

# Структура РНК-полимеразы и основные механизмы транскрипции у бактерий

д.б.н. А.В. Кульбачинский

## Раздел 1. Общие сведения о РНК-полимеразе и механизмах транскрипции

Транскрипция – процесс синтеза РНК на матрице ДНК, первая стадия экспрессии генетического материала. Процесс транскрипции ДНК в клетках всех организмов осуществляется ДНК-зависимыми РНК-полимеразами. Существует два основных типа РНК-полимераз – односубъединичные и многосубъединичные. Односубъединичные РНК-полимеразы участвуют в транскрипции геномов некоторых бактериофагов, а также митохондрий. Многосубъединичные РНК-полимеразы (РНКП) отвечают за транскрипцию большинства генов в клетках всех живых организмов. Структура РНКП и общий механизм синтеза РНК высоко консервативны и в основных чертах очень сходны у бактерий, архей и эукариот. У эукариот имеется несколько различных РНКП, между которыми разделяются функции синтеза мРНК, тРНК и рРНК. В клетках прокариот имеется только одна РНКП, которая синтезирует все типы РНК в клетке.

Существует две формы РНКП: кор- и холофермент. Кор-фермент имеет относительную молекулярную массу около 400 кДа и состоит из пяти субъединиц: двух  $\alpha$ ,  $\beta$ ,  $\beta'$  и  $\omega$ . Гомологичные субъединицы входят в состав всех известных многосубъединичных РНКП. Холофермент состоит из кор-фермента и  $\sigma$ -субъединицы. Транскрипционный цикл, осуществляемый РНК-полимеразой, включает в себя 3 последовательные стадии - инициацию, элонгацию и терминацию синтеза РНК. Кор-фермент РНКП обладает каталитической активностью, но не способен к инициации транскрипции. Для узнавания специфических последовательностей ДНК (промоторов) необходимо присоединение фактора инициации -  $\sigma$ -субъединицы, - и образование холофермента РНКП.  $\sigma$ -субъединица играет центральную роль в узнавании промоторов, плавлении ДНК и последующем уходе РНКП с промотора. По своей структуре кор-фермент РНКП напоминает клешню краба, одна половина которой образована в основном  $\beta'$ -, а вторая –  $\beta$ -субъединицей; между ними находится главный канал РНКП, в котором происходит связывание ДНК и РНК в процессе транскрипции. Димер  $\alpha$ -субъединиц располагается со стороны, противоположной главному каналу;  $\omega$ -субъединица контактирует с  $\beta'$ -субъединицей на внешней стороне РНКП. Активный центр фермента образован  $\beta'$ - и  $\beta$ -субъединицами и располагается в глубине главного канала. Сбоку от главного канала расположен вторичный канал, который соединяется с главным в области активного центра. Вторичный канал служит местом связывания

многих регуляторных факторов, и основным путем поступления нуклеотидов в активный центр РНКП.

## **Раздел 2. Реакции, катализируемые РНК-полимеразой**

Все реакции, осуществляемые РНКП, протекают по механизму бимолекулярного нуклеофильного замещения ( $S_N2$ ): во всех случаях РНКП катализирует перенос нуклеозидмонофосфата с молекулы-донора на молекулу-акцептор. В катализе участвуют два иона  $Mg^{2+}$ , координированных в активном центре РНКП. Первый ион  $Mg^{2+}$  (MgI) прочно связан с тремя остатками аспарагиновой кислоты абсолютно консервативного для всех РНКП мотива NADFDGD  $\beta'$ -субъединицы. Второй ион магния (MgII) связан с РНКП гораздо слабее, так как в его связывании принимает участие только один остаток аспарагиновой кислоты, но он дополнительно координируется молекулами воды.

В процессе транскрипции РНКП способны осуществлять несколько различных реакций: синтез РНК, пирофосфоролиз (перенос нуклеозидмонофосфата с 3'-конца новосинтезированной РНК на пирофосфат, в результате чего образуется НТФ), а также экзо- и эндонуклеазное расщепление РНК-транскрипта (перенос нуклеозидмонофосфата (экзонуклеазное) с 3'-конца новосинтезированной РНК, либо нескольких 3'-концевых нуклеотидов (эндонуклеазное) на молекулу воды, в результате чего образуется свободный НМФ или короткий фрагмент РНК). В отличие от ДНК-полимераз, имеющих отдельный активный центр для расщепления ДНК, в РНКП все реакции осуществляются в одном активном центре, что делает данную молекулярную машину уникальным ферментом. Биологическая роль такого многообразия реакций, катализируемых РНКП, крайне высока: расщепление РНК играет большую роль в реактивации комплексов после включения неправильного нуклеотида, а также в реактивации «арестованных» транскрипционных комплексов, образующихся в процессе элонгации транскрипции.

## **Раздел 3. Узнавание промоторов РНК-полимеразой и инициация транскрипции**

Процесс транскрипции подразделяется на несколько стадий – инициацию, элонгацию и терминацию. РНКП начинает транскрипционный цикл со связывания с особой последовательностью ДНК – промотором, плавления короткого участка ДНК и начала синтеза РНК, эти процессы вместе называются инициацией синтеза РНК. В составе бактериальных промоторов выделяют несколько консервативных элементов, необходимых для успешного взаимодействия с РНКП. Прежде всего это «-10»

(TATAAT) и «-35» (TTGACA) элементы, располагающиеся, соответственно, примерно на 10 и 35 нуклеотидов левее положения начала транскрипции и взаимодействующие с районами 2 и 4  $\sigma$ -субъединицы в составе холофермента РНКП (свободная сигма-субъединица не способна к узнаванию промоторных элементов). Важным оказывается и расстояние между этими участками промотора. В большинстве случаев оно составляет 17-18 нуклеотидов. В состав некоторых сильных промоторов входит А/Т богатый участок, расположенный на 50-70 нуклеотидов левее старта точки начала транскрипции. Этот участок, узнаваемый  $\alpha$ -субъединицей РНКП, получил название UP-элемента.

В процессе инициации транскрипции холофермент РНКП образует сначала закрытый промоторный комплекс, в котором ДНК остается в двуцепочечном виде. Затем комплекс изомеризуется, что приводит к локальному плавлению ДНК и образованию открытого комплекса. После образования открытого промоторного комплекса РНКП начинает синтез фрагментов РНК длиной 2-9 нуклеотидов, которые диссоциируют из комплекса. Синтез таких коротких молекул РНК называется абортивным. На стадии абортивного синтеза РНКП сохраняет связи с промотором, поэтому при продвижении фермента по ДНК расплетенные одноцепочечные участки выпетливаются из главного канала. Этот процесс сопровождается накоплением напряжения в инициаторном комплексе. Накопленная энергия может быть потрачена либо на диссоциацию РНК-продукта, либо на разрыв контактов с промотором. В случае диссоциации продукта РНКП возвращается на стартовое положение и начинает синтез заново. Разрыв контактов с промотором сопровождается диссоциацией  $\sigma$ -субъединицы и переходом транскрипции в стадию элонгации.

#### **Раздел 4. Работа РНК-полимеразы на стадии элонгации транскрипции**

На стадии элонгации кор-фермент РНКП осуществляет поэтапное присоединение нуклеотидов, комплементарных матричной цепи ДНК. Комплекс РНКП с ДНК и синтезируемой РНК называется элонгационный комплекс (ЭК). Механизм присоединения нуклеотидов сводится к последовательному протеканию следующих процессов: связывание входящего нуклеотида, закрывание активного центра (происходит в случае, если связавшийся нуклеотид комплементарен соответствующему нуклеотиду матричной цепи ДНК), катализ синтеза фосфодиэфирной связи, освобождение пирофосфата, запускающее последнюю стадию – транслокацию, т.е. перемещение РНКП на 1 нуклеотид вперед по матрице ДНК, после чего может произойти присоединение следующего нуклеотида. Основные элементы, принимающие

участие в описанном процессе - F-спираль (длинная  $\alpha$ -спираль, располагается между главным и вторичным каналом) и G-петля (располагается рядом с F-спиралью, имеет неструктурированный вид в несвернутом состоянии и может сворачиваться в  $\alpha$ -спираль при закрывании активного центра). Последовательные структурные перестройки высоко консервативных элементов активного центра F-спирали и G-петли способствуют присоединению нуклеотида и транслокации фермента, что обеспечивает функционирование РНКП. РНКП способна осуществлять и обратную реакцию – расщепление РНК. В этом случае РНКП прекращает синтез РНК и смещается назад по матрице ДНК (*англ.* backtracking). При смещении общие параметры транскрипционного пузыря также сохраняются, при этом 3'-конец РНК разрывает связь с матричной ДНК и активным центром фермента и выходит из активного центра во вторичный канал РНКП. В таком комплексе происходит расщепление РНК. Данный процесс протекает со сравнительно небольшой скоростью *in vitro*, но в клетке имеются специализированные факторы, ускоряющие реакцию расщепления РНК на несколько порядков (Gre белки).

## **Раздел 5. Терминация транскрипции**

Конечным этапом транскрипции является терминация. Время полужизни элонгационных комплексов обычно составляет величины порядка дней, но на особых нуклеотидных последовательностях, называемых терминаторами, элонгационные комплексы диссоциируют за секунды. Терминаторы необходимы для разграничения независимо транскрибируемых областей, а также являются важными элементами в процессах регуляции транскрипции. Существуют два механизма терминации транскрипции: Rho-зависимый и Rho-независимый. Для успешной терминации по второму пути не требуются никакие дополнительные факторы, достаточно взаимодействия терминатора с РНКП. Классический терминатор Rho-независимого пути представляет собой G/C-богатую РНК-шпильку и следующую за ней олиго-U последовательность. Взаимодействие новосинтезированного олиго-U участка РНК с РНКП приводит к возникновению транскрипционных пауз. За это время успевает сформироваться шпилька, что приводит к диссоциации элонгационного комплекса. Синтез примерно 80% транскриптов *E.coli* терминируется по Rho-независимому пути. В Rho-зависимом пути принимает участие белковый Rho-фактор – гомогексамерный белок, обладающий АТФ-зависимыми транслоказной и хеликазной активностями. Он связывается с РНК в специальных пиримидин-богатых сайтах, называемых rut (Rho utilization), и за счет транслоказной активности движется в направлении 3'-конца

молекулы, пока не догонит РНКП. После этого происходит диссоциация элонгационного комплекса.

## **Раздел 6. Механизмы, обеспечивающие точность транскрипции**

Высокая точность транскрипции является необходимым условием для правильной экспрессии генетического материала; ошибки транскрипции могут приводить к появлению нефункциональных молекул РНК и мутантных вариантов белков. Уровень ошибок при синтезе РНК многосубъединичными РНКП составляет примерно  $10^{-5}$ . Высокая точность синтеза РНК обеспечивается тем, что РНКП способна: (1) эффективно дискриминировать правильные и неправильные субстраты в процессе присоединения нуклеотидов к 3'-концу РНК; (2) детектировать неправильно присоединенный к 3'-концу РНК нуклеотид; (3) удалять неправильный нуклеотид (или несколько нуклеотидов) за счет реакции расщепления РНК. В целом, эти механизмы сходны с основными механизмами, используемыми ДНК-полимеразами для высокоточного синтеза ДНК. Селекция правильных нуклеотидов в процессе транскрипции обеспечивается тем, что фиксация субстратов в активном центре и катализ происходят с высокой эффективностью, только если геометрия образующейся пары точно соответствует геометрии +1-сайта активного центра. Тем не менее, этот механизм не позволяет полностью исключить включение ошибочных нуклеотидов, которое может происходить:

- 1) в результате неправильного спаривания входящего NTP и матричного нуклеотидного остатка в цепи ДНК;
- 2) напротив апуринных или апириимидиновых сайтов в матричной;
- 3) при транскрипции участков ДНК, содержащих поврежденные нуклеотидные остатки (например, 8-оксо-гуанин, тиминные димеры и др.)

Эффективное исправление ошибок обеспечивается следующими механизмами.

- 1) Присоединение неправильного нуклеотида на 3'-конец РНК значительно замедляет включение последующих нуклеотидов. РНКП включает в РНК различные неправильные нуклеотиды с разной эффективностью, что, вероятно, объясняется особенностями структуры +1-сайта активного центра. Чем больше эффективность неправильного включения для данной пары оснований (NTP/матричное основание в ДНК), тем меньше эффективность последующего удлинения РНК, содержащей неправильный нуклеотид. Это позволяет минимизировать эффективность включения ошибочных нуклеотидов в РНК-транскрипты.

Низкая эффективность удлинения РНК с ошибочным нуклеотидом на 3'-конце объясняется тем, что наличие ошибки индуцирует переход ЭК в неактивную конформацию. При помощи рентгеноструктурного анализа ЭК РНКП II *S. cerevisiae* зафиксировано каталитически-неактивное состояние ЭК с ошибочным 3'-концевым нуклеотидом. В таком состоянии 3'-конец РНК сдвигается в сторону от активного центра в претранслокационном ЭК, что препятствует транслокации и приводит к паузе транскрипции. Такая пауза транскрипции, вызванная сдвигом 3'-конца РНК без серьезных структурных перестроек ЭК, получила название элементарной паузы. Сайт связывания 3'-концевого нуклеотида РНК в состоянии элементарной паузы перекрывается с +1-сайтом связывания NTP. Комплекс в состоянии элементарной паузы может возвращаться в активное претранслокационное состояние, либо смещаться назад по матрице ДНК.

Элементарные паузы, вероятно, играют важную роль в регуляции транскрипции и могут вызываться не только наличием на 3'-конце РНК неправильных нуклеотидов, но и регуляторными сигналами, закодированными в РНК (шпильки определенной структуры), а также, возможно, транскрипционными факторами. В состоянии элементарной паузы G-петля в активном центре находится в специальной конформации, которая стабилизирует претранслокационное состояние ЭК и делает неэффективным удлинение РНК.

2) Ошибочно присоединенный 3'-концевой нуклеотид может далее отщепляться по механизму экзонуклеазного расщепления. Данная реакция может происходить даже в ЭК, содержащих РНК, полностью комплементарную матрице, причем этот процесс значительно стимулируется в присутствии NTP, которые не могут служить субстратами для удлинения РНК (некомплементарные, негидролизруемые NTP и dNTP). Стимулирующая роль NTP заключается в координировании иона Mg<sup>2+</sup> в активном центре РНКП, за счет его взаимодействия с β- и γ-фосфатными группами нуклеотида. Предполагается, что связывание некомплементарного нуклеотида происходит в E-сайте активного центра. Возможно, что экзонуклеазное расщепление стимулируется при наличии неправильного 3'-концевого нуклеотида в РНК, однако, физиологическое значение этой реакции в исправлении ошибок транскрипции на сегодняшний день остается неясным.

3) Основным механизмом, по которому происходит удаление ошибочных нуклеотидов является реакция экзонуклеазного расщепления РНК, которая происходит в смещенных ЭК. Смещение ЭК и последующее расщепление РНК значительно стимулируется наличием ошибочного 3'-концевого нуклеотида. При расщеплении с 3'-

конца РНК в основном удаляются динуклеотидные фрагменты, хотя возможно отщепление и более длинных фрагментов. Таким образом, расщепление РНК происходит в большинстве случаев в комплексах, смещенных назад на один нуклеотид. Первой стадией образования смещенного ЭК является переход претранслокационного ЭК в состояние элементарной паузы. Последующая обратная транслокация ЭК из такого состояния приводит к попаданию в +1-сайт предыдущего (то есть, второго с 3'-конца) нуклеотида РНК. Анализ структуры смещенного ЭК РНКП II показал, что при этом положение 3'-концевого нуклеотида РНК в +2 позиции (в так называемом редактирующем сайте, *англ.* proofreading site) похоже на его положение до начала обратной транслокации, что обеспечивается резким изгибом молекулы РНК между +1 и +2 позициями. При дальнейшем обратном смещении ЭК 3'-концевая часть РНК выходит через вторичный канал РНКП, при этом нуклеотид, находящийся в данный момент в +2 позиции, занимает положение в редактирующем сайте. Нуклеотид в +2 позиции РНК непосредственно стимулирует реакцию расщепления, участвуя в координировании иона магния и атакующего гидроксила воды.

Точная роль различных элементов активного центра РНКП в расщеплении РНК на сегодняшний день остается неизвестной. Определенные данные позволяют предположить, что в реакции расщепления участвуют несколько иные участки активного центра, чем в реакции синтеза РНК.

## **Раздел 7. Регуляция инициации транскрипции у бактерий**

Регуляция активности РНКП в процессе транскрипции обеспечивается действием разнообразных регуляторных факторов, которые включают белковые факторы, регуляторные последовательности ДНК и РНК, некодирующие РНК, малые молекулы и метаболиты, пептиды и антибиотики. Проведенные в последние годы структурные и биохимические исследования РНКП бактерий и эукариот позволили установить, что многие факторы непосредственно влияют на функциональное состояние активного центра РНКП в ЭК и способны либо активировать, либо подавлять различные реакции, катализируемые РНКП. Важно отметить, что не только клеточные белки способны модулировать активность РНКП, но и различные регуляторные белки и пептиды бактериофагов также оказывают порой колоссальное действие на бактериальную транскрипцию, полностью переключая ее на свои нужды.

Тонкой регуляции подвергаются все стадии транскрипции, как правило, регуляторные факторы действуют в основном на одну из них.

К регуляции на стадии инициации можно отнести наличие у бактерий разного количества альтернативных сигма-субъединиц, регуляцию посредством белковых факторов (Gfh, DksA и некоторых других белков), некодирующих РНК, ppGpp. Большая часть промоторов *E. coli* узнаются  $\sigma^{70}$  субъединицей в составе холофермента РНКП, другие  $\sigma$ -субъединицы (альтернативные  $\sigma$  факторы) ответственны за адаптацию к конкретным физиологическим состояниям. При постоянной концентрации холофермента и основных сигма-субъединиц в клетках количество некоторых альтернативных  $\sigma$  факторов (например,  $\sigma^{38}$  и  $\sigma^{32}$ ), необходимых для узнавания определенного типа промоторов, возрастает в определенных физиологических состояниях (например, при переходе в стационарную фазу роста или тепловом шоке), что имеет чрезвычайно важный адаптивный характер. В конце 1960-ых годов был открыт новый класс коротких РНК бактерий - 6S РНК. Функции их долгое время были неизвестны. Затем было выяснено, что данный тип РНК накапливается в течение стационарной фазы роста клеточной культуры, а также присутствует в клетках во время экспоненциального роста культуры. Лишь в 2000 году было обнаружено, что 6S РНК со-осаждаются с РНК-полимеразой. Последующие эксперименты подтвердили, что 6S РНК формирует стабильный специфичный комплекс с холоферментом РНК-полимеразы, который в стационарной фазе роста культуры является преимущественной формой фермента РНКП. Оказалось, что 6S РНК ингибирует транскрипцию. Так как 6S РНК конкурирует с ДНК за место связывания на РНК-полимеразе, то «прочтение» 6S РНК может иметь важную физиологическую роль – при синтезе продукта на 6S РНК, происходит их освобождение, таким образом снимается ингибирующее влияние на РНК-полимеразу посредством 6S РНК матрицы. 6S РНК нужна для оптимальной жизнедеятельности клеток в поздней стационарной фазе, повышает устойчивость клеток к стрессу. Это имеет определенную биологическую роль: когда запас питательных веществ в среде подходит к завершению, не имеет смысла синтезировать большое количество РНК и белка. При росте культуры, когда в среде большое количество питательных веществ, происходит активный синтез РНК на 6S РНК матрице, в результате чего освобождающаяся РНК-полимераза может начать транскрипцию клеточных генов. Конкретным сигналом может являться концентрация нуклеотидов – их небольшое количество во время стационарной фазы предотвращает использование 6S РНК в качестве матрицы, и, соответственно, её освобождение с отменой ингибирующего эффекта. Напротив, возрастающая концентрация нуклеотидов во время экспоненциального роста – сигнал для освобождения РНК-полимеразы от 6S РНК после «прочтения» последней.



В ответ на нехватку аминокислот в среде в бактериальной клетке происходит угнетение транскрипции одних генов (в том числе, генов рибосомальных и транспортных РНК) и активация транскрипции других (генов синтеза аминокислот и транспортных оперонов). Этот феномен получил название “строгого контроля” (stringent control, или stringent response). В отличие от многих других регуляторных механизмов, “строгий контроль” опосредуется не белковыми транскрипционными факторами, которые связываются с ДНК или РНКП, а малой молекулой гуанозинтетрафосфата, ppGpp. ppGpp образуется в клетке в ответ на аминокислотное голодание; он в основном синтезируется связанным с рибосомой белком RelA в том случае, если А-сайт рибосомы занят не аминоацилированной тРНК. Связываясь с РНКП, ppGpp подавляет транскрипцию одних генов и стимулирует транскрипцию других. Наиболее важная мишень действия ppGpp – промоторы генов рибосомальных РНК. Транскрипция с этих промоторов является лимитирующей стадией в процессе синтеза рибосом в клетке. В стадии логарифмического роста бактерий транскрипция с рибосомальных промоторов (7 промоторов в клетках *E. coli*) составляет более половины всей транскрипции в клетке. Промоторы генов рРНК обладают необычными кинетическими свойствами, которые отличают их от всех других промоторов. Промоторы рРНК обладают рядом структурных особенностей, благодаря которым образующиеся на них промоторные комплексы являются нестабильными. Инициация транскрипции на этих промоторах очень чувствительна к внутриклеточной концентрации двух молекул: гуанозинтетрафосфата ppGpp – негативного регулятора активности промотора и инициаторного NTP – позитивного регулятора активности промотора. Обе эти малые молекулы служат индикаторами обеспеченности бактериальной клетки пищевыми субстратами. Увеличение концентрации инициаторных нуклеотидов (прежде всего, АТФ) увеличивает стабильность промоторных комплексов и повышает вероятность инициации на промоторах рРНК. Напротив, связывание ppGpp с РНКП значительно сокращает время жизни открытого комплекса на промоторах рРНК. ppGpp в 5'-ориентации, по-видимому, способствует связыванию второго каталитического иона магния в активном центре. Анализ действия ppGpp на активность РНКП *in vitro* показал, что ингибиторный эффект ppGpp на транскрипцию оказывается в несколько раз слабее, чем *in vivo*. Это можно было бы объяснить существованием дополнительного фактора, усиливающего действие ppGpp *in vivo*. В результате генетических и биохимических исследований было установлено, что таким фактором является белок DksA. Ген *dksA* был ранее охарактеризован как мульткопийный супрессор температурочувствительных мутаций в генах *dnaK* и *dnaJ*.

В экспериментах *in vitro* очищенный DksA снижает концентрацию ppGpp, необходимую для проявления ингибиторного эффекта, до величин, соответствующих физиологическим. DksA состоит из двух доменов: N-концевого coiled-coil домена и C-концевого глобулярного домена. Структура N-концевых доменов DksA и Gre-белков практически идентична. Это тем более удивительно, что по аминокислотной последовательности эти белки совершенно не гомологичны. Сходство структур DksA и Gre-белков позволило сразу высказать предположение, что DksA, как и Gre-белки, связывается во вторичном канале РНКП. Как и у Gre-белков, на конце N-концевого домена DksA, в участке петли, соединяющей две  $\alpha$ -спирали, находятся два высококонсервативных отрицательно заряженных остатка аспарагиновой кислоты. Это позволило предполагать, что DksA может участвовать в координации ионов  $Mg^{2+}$ . В то же время, в отличие от Gre-белков, DksA не стимулирует реакции эндонуклеазного расщепления и не может реактивировать смещенные элонгационные комплексы. Это объясняется тем, что важный для функционирования Gre остаток аспарагиновой кислоты заменен на остаток глутаминовой кислоты в случае DksA.

## **Раздел 8. Регуляция элонгации транскрипции у бактерий**

Регуляция на стадии элонгации транскрипции очень важна и включает различные механизмы. Хорошо изученным примером регуляторных факторов являются Gre белки, активирующие расщепление РНК. Популяция элонгационных комплексов РНКП, возникающих на одном и том же промоторе, неоднородна. Некоторая часть ЭК приостанавливает синтез РНК и находится в состоянии «паузы». Это может происходить из-за особенностей структуры матрицы, отсутствия комплементарного нуклеотида или в результате действия регуляторных факторов. Структура большей части паузирующих комплексов не нарушена и по прошествии некоторого времени (которое может составлять десятки секунд и более) они продолжают синтез РНК. В то же время, часть паузирующих комплексов может переходить в так называемое «арестованное» (arrested) состояние, в котором они не способны продолжить синтез РНК. По всей видимости, одной из основных причин образования арестованных элонгационных комплексов является смещение РНКП назад по матрице ДНК – backtracking, – в результате чего активный центр теряет контакт с 3'-концом новосинтезированной РНК. Реактивация арестованных смещенных ЭК происходит при отщеплении 3'-конца РНК-транскрипта, вытесненного во вторичный канал, в результате чего в активном центре РНКП вновь оказывается 3'-концевой рибонуклеотид, комплементарно связанный с ДНК-матрицей. Хотя эндонуклеазное

разрезание транскрипта осуществляется в активном центре самой РНКП, скорость этой реакции в отсутствие специальных белковых факторов очень низка. У *E. coli* имеется 2 фактора, которые способствуют разрезанию транскрипта – белки GreA и GreB. Оба белка имеют молекулярную массу около 19 кДа и гомологичны по аминокислотной последовательности. Гомологи Gre-факторов были обнаружены у большинства видов бактерий (в частности, у *T. aquaticus* и *T. thermophilus*), что свидетельствует о важности катализируемой ими реакции. Как было показано, роль данных факторов заключается в стимуляции связывания каталитических ионов магния в активном центре РНКП. Согласно данным рентгеноструктурного анализа, белки GreA и GreB состоят из двух примерно равных по размерам доменов, соединенных между собой гибкой междоменной петлей. N-концевой домен белка представляет собой вытянутую coiled-coil структуру из двух антипараллельных  $\alpha$ -спиралей. С-концевой домен белка является глобулярным, и состоит из  $\alpha$ -спирали и  $\beta$ -листа ( $\beta$ -sheet), образованного пятью слоями. Ни один из доменов по отдельности не способен стимулировать реакцию разрезания транскрипта. Моделирование комплекса РНКП с Gre-факторами, проведенное на основе имеющихся биохимических данных, показало, что связывание Gre-факторов, вероятно, происходит внутри вторичного канала РНКП, при этом глобулярный С-концевой домен располагается у входа во вторичный канал, а  $\alpha$ -спиральный N-концевой домен проникает вглубь канала и направлен в сторону активного центра РНКП. Два абсолютно консервативных аминокислотных остатка (Asp41 и Glu44 у *E. coli*), расположенных на конце N-концевого домена Gre-факторов, играют ключевую роль в реакции эндонуклеазного расщепления транскрипта. Моделирование структуры комплекса РНКП с GreB показывает, что эти отрицательно заряженные аминокислотные остатки участвуют в координации второго иона магния в активном центре, необходимого для катализа реакции расщепления. Данное предположение подтверждается тем, что замены данных аминокислотных остатков на незаряженные приводят к потере функциональной активности белка. Так как Gre-факторы взаимодействуют со смещенными ЭК, описанный выше механизм взаимодействия предполагает, что GreA/B и 3'-конец новосинтезированной РНК могут находиться во вторичном канале одновременно. Действительно, на одной из сторон N-концевого домена белков GreA/B находится кластер положительно заряженных аминокислот. По-видимому, именно этой частью N-концевой домен Gre-белков взаимодействует с РНК, находящейся во вторичном канале в смещенном ЭК. Взаимодействия Gre-белков с РНК, вытесненной во вторичный канал, вероятно, стабилизируют связывание Gre-белков с ЭК. Можно предположить, что эти взаимодействия обеспечивают

специфичность действия Gre-белков *in vivo*, за счет стимуляции их связывания именно со смещенными ЭК.

Еще одной возможной, но пока недостаточно изученной функцией Gre-факторов *in vivo*, вероятно, является активация перехода от инициации к элонгации транскрипции. Экспериментально показано, что Gre-белки значительно повышают эффективность ухода РНКП с промотора и снижают уровень abortивной инициации. Хотя точный механизм этого процесса пока остается невыясненным, можно предположить, что связывание Gre внутри вторичного канала в ходе инициации может каким-то образом стабилизировать 3'-конец коротких РНК-транскриптов в активном центре РНКП, препятствуя их диссоциации из инициаторного комплекса.

Поиск гомологов Gre-белков *E. coli* в геноме *T. thermophilus* позволил выявить белок Gfh1, гомологичный гену GreA. Gfh1 (от *англ.* Gre factor homolog 1) обладает в несколько раз большей аффинностью к РНКП, чем GreA. Gfh1 в разной мере ингибирует все каталитические активности РНКП. Самый выраженный эффект Gfh1 оказывает на реакции синтеза РНК-транскрипта и пирофосфоролиза (50-кратное ингибирование); действие Gfh1 на экзо- и эндонуклеазное расщепление РНК менее выражено. Кроме этого, Gfh1 конкурирует с Gre-белками за связывание с РНКП, а так как Gre-белки осуществляют реактивацию арестованных элонгационных комплексов и способствуют переходу от инициации к элонгации транскрипции, данная конкуренция может оказывать дополнительный ингибирующий эффект на транскрипцию в целом. В 2006 году структуры белка Gfh1 из *T. thermophilus* и его гомолога из *T. aquaticus* были расшифрованы тремя независимыми группами исследователей. Структура данного транскрипционного фактора очень похожа на структуру Gre-белков. Gfh1 имеет несколько структурных отличий от GreA и GreB, которые, по-видимому, являются существенными для его функции. Наиболее заметным из них является иное взаимное расположение N- и C-концевого доменов: по сравнению с Gre-факторами, в Gfh1 C-концевой домен повернут почти на 180° относительно N-концевого домена. Если принять, что практически идентичные по структуре C-концевые домены Gre и Gfh взаимодействуют с РНКП одинаковым образом, то N-концевой домен Gfh1 в данной конформации просто геометрически неспособен войти во вторичный канал. C-концевой домен Gfh1 может поворачиваться относительно N-концевого домена так, что конформация белка оказывается практически идентичной Gre. Gfh1 способен «переключается» между активной и неактивной конформациями в зависимости от значения рН среды, причем «точка перехода» лежит в области физиологических значений рН (от 6,4 до 7,5). Физиологическое значение данного регуляторного

механизма пока не до конца ясно. В то же время, описанный случай представляет собой яркий пример участия вторичного канала РНКП в тонкой регуляции активности фермента и показывает, что РНКП способна дискриминировать активную и неактивную конформации транскрипционных факторов. Сравнение аминокислотных последовательностей Gfh1 и Gre-факторов показывает, что наибольшее количество различий между ними сосредоточено в петле, соединяющей две  $\alpha$ -спирали в дистальной части N-концевого домена и в прилегающих к ней участках. Как было показано ранее, именно этот участок ответственен за стимуляцию связывания ионов магния в активном центре РНКП Gre-белками. Вероятно, которой Gfh1 ингибирует активность РНКП, координируя второй ион магния в активном центре таким образом, что возможность последнего участвовать в катализе значительно снижается.

Еще один белок *E. coli*, гомологичный Gre-факторам – Rnk (regulator of nucleoside kinase). Мутации в гене, кодирующем данный белок, снижают уровень нуклеозиддифосфат-киназы в клетках *E. coli*. Рентгеноструктурный анализ белка Rnk показал, что структура его С-концевого домена очень сходна со структурой С-концевого домена у белков Gre и Gfh1, в то время как N-концевой coiled-coil домен имеет гораздо меньшую длину и не содержит отрицательно заряженных аминокислотных остатков на своем конце. На основании сходства структур С-концевых доменов было предположено, что Rnk, вероятно, связывается с тем же участком в области вторичного канала, что и факторы Gre и Gfh1. Rnk подавляет стимуляцию нуклеазной активности РНКП Gre-факторами, а также дестабилизацию промоторных комплексов фактором DksA. Однако Rnk сам по себе не влияет на транскрипционную активность РНКП и свойства промоторных комплексов. В отличие от Gre-факторов, Rnk не взаимодействует с 3'-концом РНК в смещенных элонгационных комплексах, что, по-видимому, объясняется меньшей длиной N-концевого домена. Таким образом, функцией Rnk может являться регуляция взаимодействий с РНКП различных факторов, связывающихся в области вторичного канала, в том числе Gre-белков. Функциональная роль данной регуляции *in vivo* остается неизвестна.

## **Раздел 9. Регуляция терминации транскрипции у бактерий**

Терминация транскрипции также тонко регулируется. Rho-независимый путь терминации является точкой действия многих регуляторных факторов в том числе и фаговой природы. Основной регуляторный процесс на этом этапе – антитерминация. Факторы антитерминации могут выполнять свою функцию, взаимодействуя с РНК и препятствуя образованию РНК-шпильки или работе Rho-фактора (пассивная

антитерминация), а могут взаимодействовать с РНКП (активная, или процессивная антитерминация). Последние не проявляют специфичности действия по отношению к каким-либо определенным терминаторам и позволяют преодолевать несколько последовательно расположенных терминаторов. Пассивная антитерминация играет важнейшую роль в регуляции экспрессии многих оперонов бактерий. 5'-концевая (лидерная) последовательность РНК таких оперонов способна к формированию двух альтернативных структур: G/C-богатой терминирующей и антитерминирующей шпилек. Первая представляет собой классическую терминирующую структуру, а антитерминирующую шпильку РНКП способна преодолевать. На конформацию лидерной последовательности могут влиять различные специфические агенты: белки, рибосомы, тРНК, малые молекулы. Известен белок NutP, регулирующий транскрипцию оперона утилизации гистидина у *B. subtilis*. У многих грамотрицательных бактерий рибосомы участвуют в регуляции транскрипции оперонов синтеза аминокислот. Сигналом для повышения экспрессии в этом случае служит понижение содержания соответствующих аминоацил-тРНК в клетке. Триптофановый оперон *E. coli* содержит транслируемую лидерную последовательность, в которой подряд расположены два триптофановых кодона. Если концентрация трп-тРНК в клетке высока, то рибосома быстро проходит этот участок, образуется терминирующая шпилька и элонгационный комплекс диссоциирует. В случае пониженной концентрации трп-тРНК рибосома задерживается на этом участке. За это время успевает образоваться антитерминирующая шпилька и РНКП транскрибирует оперон. Концентрация ионов и малых молекул, таких как флавин мононуклеотид, тиамин пирофосфат, S-аденозил-L-метионин, также может влиять на конформацию лидерной последовательности мРНК и стабилизировать терминирующую или, что чаще, антитерминирующую конформации. Процесс транскрипции в бактериальных клетках тесно сопряжен с процессом трансляции. Если на синтезирующейся молекуле мРНК не идет трансляция, то транскрипция терминируется Rho-зависимым путем. Согласно одной модели рибосомы механически предотвращают связывание Rho-фактора с мРНК. Согласно другой – рибосома образует контакты с белком NusG, ассоциированным с РНКП, предотвращая его взаимодействие с Rho-фактором, что важно для терминации транскрипции. Помимо этого, рибосома может толкать РНКП вперед, препятствуя возникновению транскрипционных пауз и смещенных элонгационных комплексов.

## Раздел 10. Антитерминаторные факторы транскрипции

Пассивная антитерминация позволяет РНКП преодолеть конкретный терминатор, в то время как процессивный механизм повышает вероятность того, что РНКП транскрибирует весь оперон. Фаговые факторы антитерминации эффективно предотвращают терминацию транскрипции вне зависимости от её механизма, в то время как клеточные обычно обладают меньшей силой действия в отношении механизма Rho-независимой терминации.

Среди клеточных факторов элонгации транскрипции у бактерий существует группа факторов, способных существенно влиять на эффективность терминации. К ним относятся белки Nus (N utilizing substance), RfaH. Среди Nus лучше всего изучены белки NusA и NusG. Белок NusA сам по себе стимулирует регуляторные паузы в процессе транскрипции, а также усиливает Rho-независимую терминацию. В то же время, вместе с другими белками NusA способен входить в состав антитерминаторного комплекса. По-видимому, при переходе комплекса из инициаторного в элонгационное состояние сначала разрываются слабые контакты между доменом 4  $\sigma$ -субъединицы и кор-ферментом, что позволяет NusA связаться с РНКП. Таким образом, РНКП оказывается одновременно связанной с факторами  $\sigma$  и NusA. После этого сродство  $\sigma$ -субъединицы к РНКП падает, и она диссоциирует из комплекса. Влияние NusA на терминацию связано с тем, что, находясь вблизи канала для выхода РНК, он может оказывать воздействие на изменение его ширины, а также взаимодействовать с новосинтезированной РНК и связанными с ней факторами. На значение фактора NusG указывает его крайне высокая консервативность. NusG и его гомологи Spt5 у архей и эукариот обнаружены в клетках всех трех доменов живых организмов. NusG подавляет паузы транскрипции и снижает эффективность терминации. NusG состоит из двух доменов. NusG является важным регуляторным элементом как в процессе Rho-зависимой терминации, так и в сопряжении транскрипции с трансляцией. Факторы NusA и NusG играют важную роль в образовании антитерминирующего комплекса при транскрипции оперонов рРНК. В состав такого комплекса входят белки NusA, NusB, NusG и рибосомные белки S4 и S10. Помимо этого новосинтезирующиеся рРНК защищаются рибосомными белками от взаимодействий с Rho-факторами, а rut-сайты входят в состав вторичных структур, что препятствует их узнаванию Rho-факторами. Потребность в таких механизмах антитерминации при синтезе рРНК связана с тем, что в этом случае транскрипция не сопряжена с трансляцией, а это облегчает Rho-зависимую терминацию. Белок RfaH является паралогом NusG и схож с ним по строению и функционированию N-концевого домена. С-концевой домен RfaH узнает и связывает специфическую 12-нуклеотидную

последовательность на нематричной цепи ДНК. Эта последовательность находится в 5'-концевой нетранслируемой области некоторых оперонов и носит название *ops* (operon polarity suppressor). В отличие от NusG, RfaH привлекается в элонгационный комплекс при транскрипции довольно узкого круга оперонов, содержащих *ops* элементы. RfaH ингибирует Rho-зависимую терминацию тремя способами. Во-первых, RfaH вытесняет NusG из комплекса с РНКП. Во-вторых, RfaH снижает вероятность возникновения транскрипционных пауз. В-третьих, RfaH может привлекать рибосомы на синтезируемый транскрипт. Недавние исследования показали, что С-концевой домен этого белка может быть уложен двумя способами. Когда RfaH не связан с РНКП, его С-концевой домен составлен  $\alpha$ -спиралями, а при образовании контактов с РНКП вторичная структура резко меняется на  $\beta$ -тяжи. В таком состоянии С-концевой домен RfaH похож на аналогичный участок NusG и способен связываться с рибосомным белком S10.

Среди фаговых факторов антитерминации транскрипции наиболее известны белки N и Q фага  $\lambda$ . Опероны этого вируса содержат терминаторы, что определяет важную роль этих белков в процессе регуляции экспрессии генов, в частности в переключении профиля экспрессии фаговых оперонов в ходе жизненного цикла вируса. Белок N переключает транскрипцию с ранних генов на средние, а Q предотвращает терминацию транскрипции в позднем опероне фага  $\lambda$ . Белок N связывается с РНК-шпилькой, называемой *nut* (N utilization), и предотвращает терминацию транскрипции вблизи этой шпильки. Кроме того, белок N может образовывать контакты с клеточными транскрипционными факторами NusA, NusB, NusG и рибосомным белком S10, образуя способствуя формированию антитерминационного комплекса, аналогичного комплексу, формирующемуся при транскрипции оперонов рРНК. В составе этого антитерминирующего комплекса белок N препятствует диссоциации элонгационного комплекса РНКП, подавляя и Rho-зависимую, и Rho-независимую терминацию. Белок Q связывается со специфической последовательностью двуцепочечной ДНК *qut* (Q utilization) и регионом 4  $\sigma^{70}$ -субъединицы РНКП. Белки N и Q также препятствуют образованию пауз в ходе транскрипции. Таким образом, фаговые регуляторы транскрипции используют весь арсенал доступных регуляторных механизмов, изменяя активность РНК-полимеразы в соответствии с потребностями фага.