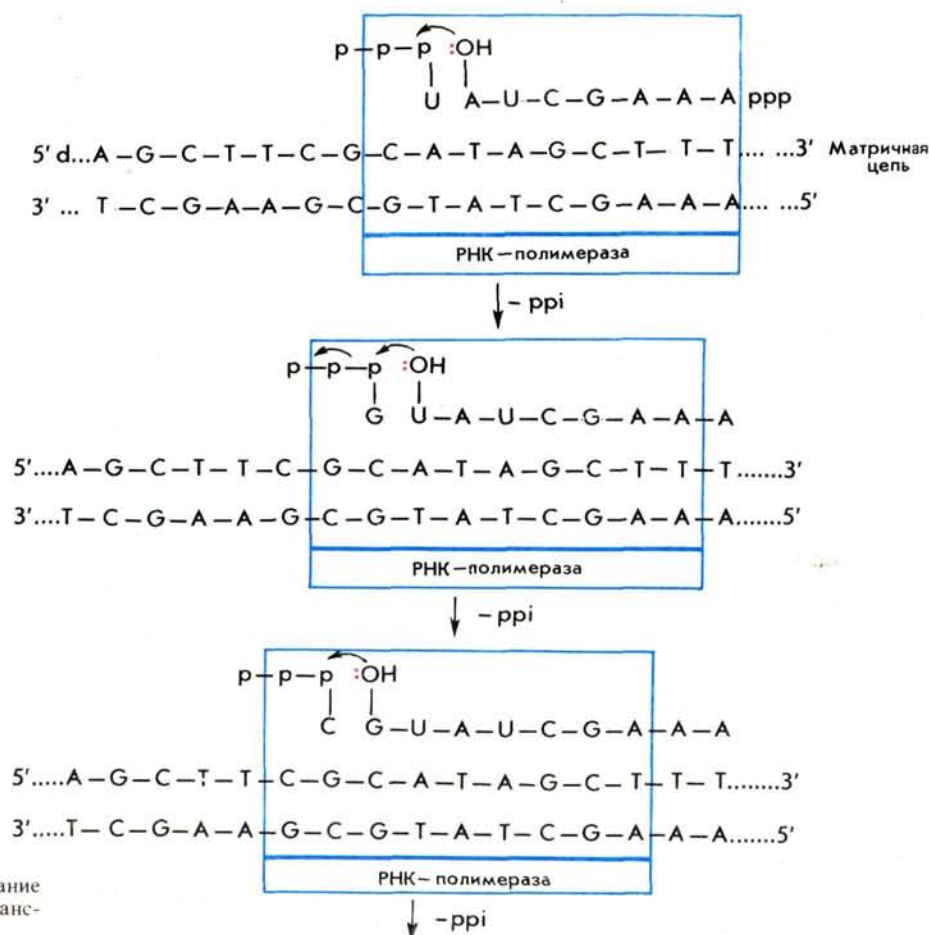


Рис. 234. Последовательное копирование одной из цепей ДНК в процессе транскрипции.

Процессы с участием
нуклеиновых кислот:
репликация, транскрипция
и трансляция



Свердлов Евгений Давидович (р. 1938), советский химик-биоорганик, член-корреспондент АН СССР (1984). Окончил Московский государственный университет (1961), работает в Институте биоорганической химии им. М. М. Шелякина АН СССР. Основные работы — в области химии нуклеиновых кислот и генной инженерии. Осуществил клонирование генов интерферона человека и получил штаммы-суперпродукты этих белков. Предложил ряд методов исследования первичной и высших структур нуклеиновых кислот. Лауреат Государственной премии СССР (1982) и Ленинской премии (1985).

Бактериальная хромосома содержит около 4000 генов, которые могут транскрибироваться как независимо друг от друга, так и координированно. Независимая транскрипция гена возможна, если он обладает собственным промотором и терминатором транскрипции. При координированной транскрипции группа генов имеет общий промотор и общий терминатор и составляет один непрерывный участок ДНК; мРНК, иницируемая на общем промоторе, содержит информацию для синтеза нескольких полипептидов и называется полицистронной. Группа генов, которые транскрибируются совместно, называется опероном.

РНК-полимеразы бактерий содержат четыре субъединицы — α , β , β' и δ , которые образуют комплекс $\alpha_2\beta\beta'\delta$, называемый полным ферментом (холоферментом). Первичная структура α, β, β' -субъединиц РНК-полимеразы *E. coli* была установлена Е. Д. Свердловым, В. М. Липкиным, Н. Н. Модяновым, Г. С. Монастырской и др. (СССР). Полный фермент $\alpha_2\beta\beta'\delta$ способен «узнать» промотор и связаться с ним с образованием прочного комплекса (рис. 235). В ряде случаев доказано, что при образовании такого комплекса в промоторной последовательности (включающей в среднем около 40 п. о.) расплетается небольшой (10—15 п. о.) участок ДНК. Вслед за этим происходит инициация синтеза РНК, в результате которой образуется первая фосфодиэфирная связь. Первый нуклеотид на 5'-конце синтезируемой РНК называется иницирующим. Чаще всего это пуриновый нуклеотид. После того как синтезируется сравнительно короткий олигорибонуклеотид, δ -субъединица отделяется от фермента, и дальнейшую полимеризацию (элонгацию) осуществляет комплекс $\alpha_2\beta\beta'$, который называется минимальным ферментом (кор-ферментом). δ -Субъединицу можно рассматривать как фактор позитивной регуляции транскрипции.

Для терминации транскрипции на ДНК имеются особые сигналы — терминаторные последовательности. В этом случае терминация осуществляется минимальным ферментом, но иногда необходим дополнительный белковый фактор, названный «ро» (ρ). Таким образом, общая схема транскрипции выглядит так, как изображено на рисунке 235.

В бактериальной клетке существует очень большое число промоторов, различающихся по нуклеотидной последовательности и эффективности инициации синтеза РНК. Часто их условно делят на сильные и слабые. Сильные промоторы обеспечивают высокий уровень синтеза РНК с оперонов, находящихся под их контролем, а слабые — соответственно низкий уровень. Использование промоторов различной силы является одним из способов регуляции синтеза РНК в клетке. Количество синтезированной РНК определяет уровень биосинтеза белка. Многие белки требуются клетке на протяжении всей ее жизни. Их количество зависит главным образом от эффективности промоторов, иницирующих синтез соответствующих мРНК. Однако у клетки существуют и более радикальные способы регуляции на уровне транскрипции. Она их использует в случае необходимости для того, чтобы включать или выключать транскрипцию некоторых генов, ускорять или замедлять ее.

Рассмотрим некоторые из них на примере регуляции лактозного оперона *E. coli*, обеспечивающего способность бактериальной клетки использовать в качестве источника углерода дисахарид лактозу. Этот оперон, называемый *lac*-опероном, явился одной из первых систем, позволивших Ф. Жакобу и Ж. Моно сформулировать концепцию регулируемого оперона.

Если клетки *E. coli* растут в среде, содержащей лактозу, то в них синтезируется фермент β -галактозидаза, гидролизующий этот дисахарид до глюкозы и галактозы, которые далее усваиваются клеткой. Оказывается, что если источником углерода для *E. coli* является глюкоза, то β -галактозидаза не синтезируется. Однако если в среду, в которой растут клетки, добавить негидролизующий аналог лактозы — изопропил- β -D-тиогалактозид, то начинается (индуцируется) синтез β -галактозидазы. Соединения, которые, подобно лактозе и изопропил- β -D-тиогалактозиду, способны индуцировать синтез белков (в данном случае β -галактозидазы), в других условиях называются индукторами.

Механизм индукции β -галактозидазы хорошо установлен. В лактозный оперон, кроме гена, кодирующего β -галактозидазу *lacZ* входят еще два структурных гена *lacY* и *lacA* (рис. 236), транскрипция которых проходит под контролем одного промотора, называемого *lac*-промотором. Ген *lacZ* кодирует β -галактозидазу, ген *lacY* — белок галактозидпермеазу, который обеспечивает транспорт галактозидов из среды в клетку, и ген *lacA* — фермент тиогалактозид ацетилтрансферазу (рис. 236, а). Перед *lac*-опероном находится еще один ген, названный *lacI*, который кодирует белок-регулятор

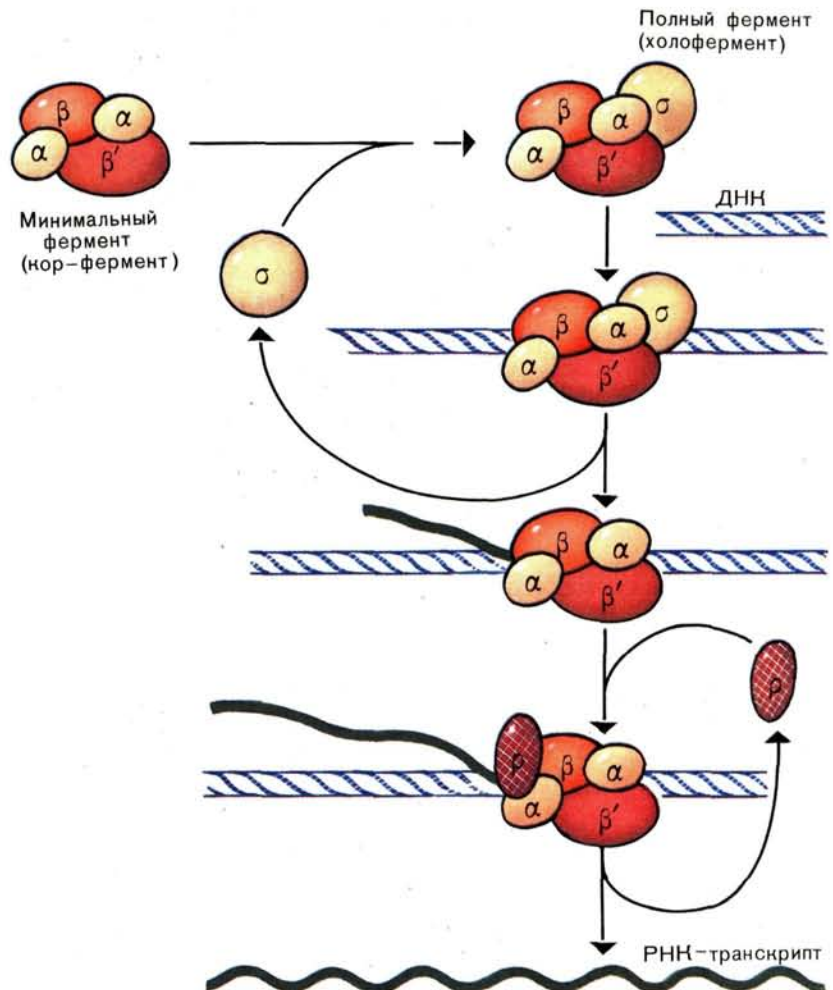


Рис. 235. Схема синтеза РНК с помощью ДНК-зависимой РНК-полимеразы.

способный прочно связываться со специальным участком ДНК — оператором. И *lac*-промотор, и *lac*-оператор находятся между геном *lacI* и *lac*-опероном. В отсутствие лактозы белок гена *lacI* связывается с операторным участком ДНК в непосредственной близости от промотора и препятствует взаимодействию с ним РНК-полимеразы. Другими словами, белок запрещает (репрессирует) транскрипцию; он назван *lac*-репрессором (рис. 236, б). Регуляция по типу запрета транскрипции называется негативной.

При появлении в среде лактозы или другого индуктора последний связывается с репрессором, образуя прочный комплекс. В результате репрессор отделяется от ДНК, освобождая промотор для взаимодействия с РНК-полимеразой. Однако в случае *lac*-оперона удаление репрессора оказывается недостаточным для того, чтобы началась эффективная транскрипция. В системе участвует еще один регуляторный элемент, который активирует транскрипцию. Активация происходит за счет взаимодействия комплекса цикло-АМР-связывающего белка САР (от англ. catabolite activator protein) и 3', 5'-цикло-АМР с участком ДНК, также примыкающим к промотору, но со стороны, противоположной оператору (рис. 236, в). Такой тип регуляции называется позитивным.

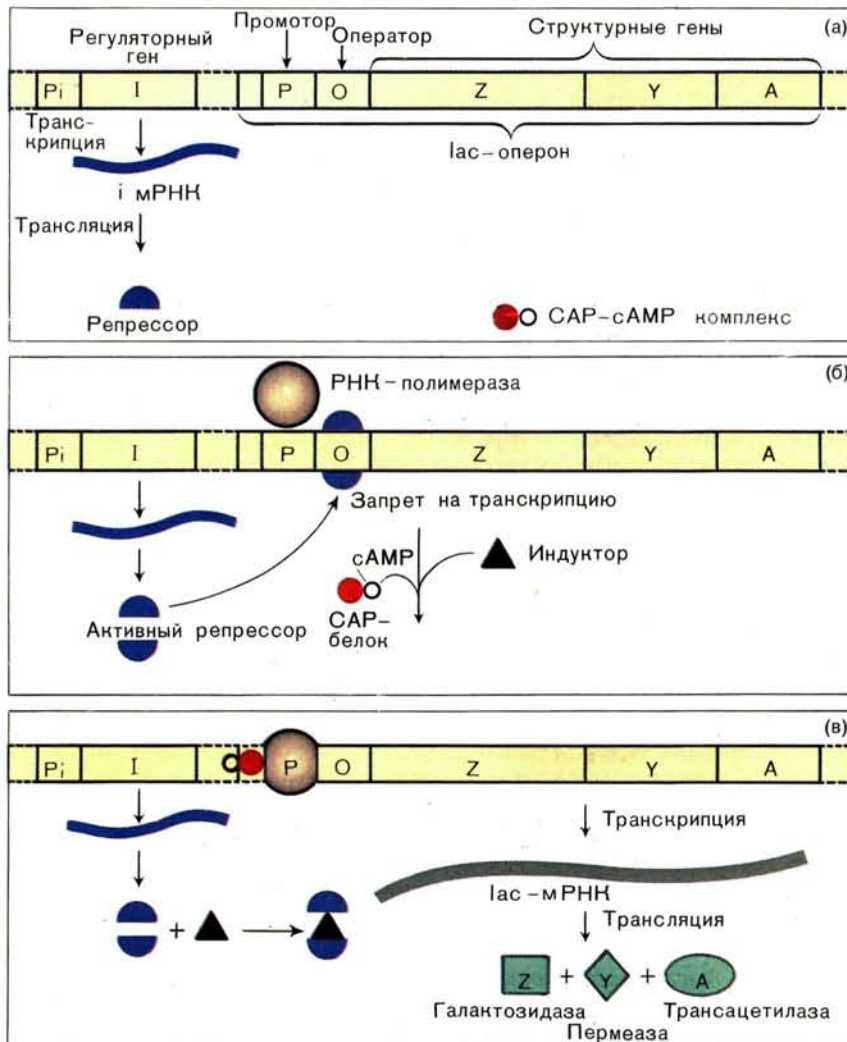


Рис. 236. Структура *lac*-оперона *E. coli* (а) и механизмы регуляции его транскрипции: репрессии (б), активации и индукции (в).

Транскрипция в клетках эукариот



Жакоб [Якоб] Франсуа (р. 1920), французский микробиолог и генетик. Окончил Парижский университет (1947), с 1965 г. — профессор кафедры генетики клетки в Коллеж де Франс. Основные работы посвящены генетике бактериальных клеток и вирусов. Предложил схему регуляции активности генов. Лауреат Нобелевской премии по физиологии и медицине (1965, совместно с А. М. Львовым и Ж.-П. Моно).

Даже самая маленькая эукариотическая клетка в 5—10 раз больше бактериальной и устроена значительно сложнее. В каждой эукариотической клетке транскрибируется только незначительная часть ДНК. В клетках различных тканей транскрипция затрагивает как общие для всех видов клеток данного организма гены, так и специфичные для данной ткани.

При клеточной дифференцировке, происходящей в процессе эмбрионального развития, транскрипция различных генов претерпевает последовательные изменения как качественного, так и количественного характера. Каждая стадия дифференциации включает в себя активацию очень большого числа структурных генов. Образование индивидуальных тканей связано с синтезом мРНК, которые кодируют белки, характерные для данной ткани. Несмотря на то, что во всех тканях одного и того же организма имеется полный набор хромосом и генов, в одних видах клеток наблюдается транскрипция тех генов, которые не транскрибируются в других. Это означает, что и в процессе дифференцировки и функционирования клеток должны существовать способы контроля транскрипции, необходимые для активации или репрессии определенных генов. Существует несколько принципиальных различий в условиях транскрипции у про- и эукариот: количество ДНК у эукариот в расчете на клетку в несколько тысяч раз больше, чем у прокариот, и если у бактерии существует одна хромосома, то у эукариотических клеток гены распределены между разными хромосомами. Кроме того, в эукариотах транскрибируется хроматин, расположенный в ядре, а синтезированная информационная РНК транспортируется в цитоплазму, тогда как у бактерий ядра нет и синтеза РНК и белка не разделены в пространстве.

Сложность регуляции транскрипции в эукариотических клетках проявляется в том, что в них различные типы РНК синтезируются с помощью различных РНК-полимераз. В эукариотической клетке обнаруживается четыре типа РНК-полимераз, три из которых — РНК-полимеразы I, II и III — локализованы в ядре и одна — в митохондриях. Клетки растений содержат еще одну полимеразу, локализованную в хлоропластах. В ядре РНК-полимераза I обнаруживается в ядрышках — структурных образованиях, где сосредоточены гены рибосомных РНК, тогда как полимеразы II и III — в нуклеоплазме. РНК-полимераза I отвечает за синтез рибосомных 18S, 28S и 5,8S РНК. Рибосомная 5S РНК и транспортные РНК синтезируются РНК-полимеразой III. РНК-полимераза II осуществляет синтез предшественников мРНК.

Для каждой из полимераз существуют свои способы контроля, которые осуществляются специфическими белками-регуляторами, взаимодействующими с полимеразами и определенными последовательностями ДНК. Кроме того, у эукариот появляется еще один новый тип контроля — контроль на уровне регуляции макроструктуры хроматина. При этом определенные участки хромосомы оказываются способными к активной транскрипции, тогда как транскрипция других запрещена.

Установлено, что структура хроматина в той области, где происходит транскрипция генов, отличается от структуры нетранскрибируемых участков. Структура хромосомы в транскрибируемой области меняется, на ней образуются утолщения (так называемые «пуффы») и т. п. Электронная микроскопия показывает, что активный хроматин значительно менее компактен, чем неактивный, и в ряде случаев даже теряет нуклеосомную структуру. При этом изменения в структуре хроматина предшествуют активации транскрип-

ции, а не наоборот. Высказывается предположение, что изменение структуры хроматина связано с действием негистоновых белков, модификацией гистонов или модификацией ДНК.

Обнаружены особые регуляторные элементы, названные энхансерами (от английского enhance — усиливать), которые резко увеличивают эффективность транскрипции эукариотических генов. Особенностью этих элементов является то, что они проявляют свою активность независимо от ориентации и положения относительно активируемого гена: они могут находиться перед геном, внутри гена или за ним.

В том случае, когда гены являются индуцируемыми, по-видимому, должен существовать механизм, включающий в себя взаимодействие регуляторных молекул со специфическими последовательностями ДНК. Примером индуцируемых генов являются гены, регулируемые стероидными гормонами. По-видимому, индукция происходит за счет связывания комплекса гормона и специфического белка-рецептора с регуляторной последовательностью ДНК, что приводит к резкой активации транскрипции.

Процессинг РНК

В результате транскрипции получается РНК, которая еще не готова к выполнению своих функций. Она содержит целый ряд «лишних» последовательностей, которые должны быть удалены, и, кроме того, не содержит минорных компонентов.

Под термином «процессинг» понимают совокупность ферментативных процессов, в результате которых синтезируемая в процессе транскрипции РНК превращается в функционально полноценную молекулу.

В наиболее изученной из прокариотических клеток — кишечной палочке *E.coli* не известны примеры процессинга мРНК. Эти молекулы синтезируются сразу в «зрелом» виде. В то же время все стабильные РНК *E.coli* (т. е. рибосомные и транспортные РНК) синтезируются в виде предшественников, которые затем превращаются в зрелые молекулы. В этот процесс вовлечена серия ферментов, в том числе эндо- и экзонуклеазы.

В отличие от прокариот, в клетках эукариот все РНК, в том числе и информационные, образуются в результате процессинга. Предшественники эукариотических мРНК называются гетерогенными ядерными РНК (гяРНК).

Основные стадии процессинга гяРНК изображены на рисунке 237. Сразу после начала транскрипции предшественник мРНК присоединяет так называемый «кэп» к 5'-концевому звену, в результате чего образуется сочленение $m^7G^5'ppp^5'Nm$, в котором концевой 7-метилгуанозин связан 5', 5'-пирофосфатной связью с метилированным 5'-концевым нуклеотидом РНК. После транскрипции в ядре к 3'-концу образовавшейся гетерогенной ядерной РНК добавляется сегмент поли(А). Практически у всех исследованных мРНК, содержащих поли(А), на расстоянии 10—25 нуклеотидов от места ее присоединения обнаруживается консервативная последовательность AAUAAA, которой приписывается роль сигнала присоединения поли(А).

Присоединение поли(А) обычно предшествует так называемому «сплайсингу» гяРНК. Термин «сплайсинг» берет свое начало от английского слова splice, которое в морской практике означает

Процессы с участием
нуклеиновых кислот:
репликация, транскрипция
и трансляция



Георгиев Георгий Павлович (р. 1933), советский биохимик, академик АН СССР (1987). Окончил 1-й Московский медицинский институт (1956), с 1963 г. работает в Институте молекулярной биологии АН СССР. Основные работы посвящены изучению механизма реализации генетической информации. Исследовал ядерные рибонуклеопротеидные частицы, содержащие про-мРНК (ядерные информосомы). Изучил в геноме животных подвижные, рассеянные по хромосомам генетические элементы. Лауреат Ленинской премии (1976) и Государственной (1983) премии СССР.

«сращивать концы тросов». Этот процесс обусловлен сложным строением большинства эукариотических генов, в которых кодирующая последовательность не является непрерывной, как в мРНК, а разбита вставками некодирующих последовательностей, называемых интронами, как это показано на рисунке 237. Таким образом, в обычном гене эукариотического организма кодирующие последовательности (экзоны) перемежаются некодирующими последовательностями (интронами). Первичный продукт транскрипции — гяРНК содержит и экзоны, и интроны. Для образования мРНК из нее необходимо убрать интроны и соединить экзоны в непрерывную кодирующую последовательность. Это и происходит в процессе сплайсинга.

Метилирование мРНК также является одной из ступеней процессинга. В результате этой модификации мРНК содержит в среднем приблизительно один остаток 6-метиладенина (m^6A) на каждые 400 остатков А.

Претерпевшая процессинг мРНК в виде нуклеопротеидного комплекса покидает ядро через поры в ядерной мембране и поступает в цитоплазму для трансляции. В эукариотической клетке вся РНК, за исключением транспортной, содержится в виде нуклео-

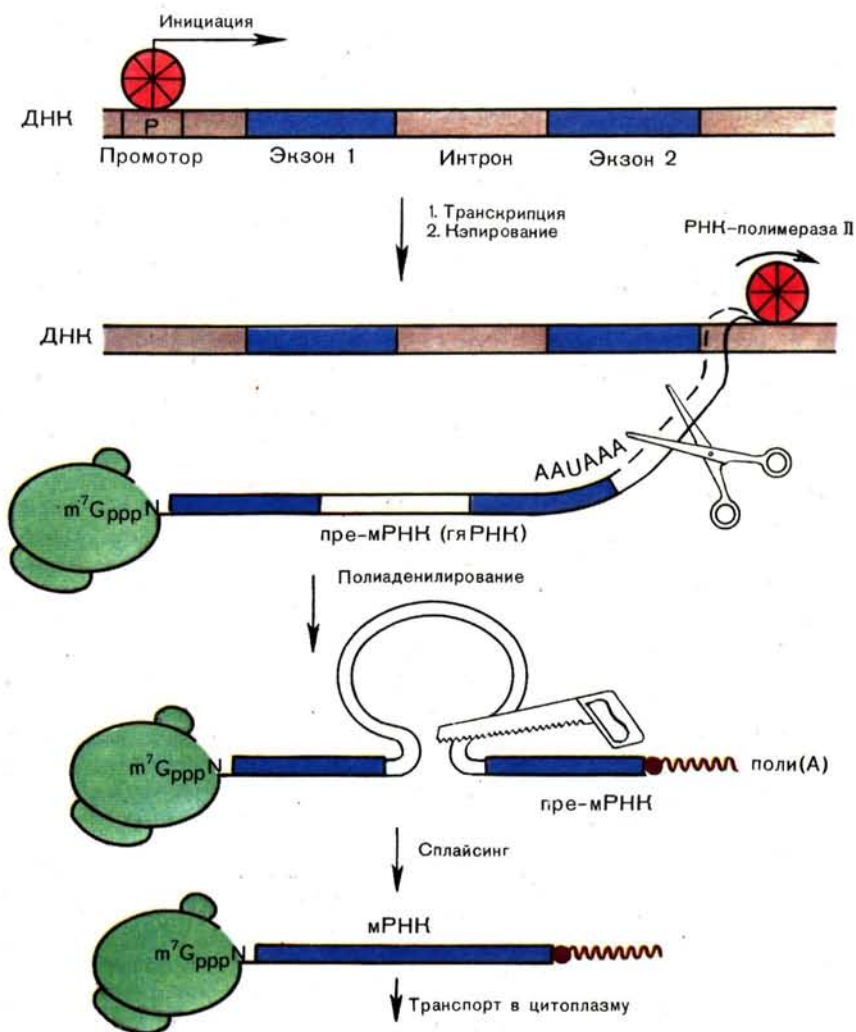


Рис. 237. Схема процессинга пре-мРНК (гяРНК) в ядре.

протеидных комплексов. Так, зрелая мРНК в ядре связана с определенным набором белков, называемых информосомами (Г. П. Георгиев), и в комплексе с ними переносится в цитоплазму, где формируются другие мРНК-содержащие рибонуклеопротеидные частицы — информосомы (А. С. Спиринов, Л. П. Овчинников).

В составе информосом мРНК защищена от клеточных нуклеаз и время жизни ее достаточно велико по сравнению с мРНК в прокариотических клетках.

Роль белков, связанных с мРНК, а также функция поли(А)-концов, как, впрочем, и «кэп»-структур на 5'-конце, еще далеко не ясна. Например, в ряде случаев наблюдается укорочение поли(А) в цитоплазме, а некоторые мРНК (например, мРНК гистонов) вообще лишены этой последовательности. Таким образом, к настоящему моменту основные пути экспрессии генов эукариот представляются только в самом общем виде и предстоит еще долгая кропотливая работа, чтобы выяснить детали механизма процесса.



Ниренберг (Nirenberg) Маршалл Уоррен (р. 1927), американский биохимик. Окончил университет в Майами (1948), с 1962 г. — заведующий лабораторией биохимической генетики Лингвского национального института сердца в Бетесде. Ему принадлежат основополагающие труды по расшифровке генетического кода. Синтезировал и испытал все 64 теоретически возможных тринуклеотида, укомплектовал кодовый словарь. Лауреат Нобелевской премии по физиологии и медицине (1968, совместно с Р. Холли и Г. Кораной).

Трансляция

Важным этапом на пути экспрессии гена является трансляция синтезированной мРНК. Трансляция — сложный многоступенчатый процесс синтеза полипептидной цепи согласно информации, заключенной в последовательности нуклеотидов мРНК. Трансляция осуществляется на рибосоме, в нее вовлечены также белковые факторы, ГТР и аминоацил-тРНК — молекулы тРНК, несущие активированные аминокислоты. Как и другие процессы матричного синтеза, процесс трансляции условно делят на три стадии: инициация, элонгация и терминация.

При инициации происходит специфическое связывание рибосомы с мРНК и с первой аминоацил-тРНК, называемой инициаторной, в результате чего образуется комплекс, способный к синтезу белка, — инициаторный комплекс. При элонгации осуществляется последовательное связывание аминоацил-тРНК с образованием пептидных связей по программе, задаваемой последовательностью кодонов в мРНК. Терминация представляет собой отщепление готовой белковой цепи от трансляционного комплекса. Все этапы синтеза белка осуществляются при участии специальных белковых факторов, которые называются соответственно факторами инициации IF, элонгации — EF и терминации RF. Энергия для процесса трансляции черпается при гидролизе ГТР.

В ходе трансляции нуклеотидная последовательность мРНК считывается в направлении от 5'-к 3'-концу. Считывание происходит по законам генетического кода, согласно которым каждой аминокислоте соответствует триплет нуклеотидов (кодон), каждый кодон кодирует только одну аминокислоту, а последовательность кодонов в мРНК определяет аминокислотную последовательность синтезируемого белка. Функцию узнавания кодона осуществляет не сама аминокислота, а молекула тРНК, к которой она присоединена и которая содержит последовательность нуклеотидов, комплементарную кодону, — антикодон. Иными словами, тРНК служит адаптором, обеспечивающим соответствие между кодоном и кодируемой им аминокислотой. Наличие адаптора — основное отличие трансляции, как процесса матричного синтеза, от транскрипции и репликации, в которых продукт имеет ту же химическую природу, что и



Очоа [Очоа] Северо (р. 1905), американский биохимик испанского происхождения. Образование получил в университетах Малаги и Мадрида, в настоящее время руководит Центром молекулярной биологии Мадридского университета. Автор фундаментальных работ по биохимии нуклеиновых кислот, изучению ферментативных превращений углеводов и жиров. Впервые осуществил (1955) ферментативный синтез РНК. Внес большой вклад в расшифровку генетического кода. Лауреат Нобелевской премии по физиологии и медицине (1959, совместно с А. Корнбергом).

матрица. Адапторная гипотеза биосинтеза белка была сформулирована Ф. Криком в 1958 г.

Генетический код един для всех организмов. Он содержит 64 кодона — число возможных сочетаний из четырех нуклеотидов по три (рис. 238). Три кодона — UAG, UGA и UAA — не кодируют аминокислоты, а являются сигналами окончания белкового синтеза. Они называются бессмысленными (nonsense) или стоп-кодонами. Остальные триплеты (61) — это смысловые кодоны, которые соответствуют 20 различным аминокислотам. История расшифровки генетического кода представляет собой цепь догадок, гипотез и открытий, связанных с именами Ф. Крика, С. Бреннера, М. Ниренберга, С. Очоа, Г. Кораны. Так как число триплетов превышает число аминокислот, генетический код является вырожденным. Большинство аминокислот кодируется несколькими кодонами. Каким образом реализуется такая ситуация? Для этого существует две возможности. Во-первых, одна молекула тРНК, специфичная к определенной аминокислоте, может взаимодействовать более чем с одним кодоном. Оказалось, что в этом случае точное образование комплементарных пар обязательно только для первых двух букв кодона, а в третьем положении допускается образование «неклассических» пар. Идея о том, что на спаривание третьего основания кодона накладываются менее жесткие ограничения, принадлежит Ф. Крику и была сформулирована в его гипотезе «качания» (wobble-гипотезе, или гипотезе неоднозначного соответствия). Согласно этой гипотезе, сейчас уже доказанной, допускается образование пар между U и G, с I (инозин) и U, C или A. Действительно, в тРНК, взаимодействующих более чем с одним кодоном, первым основанием антикодона (взаимодействующим с третьим основанием кодона) часто оказывается минорное основание инозин.

В клетке реализуется и другая ситуация, когда для одной аминокислоты существует несколько специфичных тРНК. Такие тРНК называются изоакцепторными.

	U	C	A	G	
U	UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA } Stop UAG }	UGU } Cys UGC } UGA } Stop UGG } Trp	U C A G
C	CUU } Leu CUC } CUA } CUG }	CCU } Pro CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }	U C A G
A	AUU } Ile AUC } AUA } Met AUG }	ACU } Thr ACC } ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U C A G
G	GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } Gly GGC } GGA } GGG }	U C A G

Рис. 238. Генетический код.

Активация аминокислот: тРНК и аминоацил-тРНК-синтетазы

Присоединение аминокислот к тРНК осуществляют специальные ферменты — аминоацил-тРНК-синтетазы (АРСазы). Для каждой аминокислоты существует один фермент, который узнает все изоакцепторы тРНК, способные присоединять эту аминокислоту; однако синтетазы одинаковой специфичности, но происходящие из разных источников, отличаются друг от друга. Реакция аминоацилирования тРНК протекает в две стадии, катализируемые одной и той же синтетазой (рис. 239).

Процессы с участием
нуклеиновых кислот:
репликация, транскрипция
и трансляция

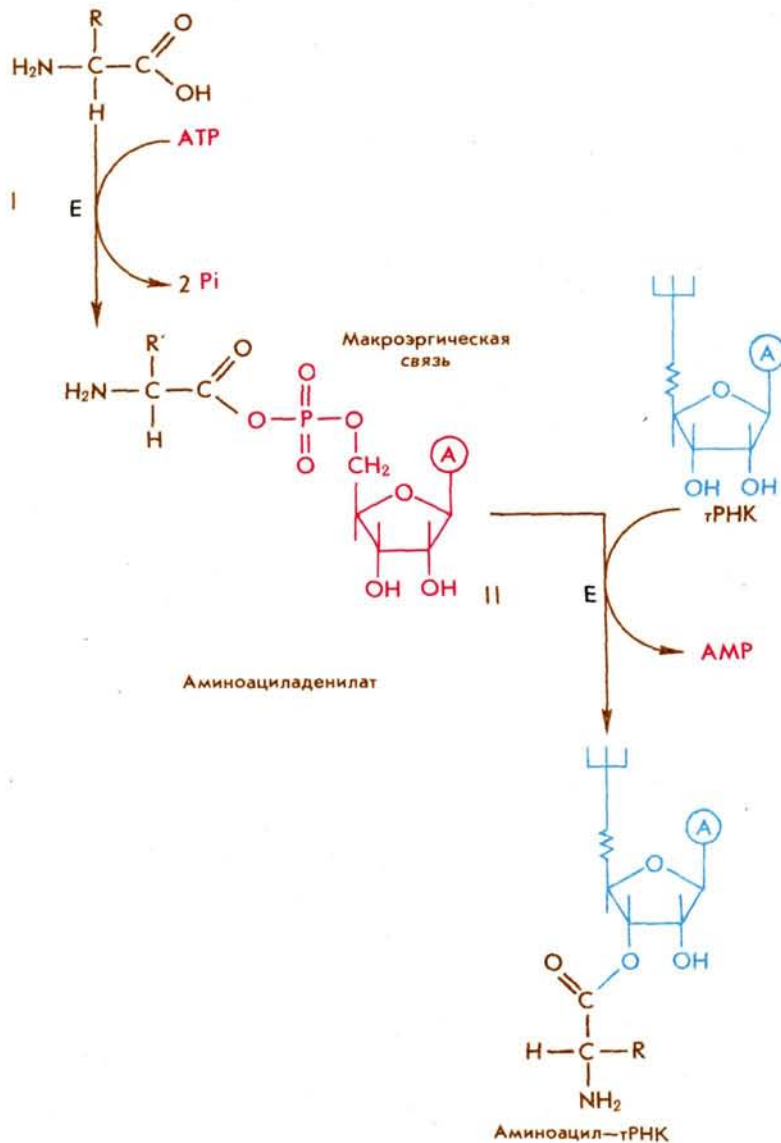


Рис. 239. Реакция аминоацилирования тРНК, катализируемая аминоацил-тРНК-синтетазой (E).

Как в прокариотических, так и в эукариотических организмах синтез белка начинается с присоединения особой (инициаторной) тРНК, которая всегда несет остаток метионина. По ряду особенностей она отличается от тРНК, несущей ту же аминокислоту для продолжения синтеза белковой цепи. Для инициаторной тРНК существует своя синтетаза. В прокариотических клетках аминогруппа метионина, связанного с инициаторной тРНК, модифицируется с помощью специального фермента. Заряженная и формилированная инициаторная тРНК обозначается как $fMet\text{-}tRNA^{Met}$; обычная метионил-тРНК — как $Met\text{-}tRNA^{Met}$. У эукариот также существует особая инициаторная тРНК, несущая метионин, но формилирование отсутствует.

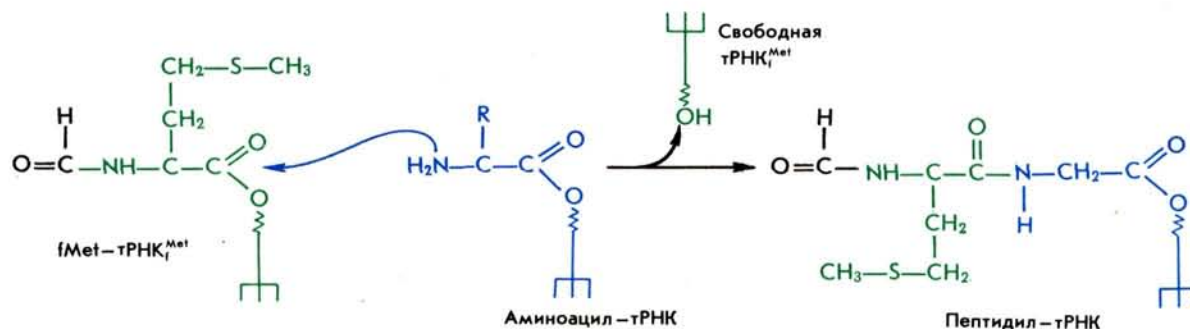


Рис. 240. Синтез первой пептидной связи.

В процессе белкового синтеза происходит последовательное присоединение аминокислот: сначала карбоксильная группа остатка формилметионина атакуется аминогруппой следующей аминокислоты с образованием пептидной связи (рис. 240), а далее процесс повторяется многократно до того момента, пока в мРНК не встретится терминирующий кодон.

Все эти события происходят на рибосоме и осуществляются с помощью факторов трансляции.

Белковые факторы, участвующие в трансляции

Помимо белков, входящих в состав рибосом, в биосинтезе белка принимает участие большое число белковых факторов, которые связываются с рибосомой только на определенных этапах трансляции и, выполнив свою функцию, отделяются. В *E. coli* в инициацию трансляции вовлечено три фактора — IF-1, IF-2 и IF-3 (IF — от initiation factor), которые представляют собой мономерные белки и обычно ассоциированы с 30S субъединицей рибосом. Факторы инициации катализируют реакции связывания мРНК (IF3) и формилметионил-тРНК (IF2) с рибосомой. IF1 стимулирует обе реакции, но играет подчиненную роль и не проявляет активности в отсутствие IF2 и IF3. У эукариот в инициацию вовлечено по мень-

шей мере 8 факторов (обозначаются EIF — eukaryote initiation factors), причем аналоги IF2 и IF3 имеют субъединичное строение. Так, EIF2, ответственный за связывание инициаторной тРНК, состоит из 9 субъединиц. Участие большого числа белковых факторов в эукариотической инициации, возможно, объясняется сложной системой контроля и регуляции этого процесса.

В *E.coli* на стадии элонгации работают три фактора — EF-Tu, EF-Ts и EF-G (EF — от elongation factor). Фактор Tu (u — от unstable, нестабильный) осуществляет транспортную функцию и переносит аминоксил-тРНК на рибосому, фактор Ts (s — от stable, стабильный) регенерирует активную форму Tu, образуя с ним комплекс Tu-Ts; EF-G участвует в транслокации, т. е. перемещении рибосом от кодона к кодону мРНК.

Такие же факторы существуют и у эукариот. Например, факторы EF-1 и EF-2 аналогичны по функции EF-Tu и EF-G соответственно.

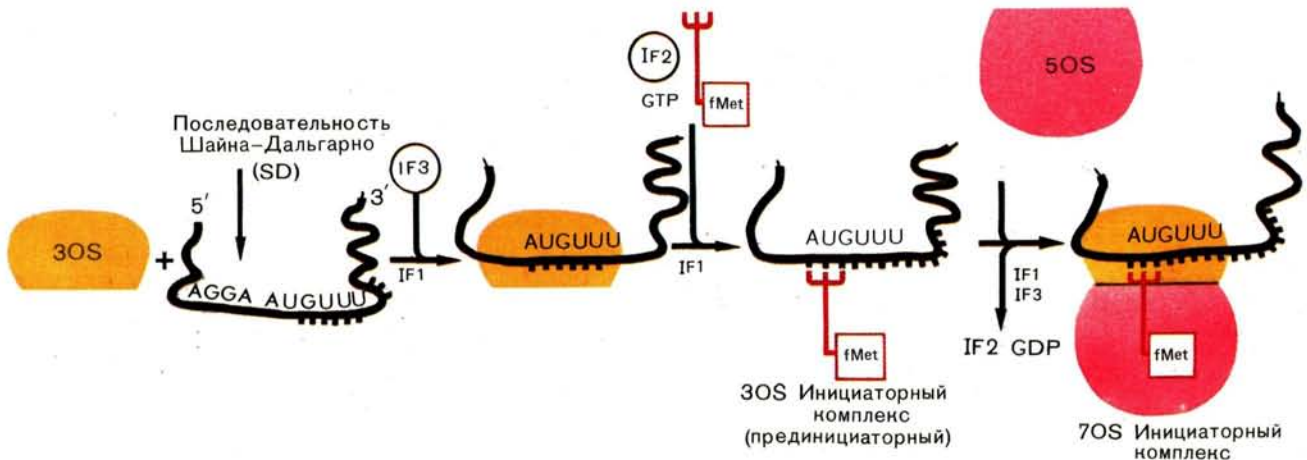
Наконец, в *E.coli* два белка катализируют на рибосоме гидролитическое отщепление полипептидной цепи от пептидил-тРНК. Эти белки проявляют специфичность по отношению к терминирующим кодоном: RF-1 осуществляет катализ, если терминирующими кодонами являются UAA и UAG, а RF-2 узнает кодоны UAA и UGA. Есть еще один фактор терминации — RF-3, который не имеет собственной каталитической активности, но стимулирует действие факторов RF-1 и RF-2. В клетках млекопитающих известен единственный RF-фактор, который узнает все три терминирующих кодона.

Процессы с участием
нуклеиновых кислот:
репликация, транскрипция
и трансляция

Инициация трансляции

Инициация трансляции в про- и эукариотических клетках имеет много общих черт. При инициации рибосомы диссоциируют на субъединицы и малая субъединица связывается с мРНК и инициаторной тРНК под действием IF-факторов. К образовавшемуся прединициаторному комплексу присоединяется большая субъединица, IF-факторы отщепляются и образуется зрелый 70S инициаторный комплекс, готовый к синтезу белка (рис. 241). Однако прежде

Рис. 241. Образование инициаторного комплекса у прокариот.



чем образовать преинициаторный комплекс, малой субъединице рибосомы приходится определить место начала синтеза белка, т. е. найти на мРНК иницирующий кодон (в общем случае это AUG, однако у прокариот иногда встречается GUG или даже AUU, причем все они связывают инициаторную формилметионил-тРНК). В структуре любой мРНК последовательность AUG встречается многократно, тем не менее субчастица узнает именно тот кодон, который соответствует первой аминокислоте синтезируемого белка. У прокариот и эукариот это узнавание происходит по-разному, но в обоих случаях информацию о начале синтеза рибосома получает из структуры мРНК.

Прокариотические матрицы полицистронны и в большинстве случаев содержат в 5'-концевой области, а также в межцистронных областях нетранслируемые последовательности, которым принадлежит важная регуляторная роль. Существует гипотеза, выдвинутая в 1974 г. Дж. Шайном и Л. Дальгарно, согласно которой рибосома определяет место инициации с помощью образования комплекса между 3'-концевым участком рибосомной 16S РНК и участком мРНК, расположенным перед иницирующим кодоном. Действительно, анализ структуры большого количества генов показывает, что на расстоянии 3 — 12 нуклеотидов от иницирующего кодона в большинстве мРНК содержатся последовательности, комплементарные 3'-концу 16S РНК. Более того, в ряде случаев из инициаторного комплекса удается после нуклеазного гидролиза выделить дуплекс, содержащий 3'-концевой фрагмент 16S РНК и комплементарный ему участок мРНК (Дж. Стейц).

Известны мутации, которые уменьшают степень комплементарности между концом 16S РНК и областью, предшествующей иницирующему кодону. В таком случае скорость синтеза белка понижается. Есть и другие аргументы, которые подтверждают значение этой комплементарности для инициации трансляции. Последовательность в мРНК, предшествующую иницирующему кодону и комплементарную 3'-концу 16S РНК, называют последовательностью Шайна — Дальгарно, она обозначается по первым буквам фамилий авторов гипотезы SD. Однако для определения места и эффективности инициации недостаточно только сочетания последовательности Шайна — Дальгарно и иницирующего кодона. Несомненно, в этом процессе участвуют и другие элементы структуры мРНК. Существуют серьезные доводы в пользу того, что большую роль играет здесь вторичная структура мРНК.

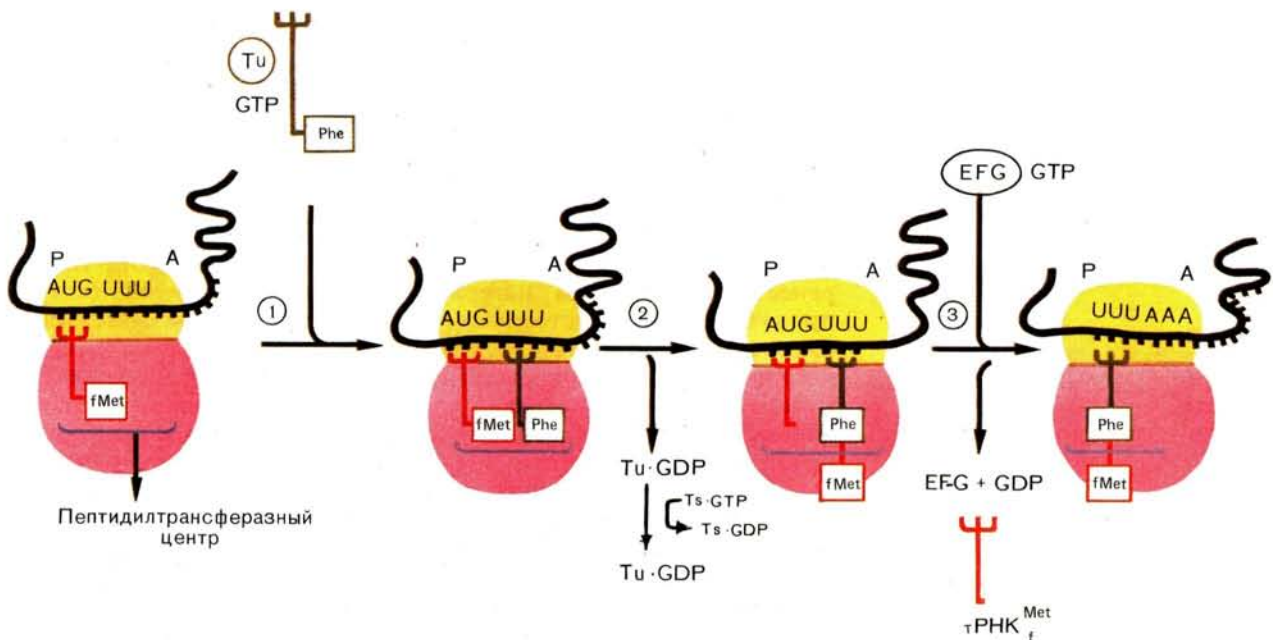
Последовательность, аналогичная Шайна — Дальгарно, как правило, отсутствует в эукариотических мРНК. Эукариотические матрицы моноцистронны и содержат на 5'-конце «кэп»-структуру, узнаваемую одним из факторов инициации эукариотической трансляции. Предполагается, что малая 40S субчастица рибосом связывается непосредственно на 5'-конце мРНК и как бы «едет» по ней в 3'-направлении до тех пор, пока не встретится первый AUG кодон, который в большинстве случаев и является иницирующим. Далее происходит присоединение 60S субчастицы, сборка полного инициаторного комплекса и начинается синтез полипептидной цепи. Эта так называемая «сканирующая модель» инициации трансляции у эукариот была предложена несколько лет назад М. Козак. Однако даже в этом случае иницирующий кодон, как было показано путем анализа известных 5'-концевых структур эукариотических матриц, находится в особом окружении, которое, по-видимому, и является сигналом остановки малой субчастицы и начала сборки полной транслирующей рибосомы.

Элонгация, в свою очередь, состоит из трех основных стадий: связывание аминоацил-тРНК, образование пептидной связи и транслокация (рис. 242). В рибосоме условно выделяют два функ-

циональных центра — А-центр, в который помещается каждая вновь вступающая в реакцию аминоацил-тРНК, и Р-центр, в котором располагается тРНК с присоединенной к ней растущей пептидной цепью (пептидил-тРНК). Фактор элонгации EF-Tu образует комплекс с GTP, который, в свою очередь, взаимодействует с аминоацил-тРНК с образованием тройного комплекса. Комплекс связывается с рибосомой, так что аминоацил-тРНК оказывается в А-центре, и ее антикодон образует комплементарный комплекс с кодоном мРНК. При этом происходит гидролиз GTP и комплекс Tu-GDP отделяется от рибосомы. В результате аминоацил-тРНК в А-центре располагается рядом с пептидил-тРНК, которая находится в Р-центре и взаимодействует с предыдущим кодоном. Фермент пептидилтрансфераза, являющийся структурной частью 50S субчастицы, переносит пептидную цепь на новую аминокислоту, в результате чего вновь вошедшая аминоацил-тРНК превращается в пептидил-тРНК, а прежняя пептидил-тРНК оказывается незаряженной (рис. 242).

После этого происходит транслокация: новая пептидил-тРНК перемещается в Р-центр, незаряженная тРНК отделяется от рибосомы, рибосома передвигается по мРНК на один триплет и в А-центре оказывается следующий кодон мРНК. Рибосома готова к приему следующей аминоацил-тРНК. Транслокация катализируется фактором EF-G.

Рис. 242. Стадия элонгации. Синтез первой пептидной связи: 1 — EF-Tu — зависимое связывание аминоацил-тРНК с А-участком рибосомы; 2 — пептидилтрансферазная реакция; 3 — EF-G — зависимая транслокация.



Терминация

Наконец, последняя стадия — терминация осуществляется, когда рибосома достигает терминирующего кодона — UAA, UAG или UGA, который оказывается в А-центре. В этой ситуации к рибосоме

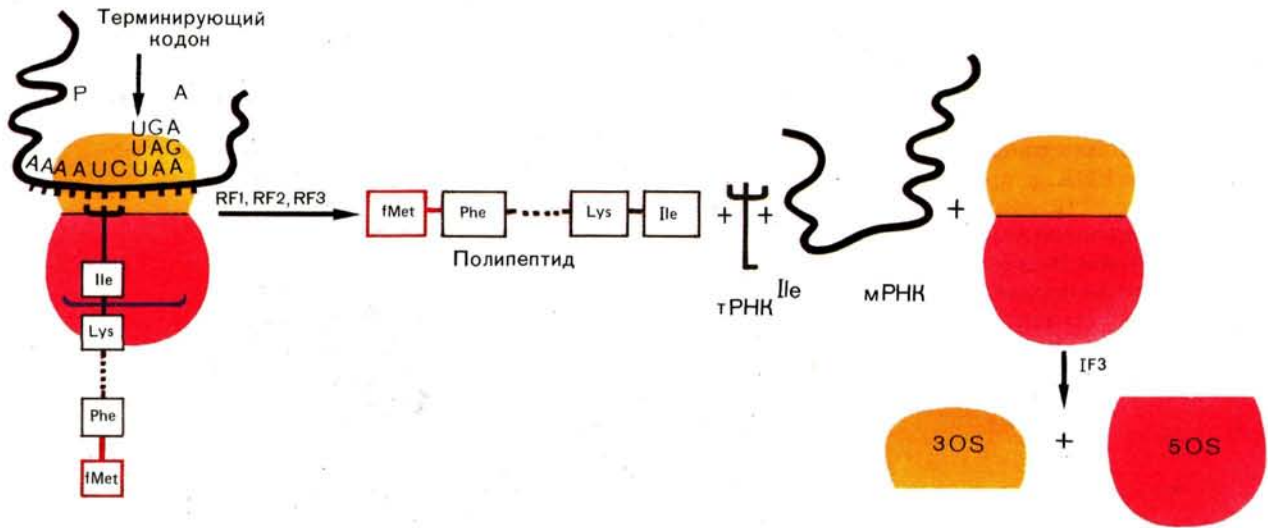


Рис. 243. Терминация белкового синтеза.

присоединяется один из RF-факторов, в результате чего происходит гидролиз сложноэфирной связи, соединяющей полипептид и тРНК. От пептидил-тРНК отщепляется готовая белковая цепь, незаряженная тРНК покидает рибосому, и процесс трансляции завершается (рис. 243).

Генная инженерия

Генная инженерия, или техника рекомбинантных ДНК, — это совокупность приемов, позволяющих путем операций *in vitro* перенести генетический материал из одного организма (который принято называть источником генов) в другой (называемый хозяином или реципиентом) таким образом, чтобы обеспечить наследование этих генов в новом для них организме. Перенос генов методами генной инженерии дает возможность преодолевать межвидовые барьеры и передавать отдельные наследственные признаки одних организмов другим (например, от человека или животного — бактерии и т. п.). В генной инженерии широко используются подходы и методы биоорганической химии.

В самом общем виде принцип генной инженерии изображен на рисунке 244. В какой-либо из заранее выделенных репликонов хозяина вводят *in vitro* нужный ген из другого источника. Полученную таким образом рекомбинантную молекулу, сохраняющую

свойства репликона, снова внедряют в клетку-хозяина, в которой он будет реплицироваться и стабильно передаваться клеткам-потомкам при делении.

Для реализации этой идеи требуется соответствующий репликон, который может легко выделяться, сохранять свои свойства и способен внедряться в клетку после введения в него чужеродной ДНК и, наконец, стабильно наследоваться. Репликоны, предназначенные для введения чужеродных ДНК в клетки, называются векторами.

В распоряжении исследователя должны быть способы соединения исходных генов с вектором и введения рекомбинантных репликонов в клетку-хозяина. Необходимо также уметь отличать клетки, содержащие рекомбинантный репликон, от исходных реципиентных клеток.

После введения нужного гена в реципиентную клетку задачи в зависимости от поставленной цели могут быть различными. Иногда оказывается достаточным просто внедрить чужеродный генетиче-

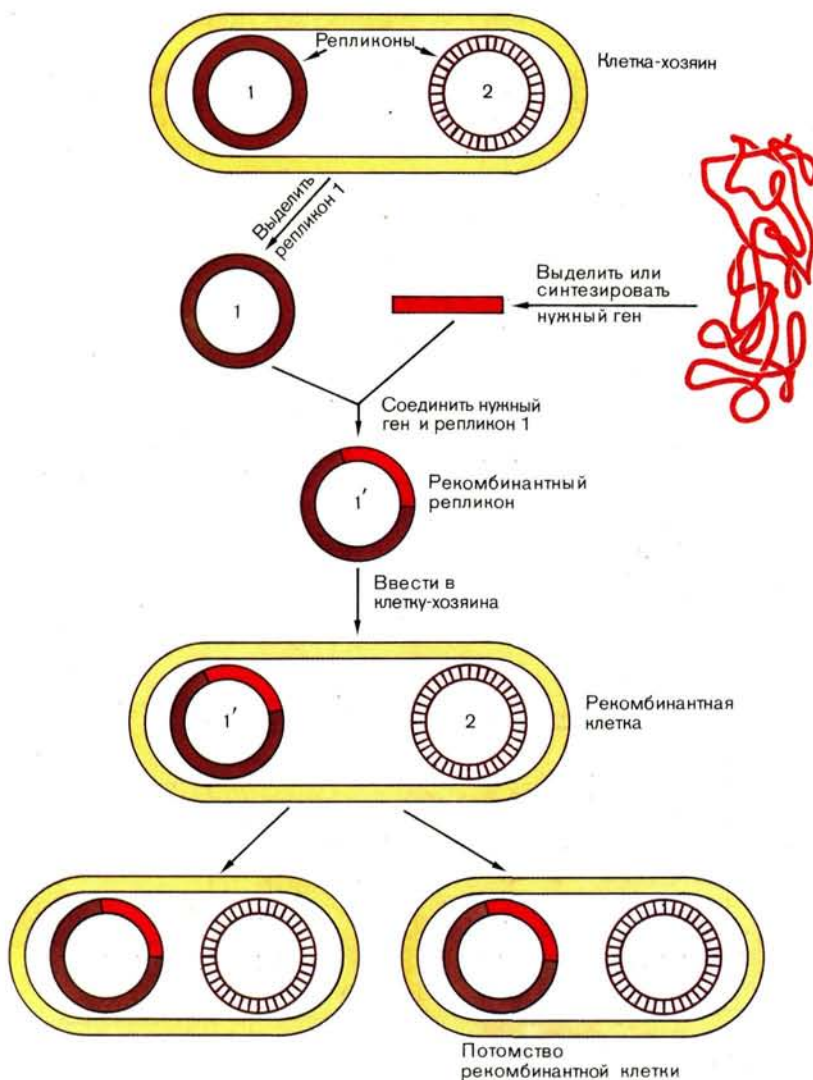


Рис. 244. Общий принцип генной инженерии.

ский материал, в более сложном случае требуется осуществить его экспрессию в новом хозяине.

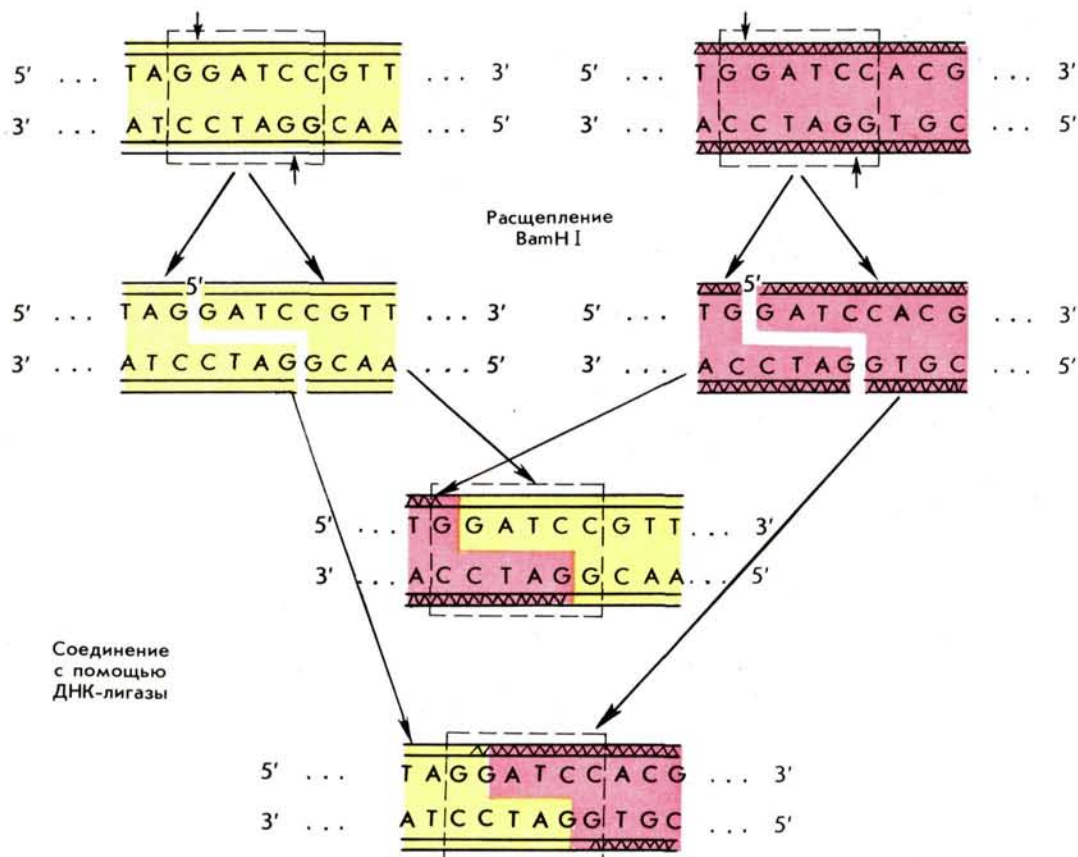
Методы решения всех этих задач целесообразно рассмотреть на примере наиболее хорошо изученного реципиента — кишечной палочки *E. coli*.

Способы соединения фрагментов ДНК, используемые в генной инженерии

Многие рестрикционные эндонуклеазы, расщепляя ДНК, дают *липкие концы* (см. с. 353) типа *EcoRI*, *HindIII*, *BamHI*, *SalI*, *PstI* и др. При расщеплении одной и той же рестриктазой различных ДНК образуются одинаковые липкие концы и полученные фрагменты соединяются друг с другом по этим концам ДНК-лигазой (рис. 245). Различные ДНК могут быть соединены при помощи ДНК-лигазы и по *тупым концам*.

Рис. 245. Расщепление молекул ДНК рестриктазой *BamHI* и соединение полученных фрагментов ДНК-лигазой.

Для обеспечения возможности расщепления рекомбинантных ДНК по местам соединения вектора и вставки используются также



специальные приемы, наиболее распространенным из которых является метод линкеров (см. с. 377).

Нередко применяется *коннекторная техника* получения рекомбинантных ДНК. Она заключается в том, что с помощью концевой нуклеотидилтрансферазы (см. с. 351) к 3'-концу одного фрагмента присоединяется гомополинуклеотид, например поли(dT) или

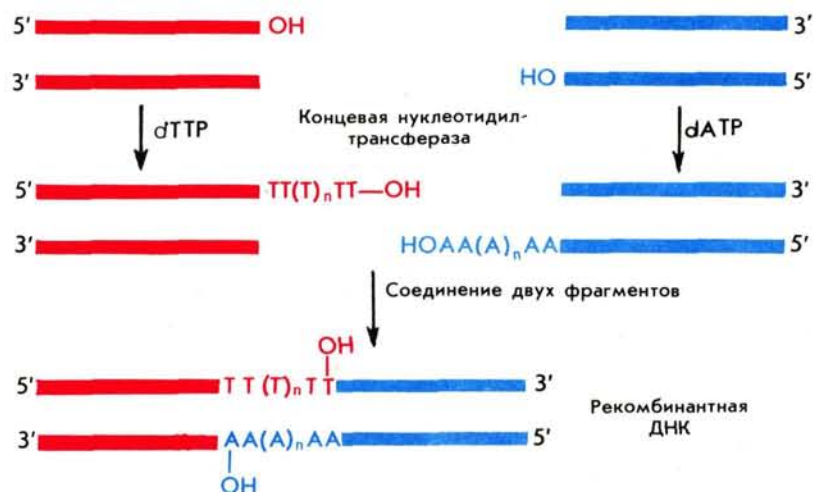


Рис. 246. Соединение двух фрагментов ДНК коннекторным методом.

Таблица 15.

Эндонуклеазы рестрикции, наиболее широко используемые в генной инженерии

Bam H I	G↓GATC C C CTAG↑G	AhaIII	TTT↓AAA AAA↑TTT
Bgl II	A↓GATC T T CTAG↑A	AluI	AG↓CT TC↑GA
EcoR I	G↓AATT C C TTAА↑G	HaeIII	GG↓CC CC↑GG
Hind III	A↓AGCT T T TCGA↑A	HincII	C A GTT↓GAC CAA↑CTG G T
Hpa II	C↓CG G G GC↑C	HindII	GTP _γ ↓PuAC CAP _u PyTG
Kpn I	G GTAC↓C C↑CATG G	HpaI	GTT↓AAC CAA↑TTG
Pst I	C TGCA↓G G↑ACGT C	PalI	GG↓CC CC↑GG
SaL I	G↓TCGA C C AGCT↑G	PvuII	CAG↓CTG GTC↑GAC
Tag I	T↓CG A A GC↑T	SmaI	CCC↓GGG GGG↑CCC

поли(dC), а к 3'-концу другого фрагмента — полинуклеотид, комплементарный используемому в первом случае, т. е. поли(dA) или поли(dG) (рис. 246). При добавлении фрагментов друг к другу они комплексуется за счет образования комплементарных пар оснований между концевыми гомополинуклеотидами, а затем соединяются в клетке-хозяине с помощью ее ферментативного аппарата.

Векторы, используемые в *E. coli* — плазмиды и бактериофаги

Молекула вектора должна отвечать определенным требованиям: быть небольшой, содержать уникальные участки расщепления рестрикционными эндонуклеазами, иметь маркеры для оптимальной селекции рекомбинантных клеток. В качестве векторов для включения чужеродной ДНК в клетки *E. coli* используются плазмиды (термин предложен Дж. Ледербергом) и бактериофаги. И те и другие являются репликаонами.

Многие плазмиды, представляющие собой двухцепочечные кольцевые молекулы ДНК, содержат гены, которые придают содержащим их бактериям некоторые фенотипические признаки, такие, как устойчивость к антибиотикам, солям тяжелых металлов и т. д.

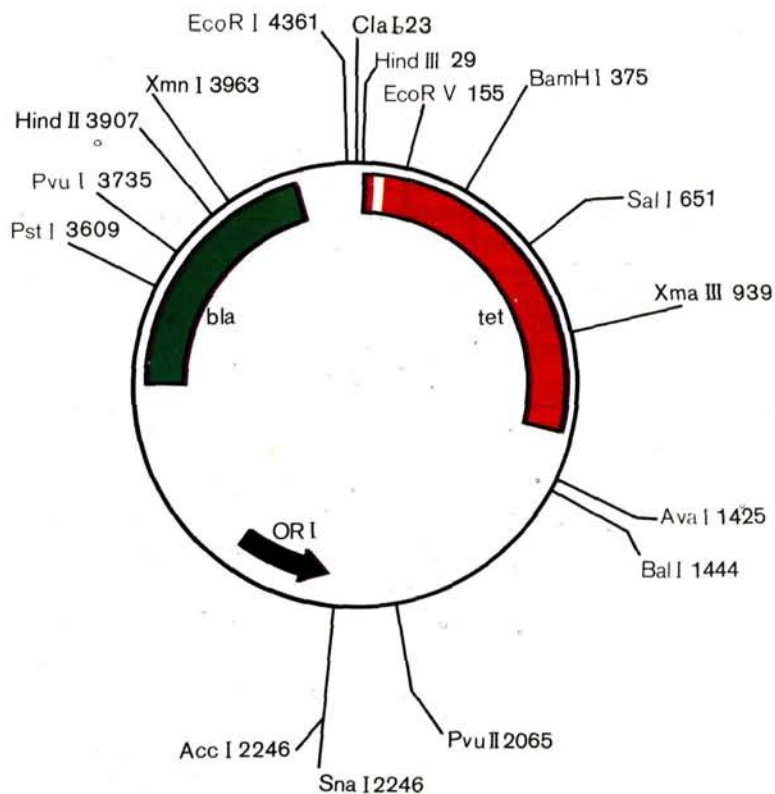


Рис. 247. Плазмида pBR322.

Если основная бактериальная ДНК имеет длину около $5 \cdot 10^6$ п. о., то размеры плазмид составляют всего несколько тысяч пар оснований. Они легко выделяются и очищаются.

Специально сконструированы векторные плазмиды, позволяющие отличать клетки, содержащие рекомбинантные молекулы, от исходных бактериальных клеток. В качестве примера на рисун-

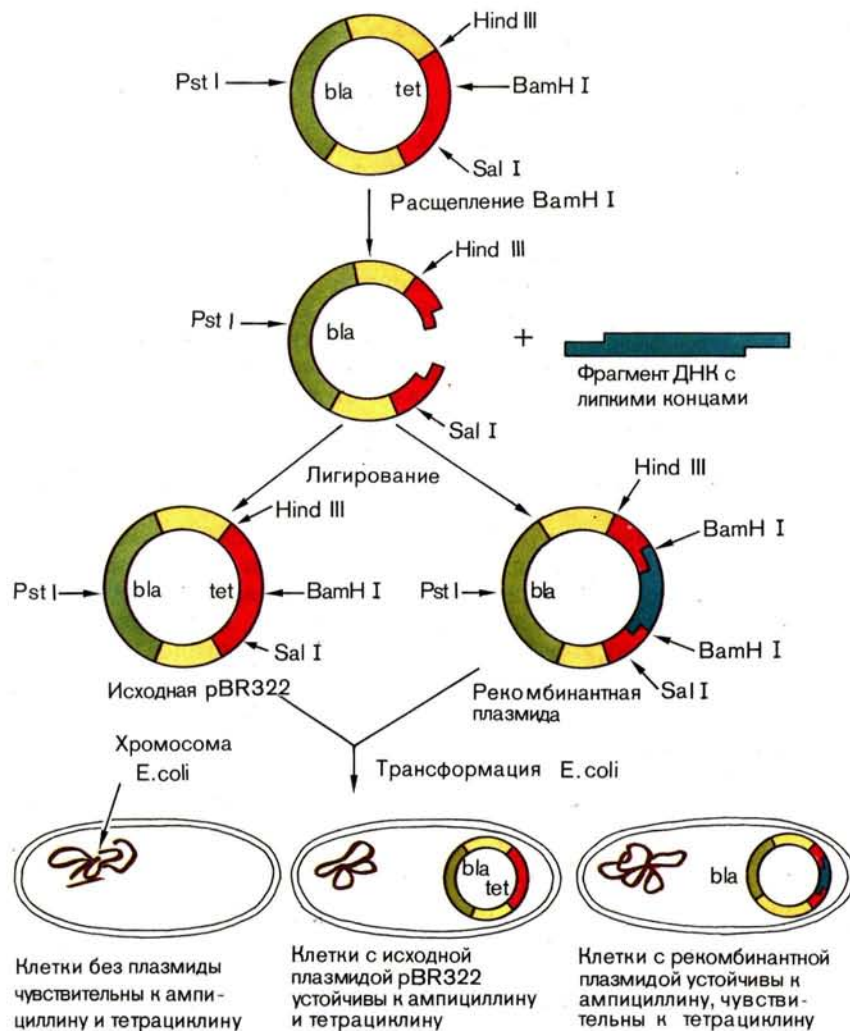


Рис. 248. Схема использования плазмиды pBR322 для отбора клеток, содержащих рекомбинантные плазмиды.

ке 247 приведена плазмиды pBR322, она содержит два гена, программирующие устойчивость к двум различным антибиотикам — тетрациклину (ген *tet*) и ампициллину (ген *bla*). В гене *tet* находятся уникальные участки расщепления рестриктазами HindIII, BamHI и SalI, а в гене *bla* — участок расщепления PstI. Если «разрезать» плазмиду любой из рестриктаз, участок расщепления которой находится в гене *tet*, и соединить ее методом липких концов с чужеродным фрагментом ДНК, то в полученной рекомбинантной молекуле останется нетронутым только ген *bla*, а ген *tet* утрачивает свою активность, поскольку его последовательность разрывается

вставкой. Наоборот, при разрезании плазмиды рестриктазой PstI и внедрении в этот участок фрагмента ДНК инактивируется ген bla, тогда как ген tet продолжает обеспечивать синтез белка, придающего *E. coli* устойчивость к тетрациклину. Использование такой методологии позволяет быстро идентифицировать клетки, содержащие рекомбинантные плазмиды. Применение подобного приема показано на рисунке 248. Расщепление исходной плазмиды pBR322 по сайту BamHI и последующее лигирование с фрагментом ДНК приводит к смеси исходных и рекомбинантных плазмид. При введении этой смеси в *E. coli* образуются клетки трех типов — не содержащие плазмиду, содержащие исходную pBR322 и клетки с рекомбинантной плазмидой. Их легко отличить по различной устойчивости к антибиотикам.

Анализируемую ДНК можно вводить и в другие репликоны, способные размножаться в клетках *E. coli*, например в бактериофаги. Из множества известных фагов чаще всего в качестве векторов используют сконструированные производные фага λ и фагов M13 и fd.

Большое разнообразие векторов существует на основе бактериофага λ , в них используется особенность фага, состоящая в том, что значительная часть его ДНК не нужна для размножения фага в клетке (рис. 249). Это позволяет вводить чужеродную ДНК в ДНК фага λ при использовании его в качестве вектора, причем длина вставляемого фрагмента может быть существенно больше величины фрагмента, встраиваемого в плазмиду.

Важную группу векторов, широко используемых при установлении первичной структуры ДНК, составляют нитевидные бактериофаги, такие, как M13, fd и f1. Фаг M13 представляет собой одноцепочечную циклическую ДНК длиной около 6500 нуклеотидов. После инфицирования бактериальной клетки одноцепочечная ДНК фага превращается в двухцепочечную репликативную форму (RF), которая во всех отношениях подобна плазмиде. Кроме того, фаговая ДНК содержит короткий участок (500 нуклеотидов), названный межгенной последовательностью (МП), несущественный для ее жизнеспособности (рис. 250).

Таким образом, выделив репликативную форму ДНК и расщепив ее в области несущественного участка, можно с помощью лигазы вставить в место разрыва чужеродную ДНК. Введение рекомбинантной двухцепочечной молекулы в клетку *E. coli* приводит к ее

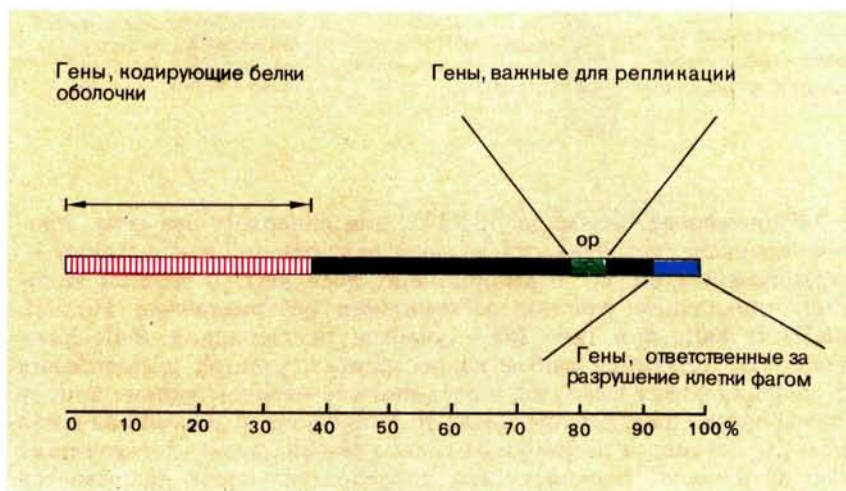


Рис. 249. Упрощенная генетическая карта фага λ (зачерченный участок соответствует генам, несущественным для размножения фага).

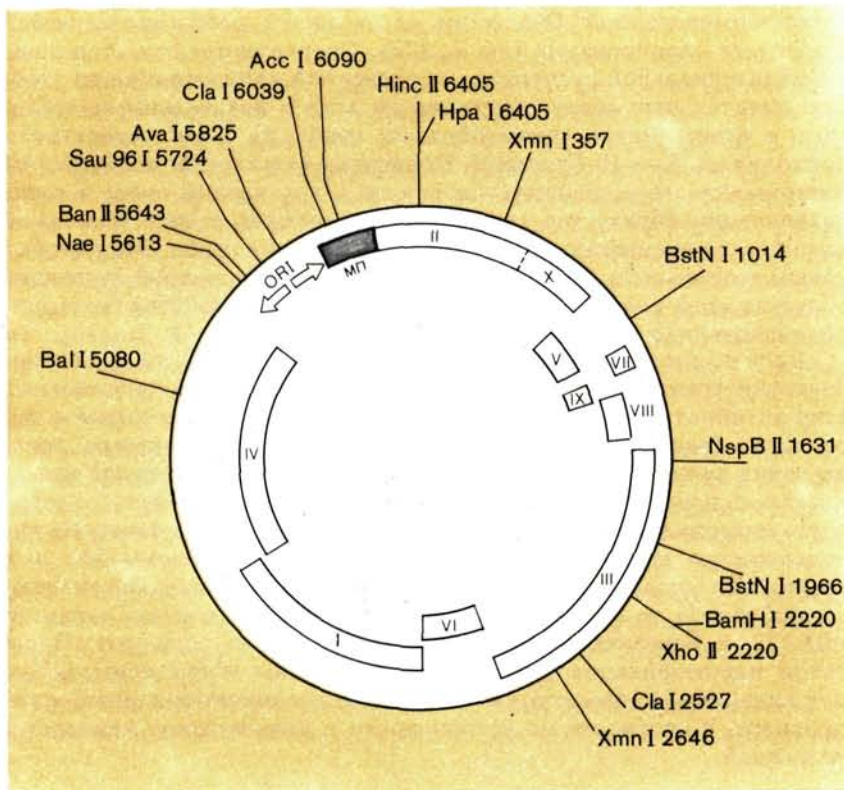


Рис. 250. Фаг M13 (темный участок МП соответствует генам, несущественным для его жизнеспособности).

репликации, синтезу (+)-цепи, упаковке последней в белковый чехол и выделению фага в среду. Далее инфицированная нитевидным фагом клетка, продолжая делиться, хотя и с замедленной скоростью, постепенно выделяет в окружающую среду большое количество фага. Этот фаг содержит в вирионе одноцепочечную циклическую ДНК, в которую встроена одна из цепей чужеродной ДНК.

Трансформация, трансфекция, клонирование и селекция

Процесс введения плазмиды в клетку, вызывающий наследственные изменения в ней, называется *трансформацией*. Процесс инфекции клеток с помощью чужеродных ДНК, приводящий к образованию зрелого фагового потомства, называется *трансфекцией*. Практически наиболее общий способ трансформации и трансфекции основан на том, что при обработке клеток бактерий CaCl_2 их мембрана становится проницаемой для ДНК. Эффективность проникновения экзогенной ДНК в клетку низка. Для плазмид типа pBR322 можно получить 10^6 — 10^7 трансформантов при добавлении 1 мкг плазмиды к обработанным CaCl_2 клеткам, т. е. из каждых 10^4 — 10^5 плазмид в клетки попадает только одна. Поэтому среди бактерий, подвергшихся трансформации, только небольшая часть оказывается



Рис. 251. Чашка Петри с колониями *E. coli*.

трансформированной. Отделение ее от общей массы выполняется в процессе клонирования (см. с. 372). Операция заключается в посеве бактериальной суспензии определенной концентрации на твердую питательную среду, например на агар с питательными добавками в чашке Петри таким образом, чтобы на 1 см^2 поверхности приходилось 5 — 10 бактерий. Бактериальная клетка, попавшая на поверхность агара, начинает делиться, и в конечном счете в точке локализации образуется семейство ее потомков в виде маленькой колонии, по внешнему виду похожей на шляпку гриба. Эта колония называется *клоном* (рис. 251). Каждая клетка исходной суспензии образует свой клон, все клетки которого имеют свойства бактерии-родоначальника (позитивная колония).

Если в качестве вектора служит плаزمида pBR322, то для отбора бактерий-трансформантов используют добавление в питательный агар антибиотика — ампициллина или тетрациклина, в зависимости от того, какой из генов, *bla* или *tet*, остался интактным после введения чужеродной ДНК. На такой среде клоны образуют только клетки с плазидами. Для отделения рекомбинантных бактерий часть материала каждого клона переносится на другую чашку Петри, содержащую антибиотик, ген устойчивости к которому был разрушен при создании рекомбинантов. На этой чашке Петри дают клоны только те бактерии, которые содержат исходную плазмиду pBR322, а рекомбинантные бактерии клоны не образуют. После такой идентификации рекомбинантные клоны могут применяться для дальнейшего анализа. Отбор по определенному биологическому признаку, в частности по устойчивости к антибиотику, называется *селекцией*.

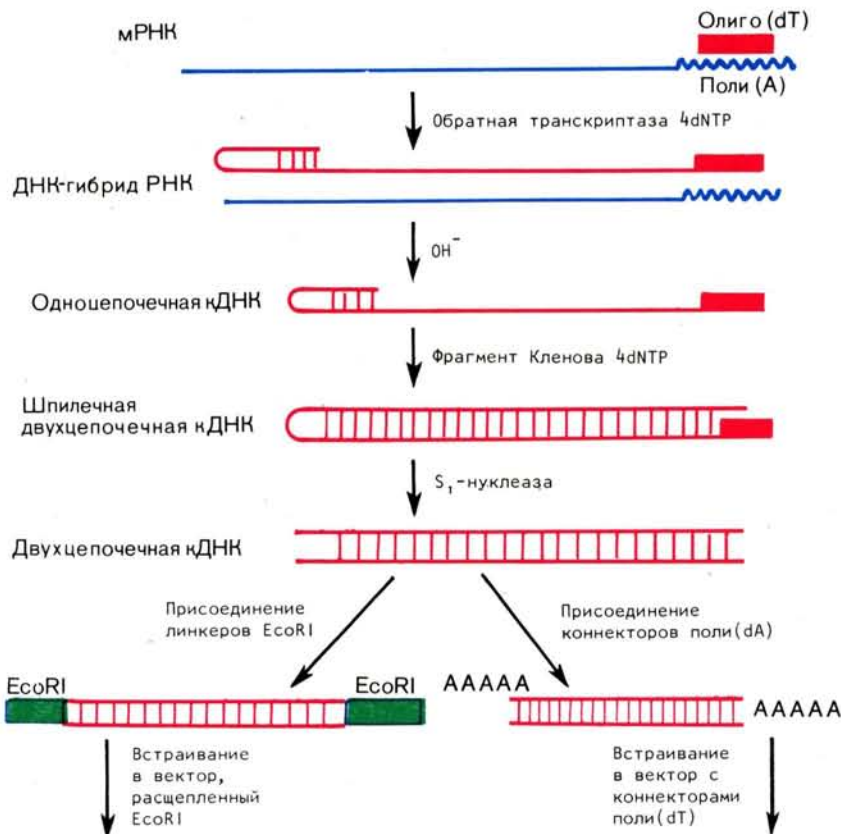


Рис. 252. Схема подготовки кДНК для клонирования.

Клонирование при использовании в качестве векторов бактериофага λ осуществляется следующим образом: к суспензии бактерий, обработанных CaCl_2 , добавляется ДНК фага и осуществляется посев на чашку Петри с питательным агаром так, чтобы на 1 см^2 поверхности попадало 5 — 10 клеток с проникшей в них ДНК фага. Клетки, не содержащие ДНК фага, размножаются на агаре и дают мутный ровный «газон», заполняющий всю поверхность чашки. Однако в тех точках, куда попали бактерии с ДНК фага, остаются прозрачные круглые просветы, называемые бляшками или негативными колониями. Образование бляшек связано с тем, что ДНК фага реплицируется в клетке и дает зрелые фаговые частицы. Затем фаговые частицы попадают в среду, заражают окружающие клетки и разрушают их. В результате каждая бляшка состоит из потомков фагов, содержащих одну и ту же ДНК — копию ДНК, попавшей в клетку.

Фаги типа M13 не дают негативных колоний, поскольку они не разрушают клетки. Однако в месте их размножения клетки *E. coli* замедляют деление, что приводит к появлению на общем мутном газоне более прозрачных бляшек. Способы селекции основаны либо на том, что в несущественную область фага (МП) вводят ген, программирующий устойчивость к антибиотикам и разрушающийся при внедрении в него чужеродной ДНК, либо на различии цветных реакций, даваемых клетками, зараженными исходным и рекомбинантным фагами.

Получение ДНК для клонирования. ДНК для клонирования может быть получена химико-ферментативным синтезом, обратной

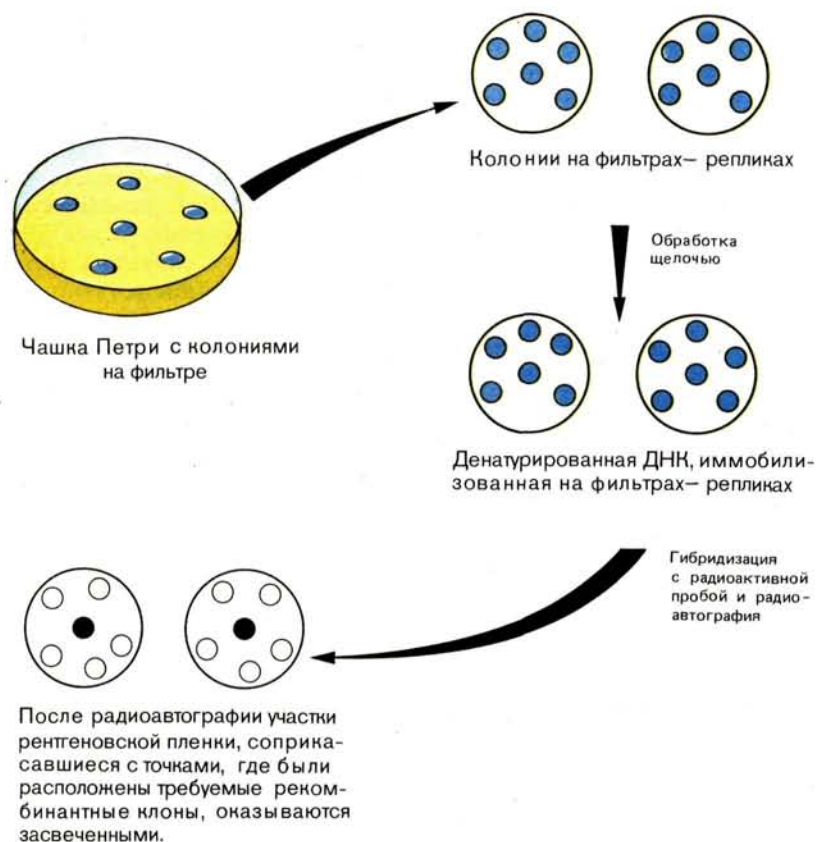


Рис. 253. Метод радиоавтографии в применении к поиску рекомбинантных клонов.

транскрипцией мРНК или путем непосредственного расщепления геномной ДНК нужной рестрикционной эндонуклеазой.

При обратной транскрипции эукариотической мРНК чаще всего используется тот факт, что на ее 3'-конце обычно содержится поли (А)-последовательность, благодаря которой в качестве затравки для обратной транскрипции можно применять олиго (dT). На первой стадии обратная транскриптаза синтезирует одноцепочечную ДНК, комплементарную мРНК. Эта ДНК обычно содержит на 3'-конце «шпильку». После удаления РНК обработкой щелочью или РНКазами образуемая одноцепочечная ДНК служит затравкой-матрицей для фрагмента Кленова ДНК-полимеразы I, который при наличии четырех dNTP достраивает вторую цепь. В результате образуется шпильчатая структура, превращаемая в истинную двухцепочечную кДНК обработкой S₁-нуклеазой (рис. 252).

Далее полученную кДНК можно встроить в соответствующий вектор с помощью линкеров или коннекторной техники.

Идентификация клонов. Если вставка содержит гены, способные к экспрессии в новом хозяине, рекомбинантные клоны могут быть идентифицированы по синтезируемому ими продукту. Однако чаще приходится идентифицировать непосредственно нуклеотидную вставку, для чего используют методы гибридизации. Бактериальные (или фаговые) колонии выращиваются на нитроцеллюлозных фильтрах, помещенных на чашку Петри с питательной средой (рис. 253). После этого приготавливаются так называемые реплики — к фильтру с исходными колониями прижимается свежий нитроцеллюлозный фильтр, который затем переносится на чашку Петри с плотной питательной средой, где на нем образуются колонии, идентичные первым.

Далее фильтр-реплику подвергают щелочной обработке, клетки в колониях подвергаются лизированию, и денатурированная ДНК из клеток связывается с нитроцеллюлозой в том месте, где была расположена соответствующая колония. Если в распоряжении исследователя имеется радиоактивная (³²P- или ¹²⁵I-меченная) ДНК или РНК, комплементарная требуемой вставке, то при выдерживании фильтра в растворе, содержащем радиоактивный полинуклеотид, последний гибридизуется с комплементарными последовательностями. В результате те части фильтра, в которых находились рекомбинантные клоны с требуемой вставкой, оказываются радиоактивными и идентифицируются радиоавтографически.

Радиоактивные «пробы» получают химическим синтезом (в тех случаях, когда известна последовательность искомой вставки или белка, который она кодирует), выделением индивидуальных или сильно обогащенных мРНК с последующим их иодированием ¹²⁵I, а также ферментативным синтезом соответствующих кДНК с использованием ³²P-меченых дезоксирибонуклеозидтрифосфатов.

Проблемы экспрессии чужеродных генов

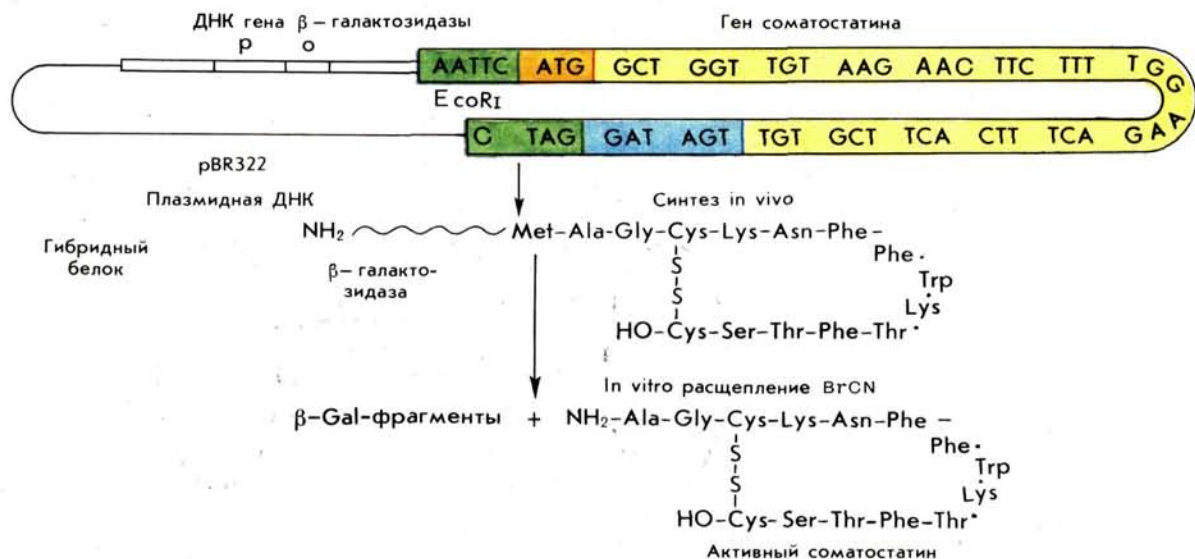
При включении бактериальных генов вместе с их регуляторными участками в *E. coli* они, как правило, экспрессируются, давая мРНК и белок. Это происходит потому, что в сигнальных последовательностях, управляющих транскрипцией и трансляцией в различных прокариотических организмах, много общего. Однако экспрессия генов эукариот в бактериях наблюдается очень редко, если не создавать специальные условия. Регуляторные сигналы эукариот сильно отличаются от регуляторных сигналов бактерий

и не узнаются бактериальными РНК-полимеразами и рибосомами. При клонировании не кДНК, а геномной ДНК эукариотической клетки экспрессия не происходит, поскольку в бактериальной клетке отсутствует система «сплайсинга».

Поэтому для осуществления экспрессии эукариотического гена соответствующая кДНК или синтетическая ДНК, содержащая кодирующую последовательность, присоединяется в составе векторной молекулы к регуляторным элементам бактерии — промотору и рибосом-связывающему участку.

Иногда присоединение ведется таким образом, что кодирующая часть эукариотического гена присоединяется к кодирующей части бактериального гена так, чтобы сохранялась «рамка» считывания. В последнем случае образуются гибридные белки, содержащие в N-концевой части бактериальный белок или его часть, а в С-концевой части — эукариотический белок. Примером использования такого способа экспрессии является получение гормона соматостатина в клетках *E. coli* (К. Итакура, Г. Бойер) (см. с. 267). Аминокислотная последовательность гормона известна, и исходя из нее, согласно генетическому коду, была выведена структура соответствующего искусственного гена. Осуществлен его синтез: ген содержал на N-конце кодон АТГ, кодирующий метионин, и липкий конец, соответствующий расщеплению рестриктазой *EcoRI* ААТТ, а на С-конце — два стоп-кодона (рис. 254). Этот фрагмент ДНК был присоединен к фрагменту гена β -галактозидазы *E. coli*, содержащему в С-концевой части участок расщепления *EcoRI*, и вместе со своим промотором и оператором был встроен в плазмиду рВR322. В результате такого соединения получен гибридный ген, в котором С-концевая часть гена β -галактозидазы заменена «геном» соматостатина. Между геном β -галактозидазы и собственно геном соматостатина находится кодон метионина. Транскрипция и трансляция этого гибрида осуществлялась за счет использования соответствующих сигнальных последовательностей β -галактозидазы. В результате образовался химерный белок, в котором соматостатин отделен от β -галактозидазной части остатком метионина. Поскольку соматостатин не содержит метиониновых остатков, то при расщеплении химерного белка бромцианом соматостатин был отделен

Рис. 254. Получение соматостатина через гибридный белок с β -галактозидазой.



в виде целого пептида. Такой способ экспрессии используется, когда возможно отщепление экспрессируемого пептида от N-концевой части химерного белка или когда достаточно получить гибридный полипептид, например в случае получения искусственных вакцин, где достаточно наличия необходимых антигенных детерминант.

Рис. 255. Схема расположения элементов для прямой экспрессии чужеродных генов в новом хозяине.



Для получения белков, содержащих более 150 аминокислот, чаще используется прямая экспрессия соответствующих генов, в результате которой сразу синтезируются требуемые белки лишь с незначительной модификацией. При этом кодирующая часть гена соединяется с бактериальным промотором, рибосомсвязывающим участком и иницирующим кодоном ATG (рис. 255).

Именно таким образом получены штаммы *E. coli*, продуцирующие гормон роста человека, интерфероны человека и другие белки, имеющие большое практическое значение. Присоединение к кодирующей части гена иницирующего кодона необходимо для обеспечения инициации трансляции. Это приводит к тому, что

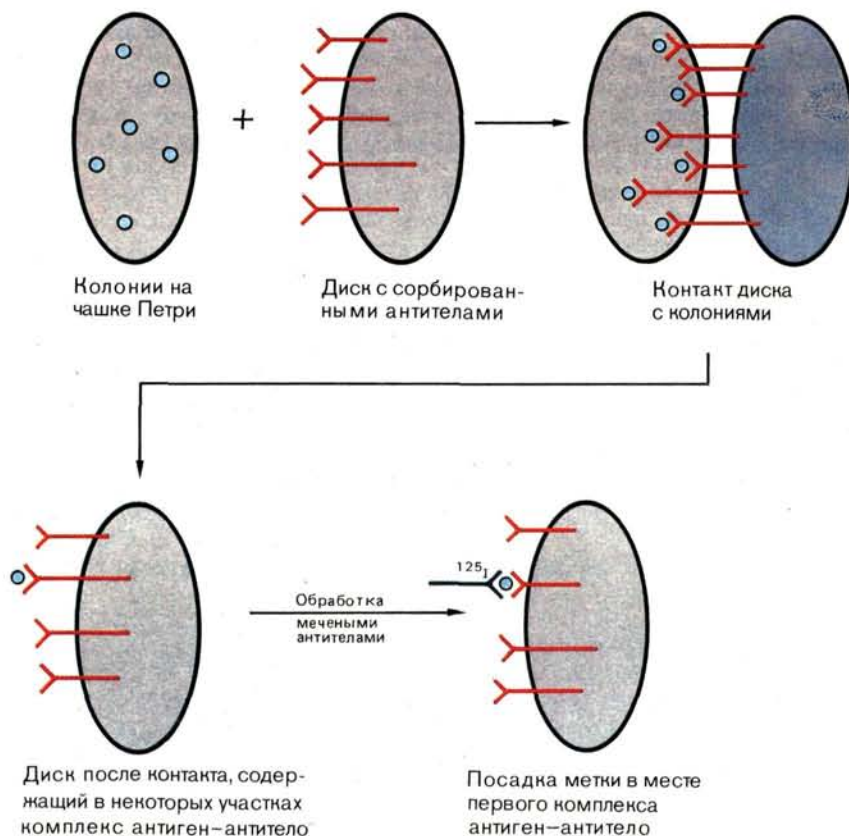


Рис. 256. Схема, демонстрирующая иммунохимическую идентификацию экспрессирующих клонов.

N-концевая часть синтезируемых таким способом полипептидов отличается от природных присутствием N-концевого формилметионина. Ферменты процессинга бактериальной клетки частично удаляют ее, однако часть синтезируемого белка все же сохраняет модификацию. Во многих случаях такая модификация не меняет физиологических свойств продукта, но при получении лекарственных препаратов необходимо учитывать возможный ее эффект.

Идентификация экспрессирующих рекомбинантных клонов. Метод идентификации экспрессирующих клонов зависит от свойств продукта экспрессии. Если этот продукт обладает собственной биологической активностью, то он может быть идентифицирован по ее проявлению. Например, если экспрессии подвергается ген, кодирующий фермент, то клоны идентифицируют по наличию в них соответствующей ферментативной активности; клоны, синтезирующие интерферон, — по противовирусной активности клеточных экстрактов и т. д.

Однако наиболее общими являются методы иммунохимического анализа, применимые как в случае прямой экспрессии, так и при синтезе гибридных белков.

Существует много вариантов иммунохимического анализа. Один из наиболее распространенных заключается в следующем. На диске из поливинилхлорида сорбируются немеченые антитела к исследуемому белку. Колонии прямо на чашке Петри лизируются в мягких условиях, и поливинилхлоридный диск вводится в контакт с поверхностью чашки. Продукты с антигенными детерминантами исходного белка образуют комплексы антиген — антитело с антителами, сорбированными на диске только в местах положения экспрессирующих клонов. Затем диск отмывается от неспецифически сорбированного антигена и обрабатывается раствором с мечеными (радиоактивно, флуоресцентно или иммуноферментно) антителами. Эти антитела образуют комплекс с антигеном за счет его других антигенных детерминант, в результате чего участки диска, содержащие первый комплекс антиген — антитело, оказываются мечеными (рис. 256). По положению меченых участков на диске легко идентифицировать экспрессирующие колонии.

Клонирование в различных организмах

В настоящее время разработаны системы клонирования в различных бактериях, дрожжах, грибах, растениях и млекопитающих. Наибольший интерес с практической точки зрения представляют системы клонирования в грамположительных бактериях, многие из которых являются промышленно используемыми, а также в дрожжах и клетках высших организмов.

Обычно векторы для клонирования в таких системах представляют собой двойные репликоны, которые могут существовать и в *E. coli*, и в той клетке-хозяине, для которой они предназначены. Это достигается созданием гибридных векторов, содержащих репликон какой-либо из плазмид *E. coli* и требуемый репликон, например плазмиды *V. subtilis* или дрожжей, что позволяет проводить первоначальное клонирование и отбор требуемых генов в хорошо изученной системе *E. coli*, а затем уже вводить выделенные рекомбинантные плазмиды в новый организм.

Такие векторы должны содержать в себе ген или гены, придающие клетке-хозяину легко тестируемый признак, например устойчивость к антибиотикам.

Клонирование в грамположительных бактериях. Наиболее широко используются методы клонирования в *B. subtilis* и стрептомицетах. Разработка систем векторов для *B. subtilis* основана на способности плазмид *Staphylococcus aureus*, несущих гены устойчивости к антибиотикам, к репликации и стабильному наследованию в сенной палочке. Для этих целей гибриды таких плазмид с плазмидами *E. coli* широко используются в качестве векторов для *B. subtilis*. Рекомбинантные штаммы несут признаки устойчивости к антибиотикам.

Стрептомицеты широко применяются в биотехнологии в качестве продуцентов антибиотиков. Конструирование векторов для клонирования в них началось с выделения плазмиды Scp2 из *Streptomyces coelicolor*. На основе этой и подобных плазмид в настоящее время сконструированы векторы, придающие стрептомицетам устойчивость, например, к таким антибиотикам, как метиленомицин А.

Клонирование в дрожжах. Наиболее широко используются штаммы *Saccharomyces cerevisiae*. Работа с дрожжами облегчается тем, что, подобно бактериям, они могут расти в жидкой среде и давать колонии на твердой, генетически хорошо охарактеризованы и имеют сравнительно короткое время генерации. *S. cerevisiae* содержит плазмиду Scp1, представляющую собой циклическую молекулу длиной 2 мкм. Ее гибриды с плазмидами *E. coli* обычно и используют в качестве векторов. Селекция дрожжевых клонов, трансформированных такими рекомбинантными плазмидами, основана на применении в качестве клеток-хозяев определенных мутантов, не способных расти на среде, лишенной какого-либо питательного компонента. Векторная плазида, в свою очередь, содержит ген или гены, которые при попадании в клетку придают ей этот недостающий признак. Трансформанты легко отбираются по их способности давать колонии на обедненной среде.

Поскольку дрожжи представляют собой эукариотический организм, можно было бы ожидать, что гены различных эукариот, в том числе и те, которые содержат интроны, будут корректно экспрессироваться в дрожжевых клетках. Однако это не так. Например, экспрессия генов β -глобина кролика в дрожжах не происходит благодаря некорректности транскрипции и последующего «сплайсинга» РНК. Тем не менее, применяя приемы, аналогичные использованным при клонировании в бактериях, удастся достичь синтеза чужеродных белков в дрожжевых клетках. Такие клетки, подобно *B. subtilis*, секретируют значительное количество белков во внеклеточную среду, что используют также для секреции чужеродных белков. С этой целью к экспрессируемому гену присоединяется участок, кодирующий сигнальный пептид, обуславливающий секрецию и отщепляемый в ее процессе. В результате в клетке синтезируется белок, содержащий на N-конце сигнальный пептид. Этот белок секретируется в окружающую среду. Таким образом были получены, например, штаммы дрожжей, секретирующие интерферон человека.

Клонирование в клетках животных. Проблема клонирования в клетках животных имеет большое значение для исследования функционирования генов высших эукариот. Предварительно клонированные гены вводят в клетку животных различными путями. Один из путей включает в себя кон-трансформацию клеток требуемым геном, соединенным с одним из генов, для которых осуществляется селекция. Примером являются клетки мыши, дефектные по синтезу тимидинкиназы (ТК⁻-клетки). Клетки трансформируются фрагментами ДНК вируса герпеса (HSV), содержащего ген тимидинкиназы, и после трансформации приобретают способность к синтезу фермента, т. е. становятся ТК⁺-клетками. Клетки ТК⁺ легко отличаются от ТК⁻, поскольку они способны расти на средах с ами-

ноптерином (блокирующим определенные стадии биосинтеза нуклеотидов), гипоксантином и тимидином.

Следовательно, при конструировании котрансформирующих векторов для трансформации клеток животных используются гибриды бактериальных плазмид с геном ТК из вируса герпеса. Предварительно клонирование и идентификация генов проводятся в клетках *E. coli* и затем рекомбинантная плазида вводится в ТК⁻-клетки. Среди образовавшихся ТК⁺-трансформантов отбираются нужные, например путем идентификации продуктов экспрессии клонированных генов.

Другой путь введения клонированных генов в эукариотические клетки аналогичен применяемому в случае *E. coli*, когда в качестве векторов используют бактериофаги. В качестве таких эукариотических векторов служат некоторые вирусы. Во многих из них существуют области для литического роста, которые могут быть заменены на чужеродные ДНК. Такие вирусы реплицируются в клетке-хозяине, экспрессируя чужеродные гены. Однако в вирусах животных размеры несущественных областей малы и не позволяют внедрить большие фрагменты ДНК. Некоторые гены животных имеют большие размеры (например, ген дигидрофолатредуктазы мыши — 42 тыс. п. о.). В большинстве случаев чужеродная ДНК замещает существенные гены, в результате чего рекомбинантные вирусы теряют способность к репликации. Для ее обеспечения используют вирусы-помощники, синтезирующие продукты недостающих генов. В присутствии помощников рекомбинантный вирус существует за счет этих продуктов. Примером вирусов, применяемых в качестве векторов, является вирус SV-40, геном которого представляет собой циклическую ДНК длиной 5243 п. о. с полностью известной последовательностью. Эукариотические векторы, использующие вирусы, способные формировать вирионы, обладают тем недостатком, что они убивают клетку-хозяина при своем размножении (так же как и бактериофаги). Поэтому постоянно делаются попытки разработать векторы, подобные плазидам. Обычно опухолевые вирусы, в том числе и SV-40, внедряют свою ДНК в хромосому клетки-хозяина. Однако вирус бычьей папилломы в трансформированных клетках существует в виде эписомы (около 100 копий на клетку) и используется в качестве основы для создания эписомных векторов.

Высокий темп исследований генной инженерии на клетках животных вселяет надежду, что в ближайшее время будут разработаны простые системы, которые позволят осуществить анализ механизмов экспрессии генов эукариот и дадут возможность создать животных, обладающих заданными свойствами.

Генная инженерия растений. Эта отрасль генной инженерии не так хорошо разработана, как в случае животных и тем более микробных клеток. Однако в настоящее время она привлекает очень большое внимание, поскольку открывает новые перспективы в растениеводстве. Обычная селекция новых сортов — процесс медленный, и кроме того, она ограничена природными видовыми барьерами. Введение новых генов с помощью техники рекомбинантных ДНК в растения могло бы ускорить этот процесс и существенно расширить его возможности. Кроме того растения обладают существенной особенностью: целое растение может быть выращено из отдельной клетки. Это не относится в равной степени ко всем растениям, например, клетки злаковых или бобовых только очень редко претерпевают такую редифференциацию, тогда как клетки табака, томата или моркови, как показала практика, подвергаются редифференциации сравнительно легко.

При растворении целлюлозной стенки растительной клетки ферментами-целлюлазами образуются протопласты, в которые

легко проникают макромолекулы, в том числе ДНК. Две различные клетки в виде протопластов соединяются с образованием гибридного протопласта (соматической гибридной клетки). Протопласты способны восстанавливать клеточную стенку и далее давать целое растение. Растение может быть получено из протопластов, включивших чужеродную ДНК, или из гибридных протопластов.

При наличии методов введения в растительные клетки определенных генов, способных к функционированию и стабильному наследованию, открываются реальные возможности создания растений с заранее заданными полезными признаками. Векторы для введения генов в растительные клетки могут быть основаны на репликациях растительных вирусов, однако до настоящего времени попытки получения таких векторов, удовлетворяющих, в частности, требованию стабильного наследования, были не очень успешны.

Наибольшее развитие получили векторы, сконструированные на так называемых Ti-плазмидах бактерий *Agrobacterium tumefaciens*. Эти бактерии инфицируют двудольные растения и вызывают образование корончатых галлов — своеобразных опухолей растений, клетки которых способны размножаться в культуре без добавления факторов роста. Было показано, что такая трансформация является следствием внедрения в геном растения части ДНК-Ti (tumor-inducing)-плазмиды, получившей название T (transforming)-ДНК. Трансформированные клетки приобретают способность синтезировать необычные аминокислоты — опины, например, нопалин или октопин, являющиеся производными аргинина и служащие источником питания *A. tumefaciens*. Какая именно аминокислота синтезируется, зависит от Ti-плазмиды, трансформировавшей клетку: одни бактерии *A. tumefaciens* несут октопиновые, а другие — нопалиновые плазмиды. И опухоловая трансформация, и синтез опинов вызваны T-ДНК. В T-ДНК содержится несколько генов, контролируемых отдельными промоторами, которые и определяют все свойства опухолевых клеток. Трансформация клеток T-ДНК стабильно наследуется; некоторые трансформированные культуры существуют без изменения уже более 20 лет. Интегрированная T-ДНК наследуется по законам Г. Менделя.

Ti-Плазмиды способны трансформировать практически все двудольные растения, что делает их перспективным вектором для введения чужеродных ДНК в растения (рис. 257).

К сожалению, трансформированные T-ДНК клетки не способны давать растения. Однако были получены мутанты T-ДНК, трансформирующие растительные клетки и не подавляющие их способности превращаться в растение. Внедрение чужеродных генов в T-ДНК, под контроль промоторов, способных функционировать в растении, которое может быть осуществлено с помощью специально сконструированных векторов, дает возможность включать их в геном растительных клеток и получать растения, содержащие новую генетическую информацию.

Для улучшения свойств сельскохозяйственных растений необходимо внедрение в них такой генетической информации, которая делала бы их устойчивыми к засухе, заморозкам, позволяла расти на засоленных почвах, придавала способность фиксировать азот и устойчивость к сельскохозяйственным вредителям. Это осуществляют переносом соответствующих генов из растений, обладающих подобными свойствами. Так, в качестве примера можно привести создание петунии (*Petunia*) или табака (*Nicotiana tabacum*), устойчивых к гербициду глифосату, путем введения в клетки растений гена, дающего резистентный к этому веществу фермент. Полученные клетки были затем превращены в целые растения. Внедрение полезной генетической информации осуществлено с помощью Ti-плазмиды.

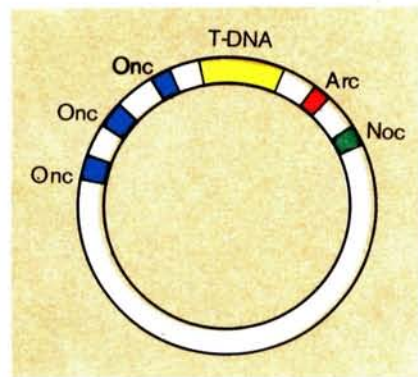


Рис. 257. Физическая карта нопалиновой Ti-плазмиды. Гены Opс ответственны за онкогенность, ген Nos — за синтез нопалина, ген Arc — за деградацию аргинина.