

# ОСНОВИ НА МОЛЕКУЛЯРНАТА БИОЛОГИЯ



Галина Яхубян

2025

## **Основи на МОЛЕКУЛЯРНАТА БИОЛОГИЯ**

Галина Яхубян

Университетско издателство Пловдивски Университет “Паисий Хилендарски”  
ISBN ... 978-619-281-026-9

2025

Онлайн версия на учебника: 1.0  
(Източник на изображенията в корицата Imagen3)

---

Този учебник има за цел да предостави основни знания за сложните механизми, които стоят в основата на живота на молекулярно ниво. Представени са основните концепции за ДНК, РНК и протеините — строителните блокове на живота, и за процесите на репликация, транскрипция, трансляция и регулиране на експресията на гените. С акцент върху класическите експерименти и съвременните технологии, учебникът има за цел да предостави цялостно разбиране за това как молекулярната биология формира нашето познание за генетиката, клетъчните процеси и биотехнологиите.

Учебникът е предназначен за студенти, които изучават биологически науки, той може да бъде използван и от студенти по медицина, както и от студенти, обучавани в биотехнологични направления.

## СЪДЪРЖАНИЕ

<b>ЦЕНТРАЛНА ДОГМА на Молекулярната биология .....</b>	<b>5</b>
<b>ГЛАВА 1 НУКЛЕИНОВИ КИСЕЛИНИ - ДНК и РНК .....</b>	<b>7</b>
<b>1.1. ДНК – носител на генетична информация .....</b>	<b>7</b>
<b>1.2. РНК – носител на генетична информация.....</b>	<b>10</b>
<b>1.3. Структура на ДНК и РНК .....</b>	<b>11</b>
1.3.1. Първична структура на ДНК и РНК.....	11
1.3.2. Вторична структура на ДНК .....	14
1.3.3. Вторична структура на РНК.....	18
<b>1.4. Динамични свойства на нуклеиновите киселини .....</b>	<b>19</b>
1.4.1. Денатурация и ренатурация на нуклеиновите киселини .....	19
1.4.2. Хибридизация.....	21
<b>1.5. Нуклеази и тяхната роля в метаболизма на нуклеиновите киселини .....</b>	<b>21</b>
<b>1.6. Хроматин – отличителен белег на еукариотната ДНК .....</b>	<b>24</b>
<b>ГЛАВА 2 РЕПЛИКАЦИЯ на ДНК.....</b>	<b>27</b>
<b>2.1. Въведение в репликацията .....</b>	<b>27</b>
2.1.1. Репликация – полуконсервативен механизъм .....	27
2.1.2. Репликон .....	28
2.1.3. Посока на репликацията.....	28
2.1.4. ДНК полимерази.....	29
2.1.5. Непрекъснат синтез на ДНК във водещата верига и прекъснат синтез на ДНК в изоставащата верига – фрагменти на Оказаци.....	30
<b>2.2. Основни етапи на репликация .....</b>	<b>31</b>
2.2.1. Инициация на репликацията.....	31
2.2.2. Удължаване на ДНК .....	34
2.2.3. Прекратяване на репликацията .....	37
<b>2.3. Точност на репликацията.....</b>	<b>39</b>
<b>2.4. Репликация на митохондриална ДНК .....</b>	<b>40</b>
<b>ГЛАВА 3 МУТАЦИИ и ПОПРАВКА на ДНК .....</b>	<b>42</b>
<b>3.1. Генни мутации – класификация на ниво ДНК, протеин и функция .....</b>	<b>42</b>
3.1.1. Мутации на ниво ДНК .....	43
3.1.2. Мутации на ниво протеин .....	44
3.1.3. Мутации на функционално ниво .....	45
<b>3.2. Спонтанни и индуцирани мутации .....</b>	<b>46</b>
3.2.1. Спонтанните мутации са следствие на естествени биологични събития .....	46
3.2.2. Индуцираните мутации възникват под въздействие мутагенни фактори.....	49
<b>3.3. Поправка на ДНК.....</b>	<b>51</b>
3.3.1. Директна поправка.....	53
3.3.2. Поправка чрез изрязване .....	53
<b>3.4. Соматични и зародишни мутации и ефекти върху потомството .....</b>	<b>58</b>
<b>3.5. Малки мутации с големи ефекти - мутации на регулаторни гени .....</b>	<b>60</b>
<b>ГЛАВА 4 СИНТЕЗ на РНК - ТРАНСКРИПЦИЯ.....</b>	<b>62</b>
<b>4.1. Въведение в транскрипцията.....</b>	<b>62</b>
4.1.1. Транскрипционна единица .....	62

4.1.2.	РНК полимерази .....	63
4.1.3.	Основни етапи на транскрипцията .....	64
<b>4.2.</b>	<b>Прокариотна транскрипция.....</b>	<b>64</b>
4.2.1.	Прокариотна РНК полимераза .....	64
4.2.2.	Бактериални промотори .....	66
4.2.3.	Инициация на транскрипцията - роля на $\sigma$ -фактора .....	66
4.2.4.	Удължаване на транскрипта.....	67
4.2.3.	Прекратяване на транскрипцията.....	68
<b>4.3.</b>	<b>Еукариотна транскрипция .....</b>	<b>69</b>
4.3.1.	Еукариотни РНК полимерази .....	69
4.3.2.	Еукариотни промотори .....	71
4.3.3.	Инициация на транскрипцията – роля на главните транскрипционни фактори за сглобяване на пре-инициаторния комплекс.....	74
4.3.4.	Удължаване на транскрипта.....	77
4.3.5.	Прекратяване на транскрипцията.....	78
<b>ГЛАВА 5</b>	<b>ПОСТ-ТРАНСКРИПЦИОННА ОБРАБОТКА на РНК .....</b>	<b>79</b>
<b>5.1.</b>	<b>Обработка на иРНК при еукариоти .....</b>	<b>80</b>
5.1.1.	Модификации на краищата на иРНК .....	80
5.1.2.	Сплайсинг (снаждане на екзони) .....	82
5.1.3.	Редактиране на иРНК.....	87
<b>5.2.</b>	<b>Обработка на рРНК и тРНК .....</b>	<b>88</b>
5.2.1.	Обработка на рРНК.....	88
5.2.2.	Обработка на тРНК .....	90
<b>5.3.</b>	<b>Продължителност на живот на РНК .....</b>	<b>91</b>
<b>ГЛАВА 6</b>	<b>СИНТЕЗ на ПРОТЕИНИ - ТРАНСЛАЦИЯ.....</b>	<b>93</b>
<b>6.1.</b>	<b>Въведение в транслацията .....</b>	<b>93</b>
6.1.1.	Генетичен код .....	93
6.1.2.	тРНК – преводачи на генетичния код .....	95
6.1.3.	Рибозомите - апаратът за синтез на протеини .....	97
<b>6.2.</b>	<b>Основни етапи на транслацията.....</b>	<b>99</b>
6.2.1.	Инициация на транслацията.....	99
6.2.2.	Удължаване на полипептидната верига .....	102
6.2.3.	Прекратяване на транслацията .....	104
<b>ГЛАВА 7</b>	<b>РЕГУЛАЦИЯ на ГЕННАТА ЕКСПРЕСИЯ.....</b>	<b>106</b>
<b>7.1.</b>	<b>Регулация на генната експресия при прокариоти.....</b>	<b>107</b>
7.1.1.	Бактериалните гени са организирани в оперони и се контролират координирано.....	107
7.1.2.	Лактозен оперон ( <i>lac</i> -оперон)- индуцируем оперон .....	108
7.1.3.	Триптофанов оперон ( <i>trp</i> -оперон) – репресируем оперон.....	111
<b>7.2.</b>	<b>Регулация на генната експресия при еукариоти .....</b>	<b>115</b>
7.2.1.	Основни цели на контрола на генната експресия при еукариоти .....	116
7.2.2.	Основни нива и механизми на регулация на генната експресия при еукариоти .....	117
7.2.3.	Координирано-регулираните гени не са свързани при еукариоти .....	121
<b>ГЛАВА 8</b>	<b>ГЕНОМНА ОРГАНИЗАЦИЯ.....</b>	<b>123</b>
<b>8.1.</b>	<b>Прокариотен геном (бактерии и археи).....</b>	<b>123</b>
<b>8.2.</b>	<b>Еукариотен геном.....</b>	<b>124</b>
<b>ПРИЛОЖЕНИЯ .....</b>	<b>129</b>	
<b>РЕФЕРЕНЦИИ .....</b>	<b>139</b>	
<b>Най-често срещани съкращения .....</b>	<b>141</b>	

## Добре дошли в завладяващия свят на Молекулярната биология!

Молекулярната биология е динамична, интердисциплинарна област на биологията, която се занимава с молекулярните основи на биологичната активност. В центъра на молекулярната биология стои изследването на структурата, функцията и взаимодействията на биомолекулите - като дезоксирибонуклеиновите киселини (ДНК), рибонуклеиновите киселини (РНК) и протеини - и как тези взаимодействия управляват процесите, съществени за живота.

Молекулярната биология не съществува в изолация - тя е тясно свързана с генетиката и биохимията. Въпреки че молекулярната биология споделя обща основа с биохимията и генетиката — всяка от които разглежда детайлите за функционирането на организмите на молекулярно ниво — тези дисциплини имат различни области на акцент. Биохимията се фокусира главно върху химическите процеси в живите организми, докато генетиката се занимава с наследствеността и вариацията в организмите. Молекулярната биология свързва тези полета, като изследва как генетичната информация се превръща в функционални молекули и как тези молекули взаимодействат в биологичните системи. Заедно с биоинформатиката, тези области допринасят за цялостното разбиране на клетъчната биология и биологията на развитието, имунологията и системната биология. Тази взаимовръзка позволява на изследователите да интерпретират структурни данни и регулаторни мрежи, които са от съществено значение за разбирането на сложни биологични системи.

Молекулярната биология е експериментална наука, където молекулярните биолози провеждат експерименти, за да разкрият структурата, функцията, обработката, регулирането и еволюционната динамика на биологичните молекули. Молекулярните биолози предимно се фокусират върху гените и протеините. Протеините са жизненоважни макромолекули, които изпълняват множество функции в клетките, включително катализиране на биохимични реакции, осигуряване на структурна подкрепа и улесняване на комуникацията между клетките. Гените от своя страна кодират информацията, необходима за синтезиране на протеини и регулиране на тяхната активност. Този фокус върху гените и протеините позволява на молекулярните биолози да разкрият сложността на клетъчните функции и техните последици за здравето и заболяванията.

В заключение, молекулярната биология служи като основополагаща област за разбирането на живота на най-фундаментално ниво. Чрез изучаване на сложните отношения между биомолекулите и техните роли в клетъчните процеси, тази област не само подобрява нашето разбиране за биологичните механизми, но също така прокарва пътя за напредък в медицината, биотехнологиите и екологичната наука. Докато продължаваме да изследваме сложността на живота през молекулярния обектив, откриваме нови възможности за иновации и открития, които могат да окажат дълбоко влияние върху нашия свят.

Надявам се този учебник да ви осигури основни знания и критично мислене, необходими за навигация в постоянно развиващата се област на молекулярната биология!

## ЦЕНТРАЛНА ДОГМА на Молекулярната биология и нейният съвременен прочит

ДНК е основната молекула, която съхранява генетичната информация и предоставя инструкциите за живота. Пълната колекция от ДНК молекули на живия организъм съставлява неговия геном.

Физически геномът може да бъде разделен на множество различни ДНК молекули, или хромозоми. Функционално геномът е разделен на гени. Генът е уникална ДНК последователност, която съдържа инструкции за производство на функционален продукт, обикновено протеинова молекула (протеин-кодиращ ген) или РНК молекула (некодиращ ген). Той има точно определена локализация върху хромозомата (locus). Генът действа като единица на наследствеността, предавана от родител на потомство и играе ключова роля в развитието, физиологията и наследяването на генетичните белези.

Сам по себе си геномът не играе активна роля в развитието на организма. Огромното количество информация, което се съхранява в ДНК, носи инструкции, които придобиват смисъл едва тогава, когато те бъдат изпълнени чрез последователния синтез на РНК и протеини в организма в подходящия момент и в съответните клетки – сложен процес, означаван като експресия, или изразяване, на генетичната информация. Протеините имат незаменима роля в развитието и функционирането на един организъм: те могат да формират част от структурата на организма (структурни протеини), да катализират метаболитни реакции (ензими), да контролират генната експресия (транскрипционни фактори), да участват в сигнализацията (рецептори) и клетъчна идентичност (мембранни протеини).

За първи път процесът, чрез който инструкциите в ДНК се превръщат във функционален продукт е предложен през 1958 г. от Франсис Крик, откривател на структурата на ДНК. Тогава той го нарича „Централната догма“, тъй като обяснява потока на генетичната информация, от ДНК към РНК, до крайния функционален продукт - протеините.

Централната догма постулира:

- √ ДНК съдържа информацията, необходима за производството на всички клетъчни протеини, а РНК е пратеник, който пренася тази информация до рибозомите;
- √ рибозомите служат като фабрики в клетката, където информацията се „превежда“, чрез прилагане на код, във функционален продукт;
- √ процесът, чрез който инструкциите на ДНК се превръщат във функционален продукт, се нарича генна експресия;
- √ генната експресия има два ключови етапа – транскрипция (синтез на РНК) и транслация (синтез на протеини);
- √ при транскрипцията информацията в ДНК на всяка клетка се преобразува в малки преносими РНК съобщения (информационна РНК, иРНК);
- √ по време на транслацията тези съобщения пътуват от мястото, където се намира ДНК, до рибозомите, където се „четат“, за да се произведат специфични протеини.



Централната догма на молекулярната биология остава основополагаща концепция, а съвременните научни постижения разшириха и усъвършенстваха разбирането ни за начина, по който протича и се регулира генетичната информация.

Първоначално се е смятало, че РНК служи само като посредник между ДНК и протеините. Днес знаем, че молекулите на РНК имат разнообразни функции извън това. Така напр. некодиращите РНК не се превеждат в протеини, но играят ключови роли в регулирането на генната експресия. Примери включват микроРНК и дълги некодиращи РНК, които могат да влияят върху стабилността на иРНК, ефективността на транслацията и дори да модифицират структурата на хроматина.

Оригиналната догма се фокусира върху потока на информация по оста ДНК - РНК - протеин. Въпреки това, откритието на обратната транскриптаза (ензим, използван от ретровируси като HIV) показва, че РНК може да се превърне обратно в ДНК, добавяйки още един слой на сложност. Този процес е основен за жизнения цикъл на ретровирусите и също се използва в биотехнологията.

Генната експресия не се определя само от последователността на ДНК, но и от епигенетични модификации като метилиране на ДНК и модификации на хистони. Тези промени могат да променят генната експресия без промяна в генетичния код и понякога могат да се предават на бъдещите поколения. Епигенетиката показва, че потокът от информация може да бъде динамично регулиран въз основа на фактори като околна среда, стрес или стадии на развитие.

Самите протеини могат да регулират генната експресия и са част от сложни обратни механизми. Например, транскрипционните фактори са протеини, които се свързват с ДНК и регулират транскрипцията на специфични гени. Това е предизвикателство в линейния поток от информация, показвайки, че протеините могат да влияят върху процеси нагоре по веригата като транскрипция на гени.

По-голямата част от човешкия геном не кодира протеини. Голямата част от тази ДНК преди се смяташе за „ненужна“, но сега знаем, че некодиращите региони играят важни роли в генната регулация, структурата на хромозомите и стабилността на генома.

Вместо да се разглежда централната догма като линеен процес, съвременната биология я тълкува в контекста на мрежи, които са обект на изучаване от системната биология. Генната експресия, РНК обработката, синтеза на протеини и клетъчната функция са част от взаимосвързани, силно регулирани системи.

## ГЛАВА 1 НУКЛЕИНОВИ КИСЕЛИНИ - ДНК и РНК

През втората половина на 19-ти век, пионерската работа на Грегор Мендел върху унаследяването на белези в грахови растения предполага, че специфични „фактори“ (днес установени като гени) са отговорни за пренасянето на белезите на организма между поколенията. Въпреки първоначално предположение, че протеините служат като наследствен материал, Ейвъри, МакЛауд и МакКарти установяват един век по-късно, че ДНК, която е открита по-рано от Фридрих Мишер, е носител на генетична информация. Тези открития проправят пътя за изследване, разкриващо химическата природа на ДНК и правилата за кодиране на генетична информация, и в крайна сметка довеждат до идеята на Уотсън и Крик за двойно-спиралната структура на ДНК.

### 1.1. ДНК – носител на генетична информация

През 1928 г. Фредерик Грифит, британски генетик, съобщава за експеримент, въз основа на който предполага, че бактериите са способни да прехвърлят генетична информация по между си чрез процес, известен като трансформация. Той установява съществуването на „трансформиращ фактор“, чрез който невирулентна бактерия се превръща във вирулентна.

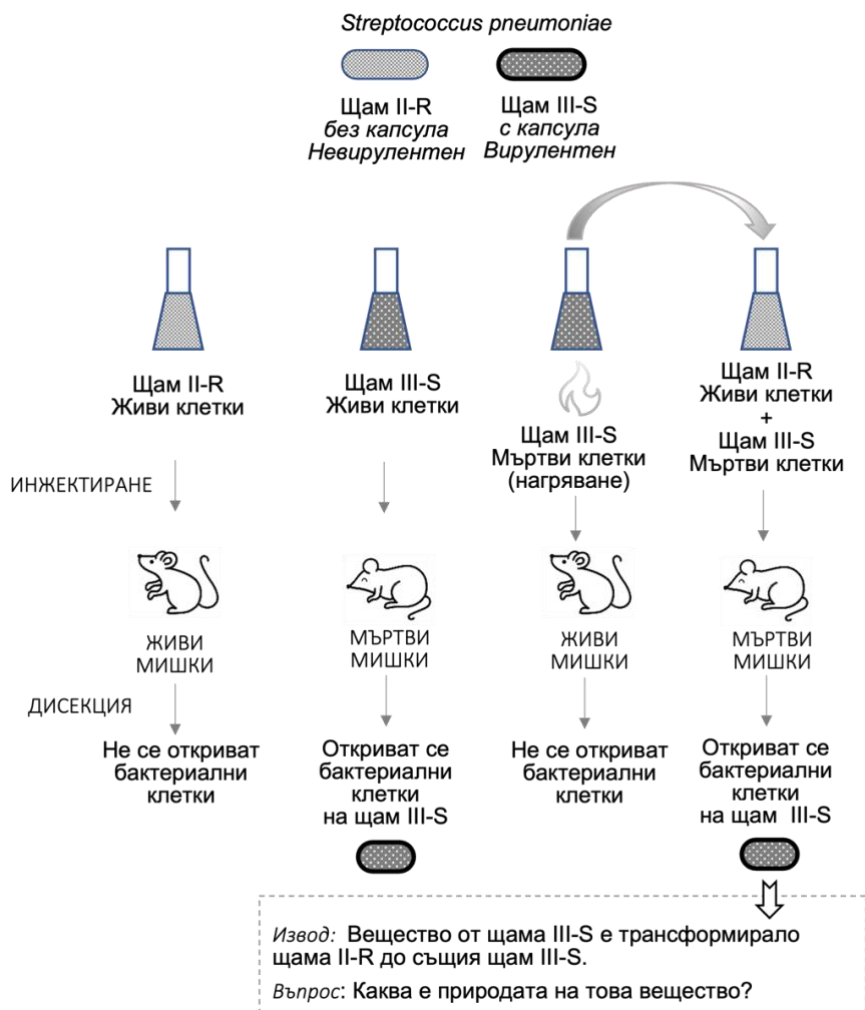
*Frederick Griffith (1928)*



Фред Нойфелд, немски бактериолог, изучаващ пневмококови бактерии в началото на 1900-те, открива 3 имунологично различни щамове на *Streptococcus pneumoniae* (типове I, II и III). Вирулентният щам (тип III) е отговорен за голяма част от смъртността по време на пандемията от испански грип в периода 1918-1920 г. Пандемия е причинена от грипния вирус H1N1 и е причинила смъртта на около 50 милиона души, най-вече защото грипната вирусна инфекция отслабва имунната система на заразените лица, което ги прави податливи на стрептококова инфекция и развитие на пневмония.

През 20-те години на миналия век Фредерик Грифит работи с два щамове *S. pneumoniae* - вирулентен (Тип III-S) и невирулентен (Тип II-R), за да проучи дали клетките на щамове Тип II могат да имунизират мишки срещу вирулентния щам. Между двата типа има една основна разлика; щамът III-S има гладка полизахаридна обвивка, която го прави устойчив на имунната система на мишки, докато щамът II-R няма тази обвивка и се унищожава от имунната система на гостоприемника. Освен това, двата щамове са лесни за разграничаване, тъй като вирулентният III-S щам расте в морфологично гладки колонии (от англ. smooth, S), докато невирулентният II-R щам образува грапави колонии (от англ. rough, R). Полизахаридната обвивка на клетките на III-S щамове прави coloniите да изглеждат гладки с невъоръжено око, докато клетките на II-R щамове II, които нямат такава обвивка, растат в колонии, които не блестят и изглеждат грапави.. Експерименталният дизайн и резултатите, които Грифит публикува през 1928 г., са обобщени на Фигура 1.1.

В експеримента мишките, инжектирани с невирулентния щам (II-R) оцеляват и в тъканите на тези мишки не се откриват бактериални клетки. Мишките, инжектирани с вирулентен (III-S) щам, умират след един ден и в кръвта им се откриват бактериални клетки. Както се очаква, инжектирането в мишки на умъртвени чрез нагряване III-S клетки, не е смъртоносно и в кръвта не се откриват бактериални клетки.



Фигура 1. 1. Експеримент на Грифит

В друг експеримент Грифит инжектира смес от живи II-R клетки и топлинно убити III-S клетки, вероятно с надеждата, че комбинацията ще предизвика имунен отговор в мишките. Можете да си представите изненадата му, когато инжектираните мишки умрат след един ден, а в тъканите им се откриват множество III-S клетки. Грифит осъзнава, че нещо важно се е случило в неговите експерименти, а именно, че в сместа от живи II-R клетки и топлинно убити III-S клетки, някои II-R клетки са се трансформирали с нещо от мъртвите S клетки.

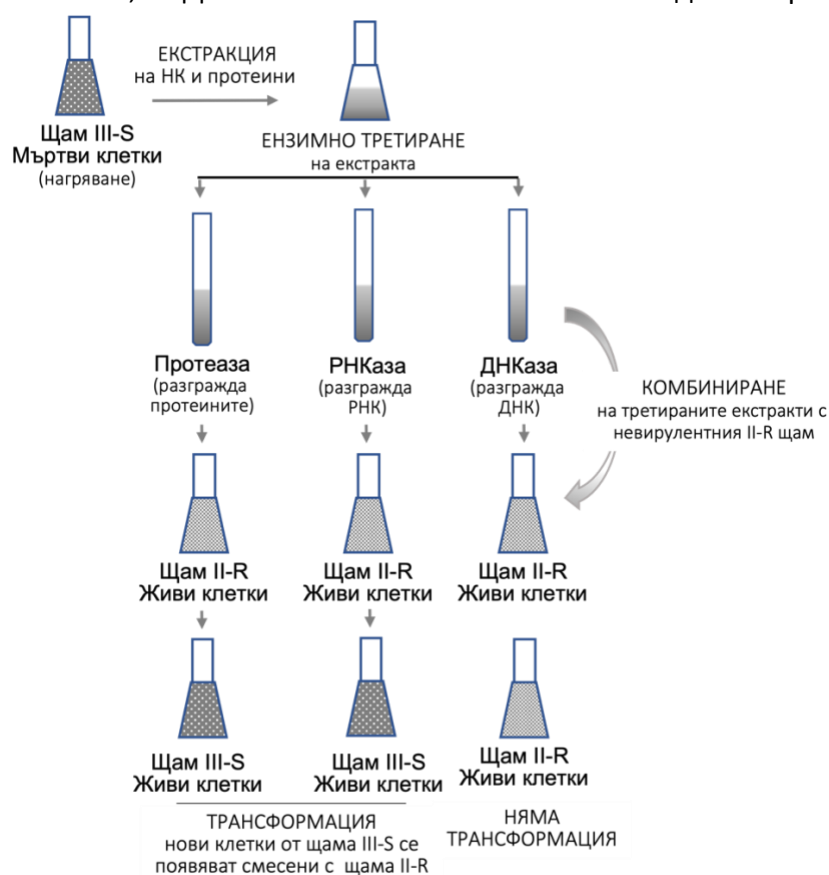
Грифит установява, че вещество (трансформиращ фактор), присъстващо във вирулентния III-S щам може трайно да промени, или трансформира, невирулентния II-R щам в смъртоносния III-S щам.

Но какъв е бил този мистериозен материал, който е превърнал безобидните бактерии в убийци? И как тази промяна е била предадена на следващите потомства?

Последващите експерименти, извършени от Ейвъри, МакЛауд и МакКарти, установяват, че ДНК е трансформация фактор, която пренася генетична информация между две бактериални клетки.

*Avery, MacLeod and McCarty (1944)*

Ейвъри и неговите сътрудници отглеждат вирулентния S щам в течна среда и след това убиват клетките с топлина (Фигура 1.2). Убитите клетки биват третираны с детергент, който разрушава клетъчната обвивка, и освобождава клетъчните компоненти в лизата. Лизатът е бил разделен на равни части, всяка от които е третирана с различен ензим. (1) Когато е прибавен ензим от групата на протеазите, белтъците в лизата се разграждат. Лизатът, в който всички белтъци са разрушени, е смесен с невирулентния R щам, и е установена трансформация на част от R клетките до S клетки. (2) Когато е прибавен ензим от групата на рибонуклеазите (РНКаза), РНК молекулите в лизата се разграждат. След смесване на лизата с R щам, отново се наблюдава трансформация. (3) Когато обаче е прибавен ензим от групата на дезоксирибонулеазите (ДНКаза), ДНК молекулите в лизата се разграждат. След смесване на лизата с R щам, не е наблюдавана трансформация на R до S клетки, което е показало, че ДНК е клетъчният компонент необходим за трансформация.



Извод: ДНК е молекулата отговорна за трансформацията

Фигура 1.2. Експеримент на Ейвъри и сътрудници

Ейвъри и сътрудници установяват, че ДНК е трансформиращият фактор и е носител на генетичната информация.

Нещо повече, след настъпването на трансформацията, капсулният полизахарид се произвежда в поколенията. Следователно трансформацията е наследствена и процесът засяга генетичния материал.

*Alfred Hershey and Martha Chase (1952)*

Хърши и Чейз са работили с бактериофага T2 (вируси, убиващи бактерии), който е изграден изцяло от протеини и ДНК. Всяка вирусна частица действа като молекулярна спринцовка, инжектираща своя генетичен материал в *E. coli*; празната вирусната капсула остава прикрепена към външната страна на клетката. За да се определят дали генетичният материал на вируса е протеин или ДНК, изследователите изолират вирусни частици и ги разделят на две части – в едната част бележат ДНК на вирусните частици с  $^{32}\text{P}$ , а в друга част бележат вирусните протеините с  $^{35}\text{S}$ . Тъй като в ДНК липсва сяра, а в протеините липсва фосфор, тези радиоактивни изотопи предоставят удобен начин за разграничаване на двата типа молекули. Изследователите инфектират *E. coli* с така белязаните вирусни частици, при което те се размножават (реплицират) в бактериалната клетка. Сместа след това е обработена чрез кратко пулсиране в блендер за да се отделят заразените бактерии от празните вирусни капсули. Когато изследователите измерват радиоактивността, те откриват, че голяма част от белязаната с  $^{32}\text{P}$  ДНК е влязла в бактериалните клетки, докато по-голямата част от белязаните с  $^{35}\text{S}$  протеини остават в празните капсули. Инжектираните ДНК молекули карат бактериалните клетки да произвеждат повече вирусна ДНК и протеини.

Проучването показва убедително, че вирусната ДНК, за разлика от вирусния протеин, влиза в клетките на бактериалните гостоприемници. Следователно генетичният материал на вируса е съставен от ДНК. Заедно с проучванията, направени от Ейвъри, МакЛауд и МакКарти, този експеримент потвърждава, че ДНК е основният носител на наследствеността.

## 1.2. РНК – носител на генетична информация

Във всички познати на човека клетки генетичната информация се съхранява трайно в ДНК. При вирусите това е различно. Вирусите не са живи същества в класическия смисъл, тъй като винаги зависят от клетка гостоприемник за своето възпроизвеждане. Независимо от това, вирусите също притежават генетичен материал под формата на нуклеинови киселини, който е необходим за тяхното възпроизвеждане в клетката гостоприемник.

*Fraenkel-Conrat and Singer (1956)*

Изследователите провеждат първият експеримент, който доказва, че РНК също могат да бъдат генетичен материал. Те изучават вирусът на тютюневата мозайка (TMV) - малък растителен вирус, който е кръстен на първото растение, в което е бил открит - тютюн още през 1800 г. Днес е ясно, че той може да заразява над 150 различни видове растения - зеленчуци, плевели и цветя. Вирусът се разпространява механично, навлизайки в растенията чрез наранявания, като симптомите на мозаечната болест включват най-често мозаечни шарки по листата в различни нюанси на зелено и жълто, забавен растеж, завиване на листата. TMV е първият вирус, който е кристализиран и чиято структура е изяснена в подробности. Той е РНК вирус, изграден от единична навита РНК, капсулирана в цилиндрична протеинова обвивка. Различните щамове на TMV се характеризират с различия в химичния състав на техните протеинови обвивки.

Френкел-Конрат разработва за първи път техники за разделяне на протеините и РНК на TMV чрез използване на химически агенти. След това използва изолатите за инфектиране на тютюневи растения, и наблюдава, че протеинът сам по себе си не причинява инфекция в тютюневите листа, докато отделената РНК молекула е достатъчна за развитие на мозаечна болест. Малко по-късно, Френкел-Конрат и Сингър използват два различни щамове на TMV (тип-

А и тип-В), на които разделят РНК от протеиновите обвивки. След това смесват протеините от единия щам с РНК на другия щам и реконструират хибридни вируси. След втриване на хибридните вирусни частици върху живи тютюневи листа, те установяват по фенотипни и генотипни белези, че се произвеждат вируси идентични с щама, от която е изолирана РНК.

Изследванията на Френкел-Конрат и сътрудници показват, че генетичната информация може да бъде съхранявана не само в ДНК, но и в РНК.

По-голямата част от растителните вируси използват РНК за съхранение на данни. В ежедневието си човек среща хиляди РНК вируси. Болести като грип, морбили, хепатит С, ебола, бяс и полиомиелит се причиняват и предават от РНК вируси.

### 1.3. Структура на ДНК и РНК

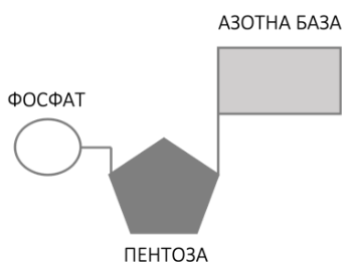
Нуклеиновите киселини са големи полимери, съставени от повтарящи се единици на едни и същи градивни елементи - мономерни, и наподобяват перлена огърлица, направена от много перли. Мономерите на нуклеиновите киселини са нуклеотидите, свързани по между си чрез ковалентни връзки. ДНК е нуклеиновата киселина, която съхранява генетична информация. Ако цялата ДНК в типичната клетка на бозайник се разтегне от край до край, тя ще се простира повече от 2 m. РНК е нуклеиновата киселина, отговорна за използването на генетичната информация, кодирана в ДНК, за производството на хиляди протеини, открити в живите организми.

#### 1.3.1. Първична структура на ДНК и РНК

Последователността или редът на нуклеотидите определя първичната структура на ДНК и РНК.

Всеки нуклеотид е изграден от три основни компонента (Фигура 1.3):

- пентоза (монозахариден остатък с пет С атома)
- азотна база (азотсъдържащо хетероциклично производно)
- фосфатна или полифосфатна група.



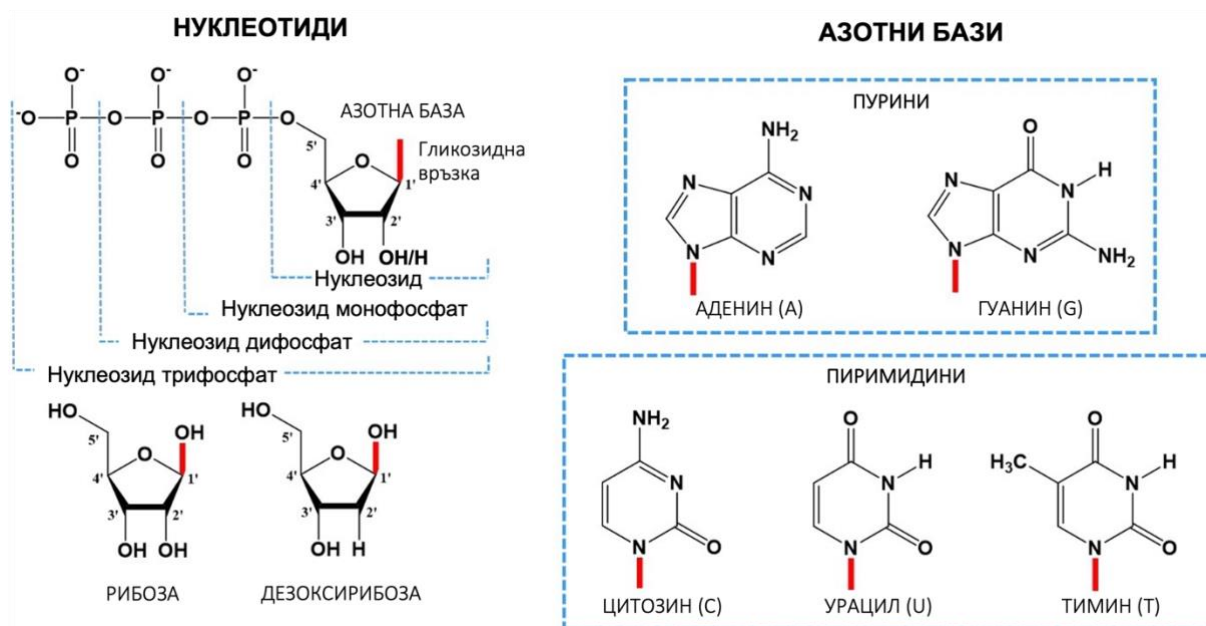
Фигура 1.3. Схематично представяне на структурата на нуклеотида

Когато пентозната захар е рибоза, нуклеотидът се нарича рибонуклеотид, а получената нуклеинова киселина е РНК, а когато е 2-деоксирибоза, нуклеотидът е дезоксирибонуклеотид, а нуклеиновата киселина е ДНК (Фигура 1.4).

Азотните бази, открити в нуклеотидите, се класифицират като пиримидини или пурини. Пиримидините са хетероциклически амини с два N атома в шестчленен пръстен и включват урацил (U), тимин (T) и цитозин (C). Пурините са хетероциклически амини, състоящи се от пиримидинов пръстен, кондензиран с петчленен пръстен с N азотни атома. Аденинът (A) и гуанинът (G) са основните пурини, открити в нуклеиновите киселини.

Азотната база се свързва посредством N1-атома (за пиримидините) или N9-атома (за пурините) от хетероциклическия пръстен с C1-атома на пентоза, посредством N-гликозидна връзка. Така се формира нуклеозид (Фигура 1.4, Таблица 1.1). Когато хидроксилната група при

C5-атома на пентозата се фосфорилира, нуклеозидът се превръща в нуклеотид (нуклеозид монофосфат).



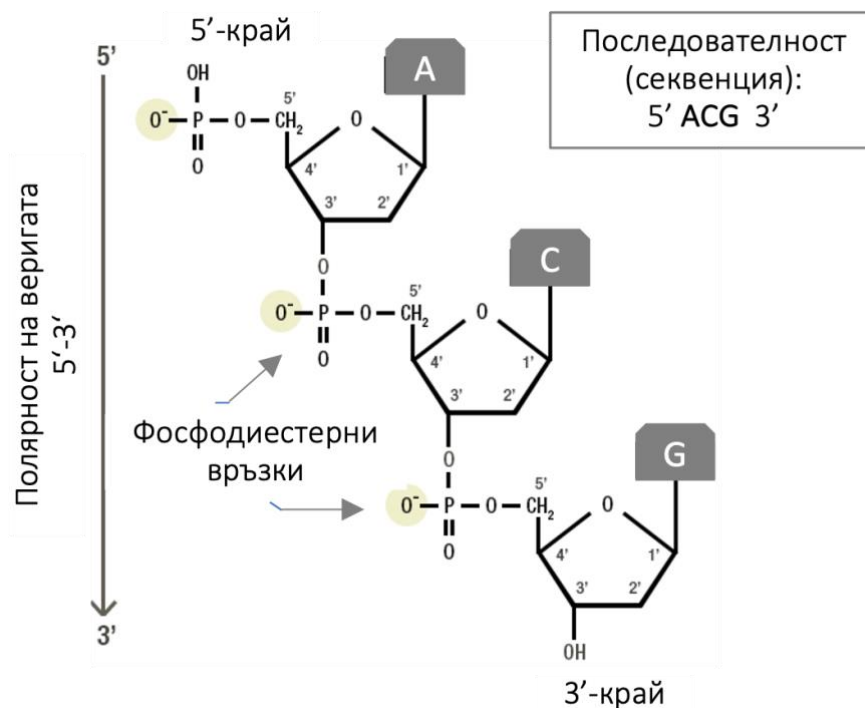
Фигура 1.4. Структура на нуклеотидите, изграждащи нуклеиновите киселини

Таблица 1.1. Наименование на нуклеотидите съгласно Международния съюз по чиста и приложна химия (The International Union of Pure and Applied Chemistry, IUPAC)

АЗОТНА БАЗА	НУКЛЕОЗИД (азотна база + пентоза)	НУКЛЕОТИД (азотна база + пентоза + фосфатна група)	Абревиатура	
Adenine	Adenosine	A	Adenosine mono-, di-, and triphosphate	AMP, ADP, ATP
Cytosine	Cytidine	C	Cytidine mono-, di-, and triphosphate	CMP, CDP, CTP
Guanine	Guanosine	G	Guanosine mono-, di-, and triphosphate	GMP, GDP, GTP
Uracil	Uridine	U	Uridine mono-, di-, and triphosphate	UMP, UDP, UTP
Thymine	Thymidine	T	Thymidine mono-, di-, and triphosphate	TMP, TDP, TTP

Нуклеотидите са свързани в полимерната верига чрез фосфодиестерни връзки (Фигура 1.5). Фосфодиестерна връзка възниква, когато хидроксилните групи във фосфорната киселина реагират с хидроксилни групи на други молекули, образувайки две естерни връзки. По-конкретно, фосфодиестерната връзка между два нуклеотида свързва C3'-атом на пентозата на единия нуклеотид и C5' -атома на пентозата на другия нуклеотид (оттук и името, 3'-5' фосфодиестерна връзка). По този начин при свързването на нуклеотидите в полинуклеотидната верига става редуване на фосфатни и захарни единици (2-деоксирибоза в ДНК и рибоза в РНК), които формират т. нар. захарозо-фосфатен скелет на нуклеиновата киселина, от който се разклоняват пуриновите и пиримидиновите бази.

Всяка фосфатна група има един кисел водороден атом, който се йонизира при физиологично рН (може да приема електрони и да отдава протони) Поради тази причина захарозо-фосфатния скелет е натоварен отрицателно, от където идва и наименованието на нуклеиновите киселини.



Фигура 1.5. Нуклеотидна верига. Представената верига включва 3 нуклеотида Аденин (А), Цитозин (С) и Гуанин (G).

Подобно на протеините, нуклеиновите киселини имат първична структура, която се определя като последователността на техните нуклеотиди (секвенция). За разлика от протеините, които имат 20 различни вида аминокиселини, в нуклеиновите киселини има само 4 различни вида нуклеотиди. При изписване на аминокиселинните последователности на протеините се започва с N-терминалната аминокиселина. При нуклеотидните последователности, крайният нуклеотид в единия край на веригата има свободна 5' фосфатна група – 5' край, докато крайният нуклеотид в другия край има свободна 3' хидроксилна група – 3' край. Последователностите на нуклеиновите киселини обикновено се записват в посока от 5' към 3' – тоест от 5' края до 3' края (Фигура 1.5).

Разбирането на първичната структура на ДНК и РНК е от съществено значение за разшифроването на генетичната информация. Прецизната последователност на нуклеотидите кодира генетичния код, който управлява биологичните функции, регулира генната експресия и в крайна сметка определя видимите характеристики на организмите. Това основополагащо знание е в основата на широк спектър от молекулярни техники, използвани за анализ и манипулиране на генетичен материал.

Появата на технологиите за ДНК секвениране революционизира молекулярната биология, позволявайки на изследователите да четат и интерпретират нуклеотидни последователности с безпрецедентна точност (Приложение 1 и Приложение 2).

### 1.3.2. Вторична структура на ДНК

В края на 50-те години на миналия век наблюдението на Ервин Чаргаф, че базите присъстват в различни количества в ДНК на различните видове, довежда до концепцията, че последователността от бази е формата, в която се пренася генетична информация.

#### *Erwin Chargaff (1944-1950)*



Чаргаф провежда поредица от иновативни експерименти (хартиена хроматография и ултравиолетова спектроскопия) през 40-те години на 20 век, фокусирани върху измерването на основния състав на ДНК в различни видове и органи. Той установява, че: (1) съотношението на пурины към пиримидини е 1; (2) съотношението на аденин (А) към тимин (Т) и на гуанин (G) към цитозин (С), съответно, също са 1; (3) съотношенията на базите на ДНК са сходни в различните органи на един и същи организъм; и (4) съставът на ДНК е видово-специфичен. Освен това той установява такава специфичност за широк спектър от видове, вариращи от бактерии до човек. Неговите резултати опровергават предишни вярвания, според които структурата на ДНК е повторение на четирите бази и следователно монотонна или неспособна да обясни биологичното разнообразие. Преди работата на Чаргаф учените са вярвали, че молекулярната основа на биологичната специфичност се намира в протеините, които са съставени от двадесет вида аминокиселини и по този начин притежават структура, по-благоприятна за отчитане на биологичното разнообразие.

Откритията на Чаргаф предполагат образуване на базовата двойки в ДНК, и помигат на Уотсън и Крик да създадат двойно-спиралния модел на структурата на ДНК.

#### Правило 1

ДНК, независимо от биологичния вид, съдържа равни количества аденин (А) и тимин (Т) и равни количества цитозин (С) и гуанин (G).

#### Правило 2

Съставът на ДНК е видово специфичен.

#### *Rosalind Franklin and Maurice Wilkins (1952)*



Розалинд Франклин и Морис Уилкинс изследват структурата на ДНК, използвайки рентгеновата дифракция като основен инструмент за анализ. При този метод, рентгеновите лъчи, облъчвайки ДНК молекула, рефрактират при контакт с атомите. Профилът на разсейване на рентгеновите лъчи се записва върху филм и се използва за изясняване на детайли от молекулярната структура

Франклин използва в изследванията си две форми на кристализирана ДНК, една от които е по-хидратирана от другата. Тя нарича по-хидратираната форма В-форма, а другата А-форма. Сега знаем, че В-формата на ДНК е основната форма на ДНК в живите клетки, където околната среда е силно хидратирана.

През 1952 г. Франклин, заедно със своя докторант Реймънд Гослинг, прави „Снимка 51“ на В-формата на ДНК с изключително ясен дифракционен профил, която придобива по-късно световна известност. Въз основа на тази снимка, както и на предходни данни от рентгено-структурните анализи на Франклин и Уилкинс, са изведени основните параметри на ДНК молекулата: (1) В-формата на ДНК е правилна спирала, която прави един оборот на всеки 3,4 nm, (2) тя има диаметър около 2 nm; (3) разстоянието между съседните нуклеотиди е 0,34 nm и по този начин трябва да има 10 нуклеотида на оборот. Снимката не показва, без математически анализ, дали има 2, 3 или 4 спирали. Снимка 51 също така загатва как са

подредени някои от химичните групи на ДНК: тежките фосфатни групи лежат от външната страна на спиралата, а базите лежат вътре в спиралата.

Кристалографска снимка на натриев тимонуклеат, тип В. "Снимка 51." Май 1952 г. Оригиналът се съхранява в документите на Ава Хелън и Линус Полинг. Създател: Розалинд Франклин, Реймънд Г. Гослинг



Кристалографска снимка на натриев тимонуклеат, тип В. "Снимка 51." Май 1952 г.

Оригиналът се съхранява в документите на Ава Хелън и Линус Полинг. Създател: Розалинд Франклин, Реймънд Гослинг (натриев тимонуклеат - ДНК, извлечена от тимусната жлеза на теле)

Като се имат предвид тези наблюдения, остават две предизвикателства: да се изработи структурата на ДНК и да се обясни как последователност от бази в ДНК може да определи последователността на аминокиселините в протеина.

#### *James Watson and Francis Crick (1953)*



През 1953 г. Джеймс Уотсън и Франсис Крик публикуват своята известна статия "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid", в която само на две страници описват двойно-спиралната структура на ДНК. В него те предлагат комплементарните бази - аденин и тимин (А-Т), гуанин и цитозин (G-C) - да се свързват помежду си в центъра на двойна спирала, поддържайки заедно двете вериги на ДНК. В самия край на статията, те коментират почти като настрана: „Не ни убягна от вниманието, че специфичното свързване на базите, което постулирахме, веднага предполага възможен механизъм за копиране на генетичния материал.“

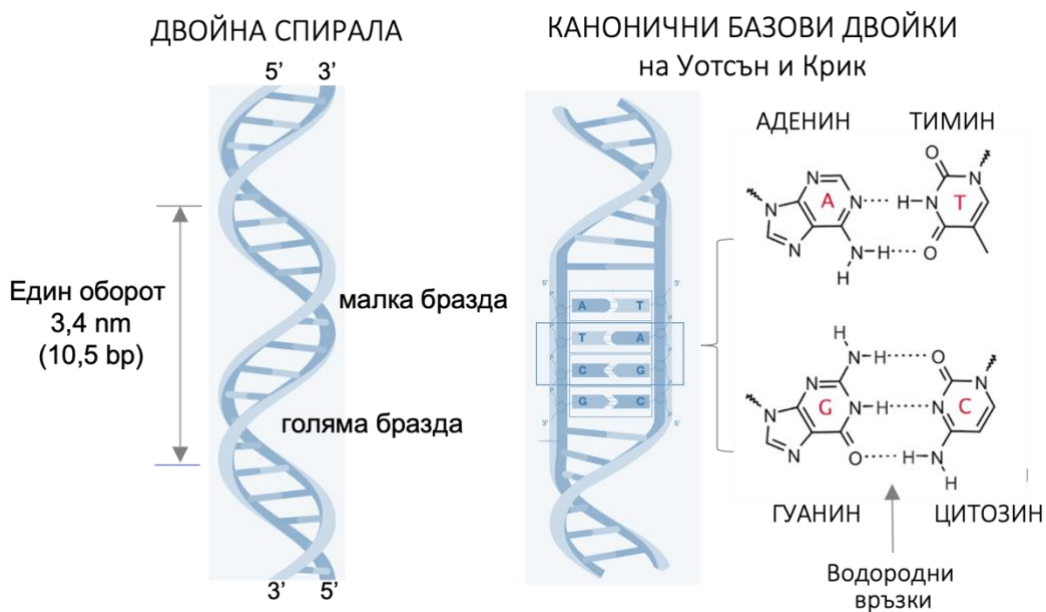
Наистина, един месец след като класическата статия се е появила в печат в списанието Nature, Уотсън и Крик публикуват втора статия, която предполага как може да се копира ДНК. В тази статия те изказват хипотезата, че двете вериги на двойната спирала да се развиват и всяка верига служи като матрица за синтеза на комплементарна дъщерна верига. В техния модел, наречен полуконсервативна репликация, всяка нова молекула на ДНК се състои от една верига, получена от оригиналната родителска молекула, и една новосинтезирана верига.

ДНК е двойно-верижна спирала, която носи генетичните инструкции и може да се самовъзпроизвежда чрез образуване на идентични копия.

Двете полинуклеотидни вериги в двойната спирала на ДНК се свързват чрез водородни връзки между азотните бази (Фигура 1.6). Обикновено G се свързва най-стабилно с C - чрез три водородни връзки, докато A се свързва най-стабилно с T – чрез две водородни връзки. Свързването на двете съответни азотни бази чрез водородни връзки се описва като образуване на базова двойка (от англ. base pair, съкратено bp) или базово сдвояване, а базите

се означават като комплементарни (от англ. complementary - допълващи се), т. е. в ДНК молекулата се образуват комплементарните двойки А-Т и G-C, наричани още двойки на Уотсън и Крик.

Във всяка базова двойка, едната база е по-обемна (пурин с два пръстена) от другата база (пиримидин с един пръстен), пространствено техните структури се напасват най-добре и се разполагат енергийно най-благоприятно във вътрешността на двойната спирала. Захарозо-фосфатният скелет е от външната страна на двойната спирала и носи отрицателни заряди върху фосфатните групи. Базовите двойки (А-Т и G-C) са от вътрешната страна на двойната спирала. Те са плоски и лежат перпендикулярно на оста на спиралата. При това подреждане всяка двойка бази има подобна ширина, като по този начин поддържа захарозо-фосфатните скелети на еднакво разстояние един от друг по протежение на молекулата на ДНК. В следствие на това всяка верига от двойна спирала на ДНК съдържа последователност от нуклеотиди, която е точно комплементарна на нуклеотидната последователност на нейната партньорска верига - А винаги съответства Т от противоположната верига, а С винаги съответства на G. Това обуславя и една от основните характеристики на двойната спирала – комплементарност на двете вериги, която е от решаващо значение, когато става въпрос за копиране и поправка на ДНК. Двете комплементарни вериги имат противоположна полярност - гледайки в една посока по протежение на спиралата, една та верига върви в посока 5'-3', а нейната комплементарна верига върви в посока 3'-5' - антипаралелност на двете вериги.



Фигура 1.6. Двойна спирала на ДНК.

Всяка базова двойка се завърта на около  $36^\circ$  около оста на спиралата спрямо следващата базова двойка, така че приблизително 10 базови двойки правят пълнен оборот от  $360^\circ$ . Усукването на двете вериги една около друга образува двойна спирала с малка бразда с диаметър около  $12 \text{ \AA}$  ( $1,2 \text{ nm}$ ) и голяма бразда с диаметър около  $22 \text{ \AA}$  ( $2,2 \text{ nm}$ ). В В-ДНК двойната спирала е „дясна“ – завъртането е по посока на часовниковата стрелка, гледано по спираловидната ос.

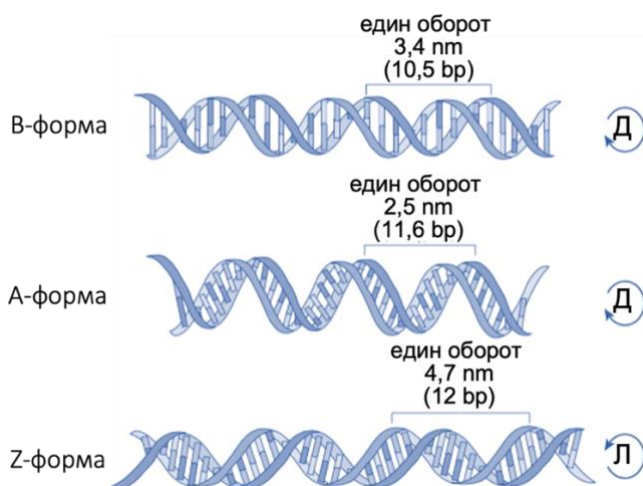
Стабилността на ДНК дуплекса се определя от два фактора - основно от водородните връзки между комплементарните бази на двете вериги, както и от т. нар. стекинг взаимодействия (от

англ. stacking – подреждане на куп) между съседните бази. Тези взаимодействия са хидрофобни и електростатични по природа и зависят от ароматната природа на базите и техните диполни моменти. Степента на стабилизация, осигурена от стекинг взаимодействията, зависи от последователността на ДНК. Някои комбинации от базови двойки образуват по-стабилни взаимодействия от други, така че взаимодействията със съседната база са важни детерминанти на стабилността на дуплекс. Този тип взаимодействия се засилват с увеличаване на солевата концентрация, тъй като високите концентрации на соли маскират дестабилизиращото отблъскване на зарядите между двата отрицателно заредени захарозо-фосфатни скелета.

- Алтернативни форми на ДНК

В допълнение към класическа структура на ДНК са наблюдавани няколко други форми на ДНК (Фигура 1.7, Таблица 1.2). Основните форми на ДНК са:

- В-форма: най-често срещаната форма, присъстваща в повечето ДНК молекули при неутрално рН и физиологична солева концентрация; това е класическата, дясно-въртяща двойно-спирална структура, описана от Уотсън и Крик.
- А-форма: дясно-въртяща двойно-спирална структура с по-голям диаметър и по-късо разстояние между базовите двойки; появява се, когато относителната влажност на околната среда е под 75%, което означава, че рядко присъства в нормално физиологично състояние; често срещана в участъци с редуващи се пурини, напр. АААААА (или пиримидини); характерна за РНК-ДНК дуплекси и РНК-РНК дуплекси.
- Z-форма: ляво-въртяща двойно-спирална структура с ясно изразен зиг-заговиден (оттук и името) модел на захарозо-фосфатния скелет; образува се в участъци с редуващи се пурини и пиримидини, напр. GCGCGC, в които G е в син-конформация (за разлика от В- и А-формите), а C – в анти-конформация; Z-ДНК е трудно да се наблюдава, тъй като е нестабилна; предполага се, че тази различна структура участва по някакъв начин в регулирането на някои клетъчни функции, като транскрипция, но все още не са налични убедителни доказателства за или против това предложение.



Фигура 1.7. Алтернативни форми на ДНК.

В-форма, А-форма и Z-форма  
Д - дясно-въртяща  
Л - ляво-въртяща

Спираловидната структура на ДНК е променлива и зависи от три основни фактора. (1) Околната среда - доказано е, че при по-високи солеви концентрации или с добавяне на не-електролити (например етанол), структурата на ДНК може да се промени от В към А форма. (2) ДНК последователността: А-формата се предпочита от определени участъци с последователни пурини (или пиримидини), докато Z-ДНК може да се образува най-лесно чрез редуващи се пурин-пиримидинови остатъци; (3) Наличието на протеини, които могат да се

свържат с ДНК в една спирална конформация и да я принудят да приеме различна конформация.

Таблица 1.2. Алтернативни форми на ДНК.

ПАРАМЕТРИ	B-форма	A-форма	Z-форма
Посока на спиралата	Дясно-въртяща	Дясно-въртяща	Ляво-въртяща
Диаметър на спиралата (nm)	2,0	2,6	1,8
Базови двойки (бд) на оборот	10,5	11,6	12
Разстояние между бд (nm)	0,34	0,26	0,37
Завъртане	+36°	+33°	-30°

### 1.3.3. Вторична структура на РНК

За разлика от ДНК, РНК са едноверижни молекули и съдържат урацил (U) вместо тимин (Т). Въпреки че РНК са едноверижни, те могат да образуват двойно-верижни вторични структури, които са важни за тяхната функция. Тези вторични структури са резултат от вътрешномолекулно базово сдвояване, което се случва между комплементарни нуклеотиди в рамките на една молекула РНК. Освен каноничните базови двойки на Уотсън и Крик (А-У и G-C), може да се образува и двойка между G и U (G-U) с две водородни връзки. Подобно на ДНК, РНК има два края – 5'- и 3'-край и нуклеотидната последователност се изписва от 5'-края към 3'-края. Всяка РНК има уникална вторична структура, определена от нейната първична структура и съдържа различни структурни мотиви (Фигура 1.8), най-често срещаните от които са: стъбло, стъбло-бримка (фуркета), вътрешна бримка, изпъкналост, псевдовъзел.

Структурата стъбло-бримка се случва, когато две области на една и съща молекула, обикновено палиндромни в нуклеотидна последователност, се сдвояват за да образуват двойно-верижна спирала, завършваща с къса несдвоена бримка. Палиндромната последователност е последователност в молекула на двуверижна ДНК или РНК, при която четенето в определена посока (от 5' към 3') на едната верига е идентично с последователността в същата посока (от 5' към 3') на комплементарната верига. Например, последователност от 5' GAATTC3' на една верига ще бъде комплементирана от 3' CTTAAG 5' на другата верига. При прочитане на последователността в посока от 5'-края към 3'-края, независимо коя е веригата, последователността е GAATTC.



Структурата стъбло-бримка е изключително разпространена и е градивен елемент за по-големи структурни мотиви, като например структурата „детелинов лист“, в която са свързани няколко стъбла-бримки, открита в транспортната РНК (тРНК). Вътрешни бримки (къса поредица от несдвоени бази в по-дълга сдвоена спирала) и изпъкналости (области, в които



Фигура 1.8. Структурни мотиви в едноверижна молекула на РНК.

едната верига на спиралата има "допълнителна" вмъкната база без комплементарна база в другата верига) също са чести.

Псевдовъзелът е вторична структура на нуклеинова киселина, съдържаща най-малко две структури стъбло-бримка, в които половината от едно стъбло е интеркалирано между двете половини на друго стъбло. Псевдовъзелите могат да образуват различни структури с каталитична активност. Например, РНК компонентът на човешката теломераза съдържа псевдовъзел, който е критичен за ензимната активност. Рибозимът на вируса на хепатит делта е добре познат пример за каталитична РНК с псевдовъзел в активното си място. Въпреки че ДНК може също да образува псевдовъзли, те обикновено не присъстват при стандартните физиологични условия.

#### 1.4. Динамични свойства на нуклеиновите киселини

ДНК е гъвкав полимер, който може да претърпи различни конформационни промени. Тази гъвкавост позволява на ДНК да адаптира формата си в отговор на свързване с протеини, промени в температурата или механични сили. Динамиката на ДНК включва специфични конформационни преходи, като отваряне на базовите двойки по време на процеси като репликация и транскрипция. Тези преходи са съществени за достъпността на генетичната информация, закодирана в ДНК. Способността на ДНК да се разплита (разделя на единични вериги) и да се повторно свързва (реформира двойни спирали) е критичен аспект от функцията ѝ в клетъчните процеси.

Динамиката на ДНК е повлияна от нуклеотидната последователност. Определени последователности проявяват различна стабилност и гъвкавост, което влияе върху това колко лесно могат да преминат през денатурация и ренатурация. Например, области с високо съдържание на гуанин (G) и цитозин (C) обикновено са по-стабилни от тези с високо съдържание на аденин (A) и тимин (T) поради по-силните водородни връзки между гуанин (G) и цитозин (C).

Динамиката на ДНК играе решаваща роля във взаимодействията между протеини и ДНК. ДНК активно участва в тези процеси, като приема специфични конформации, които улесняват или затрудняват свързването на протеините. Тази динамична природа е жизненоважна за регулирането на генната експресия и други клетъчни функции.

##### 1.4.1. Денатурация и ренатурация на нуклеиновите киселини

Денатурацията на нуклеиновите киселини е процес, при който те губят вторичната си структура без да се нарушава първичната структура - двойно-верижната спирала на ДНК се развива и разделя на две вериги чрез разкъсване на водородните връзки между базовите двойки, а РНК загубва двойно-верижните структурни мотиви.

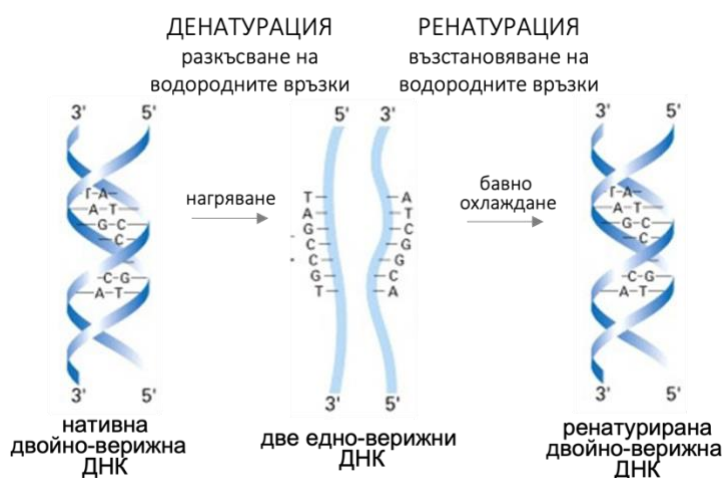
- Биологично-индуцирана денатурация

Нековалентните взаимодействия между антипаралелните вериги в ДНК могат да бъдат прекъснати, за да се "отвори" двойната спирала, когато се извършват биологично важни механизми като репликация на ДНК, транскрипция, поправка на ДНК или свързване с протеини. Областта на частично разделена ДНК е известна като „балонче“ (репликационно или транскрипционно), която може да бъде по-конкретно определена като отваряне на двойната спирала на ДНК чрез координирано разделяне на базови двойки.

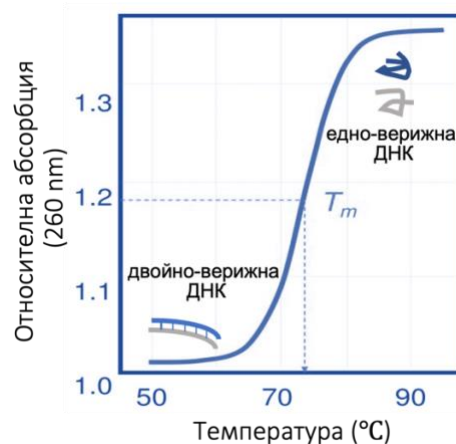
- Температурата – основен денатуриращ фактор

Основен метод за денатурация на ДНК в лабораторни условия е нагряването. Ако разтвор на ДНК се нагрее приблизително до 90°C или повече, се генерира достатъчно кинетична енергия,

която да денатурира напълно ДНК, причинявайки нейното разделяне на единични вериги. Температурно-индуцираната денатурация на ДНК се нарича още топене на ДНК (Фигура 1.9).



Фигура 1.9. Денатурация (топене) и ренатурация на ДНК.



Фигура 1.10. Крива на топене на ДНК.  $T_m$  -температура на топене

Степента на разделяне на веригата, или топенето, се измерва чрез абсорбцията на ДНК разтвора при 260 nm. Нуклеиновите киселини абсорбират светлина при тази дължина на вълната поради електронната структура на техните бази. Когато двете вериги се разделят, абсорбцията при 260 nm се повишава с 30%-40% - хиперхромен ефект (Фигура 1.10).

За множество копия на ДНК молекули, температурата на топене ( $T_m$ ) се дефинира като температурата, при която половината от ДНК молекулите са в двойно-верижно състояние, а другата половина са в едновержижно състояние. Температурата на топене зависи както от дължината на молекулата, така и от специфичния състав на нуклеотидна последователност на тази молекула.

- Химически агенти като денатуриращи фактори

Химическите агенти играят значителна роля като денатуриращи фактори в процеса на денатурация на ДНК. Тези агенти нарушават водородните връзки между базовите двойки, което води до разделяне на двойноверижната ДНК на единични вериги. Най-широко разпространените денатуриращи агенти са формамид, гуанидин, натриев салицилат, диметилсулфоксид (DMSO), пропилен гликол и урея. NaOH увеличава pH на разтвора, което нарушава водородните връзки, като отстранява протони, които допринасят за свързването на базите. Това води до бърза денатурация на ДНК.

- Ренатурация

Процесът на температурно-индуцирана денатурация на ДНК е обратим при контролирани условия на pH и йонна сила. Ако температурата бавно се понижи в разтвора, където ДНК е била денатурирана, ДНК веригите спонтанно комплементират и оригиналната двойно-спирална структура се възстановява. Първоначално комплементарните вериги могат да влязат в контакт произволно. След това обаче се подравняват правилно въз основа на комплементиращите последователности., което води до подобно на цип затваряне, където базовите двойки образуват водородни връзки и се възстановяване на структурата на двойната спирала

### 1.4.2. Хибридизация

Хибридизацията на нуклеинови киселини е процес, чрез който две комплементарни едно-верижни молекули ДНК (или РНК), произхождащи от различни източници (например от различни организми или дори синтетични последователности), образуват двойно-верижна структура чрез комплементиране на техните нуклеотидни бази.

В дните преди директното секвениране на ДНК да стане рутинно, хибридизацията на ДНК и/или РНК първоначално е била използвана за оценка на родството на различни организми, особено бактерии, където количеството на ДНК е сравнително малко. Смесването при подходящи условия на денатурирана ДНК от два различни вида води до свързване на комплементарни последователности, образувайки хибридни двойно-верижни спирали, в които едната ДНК верига произхожда от единия организъм, а другата ДНК верига – от другия организъм. ДНК от филогенетично близки видове са в състояние да образуват хибридни молекули с по-висока честота от тази на далечно-родствени видове. Например, човешката ДНК може да образува хибридни дуплекси в по-голяма пропорция с ДНК на маймуната резус, отколкото с ДНК на мишка.

В клинични условия хибридизацията може да се използва за диагностика, като например идентифициране на специфични патогени или генетични мутации чрез целева хибридизация с маркирани сонди.

#### ИЗПОЛЗВАНЕ на ХИБРИДИЗАЦИЯТА в молекулярно-биологичните техники

ДНК сонди: Хибридизацията се използва за откриване на специфични последователности в ДНК или РНК. Белязана едноверижна ДНК или РНК сонда хибридизира със своята комплементарна целева последователност, което позволява на изследователите да идентифицират или изолират специфични гени.

PCR (полимеразна верижна реакция): Молекулярно-биологична техника, използвана за амплифициране на специфични ДНК последователности – произвеждат се милиони копия на една ДНК молекула *in vitro* (в епруветка). За целта, праймери (синтетични къси ДНК последователности) хибридизират към целева ДНК матрица, за да инициират ДНК синтеза (Приложение 1).

Блот техники - Southern blotting е метод за откриване на специфична ДНК последователност в ДНК проба, докато Northern blotting е метод за откриване и качествена оценка на специфична РНК последователност (молекула) в РНК проба. И в двете техники, хибридизацията между сонда и целева нуклеинова киселина е от решаващо значение за откриване на специфични последователности.

*In situ* хибридизация: Тази техника позволява визуализиране на специфични последователности на нуклеинови киселини в тъкани или клетки чрез хибридизиране на белязани сонди към техните мишени.

### 1.5. Нуклеази и тяхната роля в метаболизма на нуклеиновите киселини

- Дезоксирибонуклеази

Дезоксирибонуклеазите (ДНКази) катализират хидролизата на фосфодиестерната връзка между нуклеотидите и водят до фрагментиране на ДНК. Те се класифицират в две класа (фамилии):

- ДНКази I - имат оптимална активност при неутрално рН (6,5-8,0) и изискват двувалентни йони като магнезий ( $Mg^{2+}$ ) и калций ( $Ca^{2+}$ ) за активиране, при разкъсване на ДНК се образуват два олигонуклеотидни крайни продукта с 5'-фосфатни и 3'-хидрокси краища.
- ДНКази II - работят най-добре при кисело рН, не изискват двувалентни йони за активиране и при разкъсване на ДНК образуват два олигонуклеотидни крайни продукта с 5'-хидрокси и 3'-фосфатни краища.

Еволюционното запазване на ДНКазите в множество биологични системи подчертава тяхното фундаментално значение за поддържане на клетъчното здраве. Те играят основна роля в метаболизма на ДНК чрез управление на извънклетъчните и вътреклетъчните нива на ДНК. Така например, при регулиране на имунната система ДНКазите разграждат извънклетъчната ДНК, освободена по време на възпалителни процеси, като ограничават потенциалните автоимунни реакции и поддържат системно равновесие. Тези ензими са от решаващо значение за програмираната клетъчна смърт (апоптоза), клетъчна диференциация и тъканно-специфични процеси на развитие.

#### КИСТИЧНА ФИБРОЗА и ДНКазно лечение

Кистична фиброза (муковисцидоза) е генетично заболяване, което причинява аномално гъст и лепкав мукус в белите дробове и други органи. Когато белите кръвни клетки се натрупат в белодробния мукус и се разпаднат, те освобождават ДНК, което увеличава вискозитета на мукуса и затруднява дихателната функция.

ДНКазните ензими, по-специално Пулмозим (Дорназа алфа), решават този проблем, като разграждат извънклетъчната ДНК. Когато се инхалират чрез небулайзер, тези ензими намаляват дебелината на мукуса, като го правят по-лесен за отстраняване от белите дробове. Лекарството работи, като се насочва към молекулярната структура на мукуса, като ефективно го "разрежда" и позволява на пациентите да дишат по-лесно.

#### • Рестрикционни ендонуклеази

Рестрикционните ендонуклеази (рестриктази) са специализирани бактериални ДНКази, които разпознават и разрязват чужда ДНК на специфични места. Тези ензими представляват фундаментален компонент на бактериалните генетични защитни системи - основната им функция е да елиминират генетичен материал от чужди организми, като например бактериофаги, които се опитват да инфектират бактериалната клетка.

Всеки рестрикционен ензим притежава уникална способност да разпознава кратка, специфична нуклеотидна последователност в чуждата ДНК. При разпознаване на тази последователност, ензимът катализира хидролизата на фосфодиестерната връзка между нуклеотидите. За да защити собствената си ДНК от разграждане, бактерията използва ензими, наречени метилази, които добавят метилови групи ( $-CH_3$ ) към аденинови (А) или цитозинови (С) бази в рамките на последователността на разпознаване, която по този начин се модифицира и защитава от ендонуклеазата.

Наименованието на рестриктазите следва строга научна номенклатура, която се извежда от рода, вида и щамата на бактерията, която ги произвежда. Класически пример е ензимът EcoRI, произведен от щам RY13 на *E. coli*.

В генетичното инженерство рестриктазите улесняват създаването на рекомбинантни ДНК молекули, като режат ДНК фрагменти с съвместими краища. Това позволява на

изследователите да вмъкват интересоващи ги гени във вектори, които са незаменими инструменти за генното клониране и създаването на генетично модифицирани организми. Рестриктазите са незаменими за производството на биотехнологични продукти като антибиотици, ензими и антитела.

В молекулярната диагностика тези ензими подпомагат откриването на генетични мутации и анализирането на генетични вариации, свързани с болести. Криминалистиката широко използва рестрикционни ензими за ДНК профилиране, където уникални ДНК фрагментни модели помагат за идентификация на индивиди.

- Рибонуклеази

Рибонуклеазите (РНКази) катализират хидролизата на фосфодиестерната връзка между нуклеотидите в РНК. Те могат да бъдат класифицирани на в два основни класа:

- Ендорибонуклеазите разрязват РНК на специфични вътрешни места и водят до образуването на по-малки РНК фрагменти (напр. RNase A и RNase III).
- Екзорибонуклеазите разрязват РНК от краищата, последователно отстранявайки нуклеотиди (RNase II).

Рибонуклеазите участват в разграждането на РНК, което не просто процес на деструкция, а сложен регулаторен механизъм, който е от съществено значение за клетъчната функция и адаптация. Чрез него клетката елиминира дефектни или аберантни РНК молекули, които биха могли да попречат на клетъчните процеси. Разграждането на РНК контролира изобилието на транскриптите чрез селективно изчистване на РНК и от тук нивата на генна експресия. Рибонуклеазите участват в пост-транскрипционна обработка (узряване) на РНК, при който предшествениците на много РНК молекули се превръщат във функционални крайни продукти като иРНК и тРНК.

Специфичните свойства на дезоксирибонуклеазите и рибонуклеазите са обобщени в Таблица 1.3.

Таблица 1.3. Основни характеристики на дезоксирибонуклеазите (ДНКази) и рибонуклеазите (РНКази).

ХАРАКТЕРИСТИКА	РНКази	ДНКази
Субстрат	РНК молекули	ДНК молекули
pH Оптимум	променлив	специфичен (неутрален/кисел)
Йонна зависимост	минимална	висока (изискват $Mg^{2+}$ , $Ca^{2+}$ )
Биологична функция	имунна защита, метаболизма на РНК	поправка на ДНК, програмирана клетъчна смърт (апоптоза)
Устойчивост	висока	средна
Клетъчна локализация	вътреклетъчна и извънклетъчна	предимно вътреклетъчна

Рибонуклеазите са изключително устойчиви ензими, които сериозно затрудняват изолирането на РНК, както и експериментите и анализите с РНК, поради своята повсеместност и способност за бързо разграждане на РНК проби. Някои РНКази могат да бъдат изключително устойчиви и инактивирането им е трудно в сравнение с неутрализирането на ДНКазите.

### 1.6. Хроматин – отличителен белег на еукариотната ДНК

В еукариотната клетка ДНК молекулите са пакетирани в хромозоми, в следствие на което те се побират в ядрото, а след като бъдат копирани, могат лесно да бъдат разпределени между двете дъщерни клетки при клетъчното деление. Сложната задача за опаковане на ДНК се постига чрез специализирани протеини, които се свързват и нагъват ДНК, което осигурява по-високи нива на организация и предотвратяват превръщането на ДНК в заплетена, неуправляема структура. ДНК е динамична структура, която търпи кондензация и декондензация – процеси, които са стриктно контролирани и позволяват на ДНК да бъде достъпна за всички ензими и други протеини, които я възпроизвеждат, възстановяват или осъществяват генната експресия.

С изключение на зародишните клетки (сперма и яйцеклетки) и високоспециализираните клетки, които нямат ДНК (като зрелите червени кръвни клетки), човешките клетки съдържат по две копия на всяка хромозома, едно наследено от майката и едно от бащата. Майчините и бащините хромозоми на една двойка се наричат хомоложни хромозоми. Единствените нехомоложни хромозомни двойки са половите хромозоми при мъжете, където Y хромозома е наследена от бащата и X хромозома от майката (жените наследяват по една X хромозома от всеки родител и нямат Y хромозома).

- Хроматин – комплекс на ДНК и хистонови белтъци

Пакетирането на ДНК се извършва от протеини, които навиват и сгъват ДНК във все по-високи и по-високи нива на организация. ДНК на интерфазните хромозоми, макар и около 20 пъти по-малко кондензирана от тази на митотичните хромозоми, все още е плътно опакована. Протеините, които се свързват с ДНК, за да образуват еукариотни хромозоми, се разделят на два основни класа: хистони и нехистонови хромозомни протеини. Хистоните присъстват в огромни количества (повече от 60 милиона молекули от няколко различни типа във всяка клетка) и тяхната обща маса в хромозомите е приблизително равна на тази на самата ДНК. Комплексът от двата класа протеини с ядрена ДНК се нарича хроматин.

Една клетка може да регулира структурата на хроматина – временно декондензира или кондензира определени области на своите хромозоми – използвайки комплекси за ремоделиране на хроматин и ензими, които ковалентно модифицират хистоновите опашки по различни начини. Хроматинът е динамична структура, която има важни ефекти върху мащабната структура на интерфазните хромозоми. Интерфазният хроматин не е равномерно опакован. Вместо това, регионите на хромозомата, които съдържат гени, които се експресират, обикновено са по-рехаво пакетирани, докато тези, които съдържат тихи гени, са по-кондензирани. По този начин структурата на интерфазните хромозоми може да се различава от един клетъчен тип до друг, което помага да се определи кои гени са експресирани. Повечето типове клетки експресират около 20 до 30% от гените, които съдържат.

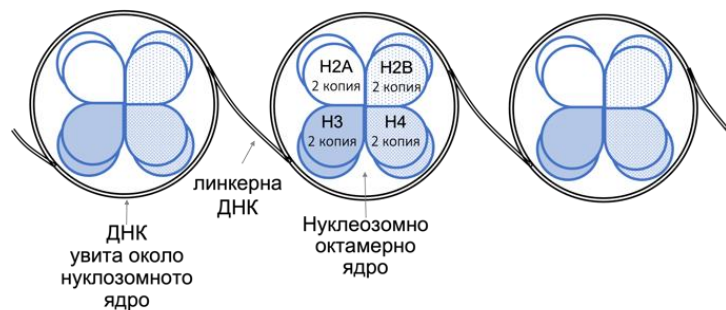
Плътно-пакетираната и кондензирана форма на интерфазния хроматин се нарича хетерохроматин (от гръцки *heteros*, „различен“ плюс хроматин). Той се класифицира в конституитивен и факултативен хетерохроматин и съставлява обикновено около 10% от интерфазната хромозома. Конституитивният хетерохроматин е богат на повтарящи се секвенции и а в хромозомите на бозайници е концентриран около центромерната област и в теломерите в краищата на хромозомите. Факултативният хетерохроматин не е богат на повтарящи се секвенции и включва гени, които се заглушават чрез различни механизми. Образоването на факултативен хетерохроматин се регулира от процеса на морфогенеза или диференциация.

По-хлabaво пакетираната и декондензирана форма на интерфазния хроматин се нарича еухроматин (от гръцки еи, „истински“ или „нормален“ плюс хроматин). Той е богат на гени, които са транскрипционно активни или могат да станат транскрипционно активни в някакъв момент по време на растежа и развитието. При човек, една от двете Х хромозоми на женския пол е инактивирана като факултативен хетерохроматин, докато другата се експресира като еухроматин.

Когато клетката се дели, тя обикновено предава своите хистонови модификации, хроматинова структура и модели на генна експресия на двете дъщерни клетки. Такава „клетъчна памет“ е от решаващо значение за установяването и поддържането на различни типове клетки по време на развитието на сложния многоклетъчен организъм.

- Нуклеозомите - основни единици на структурата на еукариотната хромозома

Хистоните са отговорни за първото и най-фундаментално ниво на пакетирание на хроматина, нуклеозомата, която е открита през 1974 г. Ако хроматинът в интерфазното ядро бъде подложен на третиране, които го карат да се разгъне частично, той може да се види в електронния микроскоп като серия от „мъниста на нишка“. Нуклеозомното ядро се състои от комплекс от осем хистонови протеина - по две молекули от хистоните H2A, H2B, H3 и H4 - и участък от двойноверижна ДНК с дължина 147 нд, която се навива около този хистонов октамер (Фигура 1.11). ДНК обвива плътно октамера, правейки 1,7 завъртания в лява намотка, и продължава върху следващата нуклеозома - линкерен ДНК участък, който може да варира от няколко нуклеотидни двойки до около 80. Образуването на нуклеозоми превръща ДНК молекулата в хроматинова нишка, която е приблизително една трета от дължината на самата ДНК молекула и осигурява първото ниво на опаковане на ДНК .



Фигура 1.11. Структура на нуклеозомата

И четирите хистони, които изграждат октамера, са относително малки протеини, с висок дял на положително заредени аминокиселини (лизин и аргинин). Положителните заряди помагат на хистоните да се свържат плътно с отрицателно заредения захарно-фосфатен скелет на ДНК. Всеки хистонов белтък има дълга N-терминална аминокиселинна „опашка“, която се простира от ядрото на нуклеозомата навън. Тези хистонови опашки са обект на няколко вида обратими, ковалентни химически модификации, които контролират много аспекти на структурата на хроматина.

Хистоните, които образуват нуклеозомното ядро, са сред най-силно консервативните от еволюционна гледна точка еукариотни протеини, отразява жизненоважната им роля в контролирането на структурата на еукариотната хромозома.

- Хромозомното пакетирание се случва на множество нива

Нуклеозомите се опаковат допълнително за да генерират по-компактна структура - хроматиново влакно. Това допълнително опаковане на нуклеозоми в хроматиново влакно

зависи от пети хистон, наречен хистон H1, за който се смята, че събира съседни нуклеозоми заедно в редовен повтарящ се масив.

По време на митозата хроматинът става толкова силно кондензиран, че отделните хромозоми могат да се видят в светлинния микроскоп. Как се нагъва хроматиново влакно, за да произведе митотични хромозоми? Отговорът все още не е известен в подробности, но е ясно, че хроматиновото влакно е нагънато в серия от бримки и че тези бримки допълнително се кондензират, за да произведат интерфазната хромозома; накрая, смята се, че тази компактна поредица от бримки преминава поне още едно ниво на опаковане, за да образува митотичната хромозома.

## ГЛАВА 2 РЕПЛИКАЦИЯ на ДНК

Репликацията на ДНК е биологичен процес на самовъзпроизвеждане на ДНК, при който в клетката се образуват две идентични копия на една ДНК молекула преди клетъчното делене, като гарантира, че всяка дъщерна клетка получава точно копие на генетичния материал. Този процес е от есенциално значение за растежа, възстановяването и възпроизводството на живите организми.

### 2.1. Въведение в репликацията

#### 2.1.1. Репликация – полуконсервативен механизъм

През 1957 г. двама млади учени, Матю Мезелсън и Франк Стал, провеждат експеримент, потвърждаващ, че ДНК се възпроизвежда чрез идентични копия, което Уотсън и Крик предвиждат по-рано въз основа на структурата на двойната спирала. Този знаков експеримент придобива незабавно известност като „най-красивия експеримент в биологията“, чиято красота е обвързана с неговата гениална простота.

*Matthew Meselson and Franklin Stahl (1958)*

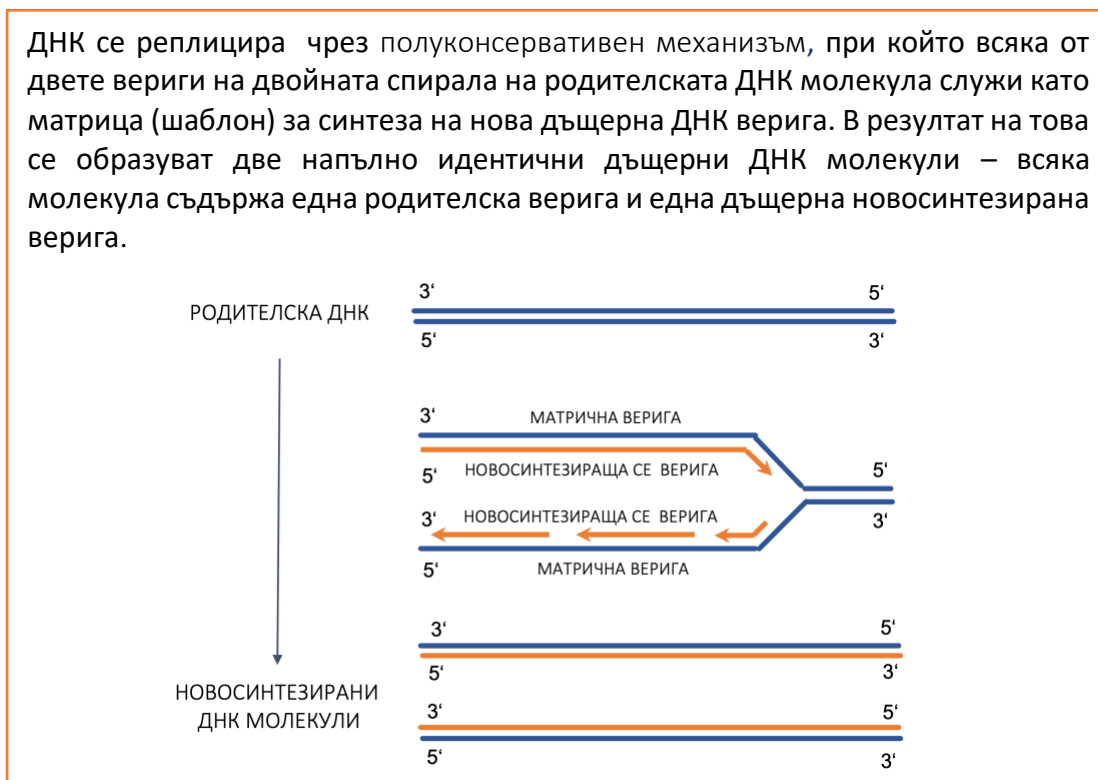


Ключът към експеримента на Мезелсън и Стал е отглеждането на клетки на *Escherichia coli* в среда, съдържаща тежък изотоп на азота,  $^{15}\text{N}$ , в продължение на много поколения. *E. coli* се нуждае от азот като хранително вещество и включва изотопа от хранителната среда в своите протеини и нуклеинови киселини.

След това те прехвърлят бактериите за различни периоди от време в среда, съдържаща обикновената форма на азота,  $^{14}\text{N}$ , така че всяка ДНК, синтезирана след трансфера, да не съдържа тежкия азотен изотоп. След всяко клетъчно делене, те изолират ДНК от клетките и разделят ДНК чрез ултрацентрифугиране в разтвор на цезиев хлорид. Цезият образува градиент на плътност при такива условия - при равновесие, ДНК молекулите в градиента се локализируют в позиция, която съответства на тяхната собствена плаваща плътност. По този начин родителската ДНК, в която и двете вериги съдържат тежкия изотоп  $^{15}\text{N}$  (тежки вериги), ще има висока плътност и ще се локализира близо до дъното на градиента. ДНК, в която двете вериги съдържат нормалния изотоп  $^{14}\text{N}$  (леки вериги), ще се локализира близо до върха на градиента. А хибридни ДНК, в които има една тежка верига е и една лека верига, ще се локализируют в междинна позиция.

Какво наблюдават Мезелсън и Стал? След първото клетъчно делене в среда, съдържаща  $^{14}\text{N}$ , те установяват, че ДНК има междинна плаваща плътност, съответстваща на хибридна ДНК. Това отговаря на очакванията на полуконсервативния модел, че всяка верига служи като матрица за синтез на нова верига. След второто клетъчно делене, те установяват, че половината от ДНК в тези клетки показва междинна плаваща плътност, съответстваща на хибридна ДНК, а другата половина показва плътност, съответстваща на ДНК, която има само леки вериги. Така след два кръга на репликация вече се генерира ДНК, в която едната верига е синтезирана през първия клетъчен цикъл, и е послужила като матрица за синтез на нова верига през втория клетъчен цикъл. Броят на леките ДНК молекули се увеличава с всяко поколение.

ДНК се реплицира чрез полуконсервативен механизъм, при който всяка от двете вериги на двойната спирала на родителската ДНК молекула служи като матрица (шаблон) за синтеза на нова дъщерна ДНК верига. В резултат на това се образуват две напълно идентични дъщерни ДНК молекули – всяка молекула съдържа една родителска верига и една дъщерна новосинтезирана верига.

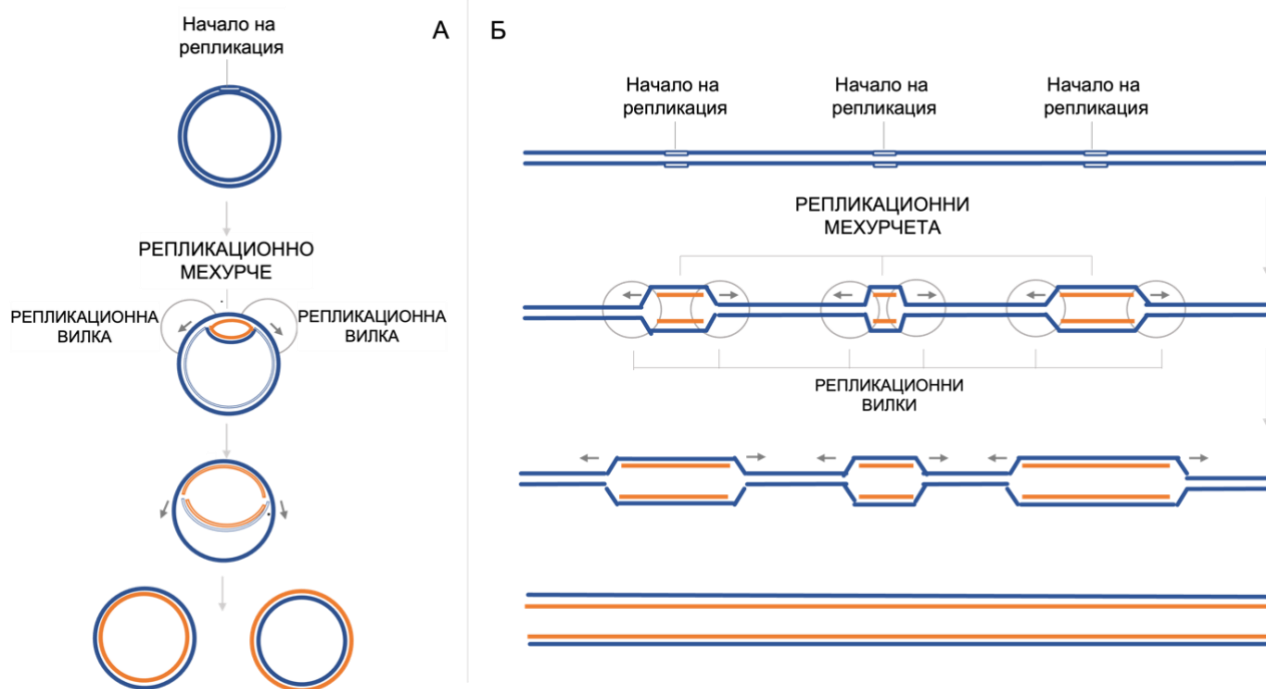


### 2.1.2. Репликон

Репликонът е област на ДНК, в която се извършва един акт на репликация (Фигура 2.1). Основен регулаторен елемент на репликацията е уникална секвенция наречена начало на репликация, в която се инициира самия процес. Бактериалната хромозома съдържа едно начало на репликация и следователно цялата бактериална хромозома представлява един репликон. Хромозомите на археи и еукариоти могат да имат множество начала на репликация и така техните хромозоми могат да се състоят от множество репликони. Броят на репликациите може да варира в генома от един в прокариоти и 500 в дрожди до няколко хиляди при животни и растения. В сравнение с прокариоти, еукариотните репликони са по-къси и скоростта на репликацията е много по-бавна. Не всички репликони започват репликацията по едно и също време, но иницирането на репликацията в различни репликони се извършва по регулиран начин през S фазата на клетъчния цикъл. Точката, в която веригите на родителската ДНК се разделят локално и се извършва репликацията (добавят се нови нуклеотиди), се нарича репликационна вилка.

### 2.1.3. Посока на репликацията

Репликацията на ДНК може да бъде еднопосочна или двупосочна, в зависимост от това дали репликацията продължава от началото на репликацията само в една посока или в двете посоки. При двупосочната репликация се формират две репликационни вилки, които се движат в противоположна посока. По този начин се извършва репликацията в бактериалните и еукариотните репликони. При еднопосочната репликация се формира една репликационна вилка, която се движи еднопосочно. Пример за еднопосочна репликация е тази на митохондриалната ДНК (мтДНК, mtDNA) при гръбначни животни, където две трети от митохондриалната ДНК се реплицира еднопосочно.



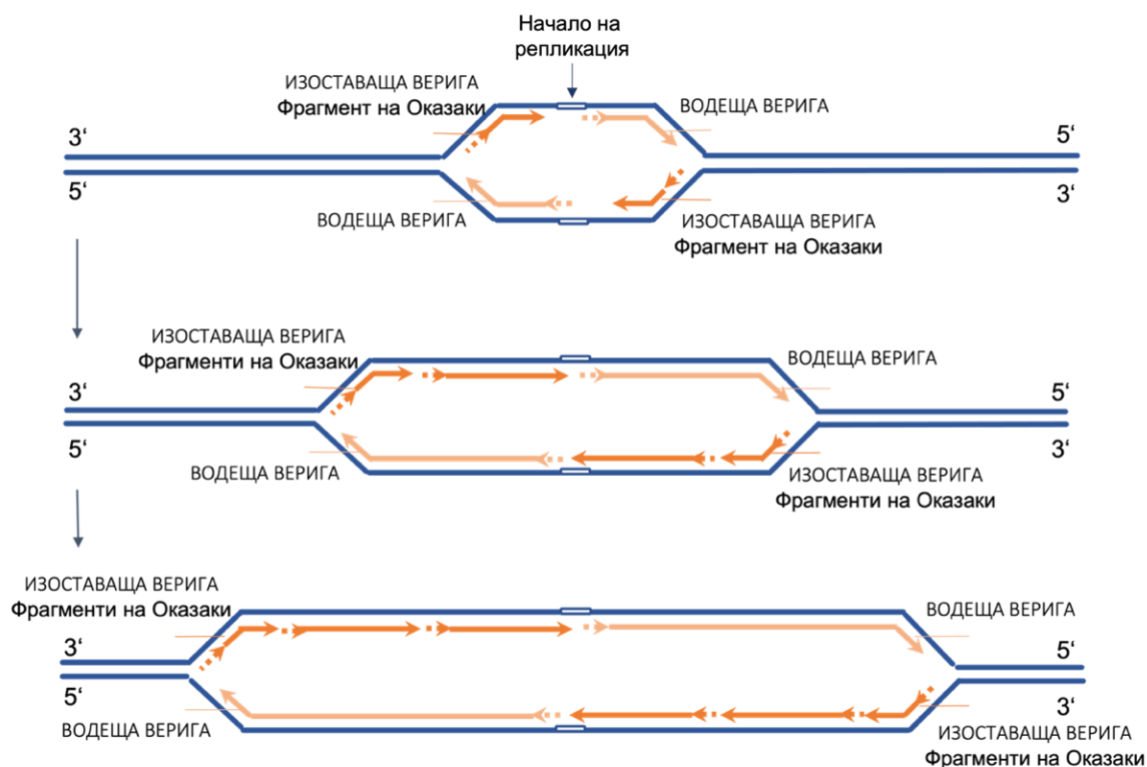
Фигура 2.1. Репликон. А. Бактериалната хромозома има едно начало на репликация и един репликон. Б. Еукариотните хромозоми имат множество начала на репликация и репликони.

#### 2.1.4. ДНК полимеризи

ДНК-зависимите ДНК полимеризи (от тук нататък наричани ДНК полимеризи) са група ензими, които катализират синтеза на ДНК върху матрица ДНК. В бактерии са изолирани пет типа ензими, като всеки е обозначен с римски цифри (I - V). При еукариоти са известни 14 ензима, пет от които имат основна роля по време на репликацията - pol  $\alpha$ , pol  $\beta$ , pol  $\gamma$ , pol  $\delta$  и pol  $\epsilon$ . Всички ДНК полимеризи са мултисубединични комплекси, образувани чрез свързване на различни субединици. Те се различават по своите свойства и функции; всички те обаче споделят следните характеристики:

- Използват като субстрати за синтез на ДНК четирите дезоксирибонуклеозид трифосфата (дНТФ, dNTPs) – АТФ (АТР), ГТФ (ГТР), ЦТФ (СТР) и ТТФ (ТТР). Образоването на нова ДНК верига се извършва чрез свързване на дезоксирибонуклеозид монофосфати (дНМФ, dNMP).
- Нуждаят се от свободна верига ДНК, която служи като матрица (матрична верига). Те синтезират нова верига върху матрицата. Матричната верига определя първичната структура на новата верига съгласно правилото за комплементарност на свързване на азотните бази - нуклеотидът, който се свързва към новосинтезиращата се верига, трябва да е комплементарен на нуклеотида на съответната позиция в матричната верига (А-Т и G-C).
- Те не са в състояние да свържат свободни нуклеотиди и да започнат нова верига, а могат само да удължат вече съществуваща верига (праймер), правилно сдвоена чрез водородни връзки с базите на матричната верига.
- ДНК синтезът се извършва чрез удължаване на 3'-края на новосинтезиращата се веригата (Фигура 2.2). Синтезът на ДНК се извършва в посока от 5'-края към 3'-края чрез присъединяване на дНМФ към 3'-края на новосинтезиращата се веригата. 3'-ОН групата на





Фигура 2.3. Синтез на ДНК – непрекъснат и прекъснат синтез. Водещата верига се синтезира непрекъснато, а изоставащата верига се синтезира чрез фрагментите на

## 2.2. Основни етапи на репликация

Инициация - включва разпознаване на репликационното начало от комплекс от протеини, разделяне на родителските вериги и стабилизирането им в едноверижно състояние, при което се създава репликационно мехурче.

Удължаване - извършва се от друг комплекс от протеини, означаван като - реплизома. Тя не съществува като независима единица (например, аналогична на рибозомата), а се сглобява *de novo* в репликационното начало при всеки цикъл на репликация. При придвижването на реплизомата по протежение на ДНК, родителските вериги се развиват и дъщерните вериги се синтезират от ДНК полимеразата, която катализира добавянето на нуклеотиди към нарастващите ДНК вериги.

Прекратяване (или терминация) - бележи края на репликацията на ДНК. В кръговите бактериални хромозоми репликационните вилки се движат двупосочно, докато не се срещнат в определен терминаращ регион. В еукариотните клетки множество репликационни вилки се движат по линейните хромозоми и в крайна сметка се срещат, завършвайки репликацията.

### 2.2.1. Инициация на репликацията

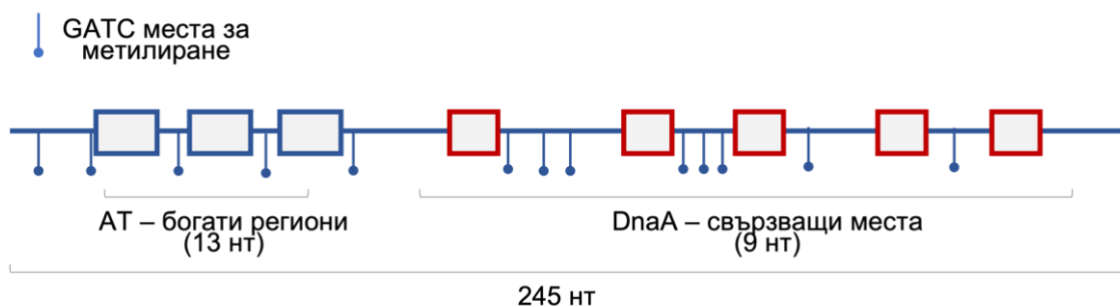
Двойната спирала на ДНК обикновено е много стабилна: двете ДНК вериги са здраво свързани заедно от големия брой водородни връзки между азотните бази. За да се използва като

матрица обаче, двойната спирала трябва първо да се отвори и двете вериги да се разделят, за да се открият несдвоени азотни бази. Инициацията е първият етап на репликацията, при който става присъединяване на инициаторни протеини в областите на начало на репликация, локално разделяне на двете вериги и формиране на репликационните вилки.

#### ✓ Прокариоти

Хромозомата на *E. coli* представлява единичен репликон с начало на репликация, идентифицирано като генетичен локус *oriC*. Представлява област с дължина 245 bp (Фигура 2.4), която включва:

- 5 копия на секвенция с дължина 9 бд (9-мерна секвенция), наречени *dnaA box* - към нея се свързва инициращия протеин *DnaA*;
- 3 копия на секвенция с дължина 13 бд (13-мерна секвенция), богата на А:Т двойки, където се извършва първоначалното разделяне двете вериги на ДНК за да инициране на репликация;
- 11 копия на 5'-GATC-3'. GATC е псевдопалиндром и присъства и в двете вериги на ДНК. Част от местата се припокриват с местата за свързване на *DnaA*. Метилазата *Dam* (DNA adenine methylase) на *E. coli* разпознава GATC секвенцията в ДНК и прехвърля метилова група (от S-аденозил метионин) към аминокгрупата на позиция 6 на аденина (A) в тази последователност. По този начин GATC секвенцията в дуплексна ДНК може да бъде неметилирана върху двете верига (неметилирана), метилирана само върху една верига (хемиметилирана) или метилирана върху двете вериги (напълно метилирана).



Фигура 2.4. Репликационно начало *oriC* на *E.*

Инициращият протеин *DnaA* е активен само когато е свързан с АТФ. Първата стъпка на инициацията е свързването на 10-20 мономера на *DnaA*-АТФ към 9-мерните секвенции, когато *oriC* областта е напълно метилирана. След това *DnaA* претърпява конформационни промени, които предизвикват торзионно напрежение в съседните 13-мерни секвенции и разделяне на двете вериги в тази област. Тогава хеликазата *DnaB* се свързва с *DnaC*, която я придружава до развитата ДНК. *DnaB* е ключовата хеликаза, която стимулира репликацията на ДНК чрез развиване на двойната спирала, но тя свързва ДНК слабо при отсъствието на *DnaC*. След като *DnaB* се присъедини към ДНК, *DnaC* се освобождава и хеликазата може да започне да развива ДНК, при условие че присъстват АТФ, протеина *SSB* и ДНК жиразата. *SSB* (от англ. *single-strand binding* - свързващ единична верига) е протеин, който стабилизира локалното развиване на спиралата чрез свързване на много негови копия към разделените единични вериги. ДНК жиразата е топоизомераза, която премахва усукването, което се генерира, когато двете вериги на двойната спирала са разделени.

Инициацията на репликацията в бактериалните клетки е силно контролирана, за да се гарантира, че репликацията на ДНК се случва само веднъж по време на всяко клетъчно делене. Нов кръг на репликация ще започне върху хромозомата на *E. coli* в *oriC* само тогава, когато условията на растеж го позволяват.

По време на репликацията на ДНК, родителската ДНК верига е метилирана в GATC последователности (включително тези в *oriC* региона) от Dam метилазата. Въпреки това, новосинтезираната дъщерна верига е неметирана в GATC сайтове. В резултат на това, веднага след репликацията областта на репликационното начало (*oriC*) е хемиметирана.

Хемиметираното състояние на *oriC* предотвратява незабавното повторно инициране на репликацията на ДНК. Протеинът SeqA се свързва с хемиметирана ДНК при *oriC*, инхибирайки свързването на инициаторния протеин DnaA и изолира репликационното начало, като гарантира, че клетката няма да иницира друг кръг на репликация преждевременно.

След определен период от време Dam метилазата напълно метилира новосинтезираната дъщерна ДНК верига в GATC последователностите, включително тези в *oriC*. Пълното метилиране възстановява *oriC* до напълно активно състояние, тъй като SeqA вече не се свързва с напълно метилирана ДНК и DnaA може отново да се свърже, за да започне нов кръг на репликация в следващия клетъчен цикъл. Началото на фазата на инициране на репликацията на ДНК се определя от концентрацията на DnaA. DnaA се натрупва по време на растежа и след това задейства иницирането на репликация.

#### ✓ Еукариоти

Репликацията на еукариотните хромозоми се иницира в множество начала на репликация. Това позволява на хромозомите да се реплицират много по-бързо, отколкото би могло да бъде, ако имаше само едно репликационно начало. Най-добре проучените репликационни начала при еукариоти са тези на дрожди. Репликацията в *S. cerevisiae* започва от 250-400 начала на репликация, разпределени върху 16 хромозоми. Много, ако не и всички, представляват автономно реплициращи се последователности - ARS (от англ. autonomously replicating sequence). Първоначално те са били изолирани от дрожди благодарение на тяхната способност да поддържат репликацията на всяка една секвенция, която е прикрепена към тях, откъдето произлиза и наименованието им. При висшите еукариоти изглежда няма консенсусна последователност в репликационните начала. Вместо това протеините за инициране на репликация могат да идентифицират и да се свържат със специфични модификации на нуклеозомите в репликационните начала.

Комплексът за разпознаване на репликационното начало ORC (от англ. origin recognition complex) е хетеромерен белтъчен комплекс от шест субединици. ORC протеините (Orc1 до Orc6) се комбинират по специфичен за вида начин, за да образуват ORC. Например, ORC на дрождите *S. pombe* се състои от шест ORC протеина, по два от Orc1, Orc2 и Orc4. В пъпкуващите дрожди ORC разпознава специфичните елементи на репликационните начала ARS, докато при по-висшите еукариоти изборът на начало не зависи от специфична ДНК последователност.

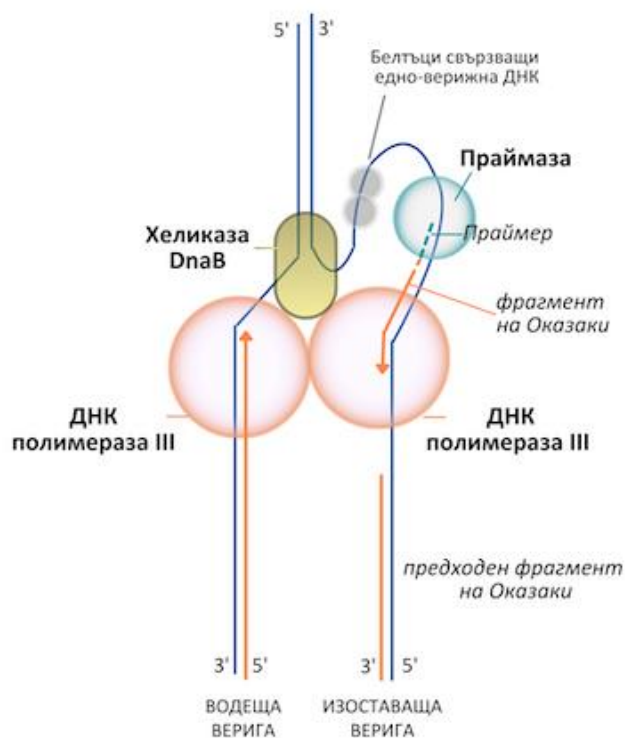
ORC се свързва с ДНК в репликационните начала по АТФ-зависим начин и служи като платформа за сглобяването на други ключови фактори за инициране. Първите два протеина, които се свързват с ORC през G1 фазата на клетъчния цикъл, са Cdc6 и Cdt1. Заедно ORC, Cdc6 и Cdt1 свързват белтъчния комплекс MCM (от англ. minichromosome maintenance) към репликационното начало. MCM протеините са основни фактори за инициране на репликация, първоначално идентифицирани като протеини, необходими за поддържане на минихромозома в *S. cerevisiae*. Най-известните сред тях са семейство от шест структурно

свързани протеини, MCM2-7, които са еволюционно запазени във всички еукариоти. ORC, Cdc6 и MCM образуват пред-репликационния комплекс. Когато клетъчният цикъл преминава от G1 фаза към S фаза, MCM, който има хеликазна активност, се активира и започва да развива ДНК, поставяйки началото на репликация. Едновременно с това Cdc6 и Cdt1 се дезактивират. Cdc6 и Cdt1 изпълняват ролята на фактори за лицензиране. Те се произвеждат само във фаза G1. Те могат еднократно да стартират репликация, тъй като след като след тяхното отстраняване от пред-репликационния комплекс от други протеини, клетката навлиза в S фаза, по време на която те се разграждат без да се репродуцират. Така те действат като лицензиращи фактори, но само заедно. Предполага се целият пред-репликационен комплекс може да се нарече лицензиращ фактор, тъй като той като цяло е необходим за инициране на репликация.

ORC остава свързан с репликационните начала през целия клетъчен цикъл, но с напредване на цикъла, една или повече от ORC субединиците се модифицират по такъв начин, че ORC активността да бъде инхибирана до приключване на митозата и сглобяването на ядрена мембрана.

### 2.2.2. Удължаване на ДНК

През етапа на инициация се осъществява първоначалното разделяне на двете ДНК вериги в областта на началото на репликация и се образува репликационната вилка, включваща хеликазата и жиразата. В етапа на удължаване към тях се присъединяват ДНК полимеразата и още няколко ензима (праймаза, нуклеаза и лигаза) за да формират апарата за репликация на ДНК, означаван като реплизома (Фигура 2.5).



Фигура 2.5. Реплизома при прокариоти

## ✓ Прокариоти

В бактерии са изолирани пет ДНК полимерази, обозначавани с римски цифри (I - V):

- ДНК полимераза I участва в отстраняването на РНК праймерите от фрагментите на Оказаци на изоставащата верига по време на репликацията на ДНК, както и в поправката на ДНК;
- ДНК полимераза II играе роля в поправката на ДНК и при рестартиране на блокирани репликационни вилки;
- ДНК полимераза III отговаря за синтеза на ДНК по време на репликация.
- ДНК полимерази IV и V имат роля в поправката на ДНК.

ДНК полимераза III е основният ензим, участващ в репликацията на ДНК в *E. coli*. Той представлява мулти-субединичен комплекс (900 kD), съдържащ 10 различни протеина, организирани в 4 типа субкомплекси:

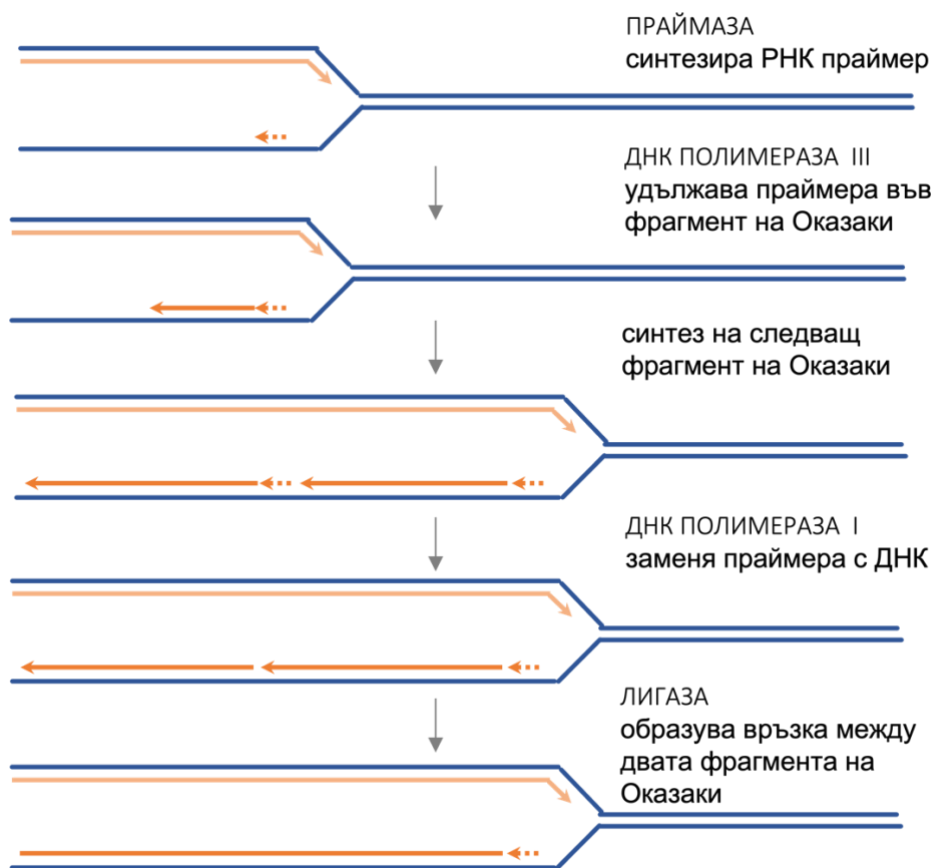
- 2 копия на каталитичната сърцевина, всяко от които съдържа  $\alpha$ ,  $\epsilon$  и  $\theta$  субединици,  $\alpha$  субединицата (кодирана от *dnaE* гена) има 5'-3' полимеразна активност,  $\epsilon$  субединицата (*dnaQ*) има 3'-5' екзонуклеазна активност,  $\theta$  субединицата (*holE*) стимулира  $\epsilon$  субединицата;
- 2 копия на „ $\beta$ -скобата“, всяко копие е хомодимер на субединицата  $\beta$  (*dnaN*), те разполагат каталитичните сърцевини върху матрични вериги и действат като плъзгащи се скоби;
- 2 копия на димеризиращата субединица  $\tau$  (*dnaX*), които свързват двете каталитични сърцевини;
- 1 копие на  $\Upsilon$  комплекса, натоварващ „ $\beta$ -скобата“, който съдържа 7 протеина.

Синтезът на ДНК веригите се стартира от праймаза - РНК полимераза, която синтезира праймер с дължина около 11 нт (Фигура 2.5). Праймерът предоставя 3'-ОН край, който се удължава от ДНК полимераза III.

Димерният комплекс на ДНК полимераза III се натоварва с помощта на „ $\beta$ -скобите“ върху ДНК. При това единият мономер на ензима се разполага върху едната матрична верига (която има полярност 3'-5') и синтезира непрекъснато водещата верига, а другият мономер се разполага върху другата матрична верига (която има полярност 5'-3') и синтезира изоставащата верига прекъснато чрез фрагментите на Оказаци (Фигура 2.6). След като е синтезиран фрагмент на Оказаци, „ $\beta$ -скобата“ се отделя от матрицата, и се асоциира отново към нея, така че да разположи каталитичната сърцевина в 3'-края на нов праймер за синтез на следващ фрагмент на Оказаци.

Нови данни показват, че бактериалната реплизомата вероятно включва три копия на каталитичната сърцевина, всяко от които има собствена  $\beta$ -скоба, а разполагането им върху матрицата се осъществява координирано от един натоварващ комплекс, който е прикрепен към хеликазата, разплитаща ДНК. *In vitro* и *in vivo* проучвания показват, че две копия на каталитичната сърцевина функционират върху изоставащата верига – това вероятно е от полза за репликационния процес, като се имат предвид множеството праймерни места, в които се инициират фрагменти на Оказаци.

След приключване на синтеза на водещата верига, както и на фрагментите на Оказаци, екзонуклеазата ДНК полимераза I премахва праймерите (5'-3'екзонуклеазна активност) и запълва празнините, останали след премахване на праймерите (5'-3'полимеразна активност), а ДНК лигаза свързва фрагментите на Оказаци (Фигура 2.6). Скоростта на ДНК репликация в *E. coli* е приблизително 1000 бд/сек и значително по-голяма в сравнение с тази при еукариоти, като една от причините за това е липсата на нуклеозоми върху бактериалната хромозома.



Фигура 2.6. Синтез на изоставащата верига чрез фрагментите на Оказаки

#### ✓ Еукариоти

Синтезът на ДНК веригите се стартира от комплекс на ДНК полимеразата  $\alpha$  и праймазна субединица, които колективно се означават като Pol  $\alpha$ /Primase. Праймазната субединица синтезира РНК секвенция с дължина около 7 нт, към която ДНК полимеразата прибавя 20-25 нт, образувайки т. нар инициаторна ДНК (идНК, iDNA). Инициаторната реакция е рискова тъй като ДНК полимеразата  $\alpha$  не притежава редактираща активност за поправка на грешни нуклеотиди. РНК праймерът и по-голяма част от идНК се подменят на по-късен етап.

След като идНК е синтезирана, към нейния 3'-край се свързва пентамерния белтъчен комплекс RF-C (от англ. Replication factor C), който използва енергията на АТФ за да натовари тримерния белтъчен комплекс PCNA (от англ. Proliferating cell nuclear antigen) върху ДНК матрицата. Свързването на RF-C и натоварването на PCNA изместват Pol  $\alpha$ /Primase от ДНК, и тогава PCNA опосредства свързването на ДНК полимеразата  $\delta$  и ДНК полимеразата  $\epsilon$ . Трите мономера на PCNA обхващат като скоба всяка от двете матрични ДНК вериги и служат като плъзгаща се платформа за ДНК полимеразите. ДНК полимеразата  $\epsilon$  (Pol  $\epsilon$ ) синтезира водещата верига. ДНК полимеразата  $\delta$  (Pol  $\delta$ ) синтезира изоставащата верига чрез фрагментите на Оказаки, като всеки се иницира от Pol  $\alpha$ /Primase. Хроматин-ремоделиращи белтъчни комплекси отделят нуклеозомите от ДНК пред репликационната вилка, и ги възстановяват върху новосинтезираната ДНК. Финалните стъпки на репликацията включват премахване на РНК праймерите (най-вероятно и на идНК) с помощта на екзонуклеазата RNase H и ендонуклеазата Fen1, и свързване на фрагментите на Оказаки с помощта на ДНК лигаза I.

И двете ДНК полимерази (Pol  $\delta$  и Pol  $\epsilon$ ), имат екзонуклеазна активност, което прави възможна редакцията на новосинтезираната ДНК верига и коригирането на грешните нуклеотиди, които могат да присъстват във веригата в следствие на неправилно комплементиране. ДНК репликацията има удивителна точност (1 грешка на  $10^9$  бд).

Белтъчните фактори, които участват в репликацията при прокариоти и еукариоти, са обобщени в Таблица 2.1.

Таблица 2.1. Репликация - белтъчни фактори: Прокариоти vs еукариоти

ФУНКЦИЯ	Прокариоти	Еукариоти
Инициатор, свързващ началото на репликация	DnaA	ORC
Натоварване на хеликаза	DnaC	Cdc6 и Cdt1
Хеликаза	DnaB	MCM комплекс
Свързване на единична верига	SSB	RPA
Синтез на праймери	DnaG праймаза	Pol $\alpha$ /primase
Репликативна ДНК полимераза	ДНК полимераза III	ДНК полимераза $\delta$ , ДНК полимераза $\epsilon$
Натоварване на скобата	$\Upsilon$ комплекс	RF-C
Скоба	$\beta$ -скоба	PCNA
Премахване на праймери	ДНК полимераза I	RNase H / Fen1
Лигаза	ДНК лигаза 1	Лигаза 12

Хроматинът (комплексът между ДНК и протеини) претърпява химични модификации, така че ДНК да може да се освобождава от свързаните с нея белтъци и да бъде достъпна за ензимите на машината за репликация на ДНК. Хистоните трябва да бъдат отстранени по време на процеса на репликация и след това възстановени върху новообразуваните ДНК молекули, което е една от причините за по-ниската скорост на репликация при еукариотите.

### 2.2.3. Прекратяване на репликацията

При прокариоти, прекратяването, или терминация, на репликацията изисква движението на репликационната вилка да спре или да бъде блокирано. Тъй като бактериите имат кръгови хромозоми, прекратяването на репликацията настъпва, когато двете репликационни вилки се срещнат една с друга на противоположния край на родителската хромозома В случаите, когато прекратяването се извършва в специфичен локус (терминираща секвенция), взема участие протеин, който се свързва с тази секвенция, за да спре физически репликацията на ДНК. При различни бактериални видове това е Ter протеина (от англ. DNA replication terminus site-binding protein).

При еукариоти, тъй като репликацията на ДНК протича само в посока от 5'-3', изоставащата верига трябва да бъде синтезирана под формата на прекъснати ДНК фрагменти, всеки от които започва с РНК праймер, поставен от праймаза. Когато репликационната вилка се приближи до края на линейната хромозома възниква сериозен проблем: въпреки че водещата верига може да бъде реплицирана чак до края на хромозомата, изоставащата верига не може. Когато последният РНК праймер върху изоставащата верига се отстрани, няма начин той да бъде заменен (Фигура 2.7). Без стратегия за справяне с този проблем, изоставащата верига ще става по-къса при всяка следваща репликация на ДНК; след множество клетъчни деления, хромозомите ще се скъсят толкова много, че ще загубят ценна генетична информация.

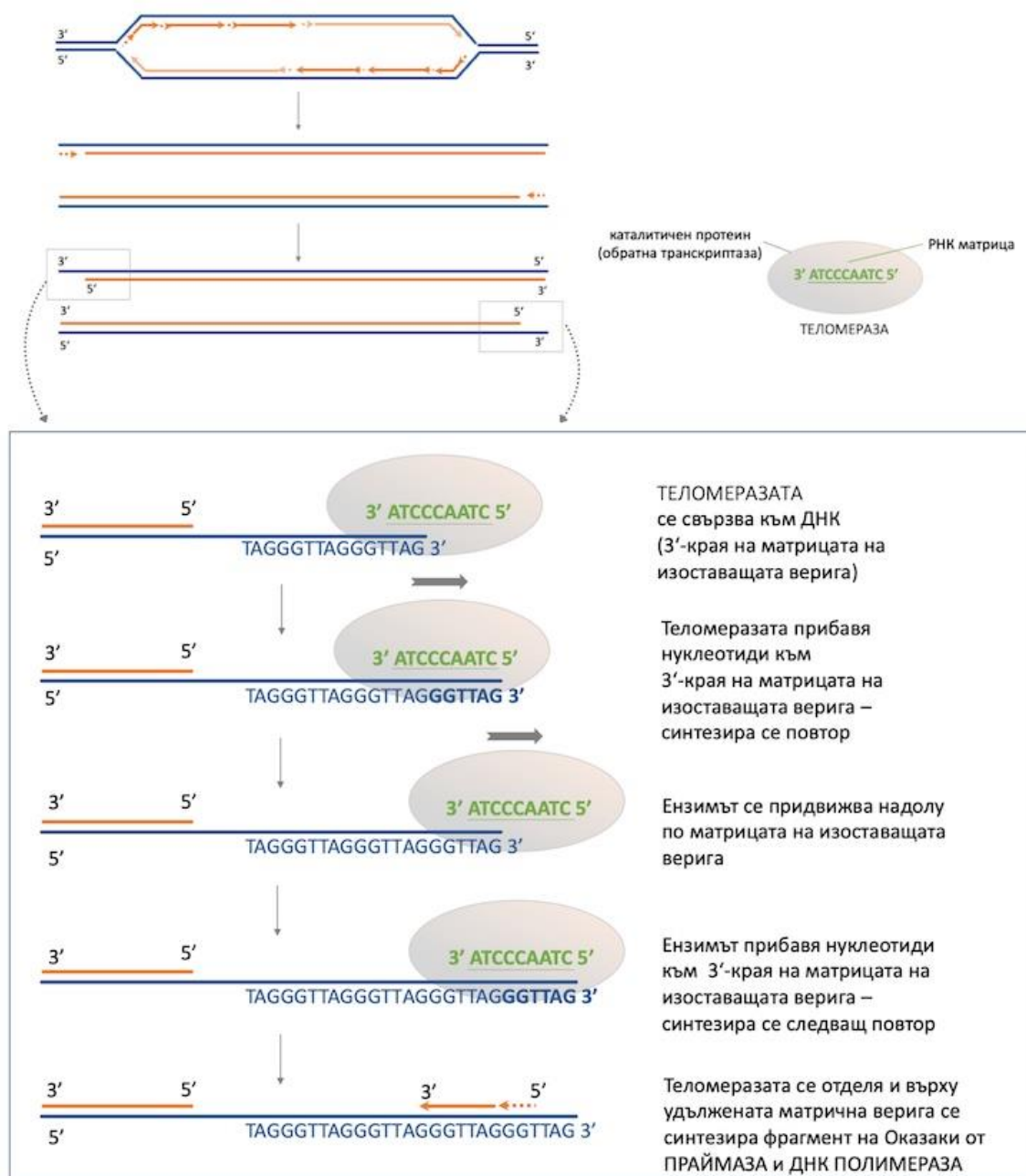
Еукариотите могат да разрешат проблема, тъй като хромозомите завършват в двата края с теломери, съставени от повтарящи се нуклеотидни последователности. При човек, теломерите включват последователността 5'-TTAGGG-3', копията на която в някои човешки клетки могат да достигнат до 1500-2000 броя. Ключовият ензим, който предотвратява скъсяването на теломерите, е теломеразата. Ензимът и неговия механизъм на действие са открити от Елизабет Блекбърн, Карол Грейдър и Джак Шостак в едноклетъчния ресничест еукариотен организъм *Tetrahymena*, за което получават Нобелова награда за медицина и физиология през 2009 г. Молекулният състав на човешкия теломеразен комплекс, описан от Скот Коен и неговия екип, се състои от две молекули, всяка съставена от човешката теломеразна обратна транскриптаза (TERT), теломеразна РНК с дължина 451 нт (TERC) и дискерин (DKC1).

Теломеразите принадлежат към ензимния клас на обратните транскриптази, които създават едноверижна ДНК, използвайки едноверижна РНК като матрица. Ензимът носи къса молекула РНК, комплементарна на повтарящата се последователност в теломерите; тази РНК действа като матрица за синтеза на теломерна ДНК. Използвайки РНК матрицата, теломеразата добавя няколко копия на повтарящата се последователност към края на матрична верига. Това позволява репликацията на изоставашата верига да бъде завършена върху удължената матрица чрез конвенционална ДНК репликация (Фигура 2.7).

**Теломераза и клетъчна смърт:** Теломеразата обикновено е активна в зародишните клетки и възрастните стволни клетки, но не и в соматични клетки. Броят на повторенията определя максималния живот на клетката. Всеки път, когато една клетка се дели, част от теломерите се губи (обикновено 25-200 бд на деление). След като теломерите бъдат скъсени до определен размер, клетката достига до критична точка, при която следващо делене е невъзможно. В резултат на това клетката умира чрез апоптоза (програмирана клетъчна смърт).

**Теломераза и рак:** Раковите клетки са вид злокачествени клетки, които се размножават неконтролируемо и образуват тумори. Теломеразата е открита в човешки ракови клетки, където нейната активност е 10-20 пъти по-висока отколкото в нормалните телесни клетки. Това осигурява селективно предимство на растежа на много видове тумори. Много ракови клетки се считат за „безсмъртни“, тъй като активността на теломеразата им позволява да живеят много по-дълго от всяка друга соматична клетка, което в комбинация с неконтролируема клетъчна пролиферация, е причината да се образуват тумори. Добър пример за безсмъртни ракови клетки са клетките HeLa, които се използват в лабораториите като моделна клетъчна линия от 1951 г.

Ако теломеразната активност можеше да бъде изключена, тогава теломерите в раковите клетки биха се скъсили, точно както правят в нормалните телесни клетки. Това би попречило на раковите клетки да се делят неконтролируемо в ранните им етапи на развитие. В случай, че туморът вече се е развил напълно, той може да бъде отстранен и може да се приложи антителимеразна терапия, за да се предотврати рецидив. По същество предотвратяването на изпълнението на функцията на теломеразата би променило раковите клетки от „безсмъртни“ в „смъртни“.



Фигура 2.7. Механизъм на действие на теломеразата

### 2.3. Точност на репликацията

Точността на репликация е от решаващо значение за поддържане на генетична стабилност през поколенията, като се гарантира, че ДНК се копира с минимални грешки по време на клетъчното делене. Репликацията на ДНК е много точен процес, постигнат чрез няколко ключови механизма, които заедно осигуряват забележителна прецизност, обикновено с нисък процент грешки като една грешка на милиард реплицирани нуклеотиди:

- Селективност на нуклеотидите: ДНК полимеразите проявяват силна предпочитание за вграждане на правилния дНТФ въз основа на правилата за комплементарно свързване на базите. Този механизъм значително намалява вероятността от вграждане на несъответстващи бази, като увеличава точността приблизително 100 пъти.

- Модел на индуцирано прилягане: ДНК полимеразите преминават през конформационни промени при свързване с правилния дНТФ за да осигурят оптимална катализата. Обратно, ако е налице неправилен дНТФ, това оптимално позициониране се нарушава, което води до намалена скорост на образуване на връзки и по този начин предотвратява вграждането на неправилната нуклеотид.
- Коригираща активност: Много ДНК полимерази притежават 3'-5' екзонуклеазна активност, която им позволява да отстраняват неправилно сдвоените нуклеотиди веднага след добавянето им. Ако е включен неправилен нуклеотид, полимеразата спира, неправилният нуклеотид се изрязва и синтезът се възобновява с правилния нуклеотид. Този механизъм може да увеличи общата точност с допълнителен фактор от 100 до 1000.
- Структурни механизми: Структурната конфигурация на ДНК полимеразите играе решаваща роля в тяхната способност да разграничават правилни и неправилни бази. Изследванията показват, че когато е вградена несъответстваща база, последващите опити за добавяне на друг нуклеотид са неефективни, тъй като ензимът не може да се ориентира правилно за каталитичната реакция. Тази неефективност насърчава корекцията и допълнително намалява грешките по време на репликацията.
- Поправка на несъответствията: В допълнение към корекцията, клетките използват система за поправка на несъответствията, която действа върху грешки, които са избегнати от първоначалната корекция от ДНК полимеразите..

Чрез механизми като селективност на нуклеотидите, конформационни промени при индуцирано прилягане, способности за корекция и пост-репликативни системи за корекция на несъответствията, ДНК полимеразите ефективно разграничават неправилните бази по време на репликацията. Тези процеси колективно осигуряват, че репликацията на ДНК протича с забележителна точност, което е от съществено значение за поддържането на генетичната стабилност и предотвратяването на мутации, които могат да доведат до заболявания като рак.

#### 2.4. Репликация на митохондриална ДНК

Репликацията на мтДНК е различна от репликацията на ядрената ДНК, тя следва уникален процес, адаптиран към нейната структура и функция и необвързан с клетъчния цикъл, както е при ядрената ДНК.

МтДНК е кръгова, двуверижна молекула, подобна на бактериалната ДНК, което отразява еволюционния произход на митохондриите. Тя има две вериги, наречени тежка верига (H-верига, от англ. heavy) и лека верига (L-верига, от англ. light), в зависимост от състава им (тежката верига е богата на гуанин). Репликацията на мтДНК започва от специфични начала на репликация за всяка верига - OriH (начало на репликация на тежката верига) и OriL (начало на репликация за леката верига). Основните стъпки на репликацията на мтДНК включват:

- Инициация в OriH и репликация на тежката верига: Репликацията започва в OriH, където се синтезира кратък РНК праймер от митохондриалната РНК полимераза. Праймерът се удължава от митохондриалната ДНК полимераза ( $\gamma$  (Pol  $\gamma$ )), която започва синтеза на новата верига.
- Изместване на веригата: Докато тежката верига се реплицира, тя измества леката верига като образува примка (D-примка, от англ. D-loop). Изместената лека верига остава едноверижна, докато репликацията в OriL не започне.
- Репликация на леката верига: След като значителна част от тежката верига е реплицирана, репликацията на леката верига започва в OriL. След синтеза на праймер, ДНК полимераза

$\gamma$  започва репликация на леката верига, движейки се в противоположна посока на репликацията на тежката верига.

- Прекратяване: Репликацията на двете вериги продължава, докато цялата митохондриална ДНК не бъде копирана. Всяка нова митохондрия получава пълно копие от кръговата митохондриална ДНК.

Освен ДНК полимераза  $\gamma$ , в процеса важна роля играят хеликаза Twinkle, която развива мтДНК, като дава достъп на репликационния апарат до веригите; митохондриалния протеин свързващ едноверижна ДНК (mtSSB), стабилизиращ едноверижната ДНК по време на репликацията; и рибонуклеазата RNase H1, която премахва РНК праймерите.

Репликацията на мтДНК не е синхронизирана с клетъчния цикъл, както при ядрената ДНК. Вместо това, репликацията на мтДНК може да се случва непрекъснато и независимо, за да отговаря на енергийните нужди на клетката. При повечето бозайници, мтДНК се унаследява по майчина линия. По време на оплождането, митохондриите от сперматозоида обикновено се разграждат и само митохондриите от яйцеклетката се унаследяват от клетките на потомството.

## ГЛАВА 3 МУТАЦИИ и ПОПРАВКА на ДНК

Унаследяването на белезите на едно поколение от следващото поколение разчита на точното предаване на генетична информация, кодирана в ДНК. ДНК, като молекула на наследствеността, носи инструкциите за развитието, функционирането и възпроизводството на живите организми. Въпреки че е достатъчно стабилна, за да поддържа генетичната цялост през поколенията, ДНК също е динамична и адаптивна молекула, която може да претърпи промени. Тези промени, известни като мутации, играят централна роля в еволюцията и биологичното разнообразие.

Мутациите, които представляват промени в ДНК последователността, неизбежно водят до промени в генотипа на организма. Въпреки това, не всички мутации водят до видими фенотипни промени, тъй като някои от тях могат да се случат в некодиращи участъци на ДНК, или може да не повлияят значително на функцията на гените. Когато мутациите все пак засягат функцията на гените, те могат да доведат до промени във фенотипа, засягайки физическите характеристики на организма, биохимичните процеси или поведението му.

В генетичните изследвания учените обикновено използват генотипа на дивия тип като отправна точка. Дивият тип представлява най-често срещания генотип или фенотип в естествената популация и служи като стандарт, спрямо който се сравняват всички вариациите, най-вече мутантните генотипове. Този сравнителен подход позволява на изследователите да оценят специфичните ефекти на всяка мутация и да определят как тя се различава от дивия тип както по генотип, така и по фенотип.

Въпреки че повечето мутации не засягат организма по някакъв забележим начин, някои имат дълбоки последици. Промените в ДНК последователността могат да доведат до малки вариации, които са в основата на разликите между индивидите от един и същи вид. Промени, които дават предимство в естествения отбор и се натрупват в продължение на милиони години, осигуряват разнообразието в генетичния материал, което прави един вид различен от друг. Мутацията е първата стъпка в еволюцията.

Черната окраска при леопардите е причинена от рецесивна генна мутация. Тя води до генетично състояние, известно като меланизъм, при което меланинът в козината е в излишък.

Меланистичните форми са по-често срещани в гъсти, влажни горски хабитати, където тяхната тъмна окраска може да осигури по-добра камуфлажност срещу хищници и при лов.



### 3.1. Генни мутации – класификация на ниво ДНК, протеин и функция

Мутациите могат да бъдат класифицирани в зависимост от техния размер и естеството на промените, които предизвикват в ДНК на хромозомни мутации и генни мутации.

Хромозомните мутации включват изменения в структурата или броя на хромозомите. Те могат да произтичат от грешки по време на клетъчно делене, кросинговър или оплождане. Тези мутации могат да доведат до значителни генетични нарушения поради големите промени, които предизвикват. Най-чести типове хромозомни мутации включват: делеции (загуба на сегмент от хромозомата), дубликации (копиране на сегмент от хромозомата), инверсии (обръщане на сегмент в рамките на хромозома), и транслокации (сегмент от една хромозома се прехвърля на друга хромозома). Тези мутации могат да засегнат множество гени и често са свързани с сериозни здравословни състояния, като рак или генетични нарушения, например синдром на Даун, който е налице допълнителна копие на хромозома 21. Генните мутации, от друга страна, обикновено включват по-малки промени на ниво ДНК.

Хромозомните мутации включват по-големи структурни промени, които могат да засегнат много гени едновременно, и да доведат до по-широки, системни ефекти. Генните мутации също могат да имат значителни биологични последствия. Дори малки промени в ДНК последователността могат да изменят структурата и функцията на протеините или да повлияят на генната регулация. Тези мутации са основен двигател на еволюционните промени, тъй като предоставят материал за естествения отбор, насърчавайки адаптацията и еволюцията на видовете. Освен това те са основната причина за генетични заболявания.

По-долу са разгледани генните мутации чрез последователен анализ на промените на ниво ДНК, как тези промени се отразяват на ниво протеините и какви могат да бъдат функционалните ефекти.

### 3.1.1. Мутации на ниво ДНК

- Точкови мутации (замени, субституции)

Всяка базова двойка ДНК може да претърпи мутация. Точковата мутация променя само една двойка бази и може да бъде разделена на два типа в зависимост от естеството на заместването на азотната база (Фигура 3.1):

- Транзиция: Най-често срещаният тип точкова мутация, която е резултат от заместването на един пиримидин с друг пиримидин или на един пурин с друг пурин. Напр. замяна на А с Г, която при репликация ще доведе до замяна на А-Т с G-C.
- Трансверсия: По-рядко срещаният тип точкова мутация, при която пурин се заменя с пиримидин или обратното. Напр. замяна на А с Т, която при репликация ще доведе до замяна на А-Т с Т-А.

✓



Фигура 3.1. Точкови мутации

- Инсерции (вмъквания) и делеции (загуби) на последователности

Вмъкванията на нуклеотиди, или инсерции, са мутации, при които един или повече нуклеотиди се добавят към ДНК последователността. Загуба на нуклеотиди, или делеции, са

мутации, при които един или повече нуклеотиди са отстранени от ДНК последователността (Фигура 3.2).



Фигура 3.2. Инсерции и делеции

Някои ДНК последователности са по-склонни към мутации поради комбинация от присъщи свойства на ДНК – т.нар. горещи точки. Известно е, че специфични последователности, като CpG динуклеотидите, са особено променливи. В тези региони цитозинът (C) може да бъде метилиран и впоследствие дезаминиран, което води, въпреки ефективните механизми за поправка, до чести транзиция на цитозин (C) до тимин (T). Други мотиви, като CpHpG тринуклеотиди или специфични повтарящи се последователности, също показват по-висок процент на мутагенност.

### 3.1.2. Мутации на ниво протеин

Точковите мутации на ДНК могат да причинят три типа мутации на ниво протеин (Фигура 3.3):

- Мутация с промяна на смисъла (от англ. missense)

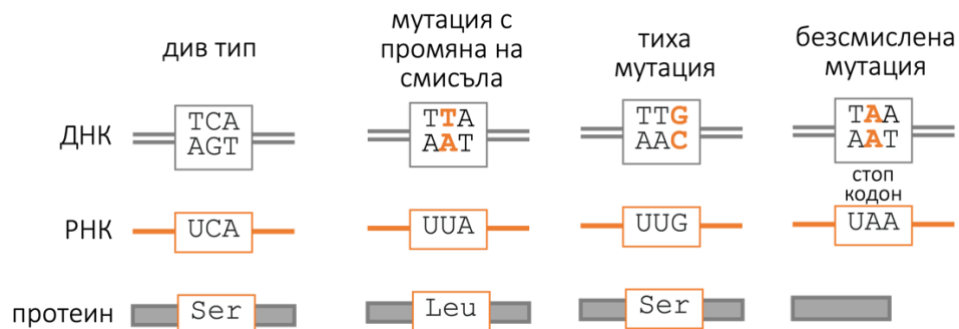
Мутация, при която промяна на единичен нуклеотид в ДНК води до заместване на една аминокиселина с друга в протеиновата молекула. Тази промяна може да има различни ефекти върху структурата и функцията на протеина, в зависимост от свойствата на включените аминокиселини и местоположението на мутацията в протеина. Тези мутациите са значими, защото могат да доведат до широк спектър от генетични заболявания. Например, при сърповидно-клетъчната анемия, точкова мутация в гена на хемоглобина (HbV) променя аминокиселината глутаминова киселина на валин. В резултат на това молекулите на хемоглобина се слепват, променяйки формата и функцията на червените кръвни клетки, което води до заболяването.

- Тиха мутация

Замяната на един нуклеотид с друг не променя аминокиселината, кодирана от засегнатия кодон, поради изроденността на генетичния код - няколко кодона могат да кодират една и съща аминокиселина, така че една нуклеотидна промяна може да кодира същата аминокиселина. Тъй като аминокиселинната последователност е непроменена, тези мутации обикновено нямат пряк ефект върху функцията на протеина, макар че в някои случаи, те могат да повлияят на стабилността на иРНК..

- Безсмислена мутация (от англ. nonsense)

Точкова мутация трансформира засегнатия кодон в стоп кодон, причинявайки преждевременно прекъсване на транслацията. Като цяло безсмислените мутации произвеждат нефункционални или непълни протеини. Това може да има сериозни последици, тъй като непълният протеин вероятно няма да функционира.



Фигура 3.3. Мутации на ниво протеин

- Мутация с изместване на рамката (от англ. frameshift).

Тъй като протеин-кодиращата ДНК (гена) е разделена на кодони с дължина три бази, вмъкването или отстраняването на една или две базови двойка може да причини мутация с изместване на рамката, променяща цялата последователност от кодони и съответно протеиновата последователност (секвенция). Вмъкването, или отстраняването, на три базови двойки може да бъде по-малко вредно от вмъкването или отстраняването на една или две базова двойка. Това е така, защото вмъкването на три базови двойки не причинява мутация с изместване на рамката.

Мутациите с изместване на рамката често имат сериозни ефекти върху протеиновата функция, тъй като нарушават цялата аминокиселинна последователност от мутацията нататък, което води до генетични нарушения или клетъчна дисфункция. Например, при болестта на Tay-Sachs възниква мутация с изместване на рамката в гена *HEXA* поради вмъкване на четири нуклеотида. Това нарушава производството на ензима, което води до натрупване на определени липиди в мозъчните клетки и причинява невродегенеративни симптоми.

### 3.1.3. Мутации на функционално ниво

- Мутации със загуба на функция (от англ. loss-of-function, LoF): мутации, които водят до намаляване или пълна загуба на генната функция.

Промяната в ДНК може да засегне структурата и функцията на протеина, кодиран от този ген. Например, ензимите са особено податливи на мутации, които засягат аминокиселинната последователност в тяхното активно място (т.е. областта, която позволява на ензима да се свърже с неговия специфичен субстрат). Това може да доведе до инактивиране на ензима.

Алелите на дивия тип обикновено кодират продукт, необходим за специфична биологична функция. Ако възникне мутация в такъв алел, функцията, която той кодира, може да се загуби. Общият термин за тези мутации е мутации със загуба на функция. Ако функцията на променения алел е напълно загубена, мутацията се нарича нулева мутация (от англ. null). Ако функцията е запазена, но под нивото на алела от див тип, мутацията се нарича пропусклива (от англ. leaky).

Мутациите със загуба на функция обикновено са рецесивни. При хетерозиготен индивид, съдържащ алел от див тип и алел със загуба на функция, нивото на експресия на алела от див тип често е достатъчно, за да произведе фенотип от див тип. Генетично това би определило мутацията със загуба на функция като рецесивна.

Например, мутациите в гена *CFTR* водят до кистична фиброза. Генът кодира протеин, който функционира като хлориден канал на повърхността на епителните клетки. Този протеин е от съществено значение за поддържане на баланса на сол и вода в различни тъкани, особено в

белите дробове и панкреаса. Над 2500 мутации са идентифицирани в *CFTR* гена с различни ефекти върху функцията на протеина. Най-честата мутация, F508del, включва делеция на три нуклеотида, което води до загуба на фенилаланин в позиция 508. Тази мутация води до неправилно нагъване и трафик на *CFTR* протеина, предотвратявайки достигането му до клетъчната повърхност, където може да изпълнява функцията си.

- Мутации с усилване на функция (gain-of-function, GoF): мутации, при които промененият генен продукт притежава нова молекулярна функция или нов модел на генна експресия.

Въпреки че може да се очаква, че повечето мутации ще доведат до загуба на функция, е възможно от мутацията да се получи нова и важна функция. В тези случаи мутацията създава нов алел, който е свързан с нова функция. При хетерозиготен индивид, новият алел често ще доминира над дивия тип алел по отношение на експресията. Това означава, че индивидът ще изрази новата функция, свързана с мутацията за придобиване на функция, което ще определи мутацията като доминантна.

От еволюционна гледна точка мутациите с усилване на функцията могат да осигурят на организмите благоприятни черти, които подобряват оцеляването и възпроизводството. Например, мутации, които позволяват на патогените да избегнат имунните отговори на гостоприемника или да увеличат тяхната трансмисивност, могат да доведат до по-успешни инфекции.

Мутациите с усилване на функцията, са ключови за разбирането на рака. Тези мутации могат да направят клетките по-склонни да станат ракови, защото активират специфични гени (наречени онкогени) или нарушават нормалните сигнални процеси в клетките. Често срещан пример включва мутации в *KRAS* гена. Този ген често се променя при ракови заболявания, особено тези на панкреаса, дебелото черво и белите дробове. Промените в *KRAS* обикновено включват единични нуклеотидни промени, които водят до аминокиселинни замени – най-често глицин се заменя от валин в позиция 12 (Gly12Val), или от аспартат в позиция 12 (Gly12Asp), или глутамин се заменя от левцин на позиция 61 (Gln61Leu). Тези специфични мутации карат протеина K-Ras да бъде "включен" през цялото време, дори когато не трябва да бъде. Обикновено K-Ras се активира само в отговор на сигнали за растеж. Но с тези мутации K-Ras сигнализира непрекъснато, сякаш е заседнал в позиция "включено". Този постоянен сигнал казва на клетката да продължи да расте и да се дели, допринасяйки за развитието и разпространението на рак.

- Мутации с доминантно-отрицателен ефект: мутации, при които генният продукт притежава променена молекулярна функция, която действа антагонистично на нормалната версия на протеина от див тип.

Това често се случва, когато протеините нормално функционират като част от комплекс. Когато един протеин в комплекса е променен чрез мутация, това може да наруши функцията на целия комплекс. Например, туморният супресорен ген *TP53* кодира протеина p53, който играе критична роля в регулирането на клетъчния растеж и предотвратяването на образуването на тумори. Някои мутации в *TP53* произвеждат дисфункционални p53 протеини. Тези мутантни протеини могат да се свържат с нормалния (див тип) протеин p53, предотвратявайки ефективното му изпълнение на функцията. В резултат на това тази намеса позволява нерегулиран клетъчен растеж, което допринася за развитието на рак.

## 3.2. Спонтанни и индуцирани мутации

### 3.2.1. Спонтанните мутации са следствие на естествени биологични събития

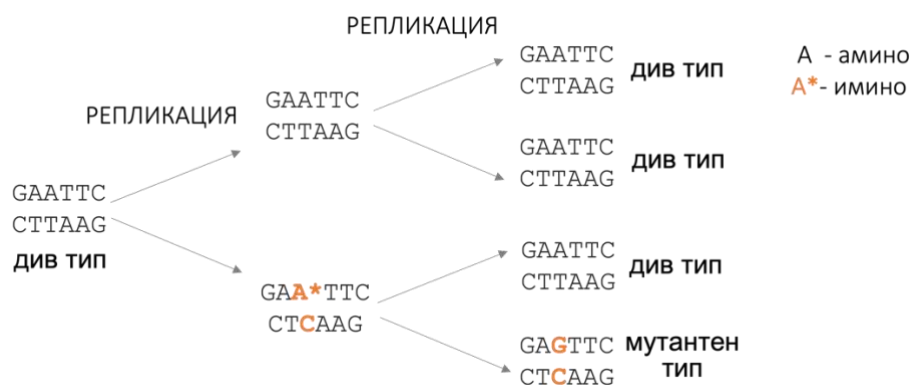
- Грешки при репликацията

При репликация, ДНК полимеразата може да катализира свързването на грешна база в полинуклеотидната верига и възникване на точкова мутации. В допълнение, грешките при репликация също могат да причинят инсерции и делеции, както и дублиране на ДНК сегменти.

- Тавтомеризация на ДНК

Азотните бази на ДНК имат способността временно да преминават в алтернативни форми, наречени тавтомери. Аденинът (А) и цитозинът (С) могат да превключват между аминок (стандартни) и имино форми. Гуанинът (G) и тиминът (Т) могат да превключват между кето (стандартна) и енолна форма. Тези алтернативни форми се различават по начина, по който образуват водородни връзки, така че могат да комплементират неправилно с други бази. Когато аденинът (А) е в своята имино форма, той може да комплементира с цитозин (С) вместо с тимин (Т). Когато гуанинът (G) е в своята енолна форма, той може да комплементира с тимин (Т) вместо с цитозин (С). Тавтомерните промени обикновено са временни. След като базата се върне в нормалната си форма, тя отново ще следва правилата за свързване на Уотсън-Крик. Но ако репликацията настъпи, докато базата е в тавтомерна форма, може да се образува грешка, което води до мутация.

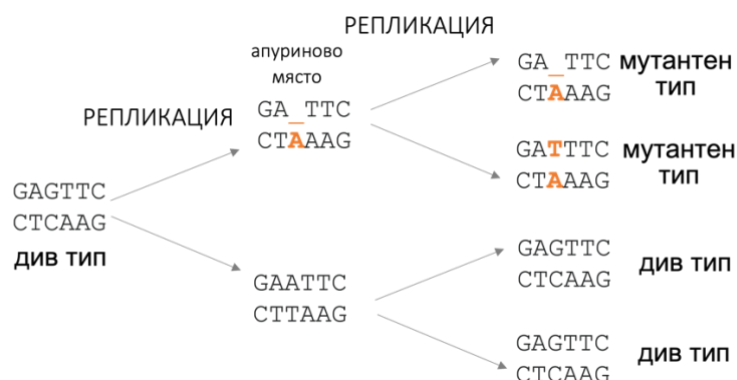
Например, ако по време на репликацията се образува рядката имино форма на аденин (А), обозначена като А\*, тя ще се свърже с цитозин (С), вместо с тимин (Т) (Фигура 3.4). Това води до несъответствие в една от двете новосинтезирани ДНК молекули: в едната ДНК ще присъства двойката А-Т, а в другата ДНК на съответстващата позиция ще присъства двойката А\*-С. При следващата репликация цитозинът (С), който е бил включен по време на предишната репликация, ще се свърже правилно с гуанин (G), образувайки правилната двойка С-G. Междувременно аденинът (А) ще се върне в своята нормална аминок форма (А) и ще се свърже правилно с тимин (Т). В резултат, след втората репликация, в една от четирите новосинтезирани ДНК молекули на дадената позиция в ДНК ще има двойката С-G, вместо А-Т. Възникнала е мутация - транзиция на аденин (А) на дивия тип в гуанин (G) при мутантния тип.



Фигура 3.4. Тавтомеризация на ДНК

- Химични модификации на нуклеотидите директно променят една база в друга база:
  - Депуриране: Точно както всяка друга молекула в клетката, ДНК непрекъснато претърпява термични сблъсъци с други молекули, често водещи до спонтанни химически промени в ДНК. Депурирането възниква, когато N-гликозидната връзка между пуриновата база и дезоксирибозата в гръбнака на ДНК се хидролизира (разкъсва), оставяйки апуриново място, където пуриновата база е била. Това води до образуване на апуриново място, където липсва пуринова база, без да се нарушава захарозо-фосфатния

скелет. Депурирането е вид спонтанна мутация, която ако не бъде поправена може да доведе до грешки по време на репликацията на ДНК (Фигура 3. 5). Аденинът се вмъква по-често срещу апуриново място, отколкото другите нуклеотиди.



Фигура 3.5. Депуриране на ДНК

- Дезаминиране: Друга често срещана реакция е спонтанната загуба на аминогрупа – дезаминиране. Най-често се извършва дезаминиране на цитозин (С), при което се получава урацил (U), който обикновено се намира само в РНК. Дезаминирането на цитозина (С) предлага обяснение защо ДНК съдържа тимин (Т) вместо урацил (U) - ако ДНК съдържаше урацил (U), тогава клетката не би могла да различи урацил (U), получен при дезаминиране на цитозин (С), от урацил (U), който нормално би присъствал в ДНК.

В клетката е възможно също така да настъпи и дезаминиране на 5-метил цитозин ( $5^mC$ ), което води до образуване на тимин (Т). Тъй като той е нормален компонент на ДНК, това създава проблем за поправящите ензими, тъй като те не могат да определят коя от двете бази на ДНК вериги е неправилната база ( $5^mC$  или Т). Поради тази причина метилираните цитозинови бази са склонни да създават горещи точки за мутации в геномите.

- Образуване на свободни радикали

Свободните радикали са нестабилни атоми или молекули, характеризиращи се с несдвоени електрони в най-външния си електронен слой, които се произвеждат естествено по време на клетъчния метаболизъм и имунните системни реакции. Докато в ниски концентрации те играят полезна роля като борба с инфекции и участие в клетъчната сигнализация, прекомерното им присъствие може да доведе значителни молекулни увреждания и да ги превърне в мутагенни фактори. В биологичните системи свободните радикали като супероксидните аниони ( $\cdot O_2^-$ ) и хидроксилни радикали ( $\cdot OH$ ) могат да взаимодействат с ДНК, потенциално предизвиквайки клетъчни мутации, както и с протеини и липиди.

Организмът разполага с естествени защитни механизми, включително ензими като супероксид дисмутаза, за неутрализиране на свободните радикали и поддържане на клетъчна хомеостаза. Въпреки това, външни фактори като радиация, замърсяване и начин на живот могат да увеличат производството на свободни радикали отвъд капацитета на тялото за неутрализация.

- Преместване на подвижни елементи (транспозони)

Транспозоните са мобилни генетични елементи, способни да променят позицията си в генома, играейки значителна роля в генетичните мутации и геномната еволюция. Те могат да индуцират мутации чрез два основни механизма: директно прекъсване на ДНК последователността, когато се вмъкнат в нови геномни местоположения, или промяна на

генната експресия чрез въвеждане или модифициране на генетични регулаторни последователности.

### 3.2.2. Индуцираните мутации възникват под въздействие мутагенни фактори

- Йонизиращото лъчение

Йонизиращото лъчение се състои от електромагнитни вълни или субатомни частици, които имат достатъчна енергия, за да йонизират атоми или молекули чрез отделяне на електрони от тях като по този начин може да предизвика промени в химичните и биологичните свойства на веществата. Видове йонизираща радиация: алфа ( $\alpha$ ) частици - състоят се от два протона и два неутрона (по същество ядро на хелий), големи и тежки, поради което имат ниска проникваща способност; бета ( $\beta$ ) частици - състоят се от високоенергийни, бързодвижещи се електрони или позитрони и имат по-голяма проникваща способност от алфа частиците; гама ( $\gamma$ ) лъчи – форма на електромагнитна радиация, подобна на видимата светлина, но с много по-висока енергия, която няма маса и може да прониква дълбоко в материалите, включително човешките тъкани; рентгенови лъчи - форма на електромагнитна радиация, чиято енергия обикновено е по-ниска от тази на гама лъчите, но все пак може да прониква дълбоко и да причинява йонизация; неутрони - свободни неутрони, освободени при ядрени реакции (като делене), които могат да изминават големи разстояния и лесно да проникват през материали, те взаимодействат с атомните ядра, което ги прави особено опасни в ядрени реактори или при ядрени взривове.

Естествени източници на йонизираща радиация включват космически лъчи, радон газ (от разпада на уран) и определени минерали, които естествено излъчват радиация. Изкуствени източници могат да бъдат рентгеновите машини, ядрените електроцентрали, медицински лечения (като лъчетерапия) и радиоактивни изотопи, използвани в научни изследвания или промишлеността.

Увреждането на ДНК, причинено от йонизираща радиация, може да се случи чрез пряк или непряк механизъм:

- Пряко увреждане възниква, когато йонизиращата радиация взаимодейства директно с молекулата на ДНК, предизвиквайки разкъсвания в захаро-фосфатния скелет на ДНК или модификации на базите. В този случай енергията от радиацията се абсорбира директно от структурата на ДНК.
- Непряко увреждане възниква, когато йонизиращата радиация взаимодейства с други молекули (особено с водата), обграждащи ДНК, като произвежда свободни радикали. След това свободните радикали могат да атакуват ДНК. Тъй като те са в състояние да дифузират на известно разстояние в клетката, не е необходимо първоначалното събитие на йонизация да се случи толкова близо до ДНК, за да причини увреждане. Установени са повече от 20 различни окислителни модификации на ДНК, от които преобладаваща е модификацията на гуанина до 8-оксо-2'-дезоксигуанозин. В много клинични проучвания концентрацията на 8-оксо-2'-дезоксигуанозин в урината се използва като чувствителен биомаркер за оксидативен стрес.

Йонизиращата радиация може да причини прекъсвания само на една от двете вериги (nick) и лесно се поправят от клетката, тъй като комплементарната верига не е увредена и може да се използва като матрица. Ако обаче настъпи прекъсване и на двете вериги, клетката изпитва много по-големи трудности при възстановяването на повредата и могат да възникнат грешки в ДНК. Това може да причини мутации и да доведе до последствия като рак или клетъчна смърт. Двуверижни скъсвания се появяват със скорост от около едно двойно прекъсване на 25 едноверижни прекъсвания. По този начин повечето

радиационни увреждания на ДНК са поправими. Йонизиращото лъчение може също да разруши имидазоловия пръстен на пурините. Последващото отстраняване на увредения пурин от ДНК чрез гликозилаза генерира апуриново място.

- UV радиация

UV радиацията обхваща диапазон от електромагнитния спектър между 200 и 400 nm. Този диапазон е допълнително разделен на късо-вълнова (200–280 nm, UV-C), средно-вълнова (280–320 nm, UV-B) и дълго-вълнова (320–400 nm, UV-A) UV светлина. UV-C светлината е с по-къса дължина на вълната носи със себе си значително повече енергия от дълговълновата UV-A светлина и е много по-вредна за ДНК. Тази увеличена способност за увреждане на ДНК от UV-C се дължи и на това, че ДНК абсорбира светлина при максимум 260 nm, което е в рамките на UV-C диапазона. За щастие обаче слънчевите UV-C лъчи не достигат до земята, защото се абсорбират от озоновия слой. Макар че UV-C лъчите представляват най-голямата опасност, UV-A и UV-B лъчите не са безвредни. При човек, UV-B и UV-A лъчите могат да проникнат съответно в епидермиса и дермата на кожата и са категоризирани като „канцерогени от клас I“ от Международната агенция за изследване на рака (IARC). Важно е, че докато UV-A светлината е по-малко енергийна от UV-B, тя е 20 пъти по-интензивна, като това съотношение варира в зависимост от географската ширина и сезона.

Основен механизъм, чрез който UV радиацията уврежда ДНК, е образуването на пиримидинови димери. Излагането на UV-B, и в по-малка степен на UV-A лъчи, води до образуването на най-често срещаните фотохимични продукти в ДНК - циклобутан пиримидинови димери. Те се образуват, когато два съседни пиримидина (тимини, ТТ или цитозини, СС) се свържат ковалентно за да формират цикличен пръстен (циклобутан), който обикновено не се среща в ДНК. Този фотохимичен продукт причинява структурно пречупване в ДНК, което пречи на пиримидините да се сдвояват със съответните бази и предотвратява репликацията на ДНК. За разлика от окислителните мутации, UV-индуцираните димеризационни мутации възникват от директната абсорбция на UV фотони. Излагането на UV лъчение може да доведе, подобно на йонизиращата радиация, до производството на свободни радикали, които след това взаимодействат с ДНК .

- ДНК алкилиране

Алкилиращите агенти представляват широк клас ДНК-увреждащи агенти, които добавят алкилни групи (метилни или етилни) към ДНК. Те могат да бъдат монофункционални (когато реагират само с една ДНК верига) и бифункционални (когато реагират с двете вериги). Монофункционалните агенти могат да прибавят алкилна група в специфично място в ДНК, най-често гуанин, причинявайки точкови мутации или неправилно свързване по време на репликация. Бифункционалните алкилиращи агенти свързват двете вериги на ДНК заедно и водят до т. нар. омрежване (от англ. cross-linking), в следствие на което пречат на тяхното отваряне и на достъпа на апаратите за репликация и транскрипция на ДНК.

В медицината, те са едни от първите химични вещества, определени като полезни при химиотерапия на рак и остават най-важните компоненти на съвременните химиотерапевтични протоколи (поотделно или в комбинация с други лекарства) поради доказаните си значими клинични противоракови действия. Причината за това е, че по-бързо пролифериращите ракови клетки всъщност имат по-малко време за възстановяване на възникнали увреждания на ДНК, преди да се разделят, отколкото другите клетки, което означава, че в присъствието на алкилиращи агенти вероятността за натрупване на увреждания на ДНК е по-голяма. В следствие на това се нарушават целостта и репликацията на ДНК, което може да спре растежа на раковите клетки или да доведе до апоптоза.

В лабораторни условия, алкилиращи агенти като метил етилметан сулфонат (EMS) и диметил сулфат (DMS) се използват за индукция на мутации за изучаване на функцията на гените и клетъчните процеси.

- Базовите аналози

Базовите аналози са химически вещества, наподобяващи азотните бази. Базов аналог може да бъде включен в новосинтезирана ДНК вместо нормална база. Примерите включват 2-амино пурин, съединение, което прилича на аденин (A), и 5-бромурацил (5 BU), съединение, което прилича на тимин (T). 5 BU може да претърпи тавтомеризация от обичайното кето състояние до рядкото енолно състояние. При репликация 5 BU в това състояние се свързва с гуанин (G) вместо с аденин (A), естественият партньор на тимина (T). Така при репликация двойката А-Т бива заместена с двойка G-C.

Базовите аналози често се използват в молекулярната биология като инструменти за изучаване на механизми за репликация и възстановяване на ДНК. Те играят ключова роля в напредъка на технологиите за секвениране на ДНК, предоставяйки нови възможности за точно и ефективно разчитане на генетична информация. Някои базови аналози имат терапевтични приложения, особено при лечение на рак, където могат да нарушат нормалния синтез на ДНК в бързо делящи се клетки.

- Интеркалиращи агенти

Това молекули, които имат хидрофобна хетероциклична пръстеновидна структура, наподобяваща структурата на азотните бази. Тези агенти се вмъкват (интеркалират) между базовите двойки на двойноверижни нуклеинови киселини в спиралата на ДНК, което в крайна сметка пречи на репликацията и води най-често до мутация с изместване на рамката.

Интеркалиращите агенти, като доксорубин и даунорубин, се използват в химиотерапията за инхибиране на репликацията на ДНК в бързо растящи ракови клетки. Те нарушават нормалните процеси на репликация и транскрипция, което води до смърт на клетките.

Интеркалиращите багрила като етидиев бромид, SYBR Gold и YOYO-1 са молекули, които могат да интеркалират в ДНК, което води до промяна на тяхната структура и усилва тяхната флуоресценция. Те се използват широко за визуализация на нуклеинови киселини в лабораторни техники като гел електрофореза и флуоресцентна микроскопия. РНК, особено в своя едноверижен вид, няма стабилната двойна спирала на ДНК и затова не се свързва толкова лесно с тези багрила. Все пак, някои от багрилата могат да се свързват с РНК при специфични условия, което позволява ограничена визуализация. Интеркалиращите багрила също така са ценни за изследване на структурните свойства на ДНК и нейните взаимодействия с различни лиганди, което е важно за разработката на нови лекарства, насочени към ДНК.

### 3.3. Поправка на ДНК

Разнообразието от живи организми и техния успех в колонизирането почти всяка част от повърхността на Земята е следствие от натрупаните генетични промени в продължение на милиони години. Някои от тези промени позволяват на организмите да се адаптират към променящите се условия и да виреят в нови местообитания. В краткосрочен план и от гледна точка на индивидуалния организъм обаче, генетичните промени могат да бъдат пагубни. За да оцелеят и да се размножават, индивидите трябва да бъдат генетично стабилни, което се постига не само чрез изключително точния механизъм за репликация на ДНК, но също и чрез работа на различни протеинови апарати, които непрекъснато сканират генома за повреди и ги поправят, когато възникне. Повечето увреждания на ДНК са само временни, защото веднага

се коригират чрез процеси, наричани общо поправка на ДНК. Основните пътища за поправка са представени на Фигура 3.6.



### Нобелова награда за откриване на механизмите за поправка на ДНК

През 2015 година Нобеловата награда по химия беше присъдена на трима учени, които разкриват как клетките поправят увредената ДНК и защитават генетичната информация.

Томас Линдал установява, че ДНК е нестабилна молекула, подложена на разпад дори при физиологични условия. Той идентифицира нова група от ДНК гликозилази, които играят роля в поправката на азотните бази на ДНК. Пол Модрих трансформира разбирането за поправка на несъответствия в ДНК, описвайки как клетките коригират грешки, които възникват по време на репликация, намалявайки честотата на грешките значително. Азиз Канкър изследва механизма на поправка на нуклеотиди, който клетките използват за отстраняване на увреждания от UV лъчение. Той също така обяснява молекулярните механизми зад фотореактивацията.

Тези открития не само че предизвикват ново разбиране за механизмите на поправка на ДНК, но и предоставиха основа за разработването на нови терапии за рак, тъй като много противоракови лекарства причиняват увреждане на ДНК и способността на раковите клетки да поправят това увреждане влияе на лечението им.



Фигура 3.7. Пътища за поправка на ДНК

### 3.3.1. Директна поправка

Директната реверсивна поправка коригира специфични типове увреждане на ДНК без необходимост от изрязване или замяна. Използва се за поправка ДНК увреждания (лезии), предизвикани от ултравиолетови лъчи и алкилиращи агенти.

UV светлината индуцира образуването на пиримидинови димери, които могат да изкривят структурата на веригата на ДНК, и да блокират репликацията и транскрипцията извън зоната на увреждане. Ензимната фотореактивация е основен механизъм за директна поправка на циклобутан пиримидинови димери, при който ензимът фотолиаза чрез използване на светлинна енергия за разрушаване на аномалната ковалентна връзка между съседни пиримидинови бази. Този тип фотореактивация не се среща при хора.

Увреждането, причинено от алкилиращи агенти, реагиращи с ДНК, също може да бъде поправено чрез директна поправка. Алкилираните бази могат да бъдат обърнати от ензими като Об-алкилгуанин-ДНК алкилтрансфераза (AGT) и диоксигенази, които премахват или модифицират съответно алкилна група.

### 3.3.2. Поправка чрез изрязване

Изрязването е общият механизъм, чрез който се правят поправки, когато една от веригите на двойната спирала е повредена. Недефектната верига се използва като матрица, като повреденият участък на другата верига се отстранява и заменя чрез синтеза на нови нуклеотиди. Поправката чрез изрязване включва три основни механизма:

- Поправка на несъответствията (от англ. MisMatch Repair, MMR)

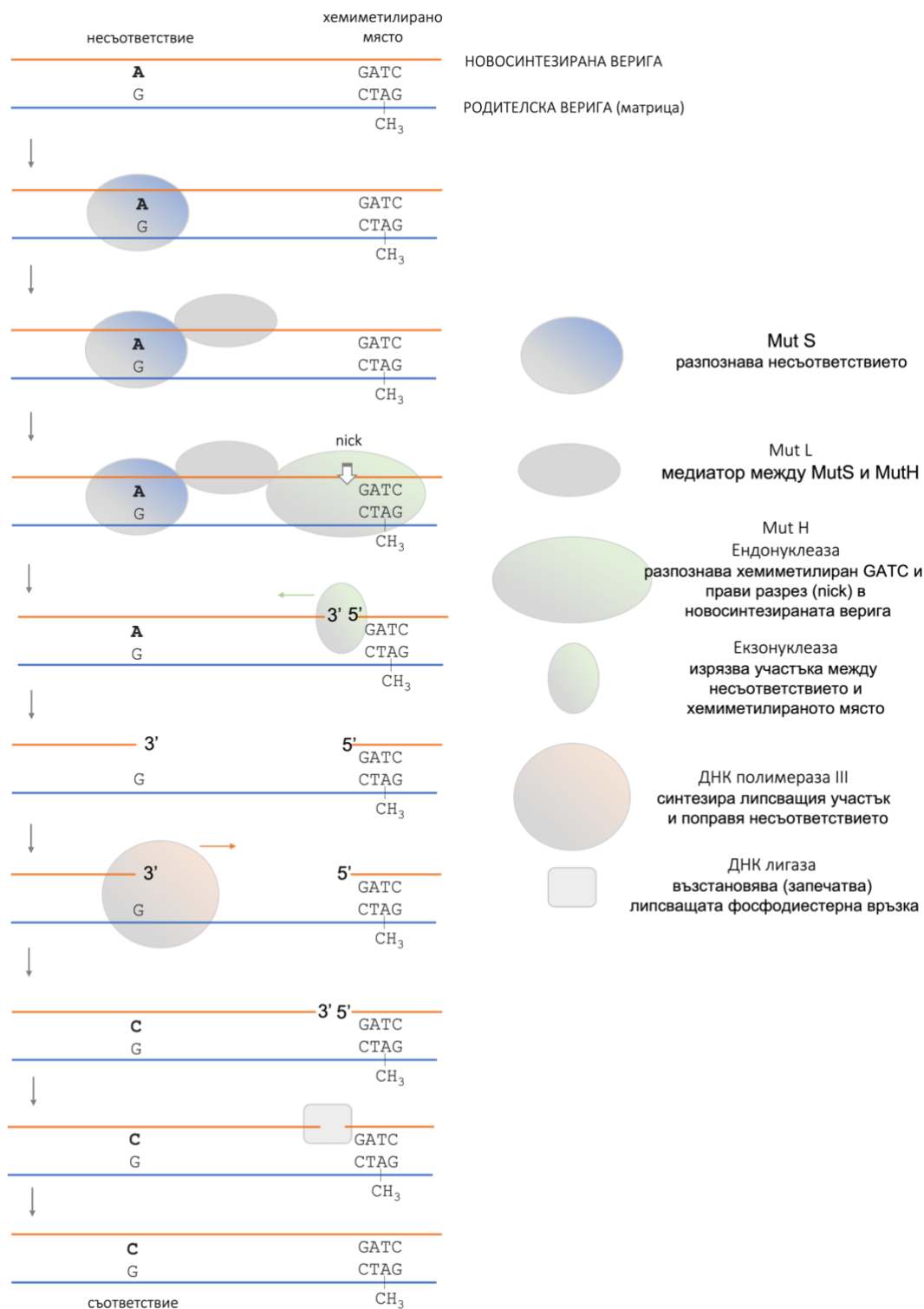
MMR е критичен процес, който коригира грешки, възникнали по време на репликацията на ДНК, по-специално несъответствия между базите, получени при свързване на некомплементарни бази в двойка, както и къси инсерции и делеции, които причиняват изпъквания на двойната спирала. Повечето от тези грешки се коригират от коригиращата активност на ДНК полимераза по време на репликация, но някои може да бъдат пропуснати и трябва да бъдат коригирани по-късно.

В *E. coli*. при поправката, неправилно свързания нуклеотид трябва да бъде изрязан от новосинтезираната верига, което изисква разграничаването и от матричната родителска верига. Важна особеност на ДНК на *E. coli* и много прокариоти е присъствието на копия на последователността 5'-GATC-3', в която адениновите остатъци са метилирани. Метилирането се катализира от ДНК аденин метилтрансфераза (Dam), която прехвърля метилова група от S-аденозилметионин (SAM) към адениновия остатък в рамките на GATC последователността. Метилирането възниква след репликация на ДНК и ензимът се нуждае от няколко минути да разпознае GATC последователността в новосинтезираната верига и да метилира аденина (A). В този кратък период, новосинтезираната ДНК е неметилирана, а матричната родителска верига е метилирана, което позволява разграничаването им по време на процесите на поправка на ДНК. MutS, MutL и MutH са основни протеини, участващи в системата за поправка на несъответствията (Фигура 3.7). Основните стъпки на процеса са:

- Разпознаване на несъответствието: MutS разпознава несъответстващи двойки бази, които се появяват по време на репликацията на ДНК; MutS образува хомодимер и преминава през конформационни промени, които му позволяват да взаимодейства с други протеини в пътя за MMR като привлича MutL, за да образува комплекс, който е критичен за активирането на процеса на поправка.
- Разграничаване на веригите и изрязване: Комплексът MutS-MutL улеснява активирането на ендонуклеазата MutH. Този ензим играе решаваща роля в метил-насочения път на поправка на несъответствието чрез разпознаване на ДНК веригата, в която GATC

последователности не са метилирани, и едноверижно разрязване (nick) в най-близката до несъответствието GATC последователност. След това екзонуклеази изрязват фрагмента между несъответствието и едноверижното разрязване при GATC.

- Синтез на правилната последователност: ДНК полимераза III синтезира правилната последователност от ДНК, използвайки родителската верига като матрица. Накрая, ДНК лигазата възстановява липсващите фосфодиестерни връзки в поправената ДНК верига.



Фигура 3.7. Поправка на несъответствия в ДНК при прокариоти (MMR механизъм)

При човек, в поправката на несъответствията участват протеини хомоложни на прокариотните:

- MutS хомолози (MSH): MSH2 - разпознава несъответствия в ДНК. Той може да образува два комплекса - MutS $\alpha$  (хетеродимер с MSH6), който участва в разпознаването на грешки в базовите двойки и малки инсерции или делеции, или MutS $\beta$  (хетеродимер с MSH3), който е по-специализиран в разпознаването на по-големи вмъквания/заличавания.
- MutL хомолози (MLH/PMS): MLH1 - посредници в процеса на поправка след разпознаването на грешката – MutL $\alpha$  (хетеродимер с PMS2), който е основният комплекс, участващ в процеса на поправка след разпознаването на несъответствията, или MutL $\gamma$  (хетеродимер с MLH3), който участва в поправката на определени типове увреждания на ДНК, включително по-големи вмъквания/заличавания.

Несъответстващият участък се разгражда от екзонуклеаза 1 (Exo1), след това ДНК полимеразата  $\delta$  синтезира правилната последователност. И накрая, ДНК лигаза „запечатва“ липсващите фосфодиестерни връзки в поправената ДНК.

При човек, мутациите в MMR гените могат да доведат до синдром на Линч, наследствено заболяване, свързано с повишен риск от рак на дебелото черво, яйчниците и други видове рак.

- Поправка чрез изрязване на база (от англ. Base-Excision Repair, BER)

BER включва разпознаване, отстраняване и замяна на единична увредена азотна база. Механизмът изисква семейство от ензими, наречени гликозилази. Ензимите премахват увредената база, образувайки апуриново място (AP място). То се поправя от AP ендонуклеаза, която изрязва захарозо-фосфатния остатък, след което нуклеотидната празнина в ДНК веригата се запълва от ДНК полимераза и се запечатва от лигаза. Пример за BER е поправката на ДНК, съдържаща урацил (U).

- Поправка чрез изрязване на нуклеотиди (от англ. Nucleotide-Excision Repair, NER)

NER се използва за поправка на пиримидинови димери, образувани от UV светлина, и увреждания, причинени от химични агенти. При *E. coli*, поправката чрез изрязване на нуклеотиди използва няколко протеина, сред които UVR протеините, и включва следните етапи:

- Разпознаване на увредения участък: UvrA сканира ДНК за изкривявания, причинени от увреждане. При разпознаване на лезия образува комплекс с UvrB. UvrB се свързва с мястото на лезията и помага за разделянето на ДНК веригите, за да се провери позицията на повредата.
- Изрязване на увредения участък: Ендонуклеазата UvrC разрязва увредената ДНК верига от двете страни на лезията, което позволява изрязване на увредения участък 8 нт преди и 4 нт след увредения участък. Хеликазата UvrD премахва увредения фрагмент от ДНК.
- Синтез на правилната последователност: ДНК полимераза I запълва празнината, използвайки комплементарната верига като матрица. ДНК лигазата възстановява липсващите фосфодиестерни връзки в поправената ДНК верига.

При човек са установени множество хомолози на UVR протеините:

- XPA и XPC- хомолози на UvrA, които разпознават увредената ДНК;
- TFIIH комплекс (XPB и XPD) - хомолог на UvrB, развива ДНК в областта на увреждането;
- ERCC1/XPF и XPG - хомолози на UvrC, които разрязват увредената ДНК;
- XPB/XPD – хомолози на UvrD, развиват увредената ДНК и премахват изрязания фрагмент.

PCNA подпомага свързването на ДНК полимеразата  $\delta$  или ДНК полимеразата  $\epsilon$  в мястото на поправка, където те синтезират нова ДНК, за да запълнят празнината, създадена след изрязването на увредената ДНК. Лигаза I или Лигаза III завършват процеса на възстановяване.

#### Пигментна ксеродерма Xeroderma Pigmentosum (XP)

XP е рядко генетично заболяване, характеризиращо се с изключителна чувствителност към UV светлина. То следва автозомно рецесивен модел на наследяване, което означава, че човек трябва да наследи две дефектни копия на съответните гени (по едно от всеки родител), за да прояви болестта. XP засяга предимно кожата, очите и в някои случаи нервната система.

XP се причинява от мутации в гени, отговорни за NER, решаващ процес на поправка на ДНК, който коригира UV-индуцирани увреждания. Някои от тях са:

- XPA - кодира протеин, който свързва увреденото място и помага за сглобяването на другите протеини, необходими за NER;
- XPB и XPD, които са част от TFIIH, някои мутации в XPB и XPD също предизвикват признаци на преждевременно стареене;
- XPF, който реже ДНК веригата от 5' страната на повредата;
- XPG, който реже ДНК веригата от 3' страната на повредата.

NER може да бъде класифицирана на два основни типа:

- Глобален геномен NER (GG-NER): Този път сканира целия геном за увреждане и е от решаващо значение за възстановяване на увреждания, които се появяват в нетранскрибирани региони на ДНК.
- Транскрипционно-свързан NER (TC-NER): Този път е насочен конкретно към увреждания в матричната верига на активни гени и се иницира, когато РНК полимеразата бъде блокирана от увреждане.

#### 3.3.3. Поправка на двойно-верижни прекъсвания на ДНК

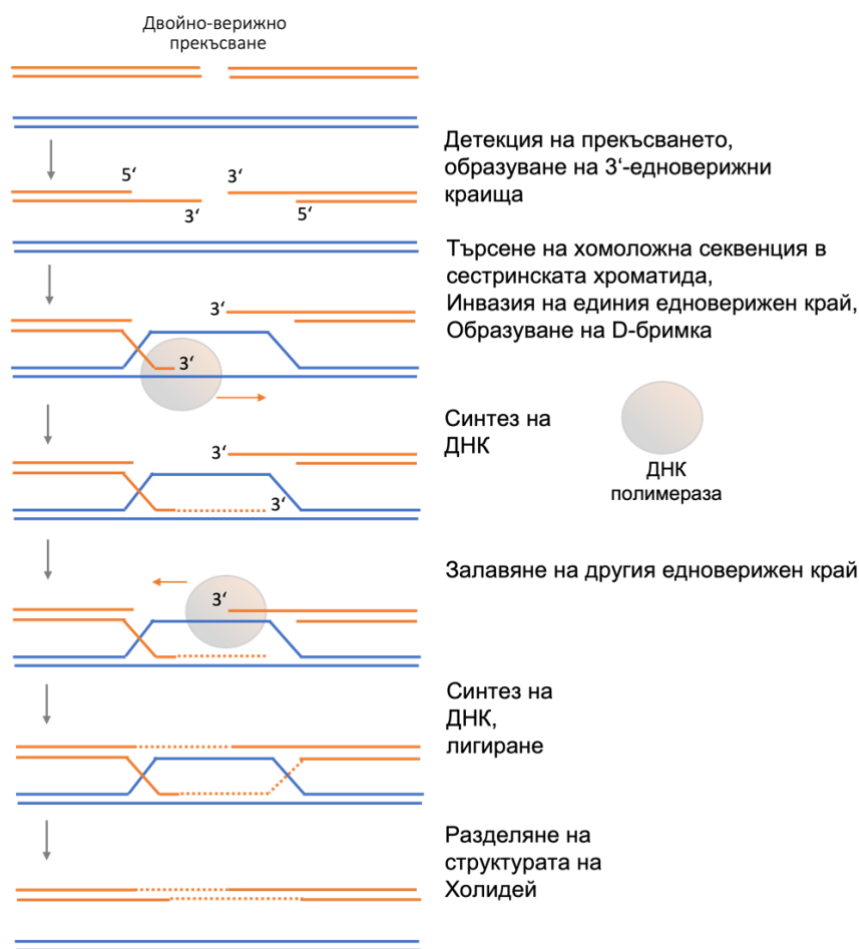
Поправката на двойно-верижни скъсвания е от решаващо значение за запазване на геномната цялост, тъй като тези скъсвания, ако не бъдат поправени, могат да доведат до хромозомни пренареждания, мутации и дори клетъчна смърт. Двуверижните разкъсвания са особено вредни, защото засягат и двете ДНК вериги, като потенциално нарушават непрекъснатостта на генома и причиняват загуба на генетична информация. Клетките разчитат на два основни механизма за тяхната поправка:

- Хомоложна рекомбинация (от англ. Homologous Recombination, HR)

Хомоложната рекомбинация е прецизен механизъм за поправка, който използва хомоложна ДНК последователност като матрица - сестринската хроматида. Основните стъпки са представени на Фигура 3.7.

При прокариоти процесът включва няколко ключови протеини:

- RecA - основен протеин за процеса на хомоложна рекомбинация, тъй като насърчава търсенето на хомоложна последователност и улеснява инвазията на разкъсаната верига в хомоложната последователност.
- RecBCD комплекс - участва в обработката на разкъсаните краища на ДНК за получаване на 3'-едноверижни краища (от англ. overhangs) и подготовката им за инвазия на вериги.
- RuvABC комплекс - разрушава структурата на Холидей, която се образуват по време на рекомбинация, и позволяват на ДНК молекулите да се разделят.



Фигура 3.7. Поправка на двойно-верижни прекъсвания чрез хомоложна рекомбинация

В еукариотните клетки хомоложната рекомбинация е най-активна по време на S, G2 и ранните M фази на клетъчния цикъл, когато е налична сестринска хроматида, която да служи като матрица за възстановяване. При хората ключовите стъпки и протеини, участващи в HR, са както следва:

- Откриване на двойно-верижното прекъсване и обработка на крайщата: Клетката разпознава прекъсването чрез сензорни протеини като MRN комплекса (Mre11-Rad50-Nbs1). Той се свързва с мястото на прекъсване и създава 3'-едноверижни краища на ДНК, които са решаващата стъпка за инициране на процеса.
- Търсене на хомоложна последователност: Протеини като Rad51 подпомагат търсенето на хомоложна ДНК матрица (обикновено върху сестринска хроматида).
- Инвазия на едноверижния край: Едноверижният край на ДНК навлиза в хомоложната последователност, образувайки D-примка, чрез сдвояване с комплементарната ДНК верига на сестринската хроматида.
- Синтез на ДНК: Нова ДНК се синтезира, използвайки неповредената матрица.
- Разрушаване на структурата на Холидей: Ендонуклеазите и хеликазите помагат за разрешаването на тези кръстовища, като възстановяват оригиналната двойноверижна ДНК и завършват възстановяването.

Дефекти в протеините, участващи в хомоложната рекомбинация, могат да доведат до онкогенеза. Например, мутации в гените *BRCA1/BRCA2*, които участват в регулирането на процеса на хомоложна рекомбинация, увеличават риска от развитие на рак на гърдата.

- Нехомоложно свързване на краищата (от англ. Non-Homologous End Joining, NHEJ)

Нехомоложното свързване на краищата не изисква хомоложни последователности за поправка, което го прави по-гъвкав, но и по-податлив на грешки в сравнение с хомоложната рекомбинация. Това може да доведе до инсерции или делеции в мястото на поправката. Макар, че механизмът е склонен към грешки, той защитава целостта на генома от възможни хромозомни транслокации, които могат да възникнат чрез хомоложна рекомбинация. Той е по-широко представен в еукариоти в сравнение с прокариоти, и при човек, включва следните основните стъпки:

- Откриване на двойно-верижното разкъсване - хетеродимерът Ku (Ku70/Ku80) се свързва с краищата на разкъсаната ДНК.
- Обработка на краищата - ензими модифицират краищата, за да ги направят съвместими за свързване, при което често се въвеждат мутации.
- Свързване на краищата - ДНК лигаза IV, заедно с XRCC4 и XLF, запечатва обработените краища, завършвайки поправката.

Нехомоложното свързване на краищата има роля в инженерството на генома. То често се използва в технологии за редактиране на генома, като CRISPR/Cas9. Когато Cas9 въвежда двойно-верижни прекъсвания, нехомоложното свързване на краищата често поправя тези разкъсвания, водейки до целеви инсерции или делеции в генома.

### 3.4. Соматични и зародишни мутации и ефекти върху потомството

Соматичните мутации могат да възникнат във всяка клетка на тялото, с изключение на зародишните клетки. Те не се предават на потомството на дадения организъм, но ще присъстват във всички потомци на мутиралата клетка в рамките на същия организъм. Много видове рак са резултат от натрупани соматични мутации.



Жълтият цвят на венчелистчето на това червено лале е причинен от соматична мутация в клетките, които образуват това венчелистче. Тъй като соматичните мутации не засягат зародишната линия (репродуктивните клетки), всички семена, произведени от това лале, няма да носят белег за жълтото оцветяване.

Мутация на зародишната линия или зародишна мутация е всяка вариация в ДНК на зародишните клетки (клетки, които, когато са напълно развити, се превръщат в сперматозоиди и яйцеклетки). Мутациите в тези клетки са единствените мутации, които могат да бъдат предадени на потомството, в случай че мутирал сперматозоид, или мутирала яйцеклетка участват в оплождане. Зиготата започва делене и формиране на клетките в новия организъм, в следствие на което мутацията ще присъства във всяка соматична и зародишна клетка на потомството.

Разбирането на разликите между зародишната линия и соматичните мутации е от решаващо значение за генетиката, медицината и изследванията на рака. Соматичните мутации са важни за разбирането на туморогенезата и развитието на различни заболявания през целия живот

на индивида. Мутациите на зародишната линия могат да имат дългосрочни последици за наследствеността, в зависимост от това могат да бъдат класифицирани като:

- Неутрални мутации, които нямат забележим ефект върху организма, например когато мутацията възниква в участък от ДНК без функция или възниква в протеин-кодираща област, но не засяга аминокиселинната последователност на протеина. Напр. вариации, като различни цветове на очите при хората, често са неутрални, тъй като не влияят на размножаването или оцеляването.
- Вредни мутации, които увреждат способността на организма да оцелее и да се възпроизвежда. Например при хората синдромът на Марфан се причинява от мутация, засягаща протеин, който е част от съединителната тъкан, което води до сърдечни проблеми и други здравословни проблеми. Пагубните мутации, известни като смъртоносни, нарушават ДНК последователности в генома, които са критични за оцеляването, и причиняват смърт на организма.
- Полезни мутации, които подобряват способността на организма да оцелее и се размножава в своята среда. Тези мутации могат да доведат до предимства, които могат да бъдат селектирани през поколенията. Напр. антибиотична резистентност при бактериите, в основата на която са мутации, които осигуряват резистентност към антибиотици и позволяват на тези бактерии да оцелеят в среди, където антибиотиците са налични. Някои насекоми са развили резистентност към пестициди като ДДТ поради специфични мутации, което им помага да оцелеят в третирани райони.

#### Полезният Apo A1M

ApoA-I Milano е естествено срещан вариант на протеина аполипопротеин А-I, който е свързан с липопротеина с висока плътност (HDL) и играе решаваща роля в метаболизма на холестерола. Мутацията ApoA-I Milano е идентифицирана през 70-те години на миналия век в малко италианско село, където е установено, че около 3.5% от населението носи тази мутация. Тя се характеризира със замяна на аргинин с цистеин на позиция 17. Въпреки по-ниските нива на HDL холестерол, носителите на мутацията ApoA-I Milano не показват увеличен риск от сърдечно-съдови заболявания. Тази парадоксална ситуация предполага, че мутираният протеин може да има подобрена функционалност в метаболизма на холестерола и атеропротекцията. По този начин италианската общност, която носи Apo A1M, е защитена от сърдечни заболявания.

ApoA-I Milano представлява уникален пример за генетична вариация, която предлага защита срещу сърдечно-съдови заболявания и има потенциал за терапевтични приложения.

#### • Казус за ефектите на мутациите: сърповидно-клетъчна анемия

Сърповидно-клетъчната анемия е генетично заболяване на кръвта, характеризиращо се с необичайно оформени червени кръвни клетки. Мутациите, които причиняват това заболяване, са широко изследвани и предоставят ясен пример за това как генетичните мутации могат да предизвикат ефекти на различни биологични нива – от ДНК до целия организъм.

- Точкова мутация в ДНК: Сърповидно-клетъчната анемия се причинява от точкова мутация в гена HBB, който кодира бета-глобиновата верига на хемоглобина. Тази мутация променя един нуклеотид в ДНК последователността, замествайки аденин (A) с тимин (T).
- Мутация с промяна на смисъла на протеиново ниво: Точковата мутация води до замяна на аминокиселината глутаминова киселина с валин на 5 позиция в молекулата на

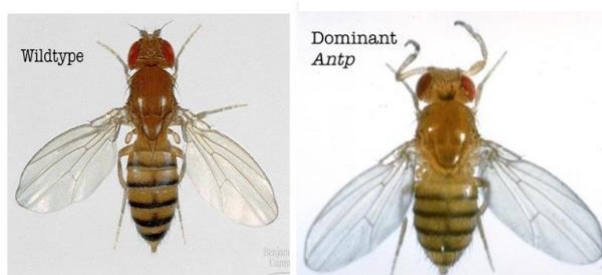
хемоглобина. Промененият хемоглобин, наречен хемоглобин S (HbS), има различна структура от нормалния хемоглобин (HbA). Поради замяната с валин, молекулите на HbS са по-склонни да се слепват при ниско съдържание на кислород, образувайки твърди влакна, които изкривяват формата на червените кръвни клетки.

- Ефекти на клетъчно ниво: Червените кръвни клетки, съдържащи HbS, придобиват „сърповидна форма“ при ниско съдържание на кислород, за разлика от нормалните кръгли червени кръвни клетки. Това сърповидно изкривяване намалява гъвкавостта на клетките, причинявайки им да се слепват и понякога да блокират кръвния поток в малките съдове, което води до увреждане на тъканите и болка.
- Отрицателни ефекти върху целия организъм: Хората със сърповидно-клетъчна анемия, особено при условия на нисък кислород, като например във високите планини или при интензивни физически натоварвания, изпитват симптоми като силна болка, умора, анемия и потенциално сериозни усложнения, дължащи се на запушени кръвоносни съдове и увреждане на органи.
- Положителен ефект върху целия организъм: Носителите на сърповидно-клетъчния признак (които имат един нормален и един мутирал бета-глобинов алел) имат селективно предимство в райони с висока заболяемост от малария. Маларийният паразит не може да оцелява в червените кръвни клетки, съдържащи HbS, което предоставя известна устойчивост на малария и обяснява високата честота на мутацията в определени популации.

### 3.5. Малки мутации с големи ефекти - мутации на регулаторни гени

Мутациите в регулаторните гени може да има значителни ефекти върху развитието, физиологията и поведението на един организъм, тъй като тези гени контролират кога, къде и как други гени се експресират. Тези гени играят ключова роля в регулирането на генната експресия и мутациите в тях могат да нарушат нормалното функциониране на цели биологични пътища, водейки до различни резултати, включително развитието на заболявания, като рак, генетични дефекти и дори еволюционни промени.

Много организми имат мощни контролни гени, които определят как е изградено тялото. Например, Нох гените се намират в много животни (включително винена мушица и човек) и определят основната структура и оформление на тялото при животните. Такива главни контролни гени помагат за насочването на изграждането на телесни „единици“, като сегменти, крайници и очи. Мутациите в Нох гените могат да доведат до значителни изменения в развитието на организма. Например, мутации в тези гени могат да предизвикат преобразуване на една част от тялото в друга, което е наблюдавано при някои видове насекоми.



Мутациите в регулаторните гени могат да трансформират една част от тялото в друга. Учените са изследвали винени мушици, носещи Нох мутации, при които израстват крака на челата им вместо антени – напр. гена *Antp* (*antennapediae*).

При човек, мутациите в регулаторни гени, особено в онкогени (гени, които стимулират клетъчния растеж) или тумор-супресорни гени (гени, които потискат клетъчния растеж, предотвратявайки неконтролирана клетъчна пролиферация и насърчавайки програмирана клетъчна смърт, когато е необходимо), са основни двигатели на развитието на рак. Например, мутациите в гена RB1, който регулира клетъчния цикъл, могат да доведат до ретинобластом – вид рак на окото. Нарушената регулация на гени, свързани с апоптоза (програмирана клетъчна смърт), клетъчно делене или поправка на ДНК, поради мутации в регулаторни гени, може да насърчи неконтролиран клетъчен растеж и рак.

Мутациите в регулаторните елементи на гените - в промотори или енхансери могат да променят силата, с която се експресира даден ген. Например, мутации в енхансери, които контролират HOX гените, могат да доведат до вродени малформации или изменено развитие. Мутации в некодиращи региони, особено в регулаторни елементи, са свързани с различни видове рак. Тези мутации могат да променят свързващата способност на ТФ, водейки до промени в генната експресия, които насърчават туморогенезата. Например, мутации в промотора на *TERT* гена, кодиращ каталитичната субединица на ензима теломераза, често се наблюдават при някои видове рак и са свързани с реактивация на теломеразата, което допринася за прогресията на заболяването.

## ГЛАВА 4 СИНТЕЗ на РНК - ТРАНСКРИПЦИЯ

Ако ДНК често се отъждествява с молекулата на живота, то генната експресия, която в крайна сметка води до изпълнение на инструкциите, записани в ДНК, със сигурност е процесът на живота. Основният клетъчен метаболизъм, клетъчното делене и невероятните процеси на клетъчна диференциация и адаптация към условията на външната и вътрешната среда, всички те зависят от контролираната експресия на специфични гени, започваща с процеса на транскрипция.

Протеин-кодиращите гени се транскрибирани в РНК молекули, които могат да кодират протеини - информационна РНК (иРНК). Много участъци от ДНК могат да се транскрибират в РНК молекули, които са функционални без да произвеждат протеин - некодиращи РНК (като рибозомна РНК, транспортна РНК, малки РНК с регулаторна функция и др.).

### 4.1. Въведение в транскрипцията

Транскрипцията е процес на пренос на генетичната информация от ДНК към РНК, който се извършва чрез синтез на РНК върху матрица ДНК. Процесът се катализира от ензима ДНК-зависима РНК полимераза (или само РНК полимераза). Ензимът се движи по едната верига на ДНК, която се нарича матрична верига и има 3'-5' поляритет, докато ДНК веригата с 5'-3' поляритет се нарича смислена верига (или кодираща) (Фигура 4.1). Новосинтезираната РНК е идентична със смислената ДНК верига с единствената разлика, че вместо тимин (Т) съдържа урацил (U).



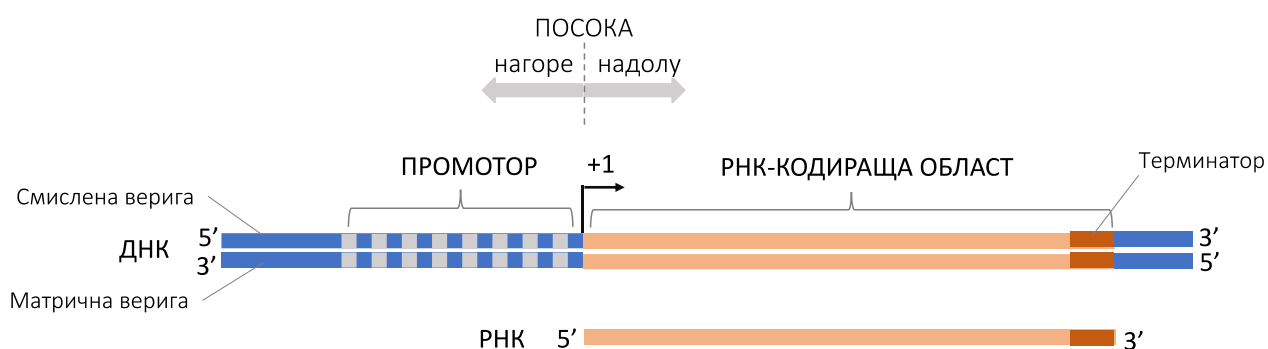
Фигура 4.1. Смислена (кодираща) верига и матрична верига на ДНК при транскрипция

#### 4.1.1. Транскрипционна единица

Транскрипционната единица е последователност от нуклеотиди в ДНК, която кодира една молекула РНК, заедно с последователностите, необходими за нейната транскрипция. Тя включва три компонента: регулаторна област - промотор, РНК-кодираща област и терминатор (Фигура 4.2). Нуклеотидът в кодиращата верига на ДНК, от който РНК полимеразата започва

транскрипцията, и който кореспондира на първия нуклеотид в транскрибираната РНК, се нарича Транскрипционно начало (от англ. Transcription Start Site , TSS) и се бележи с +1. Много често стартовата точка е пурин, най-често А - средната база на секвенцията САТ, но нейната консервативност не е достатъчна за да служи като сигнал за инициация на транскрипцията. Спрямо транскрипционното начало са дефинирани посоките в транскрипционната единица – посоката, която съвпада с посоката на движение на РНК полимеразата се означава като надолу от транскрипционното начало, а обратната на нея посока, се означава като нагоре от транскрипционното начало. При повечето гени, промоторът се разполага непосредствено нагоре от транскрипционното начало, той не се транскрибира, но има незаменима регулаторна функция, тъй като към него се свързва транскрипционната машина и се инициира транскрипция.

Обикновено транскрипционната единица обхваща един ген и продуцира една моноцистронна РНК, макар и по-рядко, тя може да обхваща няколко гена и да продуцира полицистронна РНК.



Фигура 4.2. Транскрипционна единица

#### 4.1.2. РНК полимерази

При прокариоти и археи съществува само един тип РНК полимераза. Еукариотната клетка използва три различни РНК полимерази, а в растителната клетка се откриват още две РНК полимерази. Транскрипцията в митохондриите и пластидите се извършва от специфични за органелите РНК полимерази.

- В процеса на транскрипция се синтезира РНК, която е полимерна молекула. Основната активност на РНК полимеразата е 5'-3' полимеразна активност - катализира синтеза на РНК чрез добавяне на рибонуклеотиди (нуклеозид монофосфати – AMP, GMP, CMP, UMP) към нарастващата верига на РНК, използвайки ДНК като матрица.
- Субстратите на реакцията са четирите типа нуклеозид трифосфати - АТФ, ГТФ, СТР, УТР.
- Нуклеотидната последователност на РНК веригата се определя чрез правилото за комплементарност (сдвояване на базите) между влизания в активния център на ензима нуклеозид трифосфат и ДНК матрицата.
- Каталитичният механизъм на РНК полимеразата е подобен на този на ДНК полимеразата – 3'-ОН групата на нарастващата РНК верига (или първият нуклеозид трифосфат) извършва нуклеофилна атака върху  $\alpha$ -фосфата на входящия нуклеозид трифосфат, ако той е комплементарен на матрицата, при което се освобождава пирофосфат и се формира фосфодиестерна връзка.
- Всяка РНК молекула започва с нуклеозид трифосфат – единствено при него се запазва 5'-трифосфатна група, която поставя началото на синтезиращата се РНК.

- За разлика от ДНК полимеразата, РНК полимеразата не се нуждае от предварително съществуваща верига (праймер), която да започне да удължава.
- РНК полимеразата няма високо активна и ефективна функция за корекция по време на транскрипцията.
- РНК полимеразата е мултисубединичен комплекс.
- РНК полимеразата има хеликазна активност за да разплита локално двете вериги на ДНК.

#### 4.1.3. Основни етапи на транскрипцията

**Инициация** - Започването на транскрипцията е особено критичен процес, тъй като това е основната точка, в която клетката избира кои протеини или РНК да бъдат произведени. За да започне транскрипцията, РНК полимеразата трябва да може да разпознае промотора и да се свърже здраво с ДНК на това място. Създава се **транскрипционното мехурче** чрез локално развиване на ДНК в областта на свързване на РНК полимеразата. Начинът, по който РНК полимеразите разпознават промотора, се различава между прокариоти и еукариоти.

**Удължаване** - РНК полимеразата се движи върху матричната ДНК верига и синтезира РНК в посока 5'-3' като прибавя нуклеотиди към 3'-ОН края на нарастващия транскрипт.

**Прекратяване** - синтезата на РНК се прекратява в определен участък на ДНК, наречен терминатор, и се освобождава новосинтезираната РНК и ензима.

## 4.2. Прокариотна транскрипция

### 4.2.1. Прокариотна РНК полимераза

Каталитичното ядро на бактериалната РНК полимераза съдържа четири основни субединици – 2 копия на  $\alpha$ -субединицата и по 1 копие на  $\beta$ -,  $\beta'$ - и  $\omega$ -субединиците ( $\alpha_2\beta\beta'\omega$ ). Това е основната форма на ензима, наречена сърцевинен или кор-ензим (от англ. core - сърцевина). За позициониране на кор-ензима върху правилния промотор е необходимо свързването на още една субединица, наречена сигма ( $\sigma$ )-субединица или  $\sigma$ -фактор - комплексът на кор-ензима и  $\sigma$ -субединица ( $\alpha_2\beta\beta'\omega\sigma$ ) се нарича холоензим.

Кор-ензимът е напълно способен да катализира полимеризацията на нуклеотидите в РНК, което показва, че  $\sigma$ -факторът не е необходим за основната каталитична активност на ензима. Кор-ензимът обаче има общ афинитет към ДНК – той може да се свързва към случайни места в ДНК без да разграничава конкретна секвенция. Свързването на кор-ензима към ДНК е хлабаво и ДНК остава двойно-верижна. Кор-ензимът не може да разграничи промоторната секвенция от другите секвенции на ДНК. За разлика от него,  $\sigma$ -факторът има способността да разпознава специфични секвенции в ДНК – промоторите. Благодарение на  $\sigma$ -фактора, свързването на холоензима към промотора е здраво и води до локално разделяне на двете вериги на ДНК. Изборът на промоторите е критичен елемент от транскрипцията, тъй като синтезът на функционална РНК трябва да започне в началото на гена.

Таблица 4.1. Субединици на прокариотната РНК полимераза

СУБЕДИНИЦА	ГЕН	ФУНКЦИЯ	kDa
$\alpha$	<i>rpoA</i>	необходим за сглобяване на ензима; взаимодействат и активират протеини	37
$\beta$	<i>rpoB</i>	участва в каталитичната реакция, неспецифично свързване на ДНК	150
$\beta'$	<i>rpoC</i>	участва в каталитичната реакция, неспецифично свързване на ДНК	155
$\omega$	<i>rpoZ</i>	помощна функция при сглобяване на ензима	10
$\sigma^{70}$	<i>rpoD</i>	специфично свързване към промотора	70

В бактериите има множество различни  $\sigma$ -фактори, които могат да се свържат с кор-ензима, за да помогнат на насочването на каталитичното ядро към промотора на транскрипционната единица и да се иницира транскрипция (Таблица 4.2). В *E. coli*,  $\sigma^{70}$  е основната  $\sigma$ -субединица, отговорна за транскрибирането на повечето гени в растящите клетки - тя поддържа функционирането на основни гени и метаболитни пътища. Чрез регулиране на изобилието на всеки активен  $\sigma$ -фактор, клетката може координирано да регулира групи от гени с общи функции.

Таблица 4.2. Алтернативни  $\sigma$ -фактори на прокариотната РНК полимераза.

$\sigma$ -ФАКТОР	ГЕН	ПРОМОТОР консенсусна секвенция	ФУНКЦИЯ	kDa
$\sigma^{70}$	<i>rpoD</i>	TTGACA-N17-TATAAT	транскрипция на повечето гени по време на експоненциалната фаза	70
$\sigma^{54}$	<i>rpoN</i>	CTGGCAC-N5-TTGCA	азот-регулирана транскрипция	54
$\sigma^{38}$	<i>rpoS</i>	TTGACA-N12-TGTGCTATAC	генна експресия при гладуване и стационарна фаза	38
$\sigma^{32}$	<i>rpoH</i>	CTTGAA-N14-CCCATNT	експресия на топлинно-шокови протеини	32
$\sigma^{28}$	<i>rpoF</i>	TAAA-N15-GCCGATAA	експресия на флагелатни протеини	28
$\sigma^{24}$	<i>rpoE</i>	TAAA-N15-GCCGATAA	експресия в отговор на стрес на клетъчната обвивка	24
$\sigma^{19}$	<i>fecI</i>	GGAAAT-N17-TC	транспорт на желязо	19

Ярък пример за специализация на  $\sigma$ -факторите за различни генни промотори се наблюдава при спорулацията на бактерията *Bacillus subtilis*. Тази бактерия може да съществува в два състояния: вегетативно (растящо) състояние и състояние на спорулация. Гените, необходими за образуването на спори, обикновено не се експресират по време на вегетативния растеж. Вместо това, един от първите гени, активирани при спорулацията, кодира нов  $\sigma$ -фактор, който иницира експресията на ранните гени, свързани със спорулацията. Впоследствие се произвеждат допълнителни  $\sigma$ -фактори, всеки от които активира нов набор от гени, необходими на различни етапи от процеса на спорулация.

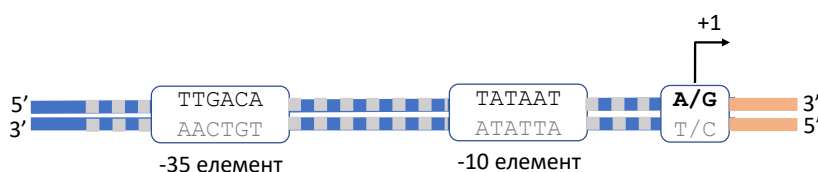
Всеки  $\sigma$ -фактор разпознава специфично промоторите на гените в своята група, без да се влияе от промоторите, които разпознават другите  $\sigma$ -фактори. Този пример илюстрира как транскрипцията може да се регулира както цис (на нивото на промотора), така и транс (чрез различни  $\sigma$ -фактори), позволявайки прецизен контрол върху генната експресия. По този начин, въпреки, че бактериите разчитат на един тип РНК полимеразата за транскрипция, използването на специализирани  $\sigma$ -фактори добавя допълнително ниво на контрол върху това кои гени се експресират в определени моменти и условия.

#### 4.2.2. Бактериални промотори

ДНК последователностите, участващи във функционирането на промотора, са били идентифицирани чрез сравняване на нуклеотидните последователности на серия от различни гени, изолирани от *E. coli*. Тези сравнения разкриват, че регионът нагоре от транскрипционното начало съдържа два елемента от последователности, които са подобни в различни гени. Тези последователности обхващат по шест нуклеотида и се наричат -10 елемент и -35 елемент, обозначаващи тяхната позиция спрямо транскрипционното начало +1 (Фигура 4.3). Последователностите на -10 елемента и -35 елемента в различни промотори не са идентични, но всички те са достатъчно сходни, за да установят консенсусни последователности - базите, които най-често се срещат на всяка позиция.

Консенсусът на -35 елемента е TTGACA ( $T_{82}T_{84}G_{78}A_{65}C_{54}A_{45}$ ), е центриран приблизително 35 нт нагоре от транскрипционното начало. Той се разпознава от  $\sigma$ -фактора, и при свързването му, РНК полимеразата се разполага върху промотора. Консенсусът на -10 елемента, наричан още Pribnow box по името на неговия откривател, е TATAAT ( $T_{80}A_{95}T_{45}A_{60}A_{50}T_{96}$ ), е центриран приблизително 10 нт нагоре от транскрипционното начало, и се разпознава от  $\sigma$ -фактора. -10 участък улеснява разплитането на ДНК матрицата.

Силните промотори са склонни да съответстват на консенсусната последователност, докато последователностите на по-слабите промотори се различават от консенсусната последователност и свързват  $\sigma$ -фактора и РНК полимеразата не толкова здраво, и съответните гени имат по-ниски нива на транскрипция.



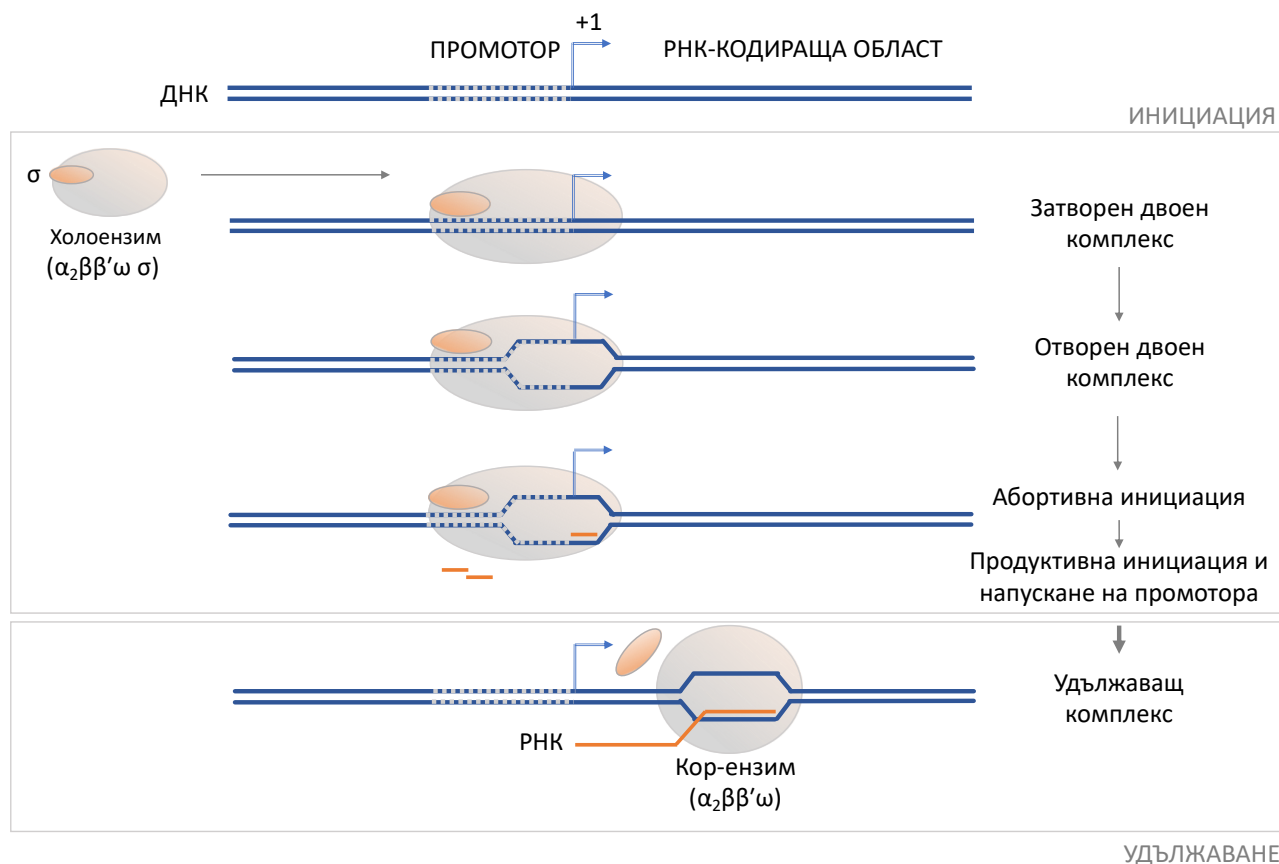
Фигура 4.3.  
Бактериален промотор

РНК полимеразата обикновено се свързва с промотора в област от приблизително 60 бд, простираща се от -40 до +20 (т.е. от 40 нт нагоре по веригата до 20 нт надолу по веригата от началото на транскрипция).  $\sigma$ -факторът се свързва специфично към -35 и -10 промоторните елементи, обосновавайки важността на тези последователности в промоторната функция.

#### 4.2.3. Инициация на транскрипцията - роля на $\sigma$ -фактора

Транскрипцията се иницира с разпознаването на промотора на транскрипционната единица от  $\sigma$ -фактора на РНК полимеразата (холоензима), при което се образува затворен двоен комплекс, включващ холоензима и промотора, в който двете вериги на ДНК са свързани (Фигура 4.4). Първоначалното специфично свързване на холоензима към промотора задейства конформационни промени, в резултат на които холоензимът огъва ДНК молекулата, а мобилните му региони я обвиват. След това РНК полимеразата развива приблизително 15 нт на ДНК в мястото на инициране за да образува отворен двоен комплекс, включващ холоензима и промотора, в който двете вериги на ДНК се разделени локално и едната верига става матрица за транскрипция – образува се т. нар. транскрипционно мехурче.

Транскрипцията се иницира от свързването на първите два рибонуклеотида чрез формиране на фосфодиестерна връзка.



Фигура 4.4. Инициация на транскрипцията и удължаване на транскрипта при прокариоти

От там нататък, нуклеотиди могат да се прибавят без движение на ензима, в резултат на което се образува РНК верига с дължина около 9 нт. При всяко прибавяне на нуклеотид, съществува вероятност холоензимът да се отдели от веригата – абортивна инициация, след което ензимът започва синтеза на РНК отново от първи нуклеотид. Абортивната инициация се случва неколкократно, при което се произвеждат серия от къси олигонуклеотиди. Инициацията завършва когато холоензимът успее да удължи веригата достатъчно – продуктивна инициация, и да напусне промотора. Абортивната инициация помага да се гарантира, че РНК полимеразата е стабилно прикрепена към ДНК матрицата и е правилно ориентирана за транскрипция, преди да се ангажира със синтеза на цялата РНК.

Дълго време се е смятало, че  $\sigma$ -факторът задължително напуска кор-ензима, след като е иницирал транскрипция, което му позволява да се свърже с друг кор-ензим и да започне транскрипция на друго място. С помощта на флуоресцентен резонансен енергиен трансфер, обаче е показано, че  $\sigma$ -факторът не напуска задължително кор-ензима. Вместо това,  $\sigma$ -факторът променя свързването си с кор-ензима през двата етапа на транскрипция – инициация и удължаване, преминавайки от състояние на здраво свързване по време на инициацията към състояние на хлабаво свързване по време на удължаване.

#### 4.2.4. Удължаване на транскрипта

След като РНК полимеразата преодолее абортивната инициация чрез успешно включване на 8–12 нуклеотида, тя влиза в режим на удължаване. В *E. coli* при преходът от режим на

инициране към режим на удължаване  $\sigma$ -факторът може или да се освободи от РНК полимеразата или да остане хлабаво свързан с нея. В резултат на това ензимът претърпява конформационни промени и се формира стабилен удължаващ комплекс, включващ РНК полимеразата, локално разплетената ДНК матрица и растящия транскрипт (Фигура 4.4). В него РНК полимеразата обхваща приблизително 35-40 нт от ДНК, в което областта на разплетената ДНК е около 12-14 нт, и съдържа РНК-ДНК хибрид с дължина около 8-9 бд.

РНК полимеразата напуска промотора, и движейки се по ДНК, извършва удължаване на РНК веригата - входящият рибонуклеозид трифосфат се прехвърля в активния център на ензима и се образува нова фосфодиестерна връзка. Ако е постъпил неправилен рибонуклеозид трифосфат, той се освобождава и процесът се повтаря, докато попадне правилният.

Когато правилният рибонуклеотид навлезе в активното място на РНК полимеразата, спонтанно се образуват водородни връзки между рибонуклеотида и комплементарната база на ДНК матрицата. Този процес се улеснява от позиционирането и ориентацията, осигурени от РНК полимеразата, което помага да се гарантира, че нуклеотидите са правилно подредени за комплементарно свързване. След възникване на водородна връзка, рибонуклеотидите се свързват ковалентно чрез фосфодиестерни връзки, които образуват захарно-фосфатния скелет на нарастващата РНК молекула. След образуването на фосфодиестерната връзка, РНК полимеразата се придвижва по матричната ДНК верига. След транслокация се добавя друг нов нуклеотид.

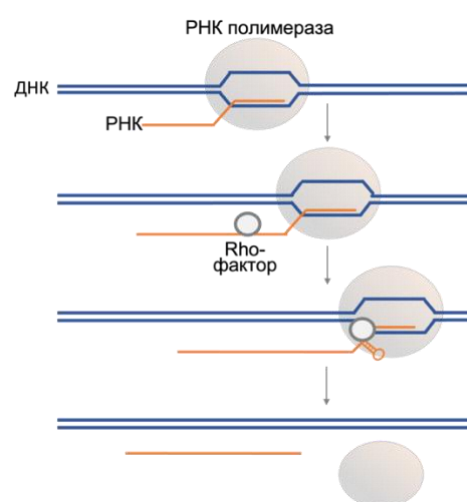
Водородните връзки, образувани в РНК-ДНК хибрида по време на транскрипцията, са преходни, което означава, че могат да се разрушат и да се реформират, докато РНК полимеразата се движи по ДНК матрицата. Тази динамична природа позволява непрекъснат синтез на РНК, като същевременно поддържа стабилно взаимодействие между РНК и ДНК веригите.

При своето движение, ензимът развива ДНК пред себе си и я преусуква зад себе си, поддържайки двете вериги на ДНК разделени в транскрипционното мехурче. При своето движение по матрицата, РНК полимеразата открива нови области за комплементарно свързване. Така по веригата на ДНК се движи един локално-разплетен хибриден участък, след който двете вериги се затварят.

#### 4.2.3. Прекратяване на транскрипцията

При прокариоти прекратяването става чрез един от двата добре характеризирани механизма, които изискват образуването на фуркетни структури във вторичната структура на РНК.

При  $\rho$ -зависимата терминация, РНК транскриптът съдържа участък за свързване за АТР-зависима хеликаза, наречена  $\rho$ -фактор (Фигура 4.5). След като  $\rho$ -факторът се свърже към транскрипта, той започва да се придвижва по него и да „преследва“ РНК полимеразата – скоростта на  $\rho$ -фактора е по-бавна от тази на ензима.  $\rho$ -терминиращото място близо до края на транскрипционната единица е причина за образуване на бримка в края на транскрипта, който кара РНК полимеразата да спре. Това позволява на  $\rho$ -фактора да настигне полимеразата и да развие



Фигура 4.5. Rho-зависимо прекратяване на транскрипцията при прокариоти

РНК:ДНК дуплекса, измествайки полимеразата от матрицата - транскрипцията спира и РНК се отделя.

Транскрипционните единици с rho-независима терминация завършват с характерен терминаращ участък – обърнат повтор, богат на G и C, който е последван от област, богата на A. Когато обърнатият повтор се транскрибира в РНК, в 3'-края на транскрипта се образува бримка с дължина 7-20 нд. Бримката забавя движението на РНК-полимеразата, водородните връзки в РНК:ДНК дуплекса (U:A) се разкъсват и РНК транскриптът се отделя от матрицата.

### 4.3. Еукариотна транскрипция

#### 4.3.1. Еукариотни РНК полимерази

В ядрото на еукариотната клетка са функционират три консервативни РНК полимерази – РНК полимеразата I, РНК полимеразата II, РНК полимеразата III. Всяка от тях осъществява транскрипцията на различен набор от гени, и може да бъде разграничена на базата на различна чувствителност към гъбния токсин алфа-аманитин, получен от някои отровни гъби от рода *Amanita*. Този токсин е цикличен пептид, който специфично се насочва към еукариотните РНК полимерази, особено РНК полимеразата II и в по-малка степен РНК полимеразата III. Неговият механизъм на действие и специфичност го правят критичен инструмент в молекулярната биология за изучаване на транскрипцията.

РНК полимеразата I транскрибира гените на рибозомната РНК (28S, 5.8S и 18S рРНК), с изключение на гена на 5S рРНК. 18S, 5.8S и 28S рРНК обикновено са кодирани в една транскрипционна единица, известна при растения като 35S рибозомна ДНК (рДНК), а при животни като 45S рДНК. Тази единица се транскрибира от РНК полимеразата I и произвежда предшественик на рРНК, която впоследствие се обработва в тези три вида рРНК. 5S рРНК се кодира от собствен ген, извън 35S/45S рДНК единицата, и се транскрибира от РНК полимеразата III.

Синтезът на рРНК се извършва в ядърцето, което е специализирана област в ядрото на еукариотните клетки. Ядърцето служи като място за транскрипция на рРНК и участва в обработката и сглобяването на рибозомни субединици.

РНК полимеразата II транскрибира всички протеин-кодиращи гени, както и множество гени за некодиращи РНК молекули като микроРНК и малки ядрени РНК (участващи в сплайсинга). РНК полимеразата II е изключително чувствителна към алфа-аманитина.

РНК полимеразата III транскрибира повечето от малките клетъчни РНК, включително тРНК и споменатата по-горе 5S РНК.

Ядрените РНК полимерази представляват големи мулти-субединични комплекси. Всяка еукариотна РНК полимеразата съдържа уникални субединици и споделени субединици със структурна и функционална роля. Уникалните субединици са от съществено значение за отделните функционалности на всяка полимеразата, позволявайки им да транскрибират различни видове РНК и да взаимодействат с различни транскрипционни фактори (ТФ). Уникалните субединици включват двете най-големи субединици, които са силно запазени през еволюцията и изграждат каталитичното ядро. Споделените между трите РНК полимерази субединици изпълняват структурна роля, която позволява правилното сглобяване и функциониране на целия комплекс на съответната РНК полимеразата, и функционална роля, тъй като участват в взаимодействия с различни регулаторни протеини и фактори на транскрипцията, което позволява адаптиране към различни клетъчни условия и нужди. Мулти-субединичният състав на трите еукариотни РНК полимерази, както и хомолжните субединици в бактериалната РНК полимеразата, са представени в Таблица 4.3.

РНК полимеразата II се състои от 12 субединици, обозначени като RPB1 до RPB12. Общото молекулно тегло на комплекса е приблизително 500 kDa до 550 kDa. Най-голямата субединица (RPB1) и втората по големина субединица (RPB2) играят критична каталитична роля. Субединици RPB3 и RPB4 образуват хетеродимер, който е важен за сглобяването на комплекса. Допълнителни субединици RPB5, RPB6, RPB7, RPB8, RPB9, RPB10, RPB11 и RPB12 допринасят за стабилността и регулирането на РНК полимеразата II.

Таблица 4.3. Еукариотни РНК полимерази.

Субединиците във всеки ред са хомоложни една на друга, представени са и хомоложните субединици в бактериалната РНК полимеразата. Споделените между трите РНК полимерази субединици са отбелязани в оцветени в оранжево редове.

ЕУКАРИОТИ			БАКТЕРИИ
РНК полимеразата I	РНК полимеразата II	РНК Полимераза III	РНК полимеразата
RPA190	RPB1	RPC160	$\beta'$
RPBA135	RPB2	RPC128	$\beta$
RPAC40	RPB3	RPAC40	$\alpha$
RPAC19	RPB11	RPAC19	$\alpha$
RPB6	RPB6	RPB6	$\omega$
RPB5	RPB5	RPB5	
RPB8	RPB8	RPB8	
RPB10	RPB10	RPB10	
RPB12	RPB12	RPB12	
RPA14	RPB9	RPC17	
RPA43	RPB4	RPC25	
RPA12	RPB7	RPC11	
RPA49		RPC53	
RPA34.5		RPC37	
		RPC82	
		RPC34	
		RPC31	

Най-голямата субединица има уникален С-терминален домен (CTD), съставен от хептапептидни повторения (YSPTSPS). Напр. броят на копията на секвенцията е 52 при мишка и човек, и 26 при дрожди. Секвенцията има много висока еволюционна консервативност, и както може да се очаква от това, е есенциална за точното функциониране на ензима, и от тук за клетъчната жизненост, макар, че броят на повторите може да бъде намален до определена степен без това да повлияе ензимната активност. CTD е от съществено значение за свързването на различни ТФ и фактори за обработка по време на транскрипция.

СТД служи като място за фосфорилиране, което е от съществено значение за функционирането на РНК полимеразата II. Дефосфорилираната форма на ензима е тази, която участва в инициацията на транскрипцията, а фосфорилираната форма е тази, която стартира удължаването на транскрипта. В допълнение на това, СТД е прицелно място за

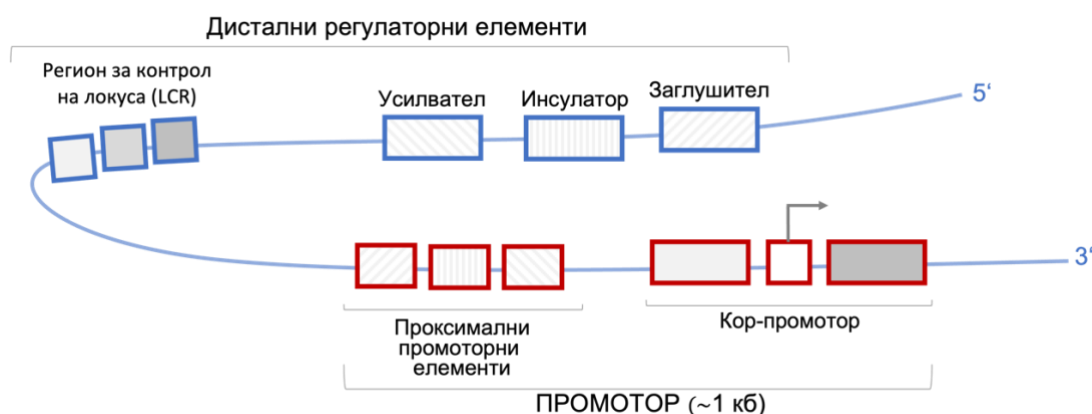
транскрипционни активатори и репресори, регулиращи транскрипционното удължаване. Нещо повече, нови изследвания разкриват, че фактори, участващи в пост-транскрипционни процеси като РНК сплайсинг, се свързват с този домен, така че растящият РНК транскрипт, продуциран от полимеразата, да бъде обработвани от фактори, като самите те са свързани с ензима.

РНК полимеразата I, II и III показват строга консервативност между организмите, което подчертава тяхното ключово значение за основните биологични процеси. Двете най-големи субединици на трите еукариотни РНК полимерази са хомоложни на  $\beta$  и  $\beta'$  субединиците на единствената РНК полимеразата на *E. coli*. Тази хомология не само илюстрира споделеното еволюционно наследство между прокариоти и еукариоти, но също така отразява критичната роля на тези ензими в транскрипцията.

Сухоземните растения притежават две допълнителни полимерази, РНК полимеразата IV и РНК полимеразата V. РНК полимеразата IV произвежда къси интерфериращи РНК с дължина 24 нт, които взаимодействат с транскрипти на РНК полимеразата V, за да контролират *de novo* метилирането на ДНК и заглушаването на транскрипцията на подвижните елементи (транспозони) в растителните геноми.

#### 4.3.2. Еукариотни промотори

Промотори на РНК полимеразата II. Гените, които се транскрибират от РНК полимеразата II обикновено съдържат два регулаторни региона: (а) промотор, който е съставен от кор-промотор и близки (проксимални) промоторни елементи, и (б) отдалечени (дистални) регулаторни елементи (Фигура 4.6). Кор-промоторът се свързва с главните ТФ, а другите регулаторни елементи се свързват с ТФ, които функционират или за усилване или за потискане на транскрипцията.



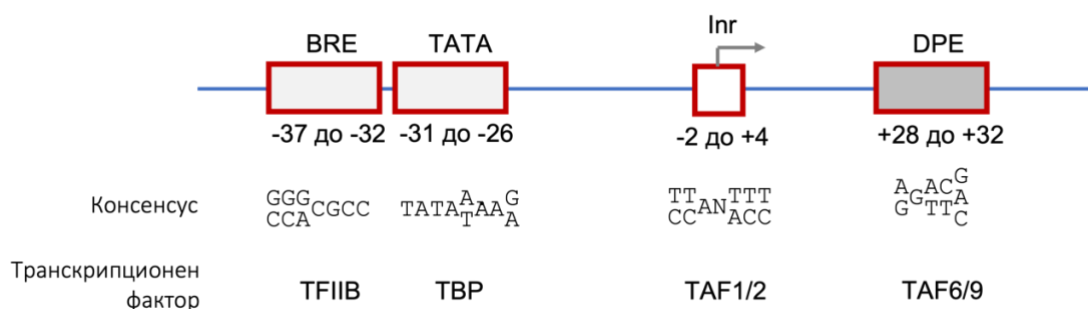
Фигура 4.6. Регулаторен регион на РНК полимеразата II – промотор и дистални регулаторни елементи

Кор-промоторът (от англ. core promoter), или основен промотор, е къса последователност, фланкираща началното място на транскрипция (обикновено  $\sim 50$  бд нагоре и  $\sim 50$  бд надолу спрямо началото). Той служи като свързваща платформа за машината за транскрипция, която включва РНК полимеразата II и нейните главни ТФ. Той е достатъчен за насочване на иницирането на транскрипцията, но обикновено има ниска базална активност, която може да бъде допълнително потисната от хроматина.

Картирането на транскрипционните начала в множество еукариотни геноми разкрива поразителни разлики между кор-промоторите. Те биват класифицирани в две групи - „фокусирани“ кор-промотори, които имат единично, добре дефинирано транскрипционно начало и „нефокусирани“ кор-промотори, които имат множество близко разположени транскрипционни начала, които се използват с подобна честота. Тези модели на инициране на транскрипция са свързани с отделни генни категории: фокусираното инициране се среща при тъканно-специфичните гени, докато нефокусирането инициране е свързано главно с конститутивните гени, свързани с основните клетъчни функции, кор-промоторът на които се припокрива с CpG острови.

Кор-промоторът на РНК полимераза II може да включва различни регулаторни елементи (Фигура 4.7):

- ТАТА (TATA-box) - разположен ~30 бд нагоре от транскрипционното начало при част от „фокусираните“ кор-промотори. ТАТА се разпознава и свързва от ТВР (от англ. TATA Binding Protein - протеин свързващ TATA-box), който е един от субединиците на TFIID. TFIID подпомага свързването на РНК полимераза II и сглобяването на ПИК, и по този начин може да определи избора на транскрипционното начало при фиксирана позиция надолу по веригата.
- Инициатор (Inr) - обхваща транскрипционното начало при друга част от „фокусираните“ кор-промотори. Човешкият Inr първоначално е дефиниран като пиримидин (С или Т), последван от пурин (А или G), позициониран така, че пуринът е първият транскрибиран нуклеотид.
- Промоторен елемент надолу по веригата (DPE) – среща се при промотори, които нямат ТАТА, и често придружава Inr като се разполага надолу от транскрипционното начало. Предполага се, че DPE се разпознава от други субединици на TFIID, наречени TAFs (от англ. TATA-box Associated Factors).
- Елементи за разпознаване от TFIIB (BREs) - регулаторен елемент, който се намира непосредствено до ТАТА-бокс и се състои от 7 нуклеотида. Срещат се два типа BRE: единият (BREu), разположен непосредствено нагоре от ТАТА, а другият (BREd) е разположен ~7 нуклеотида надолу по веригата.
- Други.



Фигура 4.7. Кор-промотор на РНК полимераза II

Изобразени са някои основни промоторни мотиви, които могат да участват в транскрипцията от РНК полимераза II, тяхната консенсусна секвенция и транскрипционния фактор, който се свързва с тях. Всеки специфичен кор-промотор може да съдържа някои, всички или нито един от тези мотиви. BRE се открива често нагоре по веригата в кор-промотори с TATA-box. В промотори без TATA-box, Inr често се съпровожда от DPE.

Интересна характеристика на ~60% от човешките гени е, че техният кор-промотор попада близо до CpG острови - относително къси участъци от ДНК, обикновено с дължина от 500 бд до 2 кб, които имат висока честота на CpG динуклеотиди в сравнение с по-голямата част на ДНК. Много от CpG динуклеотидите, които са разпръснати в генома са метилирани в петата въглеродна позиция на цитозиновата база, докато тези динуклеотиди в CpG островите обикновено са неметилирани. Те са свързани с повечето конститутивни гени, както и с много регулирани (тъканно-специфични) гени.

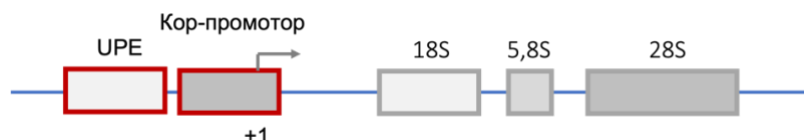
CpG островите обикновено са в неметилирано състояние в иначе силно метилирания геном, дори когато съответният ген е транскрипционно неактивен. Има, обаче, добре известни примери за CpG острови, които се метилират по време на нормалното развитие, което води до стабилно заглушаване на свързания промотор и съответния тъканно-специфичен ген. Смята се, че метилирането при CpG динуклеотидите потиска транскрипцията чрез блокиране на способността на ТФ да свързват своите разпознаващи последователности. В допълнение, специфични за метилирането свързващи протеини, като MeCP2, се свързват към метилираните CpG динуклеотиди и след това опосредстват присъединяването на хистон-модифициращи комплекси за да установят репресивна хроматинова структура.

Проксималните промоторни елементи се разполагат по-нагоре от кор-промотора (обикновено до 250 бд нагоре от транскрипционното начало). Те са свързващи места за промотор-специфични ТФ (по-често активатори и по-рядко репресори). Тези протеини обикновено съдържат два различни домена, единият отговорен за разпознаването на специфични ДНК последователности (ДНК-свързващ домен), другият изпълнява регулаторна функция (регулаторен домен). Една основна функция на регулаторния домейн е да набира кофактори, които носят дейности по ремоделиране на хроматин или могат директно да взаимодействат с транскрипционната машина на РНК полимера II.

Дисталните регулаторни елементи се разполага по-нагоре от проксималните регулаторни елементи, някои от които дори на разстояние 1Мб. Дисталните регулаторни елементи могат да включват усилватели (от англ. enhancer), заглушители (от англ. silencer), инсулатори (от англ. insulator), и региони за контрол на локуса (от англ. locus control region, LCR). Тези дистални елементи се разпознават от промотор-специфични ТФ и могат да се свържат с кор-промотора или проксималните елементи чрез механизъм, който включва огъване на ДНК. Образованата примка е резултат от взаимодействия между протеините, свързани с регулаторния елемент (напр. усилвателя), и тези, свързани с промотора.

Структурата на човешките генни промотори може да бъде доста сложна, обикновено състояща се от множество транскрипционни регулаторни елементи. Необходимостта от тази сложност става ясна, когато се вземе предвид, че човешкият геном съдържа около 20 000–25 000 гена, всеки от които може да има уникален пространствен/времеви модел на експресия. От тях, 1700–1900 гена кодират ТФ. Наличието на множество регулаторни елементи в рамките на промоторите осигурява комбинативен контрол на регулацията, което експоненциално увеличава потенциалния брой уникални модели на експресия.

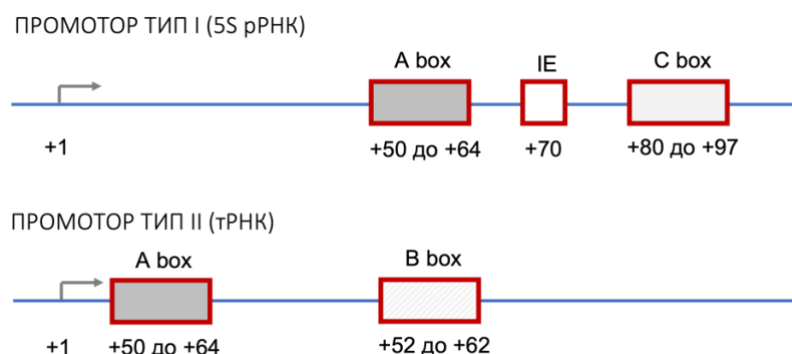
Промоторите на РНК полимера I съдържат два регулаторни региона (Фигура 4.8). Основният промотор е позициониран между -45 и +20 и се припокрива с транскрипционното начало (+1 позиция). Той е от съществено значение е за свързването на РНК полимера I. Елементът UPE (от англ. Upstream Promoter Element) е разположен между позиции -100 и -150 спрямо транскрипционното начало и има важна роля в свързването на ТФ, необходими за започване на транскрипцията.



Фигура 4.8. Транскрипционна единица на рРНК.

Промоторът включва кор-промотор и UPE и се разпознава от РНК полимеразата I. РНК-кодиращата област включва 18S, 5S и 28S рРНК.

Промоторите на РНК полимеразата III могат да бъдат категоризирани в три основни типа въз основа на тяхната структура (Фигура 4.9). Тип 1 промоторите се намират в гени, кодиращи 5S рРНК и някои тРНК и обикновено съдържат три регулаторни елемента: бокс А (A box), междинен елемент и бокс С (C box). Тип 2 промоторите контролират гените на по-голямата част от тРНК и имат два регулаторни елемента – бокс А и бокс В (B box). И при двата типа промотори регулаторните елементи на промоторите са разположени надолу от транскрипционното начало. Тип 3 промоторите са открити в гените, кодиращи други малки РНК (напр. U6 мяРНК).



Фигура 4.9. Промотори на РНК полимеразата III

#### 4.3.3. Инициация на транскрипцията – роля на главните транскрипционни фактори за сглобяване на пре-инициаторния комплекс

- Главни транскрипционни фактори

Бактериалната РНК полимеразата разпознава промотора с помощта на единствен фактор,  $\sigma$ -фактора. Въпреки, че еукариотните РНК полимеразы притежават ензимната активност, необходима за транскрипцията, те не могат да функционират самостоятелно, тъй като не могат да разпознават промотора. При еукариоти и археи няма  $\sigma$ -фактори, вместо тях РНК полимеразите използват множество главни ТФ, които взаимодействат както по-между си, така и с РНК полимеразата. Ролята на тези фактори е да разпознаят промотора и да инициират транскрипцията като организират формирането на стабилен транскрипционен апарат, който да включва РНК полимеразата и да е способен да осъществява повтарящи се цикли на транскрипция, в резултат на които да се продуцират много РНК копия на един и същи ген. Тъй като факторите са необходими за сглобяването на комплекса, но не и за самата транскрипция, повечето от тях се дисоциират, веднага щом РНК полимеразата започне транскрипция. Главните ТФ са специфични за всяка една от еукариотните РНК полимеразы. Така например, РНК полимеразата II използва шест главни ТФ – TFIIA, TFIIB, TFIID, TFIIIE, TFIIF и TFIIH.

- Сглобяване на пре-инициаторния комплекс (ПИК) на РНК полимеразата II

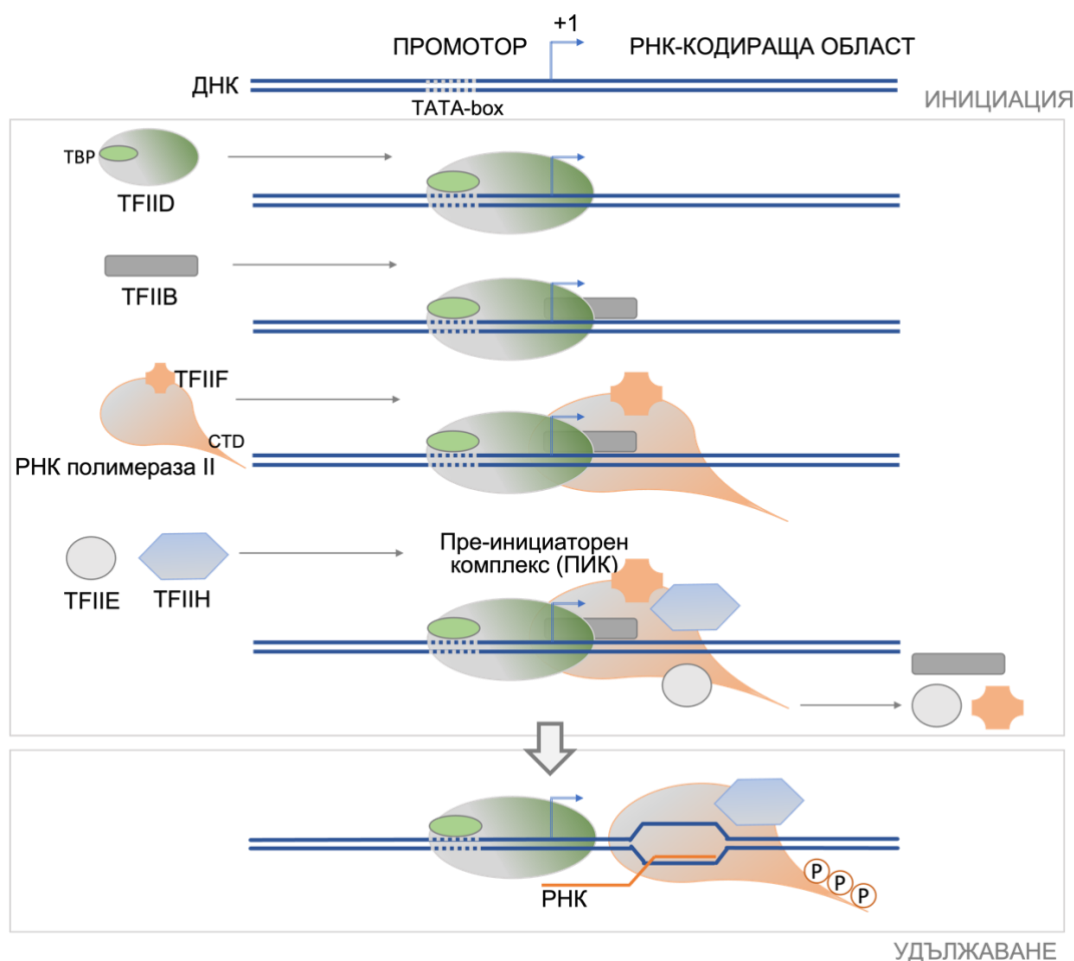
Инициацията на транскрипцията на РНК полимерата II включва поредица от събития (Фигура 4.10) на последователно свързване на главните ТФ с кор-промотора и ензима за да се образува пре-инициаторен комплекс (ПИК):

- TFIID е отговорен за първоначалното разпознаване на кор-промотора. Човешкият TFIID е съставен от TBP и 13 еволюционно запазени фактора, свързани с TBP (TAF1 до TAF13). При промотори с TATA-box, TBP разпознава TATA секвенцията и се свързва директно с нея, а при промотори без TATA-box, сглобяването на ПИК се инициира чрез асоцииране на TAF с други промоторни елементи (като Inr, DPE и др.). Заедно TBP и TAF субединиците на TFIID директно взаимодействат с промотора за да образуват, с помощта на TFIIA, платформата за сглобяване на ПИК. Свързването на TFIID води до огъване на ДНК в областта на промотора.
- Следва присъединяване на TFIIB - единичен полипептид, който контактува както с TBP, така и с ДНК нагоре от TATA-box, и подпомага в определянето на посоката на транскрипция.
- TFIIF се свързва със свободната РНК полимерата II, като по този начин я предпазва от неспецифични контакти с ДНК (извън промотора). Натоваарването на комплекса РНК полимерата II/TFIIF се подпомага TFIIB. TFIIF стабилизира РНК полимерата II, чрез контакти с TBP и TFIIB.
- Следва свързване на TFIIE, хетеротетрамер, чиято основна функция е да подпомогне натоваарването на TFIIN.
- Един от последните главни ТФ, които трябва да бъдат привлечени към ПИК, е TFIIN. Той съдържа осем субединици с различни ензимни активности. Хеликазната активност използва енергия от хидролиза на АТР за да развия кратък участък от промоторната ДНК в мястото на начало на транскрипция - преход на ПИК от затворено към отворено състояние. Това разделяне на ДНК веригите позволява на РНК полимерата II да разпознае матричната верига, да свърже комплементарните нуклеотиди и да синтезира първите няколко фосфодиестерни връзки. TFIIN също съдържа протеин киназна активност, която фосфорилира Ser-5 в хептапептидните повтори на CTD на най-голямата субединица на ензима. В инициращия комплекс, фосфорилирането на CTD го освобождава от взаимодействия с главните ТФ, което позволява на ензима да напусне промотора и да влезе в етапа на удължаване на транскрипцията. Някои от субединиците имат функции извън иницирането на транскрипция – няколко субединици на TFIIN са идентифицирани като компоненти на машината за поправка на ДНК. Установено е, че гените, кодиращи тези субединици, имат мутации в пациенти с ксеродерма пигментоза, човешко заболяване с дефекти в системата за поправка на ДНК.

ПИК осигурява само ниска (или базална) скорост на транскрипция на РНК полимерата II. ТФ, включващи активатори и репресори, заедно с всички свързани ко-активатори и ко-репресори, са отговорни за модулирането на скоростта на транскрипция.

Въпреки еволюционната дивергенция на множеството еукариотни РНК полимерази и специализацията на всяка полимеразата за уникален набор от промотори, основните механизми на транскрипция са запазени. Това запазване се отразява не само в подобни последователности на субединиците на самите полимерази, но и в присъствието на хомолози на TBP и TFIIB сред главните ТФ, които изграждат ПИК, специфични за и всеки клас полимеразата. Археите, които имат само една РНК полимеразата, съдържат както TBP, така и TFIIB, което предполага, че механизмите за инициране, използващи главни ТФ, са се развили преди дублирането на РНК полимеразите. Тъй като болшинството гени в еукариотните геноми

се транскрибират от РНК полимеразата II, по-надолу се описват в детайли особеностите на транскрипцията, извършвана от този полимеразата, а след това се посочват и особеностите при РНК полимеразата I и РНК полимеразата III.



Фигура 4.10. Инициация на транскрипцията при еукариоти  
Сглобяване на пре-инициаторния комплекс (ПИК)

- Abortивна инициация и освобождаване на промотора

След като ПИК е отворен, първият рибонуклеотид се въвежда в активното място на РНК полимеразата II за да инициира реакцията на полимеризация (в отсъствието на праймер). Това генерира зареждаща се РНК верига, която образува хетеро-дуплекс с матричната ДНК верига. Abortивната инициация се отнася до преждевременното освобождаване на кратки РНК транскрипти (обикновено с дължина 2-10 нт) и повторен синтез по време на ранните етапи на транскрипцията. Това се случва, когато РНК полимеразата иницира транскрипцията, но не успява да напусне региона на промотора, което води до синтез и освобождаване на кратки, непроизводителни РНК вериги, преди да може да се образува по-дълъг транскрипт.

Макар че abortивната инициация се възприема като пречка, тя изпълнява важни регулаторни функции. Тя действа като контролна точка, която осигурява правилното сглобяване и готовност на транскрипционната машина преди ангажиране с удължаването на транскрипта, а натрупването на abortивни транскрипти може да влияе на генната експресия.

Когато даден транскрипт достигне праговата дължина приблизително 10 нт, той навлиза в изходния канал на РНК полимеразата. Ензимът прекъсва взаимодействията си с промоторните елементи и главните ТФ на ПИК, от който вече не се нуждае. Освобождаването на промотора от ензима в еукариотната клетка изисква АТР хидролиза и, в случая на РНК полимеразата II - фосфорилиране на СТД. TFIIID може да остане свързан към кор-промотор, поддържайки повторното инициране на транскрипцията.

#### 4.3.4. Удължаване на транскрипта

След като напусне промотора и се освободи от повечето от главните ТФ, РНК полимеразата II се свързва с нови фактори, необходими за етапа на удължаване. Някои от факторите на удължаване могат да увеличат общата скорост на транскрибиране, други могат да помогнат на полимеразата да преодолее временните паузи, а трети могат да помогнат на транскрипцията през хроматина.

Удължаването на транскрипцията не е гладко движение по двойно-верижната ДНК, тъй като РНК полимеразата II, спира периодично (пауза) при определени последователности, понякога за дълги периоди от време, преди да възобнови транскрипцията. Обикновено тези паузи се извършват в началото на удължаването на транскрипта, когато РНК полимеразата II е все още близо до промотора (20-60 нт надолу от транскрипционното начало) - промотор-проксимална пауза. През тази пауза, РНК полимеразата спира своето движение, но остава стабилно свързана с ДНК. Паузата се иницира от свързването към полимеразата на негативния фактор на удължаване NELF (от англ. NEgative Elongation Factor), в сътрудничество с DSIF (от англ. DRB Sensitivity Inducing Factor). Освобождаването от действието на негативните фактори и преходът към продуктивно удължаване изисква участието на позитивния фактор за удължаване на транскрипцията, P-TEFb (от англ. Positive Transcription Elongation Factor b). P-TEFb е циклин-зависима киназа, която фосфорилира Ser-2 в хептапептидните повтори на СТД на РНК полимеразата II, както и на DSIF и NELF, с което освобождава блокираната полимеразата и съдейства за свързване към нея на белтъчни фактори, необходими за последващите процеси на сплайсинг и полиаденилиране на иРНК (описани в глава 5). Промотор-проксималната пауза е често използван механизъм за регулиране на гени, готови да бъдат експресирани бързо или по координиран начин.

Въпреки, че ензимните процеси на удължаване са по същество еднакви при прокариоти и еукариоти, еукариотната ДНК матрица е много по-сложна. Дори когато еукариотните клетки не се делят, техните гени са пакетирани с хистоновите белтъци в хроматин. За да се извършва полинуклеотидния синтез, РНК полимеразата трябва да отмести хистоновите белтъци всеки път, когато се натъкне на нуклеозома. Това става с помощта на специален димерен белтък, наречен FACT (от англ. FACilitates Chromatin Transcription – улеснява транскрипцията през хроматина). FACT частично разглобява нуклеозома надолу от РНК полимеразата като премахва два от осемте хистона - един H2A:H2B димер. Това е достатъчно да разхлаби увиването на ДНК веригата около нуклеозомното ядро, така че полимеразата да може да транскрибира разхлабения ДНК участък. След преминаването на РНК полимеразата, FACT сглобява нуклеозома като помага на липсващите хистони да се присъединят към нея.

- Точност на транскрипцията

Точността на транскрипцията е от съществено значение за продуктивната генна транскрипция. В РНК полимеразата II се намира важен домен, наречен Trigger Loop (от англ. спусък), който помага на ензима да разпознава и избира правилния нуклеозид трифосфат, съответстващ на нуклеотида от ДНК матрицата. При взаимодействие с правилния нуклеозид трифосфат, Trigger Loop преминава в конфигурация, която ускорява катализата и позволява продължаването на

синтеза. Ако обаче неправилен нуклеотид се приближи, Trigger Loop не образува стабилни връзки с него, което затруднява добавянето му и позволява на полимеразата да отхвърли грешния нуклеотид преди той да бъде включен в РНК веригата. Този механизъм на "активна селекция" играе решаваща роля за запазването на точността на генната транскрипция.

РНК полимеразата има присъща, макар и слаба, рибонуклеазна активност, която ѝ позволява да коригира грешки по време на транскрипция, повишавайки точността на транскрипцията. По този начин, подобно на ДНК полимеразите, РНК полимеразите постигат прецизност, следвайки две основни стратегии, избор на субстрат и корекция, което включва разпознаване и отстраняване на несъответстващ нуклеотид, въпреки че за разлика от ДНК полимеразите, рибонуклеазната активност на РНК полимеразите се намира в един и същ, „регулируем“ активен център, който извършва полимеризацията.

#### 4.3.5. Прекратяване на транскрипцията

При еукариоти, трите основни РНК полимерази използват различни механизми за прекратяване на транскрипцията.

РНК полимераза I: Терминирането на транскрипцията на рРНК гени от РНК полимераза I се извършва с помощта специфичен фактор за терминиране на транскрипцията, TTF-1 (от англ. Transcription Termination Factor for RNA Polymerase I). Използваният механизъм има известна прилика с  $\rho$ -зависимата терминация при прокариоти. Еукариотните клетки съдържат стотици рДНК повторения, понякога разпределени в множество хромозоми. Прекратяването на транскрипцията се случва в рибозомния интергенен регион, който съдържа няколко места за прекратяване на транскрипция, които се разпознават от TTF-1, свързването на който спира транскрипцията.

РНК полимераза II: При протеин-кодиращите гени, специфичен терминиращ сигнал не се открива. РНК полимераза II може да продължи транскрипцията след края на гена с няколко до хиляди нуклеотида. Въпреки това, транскриптът се разрязва в специфичен вътрешен сайт, още преди полимеразата да е приключила транскрипцията. Изрязаната 5'-част от транскрипта представлява първичния транскрипт (пре-иРНК), който търпи след това процесинг и се превръща в зряла иРНК, която се транслира. Остатъкът от транскрипта, се разгражда от 5'-екзонуклеаза (Xrn2 при човек), докато той все още се транскрибира от РНК полимераза II. Когато 5'-екзонуклеазата настигне полимеразата, тя подпомага освобождаването на полимеразата от ДНК матрицата.

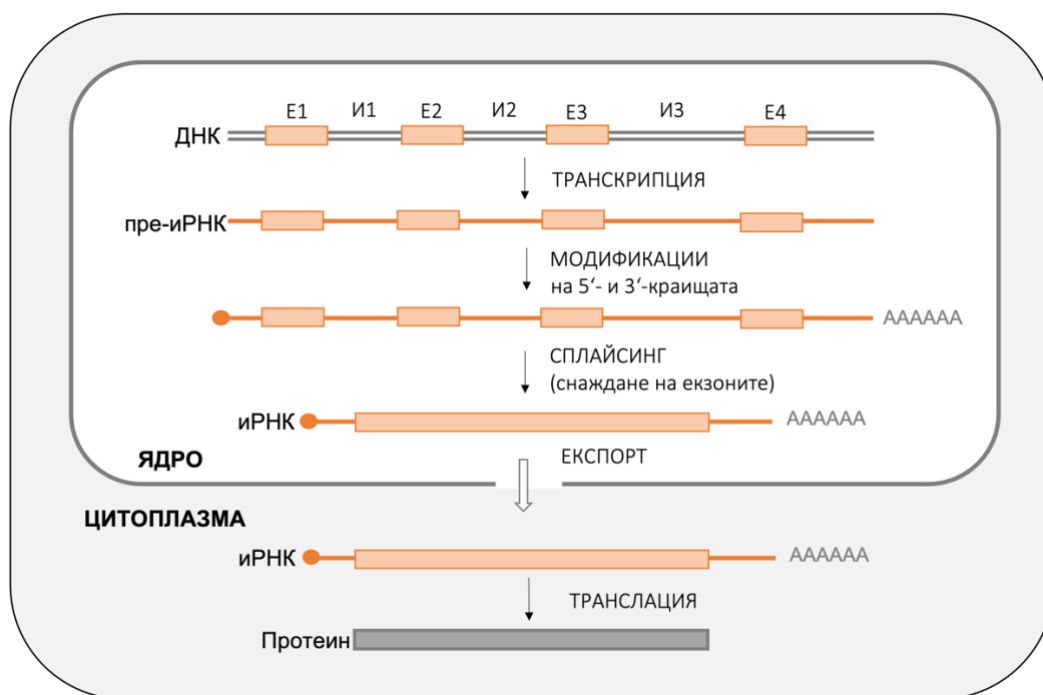
РНК полимераза III използва механизъм, който е подобен на бактериалната  $\rho$ -независима терминация, но тук дължината на уридиновия (U) участък е по-къса и няма изискване за вторична структура на РНК, която да измести РНК полимераза III.

## ГЛАВА 5 ПОСТ-ТРАНСКРИПЦИОННА ОБРАБОТКА на РНК

В резултат на транскрипцията на ДНК се образуват РНК молекули, които за да станат функционални, претърпяват пост-транскрипционна обработка и модифициране. При прокариоти, тези РНК модификации са главно фокусирани върху оптимизиране на функцията и стабилността на рРНК и тРНК. При еукариоти, пост-транскрипционните модификации засягат както рРНК и тРНК, така и иРНК. Обработката на еукариотната иРНК е обект на сложна регулация и има особено значение за осигуряване на правилната функция на белтъците в клетката.

При прокариоти, процесите на синтез на РНК и протеини се извършват в цитоплазмата. Протеин-кодиращите гени се транскрибират до иРНК, която се използва като матрица за синтез на протеини още преди да е напълно синтезирана и без да претърпява някаква обработка. Гените, кодиращи рРНК и тРНК, се транскрибират до първични транскрипти (пре-рРНК и пре-тРНК, съотв.), които претърпяват обработка за да се превърнат във функционални молекули.

При еукариоти, синтезът на РНК се случва в ядрото, а синтезът на протеини - в цитоплазмата. Двата процеса са разделени в пространството и времето. Еукариотните протеин-кодиращи гени са прекъснати – съдържат кодиращи и некодиращи области. Те се транскрибират в първичен транскрипт или предшественик на иРНК (пре-иРНК), който претърпява значителна обработка (процесинг) в ядрото за да се превърне в зряла иРНК (или само иРНК), която след това се транспортира в цитоплазмата за да послужи като матрица за синтез на протеини (Фигура 5.1). Подобно на прокариоти, гените, кодиращи рРНК и тРНК, се транскрибират до пре-рРНК и пре-тРНК, и претърпяват обработка за да се превърнат във функционални молекули.



Фигура 5.1. Пост-транскрипционна обработка (процесинг) на иРНК при еукариоти

По-надолу е описана обработката на еукариотната пре-иРНК, както и обработката на пре-рРНК и пре-тРНК, която се извършва както в прокариотните, така и в еукариотните клетки.

### 5.1. Обработка на иРНК при еукариоти

Процесингът на иРНК (Фигура 5.1) включва няколко процеса: модификации на 5'- и 3'-краищата на пре-иРНК, сплайсинг (изрязване на некодиращите области и снаждане на кодиращите области) и редактиране.

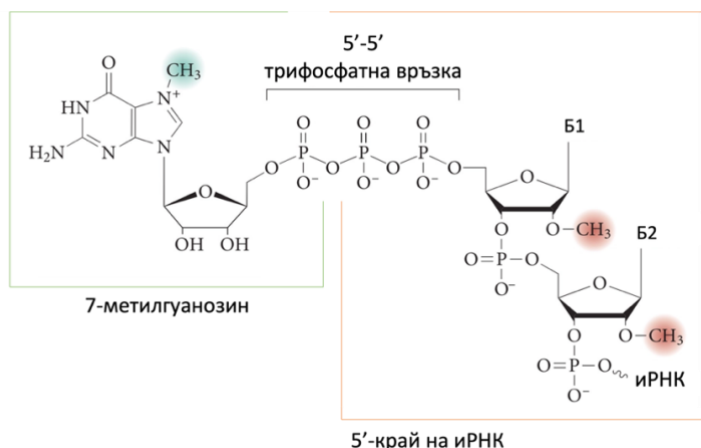
#### 5.1.1. Модификации на краищата на иРНК

Еукариотните иРНК имат по-дълъг полуживот от прокариотните иРНК поради защитни модификации в техните 5'- и 3'-краища.

- Модификации на 5'-края на иРНК

Първата модификация на първичния транскрипт се извършва докато все още се синтезира. Към неговия 5'-край се прибавя 7-метилгуанозин чрез 5'-5' трифосфатна връзка. Структурата се нарича „капаче“ (от англ. cap) и се означава с m<sup>7</sup>G (Фигура 5.2).

Образуването на „капачето“ изисква три ензимни активности. РНК 5'-трифосфатаза изрязва един фосфатен остатък от трифосфатната част при 5'-крайния нуклеотид. След това гуанилилтрансфераза катализира трансфера на гуанилилна група (GMP) от гуанозин трифосфата (GTP) към новообразувания дифосфатен 5'-край, в резултат на което се образува нетипична 5'-5' трифосфатна връзка. Метилтрансфераза прехвърля метилова група (-CH<sub>3</sub>) от S-аденозинметионин към позиция N-7 на гуанина. В многоклетъчните еукариоти и някои вируси се извършва допълнително метилиране на 2'-хидрокси групите на рибозата на първите два нуклеотида на 5'-края на иРНК. При cap-1 е метилирана 2'-хидрокси група на рибозата на първия нуклеотид, докато cap-2 има метилирани 2'-хидрокси групи на рибозите на първите два нуклеотида.



Фигура 5.2 Модификация на 5'-края на иРНК – добавяне на 7-метилгуанозин („капаче“)

Повечето клетъчни екзонуклеази нямат способност да хидролизират „капачето“, така че 5'-краят е защитен от разграждане. Освен това, „капачето“ участва в позиционирането на иРНК върху рибозомата за инициране на белтъчния синтез.

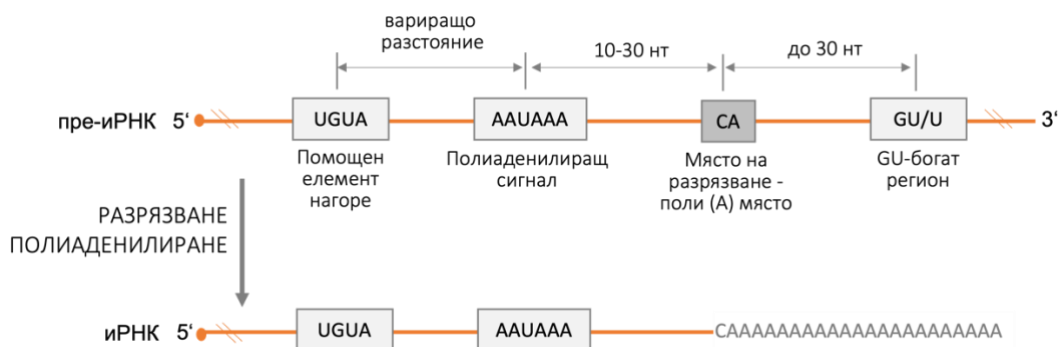
- Модификации на 3'-края на иРНК

Почти всички еукариотни иРНК, с изключение на хистоновите иРНК, имат на своя 3'-край последователност от аденозинови остатъци (A), чиито брой може да достигне 250 нт. Тази последователност, наречена поли(A) опашка, не се образува при транскрипцията на РНК

полимераза II, а се прибавя към 3'-края на иРНК от полимеризиращ ензим, който не се нуждае от матрица – Поли(А) полимераза.

При бозайници полиаденилирането се направлява от сигнални участъци, които се откриват в 3'-края на всяка една иРНК (Фигура 5.3):

- полиаденилиращ сигнал AAUAAA;
- място на разрязване/полиаденилиране или поли(А) място, което се намира от 10 до 30 нт надолу от полиаденилиращия сигнал;
- GU-богат регион, който се намира до 30 нт надолу от поли(А) мястото;
- помощен елемент нагоре (UGUA или UAUA), който се намира на различно разстояние нагоре от полиаденилиращия сигнал.



Фигура 5.3. Модификация на 3'-края на иРНК – добавяне на полиаденинова опашка

Образуването на поли(А) опашката включва две основни стъпки – разрязване и прибавяне на аденозинови остатъци и се осъществява от мултисубединичен белтъчен комплекс:

- Разрязването се извършва от CPSF (от англ. Cleavage and Polyadenylation Specificity Factor) - разпознава AAUAAA в първичният транскрипт, има ендонуклеазна активност, която разрязва пре-иРНК в поли(А) мястото;
- CstF (от англ. Cleavage stimulatory Factor) - свързва се с GU-богатия регион и подпомага разрязването;
- CFI и CFII (от англ. Cleavage Factor I, II) - разпознава помощния елемент нагоре и подпомага свързването на CPSF, дори ако AAUAAA липсва;
- След разрязването, започва полиаденилиране, катализирано от Поли(А) полимеразата. Тя изгражда поли(А) опашката чрез добавяне на аденозин монофосфатни (AMP) единици от аденозин трифосфат (АТР) към 3'-края на пре-иРНК, отцепвайки пирофосфат.

Пре-иРНК обикновено се разрязва преди прекратяването на транскрипцията. CstF се свързва както с нарастващия транскрипт, така и с РНК полимераза II. Той сигнализира на РНК полимераза II да се отдели от транскрипта.

Много транскрипционни единици имат две или повече поли(А) места, които могат да се използват алтернативно, напр. в зависимост от физиологичните условия. Алтернативното полиаденилиране се очертава като важно средство за регулиране на 3'-нетранслируемите региони на иРНК, което от своя страна предлага различни средства за регулиране на генната експресия, включително микроРНК-опосредстван контрол на трансляцията, РНК локализация и др.

Поли(А) опашката действа като място за свързване на поли(А)-свързващ протеин, който насърчава износа на иРНК от ядрото и трансляцията, и инхибира разграждането на иРНК като

по този начин увеличава нейния живот. Наличието на поли(А) опашка исторически се е използвало за изолиране на иРНК от еукариотни клетки чрез афинитетна хроматография.

В еукариотните соматични клетки, поли(А) опашките на повечето иРНК в цитоплазмата постепенно стават по-къси, а иРНК с по-къса поли(А) опашка се транслират по-слабо и се разграждат по-рано. Въпреки това, разграждането на иРНК може да отнеме много часове. Процесът на деаденилиране и разграждане може да бъде ускорен от микроРНК, които са комплементарни на 3'-некодирания регион на иРНК. В незрели ооцити, иРНК с къси поли(А) опашки не се разграждат; вместо това, те се съхраняват и са транслационно неактивни. Тези иРНК с къси опашки се активират чрез цитоплазмено полиаденилиране след оплождането по време на активацията на ооцитите.

### 5.1.2. Сплайсинг (снаждане на екзони)

Болшинството протеин-кодиращи гени при еукариоти са прекъснати - съставени са от кодиращи последователности, наречени екзони (ex-on означава, че те са експресирани) и междинни последователности, наречени интрони (Фигура 5.1). Изключение правят гените, кодиращи хистони и интерферони при гръбначните животни, които нямат интрони. Въпреки че не се превеждат, интроните изглежда имат различни функции, включително генна регулация и транспорт на иРНК.

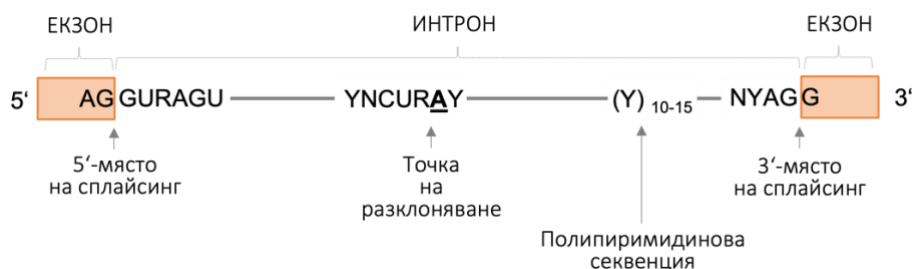
Транскрибираните РНК последователности, съответстващи на интрони, не кодират региони на функционалния полипептид и се отстраняват от пре-иРНК за се получи иРНК. Процесът на премахване на интроните от пре-иРНК и свързване на екзоните, се нарича РНК сплайсинг. Интроните се отстраняват от пре-иРНК, докато тя все още е в ядрото. Получената иРНК се транспортира в цитоплазмата и служи като матрица за протеиновия синтез.

От съществено значение е всички интрони да бъдат напълно и прецизно отстранени от пре-иРНК преди протеиновия синтез, така че екзоните да бъдат свързани заедно правилно, за да кодират функционален полипептид. Ако процесът се обърка дори с един нуклеотид, рамката на четене на последователността, създадена при свързването на екзоните ще бъде изместена и полученият полипептид ще бъде нефункционален.

- Пре-иРНК съдържа специфични секвенции, осигуряващи прецизен сплайсинг

Клетъчният апарат за сплайсинг недвусмислено идентифицира екзон/интронната граница, която определя 5'-края на интрона и интрон/екзонната граница, която определя 3'-края на интрона. Интроните се дефинират в пре-иРНК от 3 къси консенсусни секвенции. При бозайници, най-разпространени са интроните от типа GU-AG, съдържащи секвенциите (Фигура 5.4):

- 5'-мястото на сплайсинг – AG/GURAGU (екзон/интронна граница, наклонената чертичка / показва мястото на рязане, R означава пурин);
- 3'-мястото на сплайсинг – NYAG/G (интрон/екзонна граница, Y означава примидин, N може да бъде всеки един нуклеотид);
- секвенция на разклоняване YNCURAY с точката на разклоняване, която винаги е аденин (A); тя се разполага от 11 до 40 нт нагоре от 3'-мястото на сплайсинг;
- полипиримидинов участък (от 10 до 15 пиримидина, обикновено U), непосредствено нагоре от 3'-мястото на сплайсинг.



Фигура 5.4. Интрон - структурни елементи

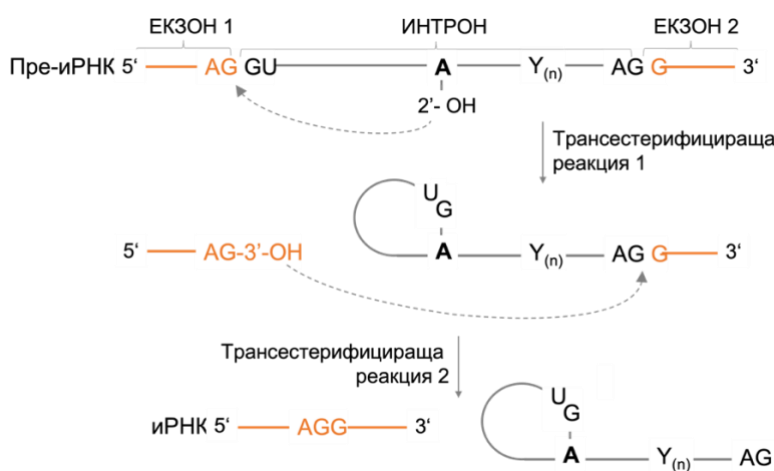
- Сплайсингът включва две координирани реакции на трансестерификация

Първа реакция на трансестерификация - 2'-ОН групата на аденина (A) в секвенцията на разклоняване извършва нуклеофилна атака върху първия нуклеотид на интрона в 5'-мястото на сплайсинг, който обикновено е гуанин (G) (Фигура 5.5). Това води до:

- разкъсване на връзката между екзон 1 и интрона - 3'-ОН край на екзон 1 е свободен,
- свързване на G на 5'-края на интрона с A в секвенцията на разклоняване - образува се структура подобна на ласо (интронът остава свързан с екзон 2).

Втора реакция на трансестерификация - 3'-ОН групата на екзон 1 иницира нуклеофилна атака върху първия нуклеотид на екзон 2 в 3'-мястото на сплайсинг. Това води до:

- свързване на екзон 1 и екзон 2,
- освобождаване на ласо-подобния интрон, който след това се разгражда.



Фигура 5.5. Сплайсингът включва две реакции на трансестерификация

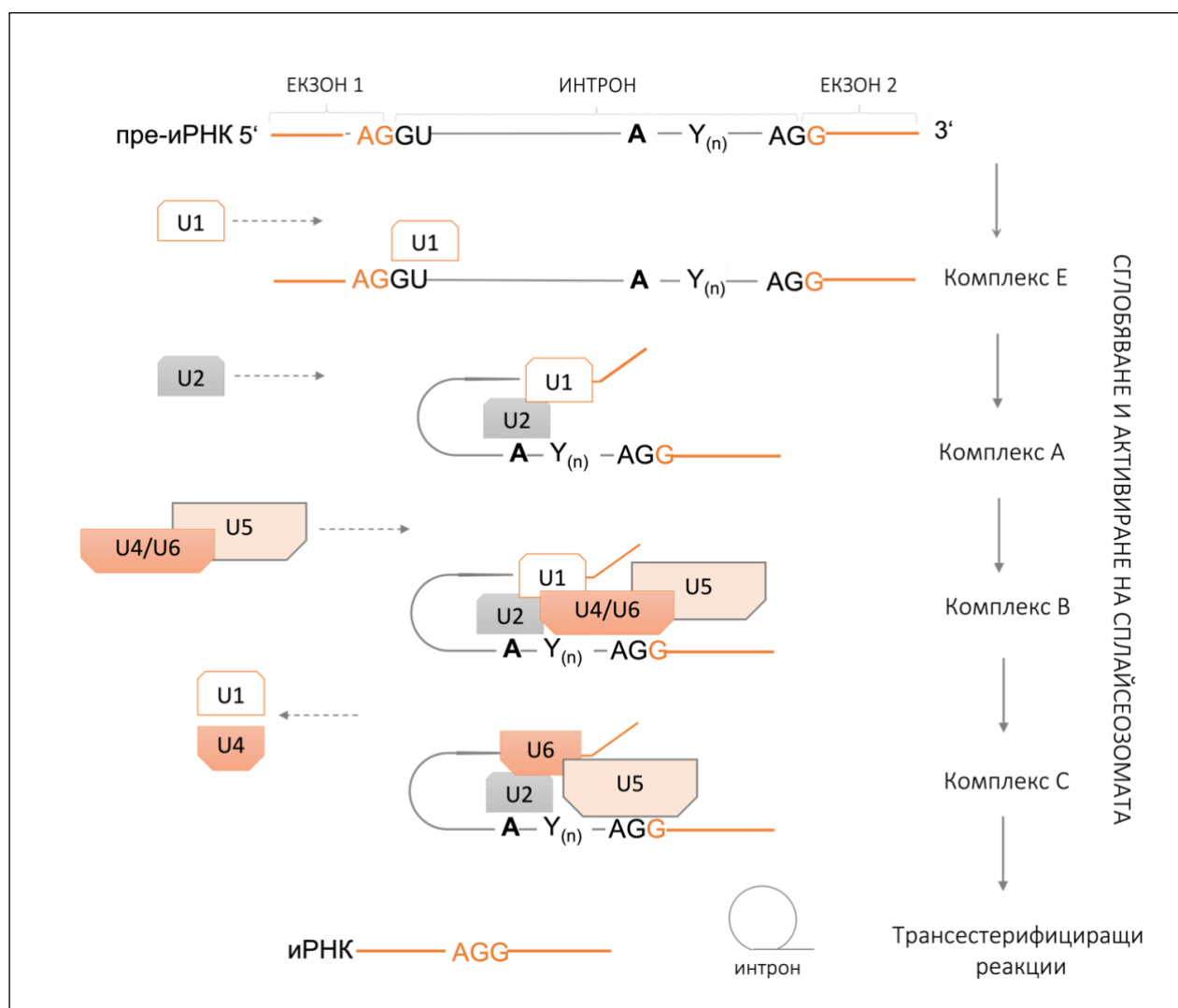
- Сплайсингът на пре-иРНК се извършва от сплайсозомата

Сплайсингът се извършва в сложен клетъчен апарат, наречен сплайсозома. По-голямата част от интрони на ядрените пре-иРНК се отстраняват от така наречената основна, или U2-зависима, сплайсозома. U12-зависимите интрони, които представляват по-малко от 0,2% от всички интрони на ядрените пре-иРНК, и досега са идентифицирани само в ограничен брой еукариоти, се изрязват от U12-зависима сплайсозома.

U2-зависимата сплайсозома, от тук нататък само сплайсозома, се състои от 5 малки ядрени рибонуклеопротеинови комплекси или малки ядрени РНП (мяРНП, small nuclear RNP, snRNP) и множество протеини. Всеки мяРНП е изграден от малка ядрена РНК (мяРНК, small nuclear RNA, snRNA) и няколко протеини. Функцията на петте мяРНП (U1 мяРНП, U2 мяРНП, U4 мяРНП, U5

мяРНП и U6 мяРНП) в сплайсозомата е да помогнат за разпознаването на специфичните нуклеотидни последователности в краищата на екзона и интрона и за позиционирането на реагиращите групи, така че да се извърши прецизно изрязване на интрона и свързване на съседните екзони. МяРНП разпознават специфичните сплайсингови последователности чрез комплементарно базово сдвояване между техните РНК компоненти и последователностите в пре-иРНК. Всички тези процеси са свързани с съществени реорганизации в сплайсозомата, които се улесняват от действието на АТР-зависими РНК хеликази.

Сглобяването на сплайсозомата включва последователно образуване на няколко комплекса (Фигура 5.6):



Фигура 5.6. Сглобяване на сплайсозомата.

Представен е примерен транскрипт, който включва два екзона и един интрон

- Комплекс Е - образува се чрез свързване на U1 мяРНП към 5'-мястото на сплайсинг. Паралелно, други протеини като сплайсинг фактор 1 (SF1) и U2 помощен фактор (U2AF) се свързват съответно с точката на разклоняване и полипиримидиновия участък, с което се постига определянето на границите на интрона.

- Комплекс А - след началните взаимодействия, U2 мяРНП измества SF1 и се свързва с последователността на точката на разклоняване и поставя основите за по-нататъшно сглобяване на сплайсозомните компоненти.
- Комплекс В - към комплекс А се присъединява тройния комплекс от мяРНП - U4/U6\*U5 за да се образува предкаталитичен комплекс; U4/U6 мяРНП е комплекс с две мяРНК, които са комплементарно свързани една с друга.

В сплайсозомата се извършва реорганизиране, което променя РНК-РНК и протеин-РНК взаимодействия и води до освобождаване на U1 мяРНП за да позволи на U6 мяРНК да се свърже с 5'-мястото на сплайсинг. U5 мяРНК контактува с краищата на двата екзона за да осигури прецизното им снаждане.

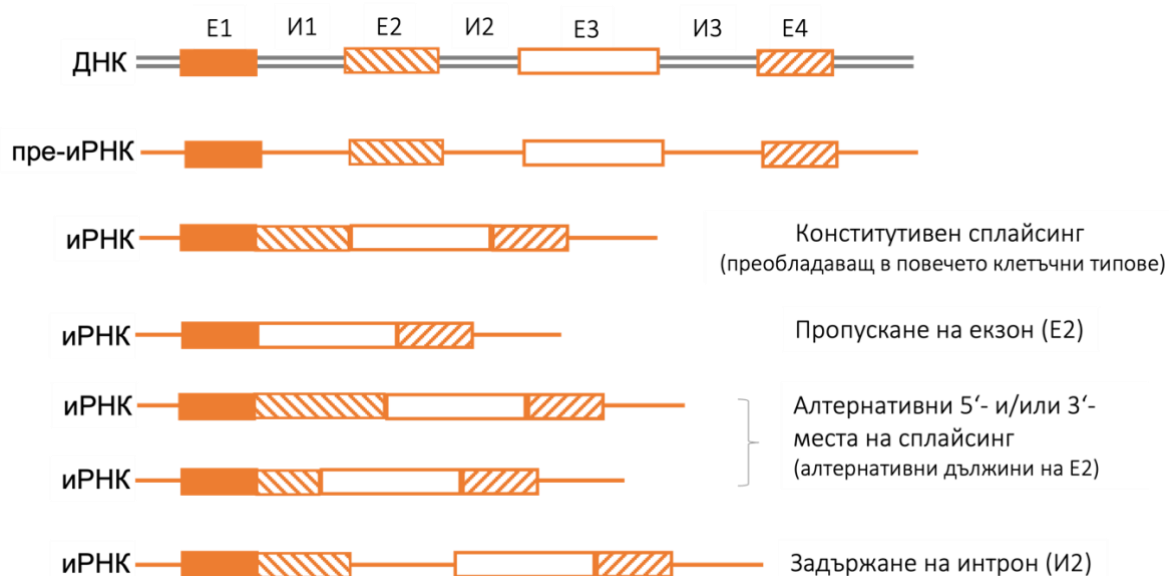
- Комплекс С - образуването на каталитично-активна сплайсозома се инициира от отделянето на U4 от U6 мяРНП – U6 мяРНК вече е свободна да влезе в контакт с U2. образувана от взаимодействието на U2-U6. Комплекс С е отговорен за изпълнението на двете трансестерифициращи реакции. В резултат се образува зрялата иРНК, която е готова да бъде експортирана в цитоплазмата за трансляция, а ласо-подобният интрон се разгражда.

- Алтернативният сплайсинг произвежда множество иРНК от една пре-иРНК

Повечето еукариотни гени се състоят от множество интрони и множество екзони. Конститутивният сплайсинг е процес на отстраняване на интрони и снаждане на екзони в реда, в който се появяват в гена. Алтернативният сплайсинг е отклонение от тази предпочитана последователност - екзони от един и същи ген се свързват в различни комбинации, което води до образуване на множество зрели иРНК с различни последователности (Фигура 5.7). Когато тези иРНК се транслират, се произвеждат множество протеинови изоформи – всички произхождащи от един и същи ген.

Най-често срещаните видове алтернативен сплайсинг са:

- Пропускане на екзони - отстраняване на определени екзони и техните съседни интрони от пре-иРНК;
- Алтернативен 5'- и/или 3'- сплайсинг - извършва чрез свързване на екзони в алтернативни 5'- и/или 3'- места на сплайсинг;
- Задържане на интрон - некодиращи части от гена се задържат в зрялата иРНК.



Фигура 5.7. Алтернативен сплайсинг

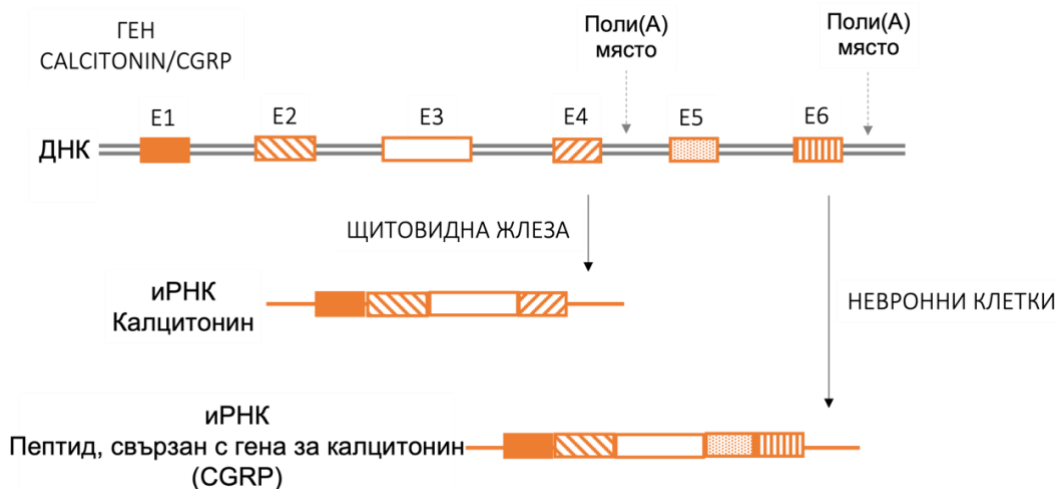
Въпреки че всеки тип алтернативен сплайсинг, описан по-горе, е различен един от друг, тези събития могат да се появят едновременно, след като се образуват пре-иРНК.

Алтернативният сплайсинг може също да добави или премахне региони от пре-иРНК, които могат да променят регулаторни елементи, засягащи трансляцията, стабилността на иРНК или субклетъчната локализация на получения протеин. Алтернативният сплайсинг може да доведе до вмъкване или заличаване на аминокиселини в протеиновата последователност, изместване на рамките на четене или дори въвеждане на нови стоп кодони.

Първият пример за алтернативен сплайсинг е описан през 1981 г. за гена CGRP (от англ. Calcitonin Gene-Related Peptide). При човек, той включва шест екзона и продуцира два различни протеина чрез алтернативна обработка на пре-иРНК (Фигура 5.8):

- Калцитонин - пептиден хормон, съставен от 32 аминокиселини, произведен от парафоликуларните С клетки (известни още като С клетки) на щитовидната жлеза. Калцитонинът функционира за понижаване на нивата на калций и фосфат в кръвта, като насърчава включването им в костната тъкан.
- CGRP - пептид с 37 аминокиселини, който действа като мощен вазодилататор и се произвежда от различни типове невронни клетки.

Генерирането на тези два протеина включва алтернативно използване на екзони и две различни места за полиаденилиране. Калцитонинът се синтезира чрез свързване на първите четири екзона и полиаденилиране при екзон 4. В контекста на CGRP, той се произвежда чрез изключване на екзон 4, свързвайки екзон 3 с екзон 5 и използвайки сигнала за полиаденилиране, разположен в екзон 6.



Фигура 5.8. Пост-транскрипционна обработка на гена calcitonin/CGRP.

Днес се знае, че повече от 95% от човешките гени претърпяват сплайсинг по начин, специфичен за развитието, тъканно-специфичен или зависим от сигналната трансдукция. Комбинацията от различни пост-транскрипционни регулаторни механизми е стратегия за осигуряване на пластичност на геномите на висшите еукариоти – един ген може да продуцира няколко генни продукта, всеки със специфична активност.

#### АЛТЕРНАТИВЕН СПЛАЙСИНГ И ИМУННА СИСТЕМА

Добър пример за размисъл тук е нашата имунна система. Имунната система на човека е еволюирала така, че за да разпознава ефективно патогени и да разграничава "себе си" и "не-себе си", тя трябва да бъде едновременно разнообразна и гъвкава.

Алтернативният сплайсинг е ключов за адаптивния имунен отговор, тъй като увеличава способността на организма да реагира на разнообразие от патогени. Той е основен механизъм в В-лимфоцитите за производството на различни класове антитела (имуноглобулини) с различни функции и специфичност.

Много автоимунни заболявания са свързани с аномални събития на сплайсинг. Алтернативният сплайсинг е от съществено значение за образуването на разнообразието на автоантигените и влияе върху производството на автоантигени. Например, специфични автоантигени, свързани с множествена склероза и диабет тип 1, се свързват със значителни събития на алтернативен сплайсинг.

Алтернативният сплайсинг не само че допринася за нормалната имунна функция, но също така играе значителна роля в патогенезата на автоимунните заболявания, като създава нови антигенни форми, които могат да предизвикат неподходящи имунни реакции.

#### 5.1.3. Редактиране на иРНК

Вариации в иРНК могат също да бъдат постигнати чрез пост-транслационни модификации, включващи промени в нуклеотидната последователност на иРНК, които имитират ефекта на мутациите - редактиране на иРНК. При по-висшите еукариоти основната форма на редактиране на РНК е заместващото редактиране на иРНК, при което базата се модифицира без промяна на дължината на пре-иРНК.

При гръбначните животни най-често срещаният тип заместващо редактиране на иРНК е деаминиране на Аденозин и Инозин (A-to-I). То се катализира от аденозин деаминази - ADAR, действащи върху РНК (от англ. - Adenosine Deaminases Acting on RNA,). Има три вида ADAR ензими, открити в гръбначните животни:

- ADAR1 катализира деаминирането на Аденозин до Инозин, основно в двуверижни РНК структури, образувани по време на обработката на иРНК на специфични целеви места. Това РНК редактиране е от съществено значение за различни биологични процеси, включително развитието на хематопоетични стволови клетки. Мутации в гена *ADAR1* при мишки са показали, че причиняват ранна ембрионална смърт, подчертавайки неговата съществена роля в развитието.
- ADAR2 участва в редактирането на иРНК за невроналния глутаматен рецептор, като конкретно модифицира аденозиновите остатъци до инозин. Тази модификация води до промени в аминокиселините, които са критични за функцията на рецептора. Недостигът на ADAR2 води до сериозни неврологични последици, включително гърчове и неонатална смърт при мишки.
- Въпреки че е по-малко разбран от ADAR1 и ADAR2, ADAR3 също е замесен в РНК редактиране, особено в мозъка. Неговите прецизни функции и приноси към неврологичното здраве са предмет на изследване.

Деаминирането на Цитидин до Уридин (C-to-U) е по-рядко срещан тип редактиране на иРНК при бозайниците, което се наблюдава главно в гена *ApoB*, който кодира човешкия

аполипопротеин В (ароВ), ключов компонент на липопротеините с ниска плътност (LDL). В чернодробните клетки транскрипцията на гена *АроВ* води до образуването на 14.1 kb иРНК, която кодира протеин с 4536 аминокиселини, известен като ароВ100. Въпреки това, в тънките черва, тази иРНК претърпява редактиране, за да произведе по-краткия протеин, наречен ароВ48, който се състои от 2152 аминокиселини. Последователността на ароВ48 е идентична на първите 2152 аминокиселини на ароВ100. Разликата в размера на протеина възниква, тъй като в тънките черва нуклеотид 6666 от иРНК се редактира чрез деаминиране на един цитидинов остатък, превръщайки го в уридинов остатък. Тази модификация променя кодона за глутамин в стоп кодон, което води до преждевременно прекратяване на транслацията и образуването на ароВ48.

Въпреки че редактирането на РНК е сравнително рядък феномен при гръбначните животни, дефектите в процеса могат да причинят няколко заболявания, свързани с централната нервна система, като амиотрофична латерална склероза, епилепсия, депресия и шизофрения.

## 5.2. Обработка на рРНК и тРНК

### 5.2.1. Обработка на рРНК

Гените на повечето прокариоти и еукариоти тРНК и рРНК (виж глава Транскрипция) се транскрибират до дълги първични транскрипти – предшественици, съотв. пре-тРНК и пре-рРНК, които претърпяват обработка (съзряване) както при еукариоти, така при прокариоти (Фигура 5.9).

При прокариоти, 30s пре-рРНК съдържа по едно копие на 16S, 23S и 5S рРНК, разделени от вътрешни транскрибирани региони (от англ. Internal Transcribed Spacer, ITS), в които могат да бъдат вмъкнати и гени за тРНК. Пре-рРНК трябва да бъде обработена, за да се освободят различните функционални РНК. Обработката на пре-рРНК изисква няколко РНКазы. Рибонуклеаза III (RNase III) разцепва пре-рРНК за да освободи 16S и 23S рРНК. Тези rRNA след това се обработват допълнително от други РНКазы, включително RNase P, която обработва tRNA прекурсори, и RNase D, която подрязва краищата на РНК молекулите.

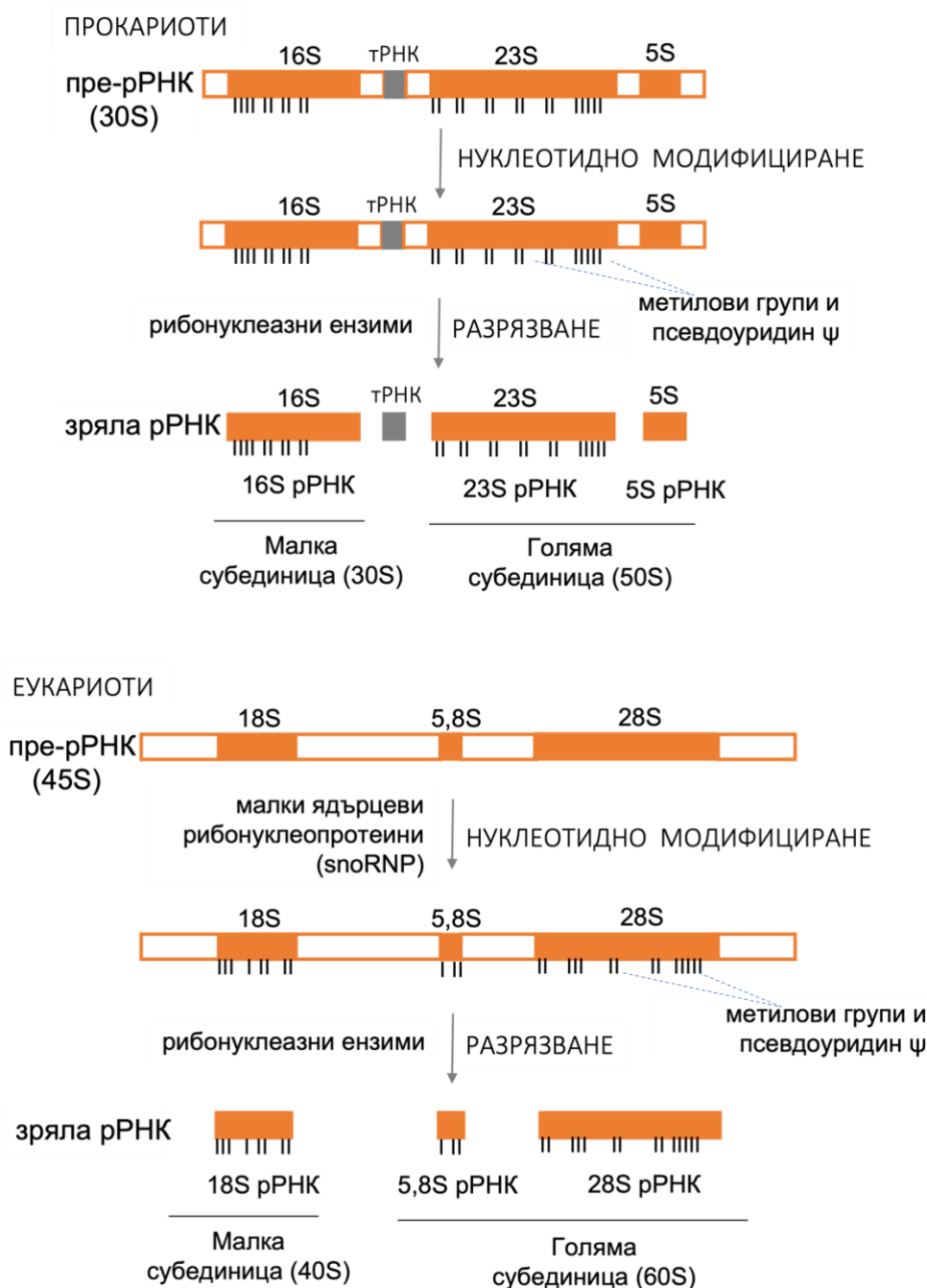
При еукариоти, пре-рРНК (45S пре-РНК) се транскрибира от рДНК от РНК полимеразата I. Този начален транскрипт обикновено е дълга полицистронна молекула, която съдържа последователности за 18S, 5.8S и 28S рРНК (при еукариотите 5S рРНК е отделно кодирана). Пре-рРНК съдържа и външни транскрибирани интервали (ETS) - не кодиращи региони в 5'- и 3'- краищата и вътрешни транскрибирани интервали (ITS), които разделят кодиращите последователности за различните рРНК.

Обработката на пре-рРНК включва серия от ендонуклеазни разрязвания, в резултат на които се премахнат ETS и ITS регионите и се освобождават зрели рРНК. При човек, пре-рРНК се разрязва на 11 различни места за да се генерират зрелите 18S, 5.8S и 28S рРНК. Обработката се извършва върху огромен рибонуклеопротеинов комплекс, наречен процесом.

По време на или непосредствено след транскрипцията пре-рРНК се модифицира от малки ядърцеви РНК (snoRNA), които насочват различни химически модификации като метилиране и псевдоуридиниране. Тези модификации са от съществено значение за правилното съгване и функция на зрелите рРНК. Зрялата човешка рРНК съдържа ~115 специфични модификации на метилиране (в повечето от тях метиловата група се добавя към рибозил-2'-хидроксилна група, давайки 2' -O-метилова модификация) и 95 специфични превръщания на уридин в псевдоуридин (U-to-ψ).

Тези модификации се въвеждат в пре-рРНК чрез взаимодействие с отделни малки ядърцеви РНК протеинови комплекси или малки ядърцеви РНП (small nucleolar RNP, snoRNP). Всеки

малък ядръцев РНП е специфичен за един или най-много няколко отделни сайта за модификация. Всеки от тях съдържа уникална водеща РНК - малка ядръцева РНК (small nucleolar RNA, snoRNA) с дължина ~60 до 300 нт, както и от една до четири протеинови молекули. Малките ядръцеви РНК съдържат силно запазени структурни мотиви, които принадлежат към две групи - C/D box или H/ACA box. Тези последователности насочват свързването на малките ядръцеви РНК към местата в пре-рРНК молекулите - малките ядръцеви РНК с C/D box подпомагат правилното позициониране на метил трансферазата, а тези с H/ACA box на псевдоуридин синтазата, по продължение на пре-рРНК за да извършат съответните модификации. Процесът е сложна структура, съдържаща повече от 100 малките ядръцеви РНК и повече от 100 отделни протеина.

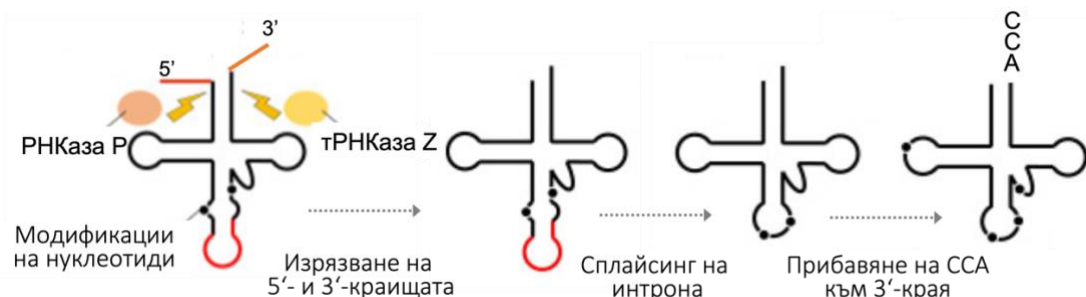


Фигура 5.9. Пост-транскрипционна обработка на рРНК

### 5.2.2. Обработка на тРНК

Пост-транскрипционната обработка на пре-тРНК до зряла тРНК включва няколко реакции (Фигура 4.10):

- изрязване на последователности на 5'- и 3'- краищата с помощта на рибонуклеази - съответно RNase P и tRNase Z;
- в случаите, когато пре-тРНК съдържа и некодираща област - интрон, тя се изрязва от специфична ендонуклеаза, а след това двата фрагмента се свързват с помощта на РНК лигаза;
- във всички еукариотни пре-тРНК, както и в някои бактериални и археални пре-тРНК, се добавя последователността ССА към 3'-края на пре-тРНК, след като оригиналният 3'-край е отрязан; добавянето на ССА става с помощта на ензим – тРНК нуклеотидилтрансфераза;
- някои бактериални пре-тРНК имат тринуклеотида ССА, кодиран в техния транскрипт непосредствено преди 3' мястото на разрязване, така че не е необходимо да добавят такъв; ССА в 3'-края на зрялата тРНК е мястото, където се присъединява специфичната за тРНК аминокиселина;
- модифициране на множество нуклеотиди в пре-тРНК, при което се променят техните азотни бази - средно около 12 нуклеотида са модифицирани на тРНК; най-честите модификации включват превръщане на аденин (А) в псевдоуридин ( $\psi$ ), на аденин (А) в инозин (I) и на уридин в дихидроуридин (D); друга важна модификация е добавяне на метилови групи към хетероцикличния пръстен на пуриновите бази, както и към 2'-ОН групата, присъстваща в рибозата.



Фигура 5.10. Пост-транскрипционна обработка на тРНК

Голямо разнообразие от модификации на тРНК се открива в антикодона – такива модификации са от решаващо значение за прецизното разпознаване на кодона и поддържането на рамката за четене, като по този начин се гарантира точен и ефективен протеинов синтез. В допълнение, другите части на тРНК също често се модифицират и по този начин стабилизират тРНК в клетката. Нови проучвания разкриват, че модификациите на тРНК могат да бъдат динамично променени в отговор на нивата на клетъчните метаболити и на стреса от околната среда. Става ясно, че нарушенията в модификацията на тРНК могат да имат патологични последици - митохондриални заболявания, неврологични разстройства и рак.

### РИБОЗИМИТЕ и ХИПОТЕЗАТА за РНК СВЕТА

Традиционното схващане, че ДНК е била първата нуклеинова киселина на Земята, бе поставено под въпрос от убедителни доказателства, предполагащи първичната роля на РНК. Учените сега предполагат, че РНК молекулите са служили като най-ранните каталитични биополимери на Земята, функциониращи едновременно като генетичен материал и биохимични катализатори преди появата на ДНК и протеините.

Няколко ключови открития подкрепят тази хипотеза за "РНК света". Първо, изследователите идентифицираха самосплайсващи РНК молекули, способни да отстраняват интрони от пре-иРНК чрез своята присъща ензимна активност, демонстрирайки каталитичния потенциал на РНК. Второ, подробните структурни и биохимични анализи на рибозомите - клетъчните машини, отговорни за синтеза на протеини - разкриха, че каталитичният център се състои от РНК, а не от протеин. Това откритие предполага, че самият протеинов синтез се е развил в свят, доминиран от РНК.

Тези каталитични РНК молекули, наречени рибозими, са могли да изпълняват основни биологични функции, включително самовъзпроизвеждане. С еволюцията на живота, ДНК се появява като по-стабилно хранилище за генетична информация, докато протеините се развиват като по-ефективни катализатори. Този преход в крайна сметка води до съвременната биологична система, където ДНК съхранява генетична информация, РНК служи като междинен посредник, а протеините изпълняват повечето каталитични функции.

### 5.3. Продължителност на живот на РНК

При прокариоти, иРНК обикновено има много кратка продължителност на живот, често само няколко минути. Това бързо обновяване се случва защото, за разлика от еукариотите, иРНК няма защитни модификации като 5'-шапчица и 3'-поли-А опашка, и биват бързо атакувани от РНКазы. Средният период на полуразпад е между 2-10 минути, въпреки че някои транскрипти могат да издържат до час – например иРНК, кодиращи протеини на външната мембрана. Това бързо обновяване позволява на бактериите да се адаптират бързо към промените в околната среда чрез бързо променяне на генната им експресия.

При еукариоти, РНК обикновено има много по-дълга продължителност на живот тъй като 5'-шапчицата и поли-А опашката защитават иРНК от разграждане. По-дългата продължителност на живот е необходима за значителните пост-транскрипционни промени, които претърпява иРНК, преди да напусне ядрото, се извършва обширна обработка. Специфични последователности в иРНК, както и РНК-свързващи протеини, могат допълнително да модулират стабилността. Средният период на полуразпад е от няколко часа до дни. Например бета-глобиновата иРНК в червените кръвни клетки има забележителен полуживот от около 16-24 часа. След като еритроцитите загубят своето ядро по време на узряването, тези стабилни иРНК продължават да произвеждат хемоглобинови протеини. Специални елементи за стабилност в 3' UTR региона помагат за защитата на тези транскрипти от разграждане.

Тази разлика в продължителността на живота между прокариоти и еукариоти отразява различните биологични нужди на тези организми. Прокариотите се нуждаят от бърз отговор на промените в околната среда, докато еукариотите изискват по-сложна регулация и устойчива генна експресия. Компартментализацията в еукариотната клетка позволява допълнителни

стъпки за контрол на качеството на иРНК, а енергийната инвестиция в обработката на еукариотна РНК прави по-дългата продължителност на живот по-ефективна

Макар, че прокариотите и еукариотите използват сложни механизми за разграждане на РНК, специфичните ензими, пътища и включени регулаторни фактори се различават значително между двете. Прокариотите разчитат в голяма степен на деградозомата и структурните елементи в иРНК, докато еукариотите използват различни екзонуклеази, ендонуклеази и малки некодирани РНК като микроРНК, за да регулират стабилността и разграждането на иРНК.

## ГЛАВА 6 СИНТЕЗ на ПРОТЕИНИ - ТРАНСЛАЦИЯ

Протеиновият синтез (или транслация) е процес, при който информацията, записана в иРНК (респ. ДНК), се използва за производството на протеинови молекули. По време на транслацията нуклеотидната последователност на иРНК се прочита с помощта на генетичния код. Транслацията изисква съгласуваното действие на иРНК с други РНК молекули - тРНК и рРНК, както и на много цитоплазмени протеини. Този процес може да се разглежда като кулминационната точка на преноса на генетична информация, тъй като крайният му резултат са протеиновите молекули - основните структурни и функционални компоненти на живите клетки.

### 6.1. Въведение в транслацията

#### 6.1.1. Генетичен код

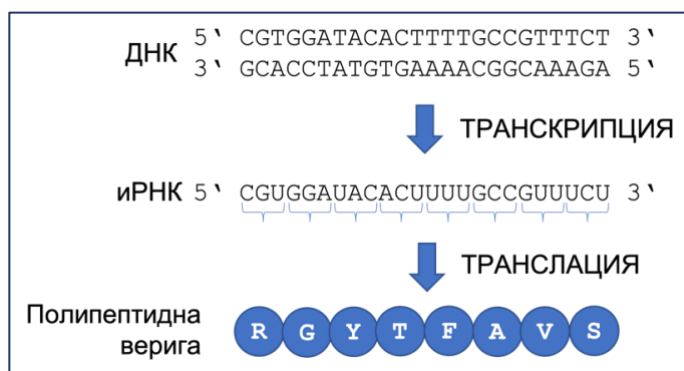
Генетичният код е набор от правила, които живите клетки използват, за да преведат информацията, кодирана в генетичния материал в протеини. Състои се от последователности от нуклеотиди, групирани в триплети, известни като кодони, които определят кои аминокиселини ще бъдат включени в протеина по време на синтеза.

- Колинеарност на нуклеотидната последователност на гена и аминокиселинната последователност на кодирувания протеин

Колинеарността се отнася до директното съответствие между нуклеотидните последователности в гените и аминокиселинните последователности в белтъците (Фигура 6.1). Тази концепция предполага, че дължината на ДНК последователността в гена е пропорционална на дължината на полипептида, който кодира. По време на транслацията, иРНК, синтезирана при транскрипцията на ДНК, служи като матрица, чиято нуклеотидна последователност се превежда от рибозомите в специфична последователност от аминокиселини чрез използването на генетичния код.

Докато колинеарността е често срещан феномен при прокариоти и много вируси, тя е по-малко ясна при еукариотни организми. При еукариоти, гените често съдържат интрони, които се изрязват по време на обработката на иРНК, с което се нарушава простото съответствие между дължината на гена и дължината на белтъка. Въпреки това основният принцип на колинеарността остава важен за разбирането на начина, по който генетичната информация се превежда в функционални белтъци.

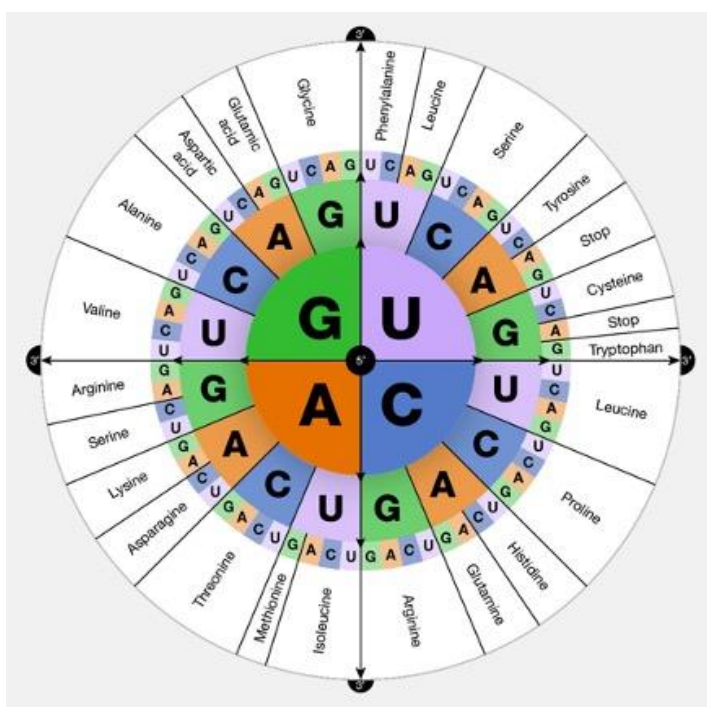
Фигура 6.1. Колинеарност на гена и полипептидната верига



- Генетичният код е триплетен

Концепцията за триплетите, или кодоните, е описана за първи път от Франсис Крик и колегите му през 1961 г.

Генетичният триплетен код използва 4 символа – А (аденин), G (гуанин), С (цитозин) и U (урацил) и може да създаде 64 различни комбинации ( $4^3=64$ ), в случая кодони. От тях, 61 кодона кодират една от 20-те аминокиселини (Фигура 6.2). Три от тях (UAA, UAG, UGA) са „стоп“ кодони, не кодират аминокиселина и се използват за сигнализиране на прекратяването на протеиновия синтез. Кодонът AUG кодира метионин навсякъде, където се появява в иРНК, но също така има уникална функция за определяне на началната точка за протеинов синтез в повечето иРНК, и затова се нарича „старт“ кодон.



Фигура 6.2. Генетичен код  
(National Human Genome Research  
Institute, [www.genome.gov](http://www.genome.gov))

- Генетичният код е изроден

Генетичният код е изроден (или дегенеративен), тъй като болшинството аминокиселини се кодират от повече от един кодон. Например, аминокиселината валин се кодира от 4 кодона - GUU, GUC, GUA и GUG. Всъщност всички аминокиселини, с изключение на метионин (AUG) и триптофан (UGG), имат повече от един кодон.

- Генетичният код е универсален, с малки изключения

Генетичният код, първоначално считан за универсален във всички форми на живот, се различава леко в определени организми и в митохондриите на някои еукариоти. При бактерии, например, както GUG, така и UUG, могат да бъдат разчетени като метионинов кодон, ако се появят близо до 5'-края на иРНК. Съществуват и малки разлики в генетичния код на митохондриалната ДНК; например в човешките митохондрии AGG и AGA функционират като стоп сигнали (кодони за аргинин в универсалния генетичен код), UGA съответства на триптофан (универсален терминиращ кодон), а AUA означава метионин (универсален кодон за изолевцин). Независимо от това, тези разлики са редки и генетичният код е идентичен при почти всички видове, като едни и същи кодони определят едни и същи аминокиселини.

- Генетичен код и рамка на четене

Генетичният код се чете в нуклеотидната последователност на ДНК, или РНК, чрез неприпокриващи се кодони – рамка на четене (Фигура 6.3). По време на транскрипцията РНК полимеразата прочита матричната ДНК верига в посока 3'→5' и синтезира иРНК в посока 5'→3'. иРНК е едноверижна и следователно съдържа три възможни рамки за четене, които могат да бъдат прочетени в тази посока (5'→3'), като всяка започва от различен нуклеотид в първия триплет. От трите възможни рамки на четене, обикновено само една има потенциал да бъде преведена в протеин – тази рамка се нарича „отворена рамка на четене“ (от англ. Open Reading Frame, ORF), тя започва със старт кодон и завършва със стоп кодон.

Фигура 6.3. Рамка на четене на иРНК.

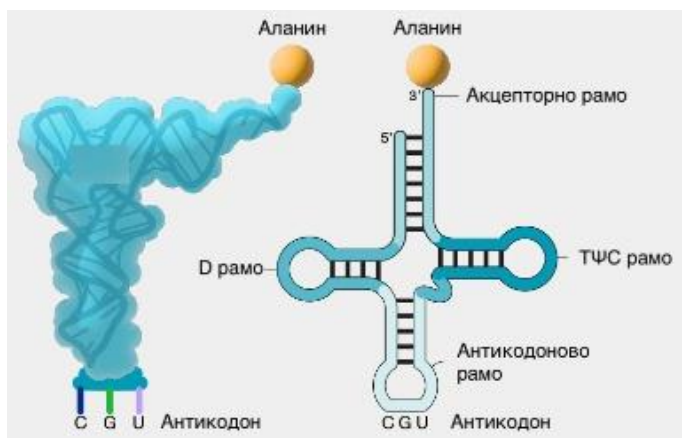
Най-вероятната отворена рамка на четене на представената иРНК е първата рамка на четене, която започва със старт кодон и завършва със стоп кодон.



### 6.1.2. тРНК – преводачи на генетичния код

- Структура на тРНК

Генетичният кодът на иРНК се превръща в аминокиселинна последователност с помощта на тРНК. Тези РНК са сравнително малки молекули, чиято дължина варира от 75 до 93 нт. Някои от нуклеотидите в тРНК се свързват един с друг чрез водородни връзки, така че тРНК се нагъва във форма, която може да бъде представена двуизмерно като детелина. Тази форма допълнително се нагъва за да образува L-образна молекула в триизмерното пространство (Фигура 6.4). В молекулата на тРНК се откриват следните структурни елементи:



Фигура 6.4. Транспортна РНК (тРНК). Вторична структура (вдясно) с форма на детелина; L-образна третична структура (вляво). (National Human Genome Research Institute, www.genome.gov)

- Акцепторно рамо: стъбло с дължина 7-9 бд, получава се при свързване на нуклеотидите на 5'-края с тези на 3'-края на тРНК. 3'-краят завършва 3'-терминалната група ССА, към която се свързва аминокиселината, за да се образува аминоксил-тРНК.

- Антикодоново рамо: стъбло с дължина 5 бд, чиято бримка съдържа триплетен антикодон, който се свързва със съответния кодон в иРНК.
- D-рамо: стъбло с дължина 4-6 бд, завършващо с бримка, която често съдържа дихидроуридин. D-рамото играе важна роля в стабилизирането на третичната структура на тРНК и влияе върху кинетиката и точността на трансляцията в рибозомата.
- ТΨС-рамо: съдържа стъбло и бримка, в която има модифицирани бази; рамото се нарича ТΨС, за да се уточни наличието на тимидин, псевдоуридин и цитидинови остатъци. Молекулите на тРНК са необичайни, тъй като съдържат голям брой модифицирани бази, както и тимидин, който обикновено се среща само в ДНК. ТΨС-рамото участва във взаимодействието на тРНК с рибозомата.
- Променлива бримка: тя се намира между антикодоновото рамо и ТΨС-рамото и, както подсказва името ѝ, варира по размер от 3 до 21 бази при различните тРНК.
- **Изоакцепторна група тРНК**

Повечето аминокиселини се кодират от множество кодони, които се разпознават в иРНК от съответните тРНК (изоакцепторна група тРНК). Поради изродеността на кода, повечето от тРНК за дадена аминокиселина имат различни последователности на антикодоните. При много видове броят на тРНК далеч надхвърля 61 и различни тРНК, носещи един и същ антикодон, могат да покажат различна ефективност при трансляцията, добавяйки ниво на регулация към процеса на протеинов синтез.

- **Кодон-антикодоново взаимодействие**

тРНК действат като преводачи на кода – осигуряват правилната аминокиселина в отговор на всеки кодон по силата на прецизно комплементарно свързване кодон-антикодон. тРНК се свързва с кодоните и последователно вмъкват своите аминокиселини в точния ред, определен от последователността на кодоните в иРНК. Този процес протича в рибозомата.

Антикодонът на тРНК се свързва комплементарно и антипаралелно със съответния кодон. За разлика от кодона, който се чете от 5'- към 3'-края в иРНК, антикодонът се чете от 3'-към 5'-края на тРНК. Например, ако иРНК кодонът е 5'-AUG-3', съответният тРНК антикодон ще бъде 3'-UAC-5'. Следователно 1-вата база на кодона се свързва с 3-тата база на антикодона, 2-рата база на кодона - с 2-рата база на антикодона, и 3-тата база на кодона - с 1-вата база на антикодона. За първите две бази на кодона свързването следва обичайното правило: аденинът (A) комплементира с урацил (U), а гуанинът (G) комплементира с цитозин (C). Въпреки това, при 3-тата база на кодона свързването се извършва съгласно набор от правила, известни като хипотеза на колебанието (от англ. wobble hypothesis).

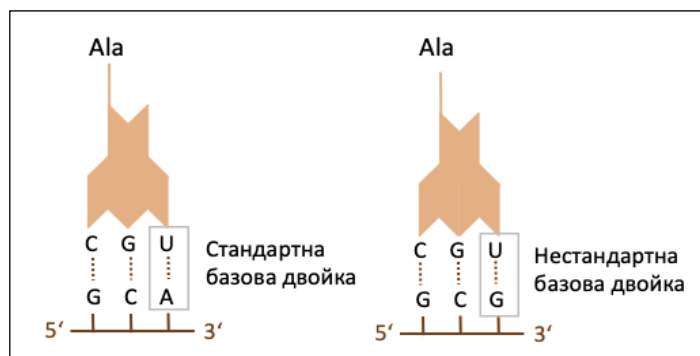
Хипотезата на колебанието, предложена през 1966 г. от Франсис Крик, постулира, че 5'-базата на антикодона на тРНК, която се свързва с 3'-базата на иРНК, не е толкова пространствено ограничена като другите две бази, и следователно може да има нестандартно свързване на бази. Пуриновата база инозин (I) се среща често вместо аденин (A) в 1-вата позиция (5') на антикодона. На това място инозинът може да образува базова двойка с която и да е от трите бази (U, C или A) в 3-тата позиция (3') на кодона (Таблица 6.1).

Хипотезата на колебанието обяснява защо множество кодони могат да кодират една аминокиселина. Дадена тРНК молекула (с една натоварена аминокиселина) може да разпознае и да се свърже с повече от един кодон, поради по-малко прецизните базови двойки, които могат да възникнат между 3-тата база на кодона и 1-вата позиция на антикодона (Фигура 6.5). Това обяснява защо съществуват повече кодони, отколкото специфични тРНК молекули. Хипотезата на колебанието също така илюстрира защо

единствената променлива между кодоните, които кодират една и съща аминокиселина, е тяхната 3-та база.

Таблица 6.1. Позволено базово свързване на 3-тата база на кодона с 1-вата база на антикодона съгласно Хипотезата на колебанието.

АНТИКОДОН 1-ва база	КОДОН 3-та база
C	G
A	U
U	A или G
G	U или C
I	U, C или A



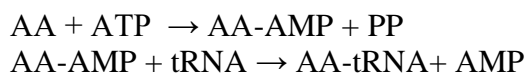
Фигура 6.5. Хипотеза на колебанието.

тРНК на Аланина (Ala), която има антикодона CGU, може да се свърже с два от кодоните, кодиращи Ala – GCA (стандартна базова двойка) и GCG (нестандартна базова двойка).

- Зареждане на аминокиселините върху тРНК – аминоксил-тРНК синтетази

Специфичното прикрепване на аминокиселините към тРНК (натоварване на тРНК) се катализира от клас ензими, известни като аминоксил-тРНК синтетази. Реакцията е известна като аминоксилация и води до образуване на аминоксил-тРНК. Всяка аминокиселина има своя аминоксил-тРНК синтетаза, която катализира нейното натоварване върху всички тРНК, които имат антикодони, съответстващи на тази аминокиселина (изоакцепторна група тРНК). Всяка аминоксил-тРНК синтетаза трябва да прави разлика между различните тРНК, така че само правилните да бъдат аминоксилирани. По този начин точното разпознаване на тРНК от тези ензими установява правилата на генетичния код.

Аминоксил-тРНК синтетазите катализират образуването на естерна връзка между 3'-хидроксилната група на аденина (A), който се намира на 3'-края на тРНК, и карбоксилната група на аминокиселината. Прикрепването на аминокиселина към тРНК е двуетапна реакция, която води до образуване на аминоксил-тРНК:



tRNA - тРНК

AA (Amino Acid) - АК (Аминокиселина)

ATP (Adenosine Tri-phosphate) – АТР

AMP (Adenosine Mono-phosphate) – АМФ

PP (pyrophosphate) – пиррофосфат

### 6.1.3. Рибозомите - апаратът за синтез на протеини

Рибозомите са макромолекулни машини, в които се извършва синтез на протеини от аминокиселини. Всяка рибозома се състои от две субединици - голяма субединица и малка субединица, изградени от рибозомни протеини и рРНК (Таблица 6.2). Рибозомите, и техните субединици, се измерват в единица Svedberg (S), която е мярка за техния размер и скоростта на утаяване в центробежно поле. Прокариотните клетки съдържат 70S рибозоми, всяка от

които се състои от малка (30S) и голяма (50S) субединица. Еукариотните клетки съдържат 80S рибозоми, разположени в цитозола, всяка от които се състои от малка (40S) и голяма (60S) субединица.

В допълнение, еукариотните клетки съдържат митохондриални рибозоми, а растителните клетки имат и хлоропластни рибозоми. Тези рибозоми значително се различават по размер и състав от другите рибозоми, открити в еукариотните клетки, и са по-сходни с тези, присъстващи в клетките на бактериите и синьо-зелените водорасли. Сходството на митохондриалните и хлоропластните рибозоми с прокариотните рибозоми обикновено се счита за силно подкрепящо доказателство, че митохондриите и хлоропластите са еволюирали от предшественици на прокариоти

Таблица 6.2. Структура на бактериалната и еукариотната рибозома.

РИБОЗОМА	Субединици	rPHK	Рибозомни белтъци
Бактериална рибозома (70S)	Голяма (50S)	23S rRNA 5S rRNA	31
	Малка (30S)	16S rRNA	21
Еукариотна рибозома (80S)	Голяма (60S)	28S rRNA 5,8S rRNA	49
		5S rRNA	
	Малка (40S)	18S rRNA	33

За да бъде доставена до рибозомата, всяка аминокиселина се натоварва върху съответна тРНК, при което се образува аминоксил-тРНК. Свързването на аминокиселината към растящата полипептидна верига става чрез взаимодействие с тРНК, доставила в рибозомата предходната аминокиселина и натоварена с растящата полипептидна верига, пептидил-тРНК. Всяка от тези тРНК заема определено място в рибозомата (виж по-надолу Фигура 6.9):

- А-мястото се заема от новопостъпилата аминоксил-тРНК, чийто антикодон се свързва комплементарно с кодона на иРНК;
- Р-мястото се заема от пептидил-тРНК, която носи растящата полипептидна верига и все още е свързана с иРНК в кодона, съответстващ на последната аминокиселина, присъединена към растящата полипептидна верига;
- Е-мястото се заема от деацилираната тРНК след образуването на пептидната връзка, но преди тя да излезе от рибозомата.

Във всички рибозоми голямата субединица взаимодейства с 3'-ССА края на тРНК и съдържа активното място (пептидил трансферазен център), което катализира образуването на пептидна връзка. Малката субединица свързва антикодона на тРНК, като наблюдава (по време на декодиране) и поддържа (по време на транслокация) взаимодействията между базите на тРНК и иРНК.

При еукариоти, рибозомните субединици се произвеждат и сглобяват в ядрото. Рибозомните протеини навлизат в ядрото и се комбинират с четирите рРНК, за да създадат двете рибозомни субединици. Те напускат ядрото през ядрените пори и в цитоплазмата остават отделени една от друга, когато не синтезират активно протеини. Двете субединици се свързват в цитоплазмата, когато се извършва синтез на протеини

Типичната клетка съдържа милиони рибозоми. Те могат да образуват групи (полирибозоми или полизоми), прикрепени към иРНК, която превеждат по време на протеиновия синтез. Рибозомите също могат да бъдат намерени прикрепени към мембраните на грапавия ендоплазмен ретикулум, където те синтезират протеини, които или се секретират от клетката, или се включват в клетъчните мембрани.

## 6.2. Основни етапи на трансляцията

Инициация – етап на сглобяване върху иРНК на малката и голямата субединица във функционална рибозома, при което се разпознава старт кодона и се свързва инициаторната тРНК, натоварена с метионин. Този етап е относително бавен и от него обикновено зависи скоростта, с която дадена иРНК се транслира.

Удължаване - включва реакциите от образуване на първата пептидна връзка до прибавянето на последната аминокиселина в полипептидната верига. Последователното присъединяване на аминокиселините е най-бързия етап на белтъчната синтеза.

Прекратяване - на този етап става освобождаване на полипептидната верига от рибозомата, която след това се отделя от иРНК.

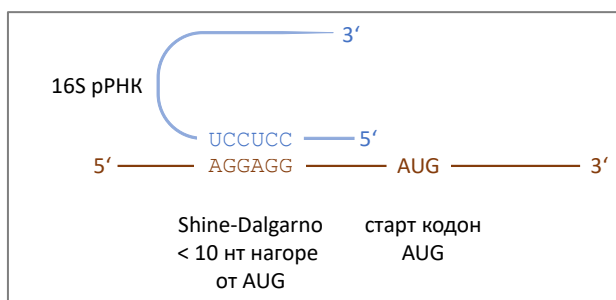
### 6.2.1. Инициация на трансляцията

По време на този етап малката субединица на рибозомата се свързва с иРНК. Инициаторната тРНК, носеща метионин (Met), се разполага в Р-мястото на малката субединица и се свързва със старт кодона AUG. След това към този комплекс се присъединява голямата субединица за да се образува функционалната рибозома.

Как рибозомата знае кой е старт кодона в иРНК, откъдето да започне протеиновия синтез? 5'-НТР на иРНК съдържат регулаторни елементи, които участват в избора на началното място за протеиновия синтез и в контрола на неговата скорост.

При прокариоти, инициацията на трансляцията е функция на малката субединица (30S) на прокариотната рибозома, прокариотните фактори на инициация (IF-1, IF-2, IF-3) и инициаторната тРНК – tRNA<sup>fMet</sup>.

В повечето прокариотни иРНК е идентифицирана последователност, която насочва рибозомата към началото на протеин-кодирания регион. В *E. coli* тази последователност (AGGAGG) е известна като област на Шайн-Дагарно (по името на нейния откривател Shine-Dalgarno). Областта е комплементарна на част от 16S рРНК в малката субединица на бактериална рибозома. Малката субединица разпознава областта и, чрез образуването на водородни връзки, се позиционира върху нея (Фигура 6.6).

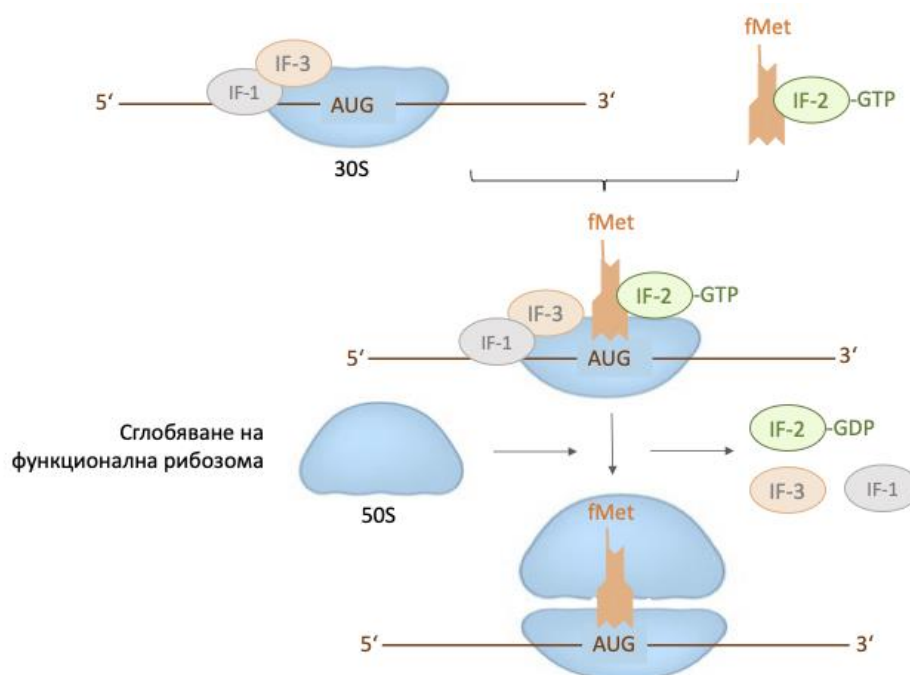


Фигура 6.6. Инициация на протеиновия синтез при прокариоти

Инициаторната тРНК ( $tRNA^{fMet}$ ) носи аминокиселината метионин, която се формилира след прикрепването към тРНК. Всяка полипептидна верига, синтезирана от *E. coli*, започва с формилиран метионин (fMet), който обикновено се отстранява след завършване на транслацията. Когато в хода на удължаване на полипептидната верига, рибозомата срещне кодона AUG, обикновена  $tRNA^{Met}$  вмъква неформилиран метионин.

При прокариоти, инициацията на транслацията е подпомогната от 3 прокариотни фактори на инициация, и включва няколко етапа (Фигура 6.7):

- Образуване на троен комплекс IF-3\*30S\*иРНК - IF-3 се свързва с 30S субединица и предотвратява свързването ѝ с 50S субединицата; IF-3 подпомага малката субединица да се свърже с иРНК в областта на Шайн-Далгарно; IF-1 се свързва с А-мястото, където предотвратява навлизането на нова аминоксил-тРНК молекула, преди да се сглоби пълната рибозома;
- Образуване на двоен комплекс IF-2(GTP)\* fMet-tRNA<sup>fMet</sup> - IF-2 (свързан с GTP) се свързва с fMet-tRNA<sup>fMet</sup> ;
- Свързване на двойния и тройния комплекс и стабилизиране на функционална рибозома - IF-2 настанява fMet-tRNA<sup>fMet</sup> в Р мястото на 30S субединицата, което е позиционирано така, че антикодонът на тРНК да се настани над старт кодона AUG в иРНК; IF-3 проверява точността на разпознаване на първата аминоксил-тРНК (fMet-tRNA<sup>fMet</sup>); следва присъединяване на 50S субединица, при което IF-2 катализира хидролизата на GTP до GDP като освободената енергия води до освобождаване на инициращите фактори и предизвиква конформационни промени за стабилизиране на функционалната рибозома.



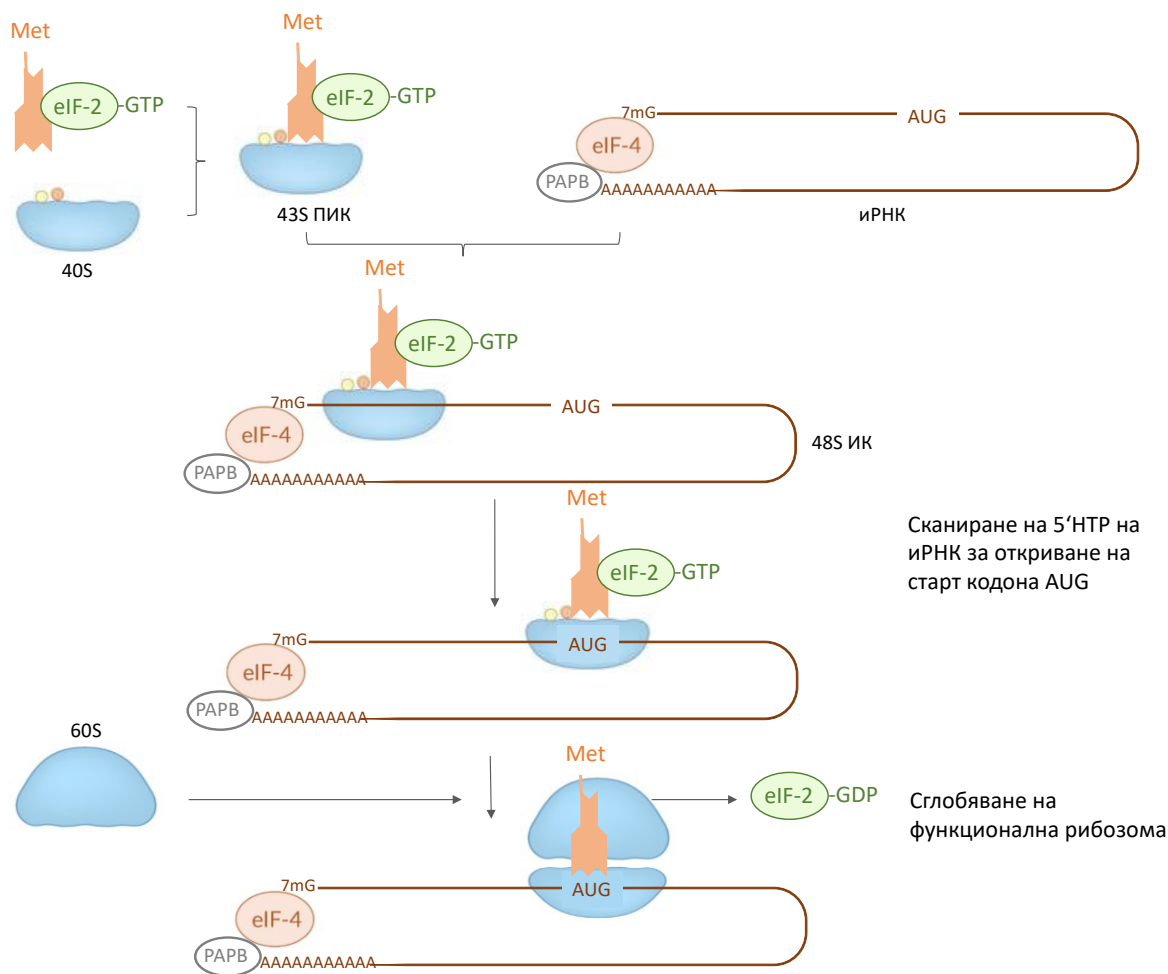
Фигура 6.7. Инициация на протеиновия синтез при прокариоти

При еукариоти, инициацията на транслацията е функция на малката субединица (40S) на еукариотната рибозома, еукариотните фактори на инициация (eIF) и инициаторната тРНК – tRNA<sub>i</sub>.

Инициацията на транслацията започва със свързването на 40S субединицата към 7-метилгуаниновата "шапчица" (7<sup>m</sup>G) в 5'-края на иРНК, след което се придвижва надолу по молекулата, сканирайки 5'-НТФ последователността, докато срещне първия AUG кодон. Много еукариотни иРНК се транслират от първия AUG, но това не винаги е така. Според правилото на Козак (от англ. Kozak's rule), нуклеотидите около AUG показват дали това е правилният начален кодон. Правилото е постулирано за гръбначни животни и гласи, че следната консенсусна последователност трябва да се появи около AUG триплет за да бъде разпознат като „старт“ кодон: 5'-gccRccAUGG-3' (където R означава пурин и показва място, което може да бъде А или G, но не може да бъде С или U). При някои вирусни и еукариотни иРНК са установени и алтернативни места за свързване на рибозома (internal ribosome entry site, IRES), които позволяват инициацията на транслацията, която е независима от „шапчицата“. Метионинът върху заредената инициаторна тРНК (Met-tRNA<sub>i</sub>) не е формиран. Въпреки това, инициаторната тРНК (tRNA<sub>i</sub>) се различава от другите тРНК за метионин (tRNA<sup>Met</sup>) по това, че само тя може да свързва eIF-2.

При еукариоти, инициацията на транслацията зависи от координираното взаимодействие на най-малко 12 еукариотни фактори на инициация, иРНК и 40S субединицата, и преминава през няколко стъпки (Фигура 6.8):

- Образуване на двоен комплекс eIF-2(GTP)\*Met-tRNA<sub>i</sub> - eIF-2 (свързан с GTP) присъединява инициаторната тРНК, натоварена с метионин (Met-tRNA<sub>i</sub>);
- Образуване на 43S пре-инициращ комплекс (43S ПИК) - двойният комплекс се свързва с 40S субединицата с помощта на eIF-1, eIF-1A, eIF-3 и eIF-5;
- Образуване на 48S инициращ комплекс (48S ИК) - eIF-4F, съставен от eIF4A, eIF4E и eIF4G, улеснява свързването на иРНК към 43S ПИК;
  - eIF4E е свързан с 5'-шапчицата (7<sup>m</sup>G); eIF4G е свързан с поли (А)-свързващия протеин (PABP), който се свързва с поли (А)-опашката на иРНК, като по този начин иРНК придобива пръстеновидна форма и се стабилизира; eIF4A функционира като хеликаза и улеснява пре-инициращия комплекс при сканирането на иРНК.;
- Сканиране в посока 5'-3' на 5'-НТФ на иРНК за откриване на AUG - използвайки енергията на АТФ хидролизата, 48S ИК сканира иРНК за да локализира първия AUG кодон;
  - Сглобяване на функционалната рибозома - при разпознаване на AUG, eIF-2 катализира хидролизата на GTP, задействайки отделянето на всички инициращи фактори и свързването на 60S субединицата на рибозома.



Фигура 6.8. Инициация на протеиновия синтез при еукариоти

### 6.2.2. Удължаване на полипептидната верига

Ключов етап на протеиновия синтез върху рибозомите е пептидиловият трансфер, при който нарастващият полипептид се удължава с една аминокиселина във всеки цикъл на удължаване според последователността на кодоните в иРНК. иРНК се транслира в посока от 5'-края към 3'-края, а протеинът нараства от N-края (със свободна амино-група) към С-края (със свободна карбокси-група), като всеки цикъл добавя една аминокиселина към С-края на полипептидната верига. Етапът на удължаване се подпомага от фактори на удължаване (от англ. elongation factor, EF) - G-протеини, които използват гуанозин трифосфат (GTP) за своите функции. GTP действа като източник на енергия по време на транслацията - както в началото на удължаването, така и по време на транслокацията на рибозомата.

Етапът на удължаване се състои от поредица от цикли, чийто общ брой се определя от дължината иРНК. Всеки цикъл включва няколко основни стъпки (Фигура 6.9):

- Натоварване на аминоксил-тРНК в А-мястото на рибозомата. При първият цикъл на удължаване Р-мястото е заето от инициаторната тРНК, натоварена с метионин, а при всеки следващ цикъл Р-мястото се заема от пептидил-тРНК (тРНК, която носи растящия полипептид). Заредена тРНК, носеща следващата аминокиселина (аминоацил-тРНК), се

свързва към свободното А-място чрез образуване на базови двойки с кодона в иРНК, който се намира в А-мястото.

При прокариоти, аминоксил-тРНК се пренася до рибозомата от термо-нестабилния фактор на удължаване EF-Tu. Тройният комплекс аминоксил-тРНК\*EF-Tu(GTP) навлиза в А-мястото. Ако към кодона в иРНК се свърже правилен антикодон, рибозомата променя конфигурацията си и активира GTP-азната активност на EF-Tu, което води до хидролиза на GTP до GDP и Pi. При хидролиза на GTP, конформацията на EF-Tu се променя драстично - EF-Tu(GDP) се отделя от аминоксил-тРНК и напуска рибозомата.

EF-Tu играе важна роля за поддържане на точността на трансляция. Рибозомата не активира GTP-азната активност на EF-Tu, ако антикодонът на аминоксил-тРНК в А-мястото не съвпада с кодона в иРНК, като по този начин преференциално се увеличава вероятността неправилната аминоксил-тРНК да напусне рибозомата. Освен това, независимо от съпадението на кодон:антикодон, след отделянето си от аминоксил-тРНК, EF-Tu предизвиква забавяне в пълното навлизане на аминоксил-тРНК в А-мястото (процес, наречен акомодация). Този период на забавяне е втора възможност за неправилно заредени аминоксил-тРНК да се освободят от А-мястото, преди неправилната аминокиселина да бъде необратимо добавена към полипептидната верига в цитоплазмата, дезактивираният EF-Tu(GDP) взаимодейства с термо-стабилния фактор на удължаване EF-Ts, който кара EF-Tu да освободи GDP. EF-Tu е в състояние да образува комплекс с GTP, което води до реактивиране на EF-Tu (GTP) за да може да се свърже с друга аминоксил-тРНК.

При еукариоти, еукариотният EF1A (еEF1A) образува троен комплекс аминоксил-тРНК\*еEF-1A(GTP). Тези тройни комплекси навлизат в празното А място на рибозомата, където GTP се хидролизира и еEF1A се освобождава, ако се развие подходящо взаимодействие кодон-антикодон между аминоксил-тРНК, влизаща в мястото А, и кодона в иРНК.

- Свързване на аминокиселината с растящата полипептидна верига. Свободната аминоксидна група на аминоксил-тРНК атакува естерната връзка, която свързва тРНК, разположена в Р-мястото, с растящия полипептид. С-края на полипептида се отделя от тРНК в Р-мястото и се свързва чрез пептидна връзка към свободната аминоксидна група на новопостъпилата аминокиселина, свързана с тРНК (аминоксил-тРНК), в А-мястото. Всъщност, образуването на пептидната връзка е свързано с прехвърляне на растящата полипептидна верига върху новата аминокиселина – в Р-мястото остава деацилирана тРНК, а в А-мястото вече се намира полипептид-тРНК (полипептидът е удължен с една аминокиселина).

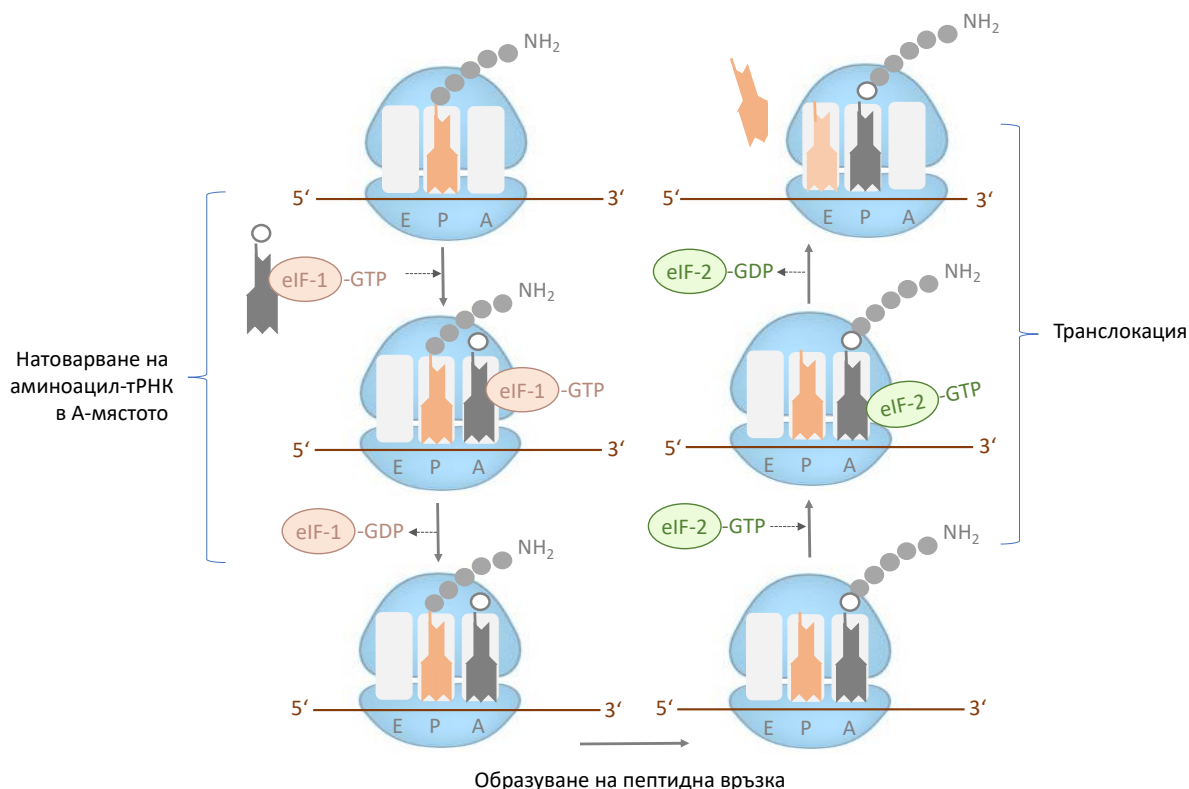
Пептидил-трансферазна активност е отговорна за образуването на пептидната връзка от естерна връзка. Активността се приписва на рибозомата, но рибозомните протеини нямат капацитета да катализират тази реакция. Кристалните структури на рибозомата показват, че пептидил-трансферазната активност е функция рРНК на голямата субединица на рибозомата (23S рРНК при прокариоти и 28S рРНК при еукариоти). По този начин рибозомата е най-големият известен РНК катализатор – рибозим.

- Придвижване на тРНК и иРНК в рибозомата (транслокация). Цикълът на удължаване е завършен, когато пептидил-тРНК, разположена в А-мястото, и деацилираната-тРНК, разположена в Р-място, се преместят съответно в Р- и Е-мястото, при което А-мястото се освобождава за приеме следващата аминоксил-тРНК. Транслокацията на двете тРНК е свързана с движението на иРНК през рибозомата.

Съгласно общоприетия модел за хибридни състояния, след пептидилното прехвърляне, акцепторните краища на двете тРНК (пептидил-тРНК и деацилираната-тРНК) могат да се движат спонтанно по отношение на голямата субединица за да образуват хибридни състояния. В хибридно състояние деацилираната тРНК заема Р-мястото в малката субединица и Е-мястото в голямата субединица (Р/Е хибридно състояние), докато пептидил-тРНК заема А-мястото в малката субединица и Р-място в голямата субединица (А/Р хибридно състояние).

Предполага се, че образуването на хибридни състояния предхожда движението на иРНК през малката субединица – при това движение контактите между кодоните на иРНК и антикодоните на двете тРНК остават непроменени за да бъде запазена рамката на четене, като комплексът иРНК-тРНК се придвижва с един кодон. Това движение се катализира от фактори на удължаване и използва енергията от хидролиза на GTP. При прокариоти, процесът се катализира от фактора на удължаване EF-G(GTP), а при еукариоти - от неговия хомолог eEF-2(GTP).

Движението на иРНК през малката субединица превръща хибридни състояния Р/Е и А/Р съответно в състояния Е/Е и Р/Р - деацилираната-тРНК е разположена вече в Е-мястото, откъдето напуска рибозомата, пептидил-тРНК е разположена в Р-мястото, а А-мястото е свободно за да приеме следващата аминоксил-тРНК.



Фигура 6.9. Удължаване на полипептидната верига при прокариоти

### 6.2.3. Прекратяване на транслацията

Прекратяване на протеиновия синтез както в еукариотите, така и в бактериалните клетки, се осъществява, когато мястото А-мястото на рибозомата достигне до един от стоп кодоните

(UAA, UAG и UGA) на иРНК. Протеините, наречени освобождаващи фактори (от англ. releasing factors, RF) разпознават тези кодони и освобождават полипептида, който е прикрепен към последната тРНК в Р-мястото. Този процес е енергийно-зависим - енергията от хидролизата на GTP се използва за освобождаването на новосинтезирания протеин от тРНК в Р-мястото. Следва дисоциация на рибозомните субединици, и освобождаване на тРНК и иРНК.

При прокариоти, има три освобождаващи фактора. RF1 и RF2 взаимодействат с терминаращите кодони – RF1 разпознава UAA и UAG, а RF2 разпознава UAA и UGA. Третият фактор RF3 стимулира скоростта, с която се освобождава протеина. При еукариоти, има два освобождаващи фактора. eRF1, разпознава и трите терминаращи кодони. eRF3 действа кооперативно с eRF1 за да гарантира бързо освобождаване на синтезирания протеин.

### ПРОТЕИНОВ СИНТЕЗ и АНТИБИОТИЦИ

Механизмът на протеинов синтез е сходен в прокариотните и еукариотните клетки, но рибозомите и белтъчните фактори на инициация, удължаване, терминация са различни, което позволява разработването на антибиотици, които селективно инхибират протеиновия синтез в бактериите. Антибиотиците, които са специфични за прокариотната 70S рибозома или нейните субединици, нямат афинитет към еукариотната 80S рибозома или нейните субединици. В Таблица 6.3 са представени едни от най-често използваните антибиотици с техните прицелни места в 70S рибозомата на прокариоти и механизмите им на действие.

Таблица 6.3. Антибиотици, техните прицелни места в бактериалната рибозома и механизъм на действия

АНТИБИОТИК	Прицелно място и механизъм на действие
Тетрациклин	30S субединица, пречи на свързването на аминокцил-тРНК в А-мястото
Аминогликозоди (стрептомицин, амикацин, гентамицин, пурамицин)	30S субединица, блокират инициацията, пречат на процеса на корекция в протеиновата и увеличават нивото на грешки
Макролиди (азитромицин, klarитромицин, еритромицин)	50S субединица, инхибират пептидилния трансфер или транслокацията
Хлорамфеникол	50S субединицата; инхибира пептидилния трансфер

## ГЛАВА 7 РЕГУЛАЦИЯ на ГЕННАТА ЕКСПРЕСИЯ

Генната експресия представлява процес, при който генетичната информация се използва за синтез на функционални продукти - протеини или РНК. Протеин-кодиращите гени служат като матрица за производството на протеини, които изпълняват разнообразни клетъчни функции, включително катализиране на биохимични реакции, структурна поддръжка на клетките, участие в клетъчната сигнализация и други. В същото време некодиращите гени произвеждат различни видове РНК, като тРНК, рРНК, микроРНК и дълги некодиращи РНК. Тези молекули играят важна регулаторна и структурна роля, която е от съществено значение за контролирането на генната експресия.

Балансът на нивата на протеините в клетката се регулира чрез синтетични процеси, като транскрипция и транслация. Допълнителна модулация се осъществява чрез обработка на РНК, посттранслационни модификации и механизми за разграждане на протеините. Контролът на тези процеси осигурява необходимата динамика за адаптиране на клетъчната функция в отговор на променящи се условия.

Генната експресия в многоклетъчните организми има специфични качествени и количествени измерения през различните етапи на развитие и при различни биологични условия. Тя варира в зависимост от функциите, необходими на организма в определен момент, което позволява адаптация към различни физиологични предизвикателства и осигурява правилното функциониране на тъканите и органите. Тази сложност подчертава важността на регулацията на генната експресия за поддържане на хомеостазата и здравето на организма.



### *François Jacob, Jacques Monod (1956)*

В края на 50-те години на XX век, Франсоа Жакоб и Жак Моно публикуват свои трудове по молекулярна биология на бактериите, в които предлагат първия теоретичен модел за регулация на генната експресия в прокариоти - така наречения „модел на оперона“. След поредица от иновативни експерименти с бактерии, те идентифицират регулаторни гени, които контролират дейностите на структурните гени, отговорни за производството на ензими, други протеини и РНК. Те откриват, че балансът между регулаторните и структурните гени в нормалната клетка позволява адаптация към различни условия.

Моно предлага идеята за съществуването на информационна РНК (иРНК, mRNA), чиято нуклеотидна последователност е комплементарна на тази на ДНК в клетката. Той постулира, че иРНК пренася „информацията“, кодирана в нуклеотидната последователност на ДНК до рибозомите, местата на протеиновия синтез, където тя се трансформира в аминокиселинна последователност.

В експериментите си Жакоб и Моно използват обикновената лабораторна бактерия *Escherichia coli*, но много от основните концепции за регулация, открити от тях, са фундаментални за клетъчната регулация във всички организми. Гените не се експресират постоянно, а се регулират – активират или потискат – по прецизен начин, за да отговорят на нуждите на организма. Например, храносмилателните ензими не се произвеждат

непрекъснато, а само след хранене; имунната система се активира при заплахата от чужди агенти. Откритията на Жакоб и Моно предоставят основа за разбирането на тези процеси.

През 1965 г. Жакоб и Моно, заедно с Андре Лвоф, получават Нобелова награда за физиология или медицина за техния принос към молекулярната биология.

### 7.1. Регулация на генната експресия при прокариоти

Основната цел на контрола на генната експресия в прокариотите е да регулира синтеза на протеини в отговор на вътрешни и външни промени в околната среда. Тази регулация е от решаващо значение за оцеляването и адаптивността на прокариотните организми, позволявайки им ефективно да управляват своите метаболитни ресурси и да реагират на различни условия.

#### 7.1.1. Бактериалните гени са организирани в оперони и се контролират координирано

При прокариоти, гените, които кодират белтъци със свързани функции, най-често ензими от един метаболитен път, са организирани в клъстери, наречени оперони. Оперонът е единица за координирана експресия на група структурни гени, които са под контрола на общи *cis*-действащи регулаторни елементи (промотор, оператор и др.). Гените на оперона се транскрибират в една полицистронна (мултигенна) иРНК молекула, и трансляцията на всеки протеин от вътрешен AUG инициращ кодон води до координирана продукция на белтъците, кодирани от индивидуалните гени.

Регулацията на прокариотната транскрипция се осъществява от регулаторни протеини, които се свързват с регулаторни елементи, разположени близо до мястото на начало на транскрипция на транскрипционна единица. При отсъствие на регулаторни протеини, РНК полимеразата все още може да се свърже с промотора и да иницира транскрипция, но на ниско "базално" ниво.

Регулаторните белтъци се кодират от регулаторни гени, те могат или да потискат транскрипцията (репресори) – негативен контрол; или да стимулират транскрипцията (активатори) – позитивен контрол. Съществуват и други нива на регулация – така например, молекула, наречена индуктор, може да се свърже с репресора и да го инактивира; или репресор може да не е в състояние да се свърже с оператора, освен ако не е свързан с друга молекула, корепресора.

Катаболитните пътища катализират разграждането на хранителни вещества за генериране на енергия под формата на АТФ. Индуцируемите оперони често съдържат гени, кодиращи ензими, които участват в катаболитни процеси. Генната транскрипция е „изключена“, но може да бъде „включена“ в отговор на малка индуцираща молекула, идентична или подобна на ензимния субстрат. Ензимите са необходими само когато субстратът е наличен, така че експресията на опероните обикновено се индуцира само в неговото присъствие.

Анаболитните пътища използват енергия под формата на АТФ и редуциращи еквиваленти под формата на NAD(P)H за да катализират синтеза на клетъчни компоненти от по-прости материали, напр. синтез на аминокиселини от малки дикарбоксилни киселини. Репресиремите оперони съдържат гени, кодиращи ензими, които участват в анаболитни процеси. Генната транскрипция е „включена“, но може да бъде „изключена“ в отговор на малка репресираща молекула, идентична или подобна на крайния продукт. Когато продуктът от биосинтетичния път започне да се натрупва в клетката, експресията на оперона се потиска. Най-простите и най-добре разбрани примери за генна регулация се срещат в бактериите и във вирусите, които ги заразяват.

### 7.1.2. Лактозен оперон (*lac*-оперон)- индуцируем оперон

Генното регулиране на лактозния оперон (*lac*-оперон) в *E. coli* е първият генетичен регулаторен механизъм, който е разбран в детайли, и става най-яркия пример за прокариотна генна регулация. Този регулаторен механизъм е разгадан от Франсоа Жакоб и Жак Моно, когато те изучават метаболизма на лактозата за да отговорят на въпроса как една биологична клетка знае какви ензими да синтезира.

Геномът на *E. coli* се състои от една кръгова ДНК молекула от около 5.0 Mb и кодира приблизително 4 300 протеина, въпреки че само малка част от тях се произвеждат в даден момент. *Lac*-оперонът е необходим за транспортирането и метаболизма на лактозата в *E. coli* и много други чревни бактерии. Въпреки че глюкозата е предпочитаният източник на въглерод за повечето бактерии, *lac*-оперонът позволява ефективно усвояване на лактозата, когато глюкозата не е налична чрез активиране на бета-галактозидазата.

#### 7.1.2.1. Структура на *lac*-оперона

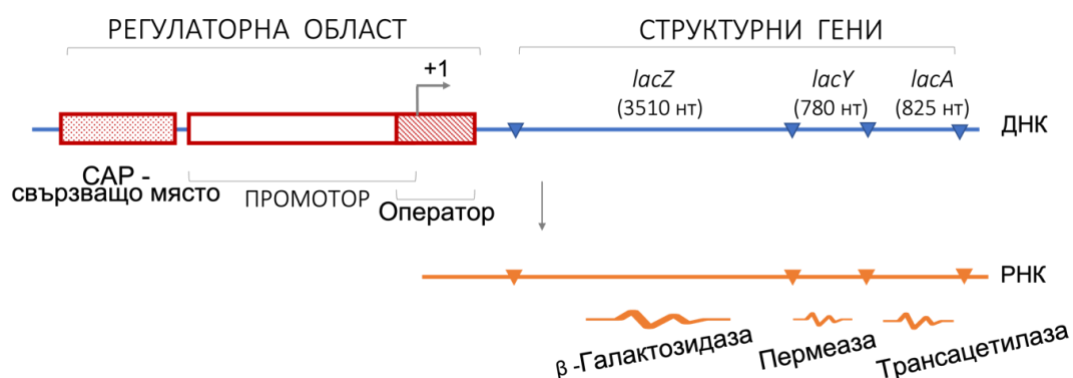
*Lac*-оперонът се състои от регулаторна област и 3 структурни гена, необходими за усвояването на лактозата (Фигура 7.1). Структурните гени са:

- *lacZ* кодира  $\beta$ -галактозидаза (*LacZ*), ензим, който разгражда дизахарида лактоза до глюкоза и галактоза.
- *lacY* кодира  $\beta$ -галактозид пермеаза (*LacY*), мембранен белтък, който транспортира  $\beta$ -галактозиди, включително лактоза, в клетката, използвайки протонен градиент.
- *lacA* кодира  $\beta$ -галактозид трансацетилаза (*LacA*), ензим, който пренася ацетилна група от ацетил-СоА към тиогалактозид.

Само *lacZ* и *lacY* са основно необходими за катаболния метаболизъм на лактозата, докато *lacA* играе по-малка роля в този процес.

Регулаторната област съдържа регулаторни ДНК последователности. Това са участъци на ДНК, към които могат да се свържат определени регулаторни белтъци за контролират транскрипцията на оперона:

- промотор, към който се свързва РНК полимеразата;
- оператор, към който се свързва репресора - *lac*-репресор или *LacI*; когато *lac*-репресорът е свързан, РНК полимеразата не може да се свърже с промотора и да започне транскрипция, тъй като операторът се припокрива частично с промотора;
- CAP-свързващо място, към което се свързва активатора (от англ. Catabolite Activating Protein, CAP) - когато CAP е свързан, той стимулира транскрипцията като помага на РНК полимеразата да се свърже с промотора.

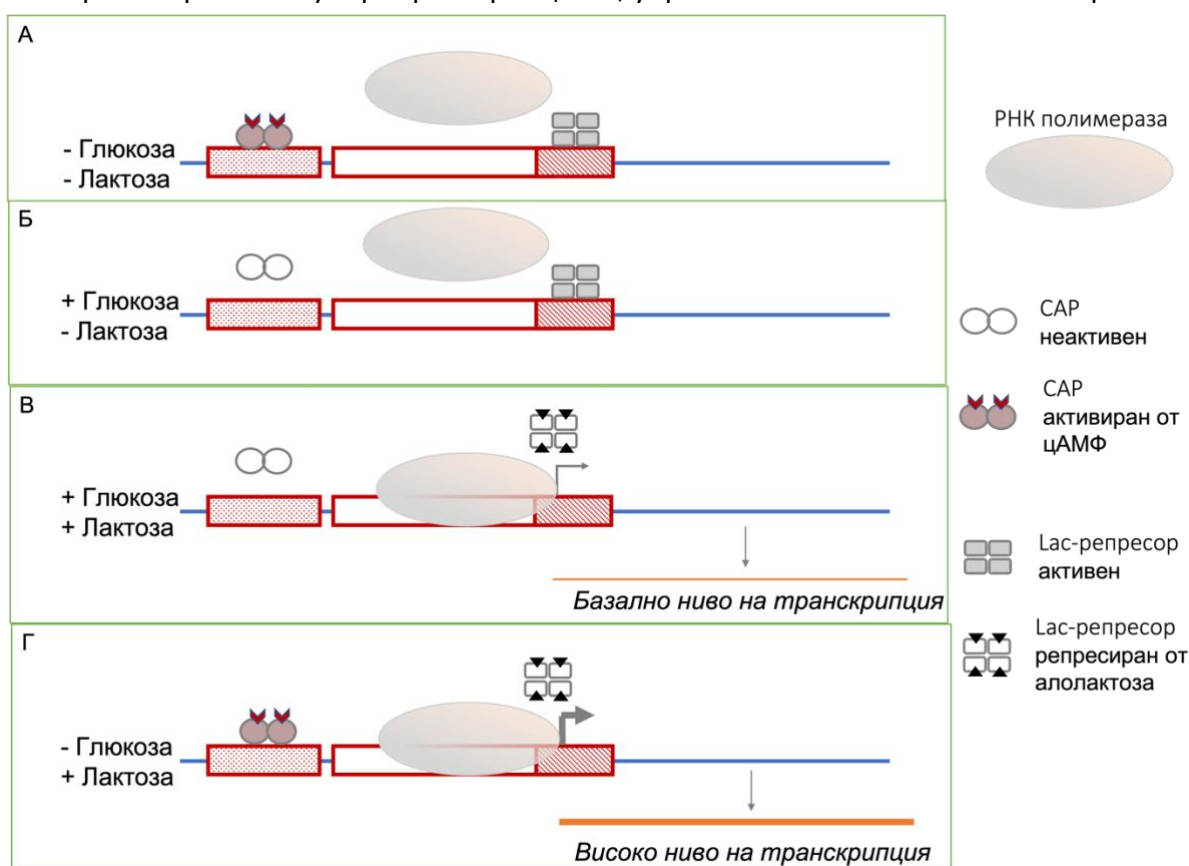


Фигура 7.1. Структура на *lac*-оперона

### 7.1.2.2. Регулация на *lac*-оперона

*Lac*-оперонът на *E. coli* се контролира от слаб промотор, който свързва РНК полимеразата неефективно и инициира транскрипция на ниско базално ниво. Базалната транскрипция е минималното ниво на генна експресия, което възниква при отсъствие на регулаторни белтъци, или когато те са в неактивна форма (Фигура 7.2 В). Тя е от решаващо значение за способността на клетката да реагира на промените в наличността на лактоза в околната среда, позволявайки бързото ѝ адаптиране, когато лактозата стане единственият източник на въглерод.

Активността на промотора, който контролира експресията на *lac*-оперона, се контролира от два регулаторни протеина. Единият протеин (*lac*-репресор) пречи на свързването на РНК полимеразата към промотора и потиска транскрипцията, упражнявайки негативен контрол, другият протеин (катаболитния активатор, CAP) подпомага свързването на РНК полимеразата към промотора и стимулира транскрипцията, упражнявайки положителен контрол.



Фигура 7.2. Регулация на експресията на *lac*-оперона

Наличието или отсъствието на захарите лактоза и глюкоза контролират нивото на експресия на *lac*-оперона. Активирането на експресия изискват наличието на лактоза (и следователно липсата на активен *Lac*-репресор) и липсата на предпочитания енергиен източник, глюкоза (и следователно наличието на активен CAP).

- Негативен контрол на *lac*-оперона

*Lac*-репресорът се кодира от *lacI* гена, който се разполага нагоре от *lac*-оперона и се контролира от собствен промотор. Експресията на *lacI* гена не се регулира и непрекъснато се синтезират много ниски нива на *lac*-репресора (конститутивна експресия). Самият репресор се образува от четири идентични субединици, които се самосглобяват, за да образуват активния тетрамер.

При липса на лактоза в хранителната среда, *lac*-репресорът е в активна форма и блокира експресията на *lac*-оперона чрез свързване към оператора (Фигура 7.2. А и Б).

В присъствие на лактоза, малка част от лактозните молекули претърпяват изомеризация до алолактоза. *Lac*-репресорът, който има голям афинитет към алолактозата, се свързва с нея, променя формата си и се инактивира. Той вече не може да се свърже с ДНК, и РНК полимеразата има достъп до промотора и може да започне транскрипция – негативна индукция на *lac*-оперона (Фигура В7.2. В и Г).

За протеини като *lac*-репресора, които променят своята форма и функционални свойства след свързване с лиганд, се казва, че се регулират чрез алостеричен механизъм. Субстратите, които карат репресорите да се отделят от техните оператори, се наричат индуктори, а гените, които се регулират от такива репресори - индуцируеми гени.

- Позитивен контрол на *lac*-оперона

Силата на даден промотор се определя от способността му да свързва РНК полимеразата и да образува отворен комплекс. Промоторът на *lac*-оперона е слаб и следователно *lac*-оперонът се транскрибира слабо при индукция в среда с лактоза. Ако в хранителната среда присъства само лактоза като въглероден източник, CAP протеинът е в активна форма, свързва се към CAP-свързващото място на оперона и огъва ДНК. Така РНК полимеразата може да се свърже по-ефективно към промотора и транскрипцията на *lac*-оперона се стимулира значително (Фигура 7.2. Г).

CAP е друг пример за алостерично регулиран регулаторен белтък. Самият CAP се активира от цикличния AMP (сAMP), който се синтезиран от АТФ с помощта на ензима аденилат циклаза. Само когато CAP е свързан с сAMP, той може да се свърже към CAP място в промотора на *lac*-оперона. За разлика от *lac*-репресора, който се свързва с оператора само в отсъствието на неговия лиганд (алолактоза), CAP се свързва с CAP-мястото на промотора в присъствието на своя лиганд (сAMP).

- Катаболитна репресия

Катаболитната репресия (наричана още глюкозна репресия) е широко разпространено явление при микроорганизмите, при което клетките, отглеждани върху глюкоза, потискат експресията на голям брой гени, които са необходими за метаболизма на алтернативни източници на въглерод. Глюкозната репресия е широко изследвано в *E. coli*, която расте по-бързо върху глюкоза, отколкото върху всеки друг източник на въглерод. Ако *E. coli* се постави върху агар, съдържащ само глюкоза и лактоза, бактериите първо ще използват глюкозата и след това лактозата.

В присъствие на глюкоза в хранителната среда (дори и да има лактоза), *lac*-оперонът е обект на катаболитна репресия (Фигура 7.2. В). Глюкозата обикновено се транспортира в клетката чрез фосфоенолпируват (PEP)-зависима фосфотрансферазна система (PTS). PTS е фосфорилираща каскада, при която серия от протеини прехвърлят фосфатна група, като в крайна сметка фосфорилират глюкозата, докато тя се транспортира в клетката. Един от компонентите на PTS, протеинът EIIA<sup>Glc</sup>, играе решаваща регулаторна роля.

Когато нивата на глюкоза са високи, EIIA<sup>Glc</sup> е предимно в своето дефосфорилирано състояние, тъй като фосфатът от PEP се използва в PTS за фосфорилиране на входящата глюкоза, поддържайки EIIA<sup>Glc</sup> нефосфорилиран. В своята нефосфорилирана форма EIIA<sup>Glc</sup> се свързва с ензима лактозна пермеаза (LacY) в клетъчната мембрана, предотвратявайки приема на лактозата в клетката. Този механизъм, известен като изключване на индуктора, гарантира, че глюкозата се метаболизира преференциално пред лактозата.

Ако нивата на глюкоза са ниски в растежната среда, по-малко глюкоза се транспортира в клетката, което води до намалено търсене на PEP и, следователно, намаляване на PTS активността. При тези условия EIIA<sup>Glc</sup> става фосфорилиран, тъй като фосфатните групи вече не се използват широко за усвояване на глюкоза. Фосфорилираният EIIA<sup>Glc</sup> не се свързва с лактозната пермеаза, което ѝ позволява да бъде активна и да улесни усвояването на лактозата в клетката. Тази промяна позволява на *E. coli* да премине към метаболизиране на лактоза, когато глюкозата е оскъдна.

При високи нива на глюкоза, нефосфорилираният EIIA<sup>Glc</sup> води и до намалена активност на аденилат циклазата, по-ниски вътреклетъчни нива на cAMP и по-ниски нива на cAMP-CAP комплекса, в следствие на което инициацията на транскрипция на *lac*-оперона е много слаба; при понижаване на нивата на глюкозата, концентрацията на cAMP започва да нараства и се активира CAP, който от своя страна се свързва с CAP-свързващото място и стимулира транскрипцията.

Дори при високи нива на глюкоза, когато транскрипцията на *lac*-оперона е изключително трудна, тя все още се случва, макар и много неефективна. Малки количества пермеаза обикновено се произвеждат поради редки, случайни събития, които могат да "стартират" процеса веднага щом глюкозата се изчерпи, при условие че лактозата е налична. Докато глюкозата има силен инхибиторен ефект върху усвояването на лактозата, малки количества лактоза понякога могат да присъстват в клетките на *E. coli* заедно с глюкозата, особено при динамични условия на околната среда или когато генетични мутации засягат метаболитната регулация.

### 7.1.3. Триптофанов оперон (*trp*-operon) – репресируем оперон

Trp-оперонът на *E. coli* контролира биосинтезата на триптофан (*trp*) в клетката от първоначалния предшественик хоризматова киселина.

#### 7.1.3.1. Структура на *trp*-оперона

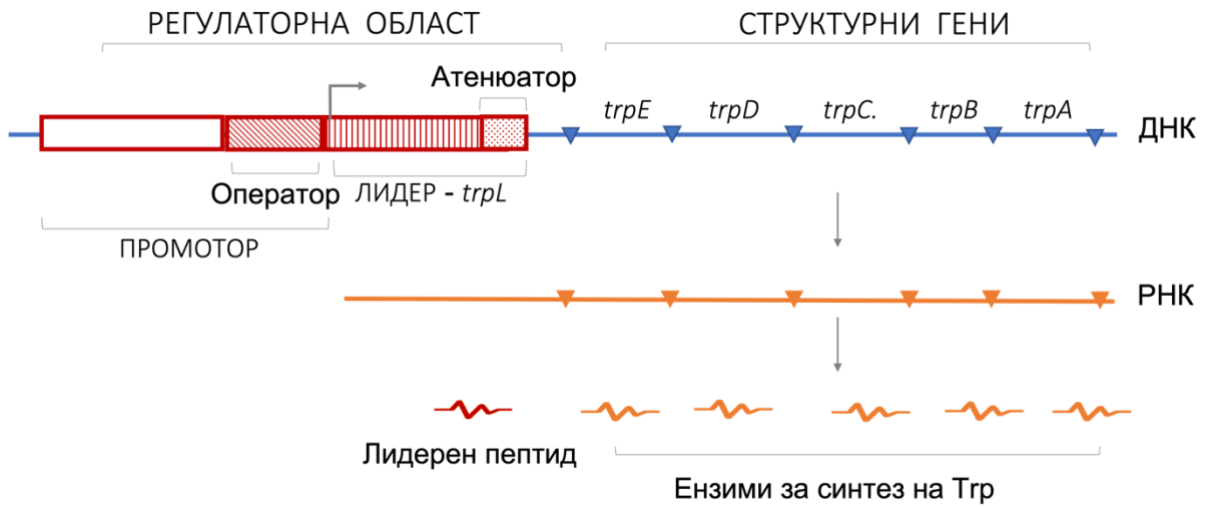
*Trp*-оперон се състои от регулаторна област и 5 структурни гена (Фигура 7.3). Този оперон съдържа гени за производството на пет протеина, които се използват за производството на три ензима. Структурните гени са :

- *trp-E* и *trp-D* кодират субединиците на антранилат синтетазата, катализираща първите две реакции в триптофановия път;
- *trp-C* кодира индол глицеролфосфат синтетеза, отговорна за катализирането на следващите две стъпки в пътя;
- *trp-B* и *trp-A* кодират субединиците на триптофан синтетазата, катализираща образуването на триптофан от индол-глицерол фосфат и серин.

Регулаторната област съдържа регулаторни ДНК последователности

- промотор, към който се свързва РНК полимеразата;
- оператор, към който се свързва репресора (наричан *trp*-репресор или TrpR) - когато *trp*-репресорът е свързан, РНК полимеразата не може да се свърже с промотора и да започне транскрипция, тъй като операторът се припокрива частично с промотора;
- *trp-L*, последователност от около 130 нуклеотида, която се транскрибира в т. нар. лидерен (водещ) транскрипт; включва четири кратки последователности, обозначени с 1, 2, 3 и 4, всяка от които е частично комплементарна на следващата, така могат да се образуват три алтернативни двуверижни структури (бримки) - 1:2, 2:3 или 3:4. Част от лидерния транскрипт кодира кратък полипептид от 14 аминокиселини, наречен лидерен (водещ) пептид. Този пептид съдържа два съседни *trp* остатъка, което е необичайно, тъй като триптофанът е доста необичайна аминокиселина (около един на сто остатъка в типичния

протеин на *E. coli* е триптофан). Последователност 1 в лидерния транскрипт обхваща областта, кодираща *trp* остатъци в лидерния пептид. Малко след тях се намира стоп кодона на водещия транскрипт.

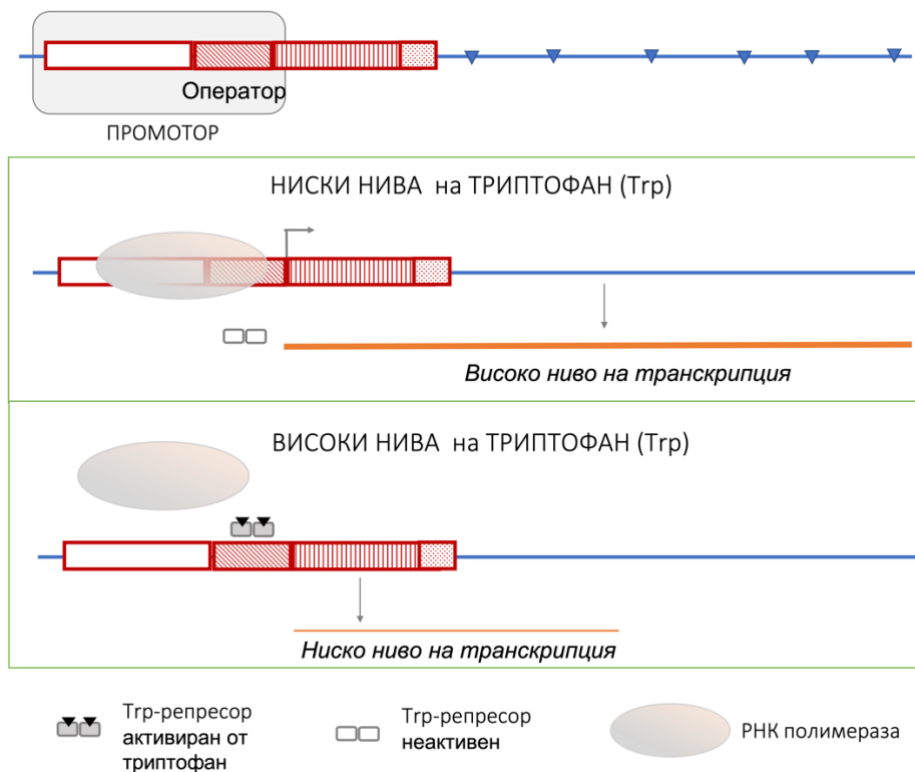


Фигура 7.3. Структура на *trp*-оперон

### 7.1.3.2. Регулация на *trp*-оперон

- Негативен контрол на *trp*-оперон

В тази система обаче, за разлика от *lac*-оперона, генът за *trp*-репресор (*trpR*) не е в съседство с промотора, а по-скоро е разположен в друга част от генома на *E. coli*.



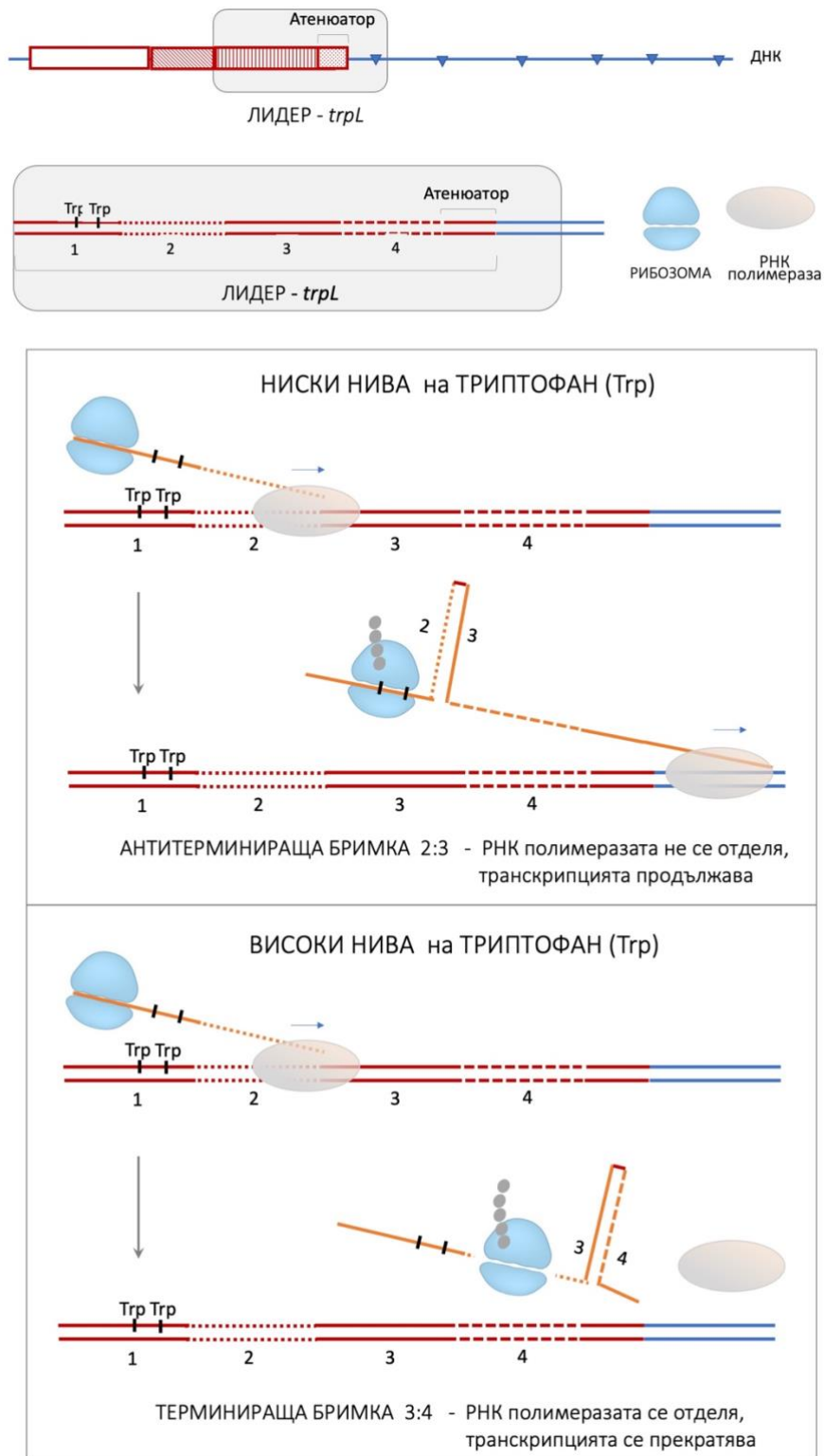
Фигура 7.4. Негативна регулация на експресията на *trp*-оперон

Подобно на *lac*-репресора, *trp*-репресор е конститутивно експресиран на ниско ниво. Синтезираните мономери на *trp*-репресор се свързват в димери. Когато присъства триптофан в клетката (Фигура 7.4), той се свързва с *trp*-репресор, причинявайки промяна в конформацията му, което му позволява да се свърже с оператора. Това пречи на РНК полимеразата да се свърже към промотора и да транскрибира оперона, така че триптофан не се произвежда от неговия предшественик. Когато триптофанът не присъства, репресорът е в неактивна конформация и не може да свърже операторния регион, което прави възможно свързването а РНК полимеразата към промотора и да започне транскрипция.

- Атенюация (или затихване) на транскрипцията от трансляцията

*Trp*-репресор понижава генната експресия като влияе на инициацията на транскрипцията, докато затихването прави това чрез промяна на процеса на транскрипция, когато вече е в ход. Докато *trp*-репресор намалява транскрипцията с коефициент 70, затихването може допълнително да я намали с фактор 10, като по този начин позволява репресия на генната експресия от около 700 пъти. Затихването става възможно, тъй като при прокариотите (които нямат ядро) рибозомите започват да транслират иРНК, докато РНК полимеразата все още транскрибира ДНК. Това позволява на процеса на трансляция да повлияе директно върху транскрипцията на оперона.

Lee и Yanofsky (1977) откриват, че ефективността на затихване е свързана със стабилността на вторичната структура, която придобива лидерния транскрипт, и формираните фуркетни структури – бримки (Фигура 7.5). Типът алтернативна бримка 2:3 или 3:4, който ще бъде формиран в лидерния транскрипт, се определя от процесът на трансляция, извършван върху лидерния транскрипт. По-нататък, в зависимост от това коя бримка се образува, се определя съдбата на транскрипцията през структурните гени на оперона – дали РНК полимеразата ще продължи транскрипцията през структурните гени, или ще се отдели от матрицата преди началото на структурните гени и ще терминира транскрипцията.



Фигура 7.5. Регулация на *trp*-оперон чрез затихване на транслацията от транскрипцията.

Образуване на антитермираща бримка 2:3 при ниски концентрации на триптофан в клетката или на терминираща бримка 3:4 при високи концентрации на триптофан.

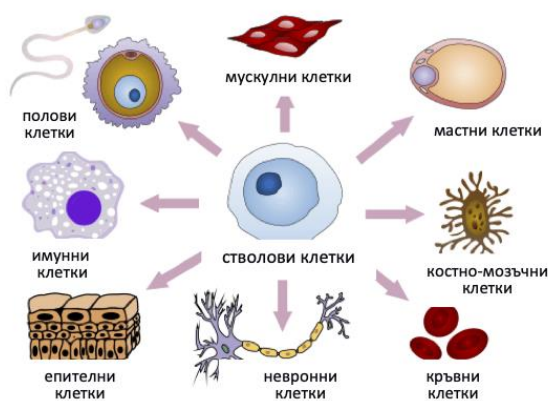
При ниски нива на триптофан в клетката, движението на рибозомата по последователност 1 на лидерния транскрипт се затруднява при двата *trp* кодона. Рибозомата се задържа на това

място и физически покрива последователност 1 на лидерния транскрипт, предотвратявайки образуването на бримка 1:2 (Фигура 7.5). Последователност 2 е свободна да хибридизира с последователност 3 – образува се бримка 2:3, която не позволява образуването на терминараща бримка 3:4, поради което бримка 2:3 се нарича антитерминираща бримка. В присъствието на бримка 2:3, РНК полимеразата продължава да транскрибира структурните гени на оперона.

При високи нива на триптофан в клетката, въпреки, че Trp-репресор възпрепятства иницирането на транскрипция от РНК полимеразата, той не елиминира напълно цялата транскрипционна активност. Пропусклива транскрипция (англ. leaky transcription) може да възникне дори когато репресорът е свързан, което означава, че малко количество РНК полимеразата все още може да се свърже с промотора и понякога да иницира транскрипция. Тази ограничена транскрипция позволява производството на водещата последователност иРНК (Фигура 7.5). Рибозомата ще транслира целия лидерен пептид без забавяне при trp-кодони (тъй като е налице изобилие на заредена tRNA<sup>trp</sup>) и ще спре само при стоп кодона на лидерния пептид за да прекрати транслацията. В този момент рибозомата покрива физически и двете последователности 1 и 2. Следователно последователности 3 и 4 са свободни да образуват терминараща бримка 3:4 (изобилна в G/C и непосредствено последвана от няколко урацилови остатъка), която прекратява транскрипцията през структурните гени на оперона.

При средни концентрации на триптофан, честотата на спиране на рибозомата ще бъде такава, че да поддържа оптималните концентрации на ензимите, кодирани от *trp*-оперон и необходими за синтез на триптофан. Атенюацията е регулаторен механизъм, който е установен освен за *trp*-оперон, и при други биосинтетични оперони на аминокиселини, например хистидин и фенилаланин. Някои бактериални оперони се регулират единствено чрез затихване без взаимодействия репресор-оператор.

## 7.2. Регулация на генната експресия при еукариоти



За разлика от прокариотните организми, многоклетъчните еукариоти имат много по-голям геном, който е организиран в множество хромозоми с много по-висока висока степен на сложност на първичната структура. Всички соматични клетки на многоклетъчния организъм носят една и съща ДНК, но тези клетки образуват тъкани с изключително различни морфология, свойства и функции. Как може да бъде обяснено това огромно разнообразие от клетъчни типове в един организъм? Отговорът на този въпрос се крие в

диференциалната генна експресия – в различните клетъчни типове са „включени“ и „изключени“ различни гени.

Друга важна разлика между еукариоти и прокариоти е организацията на гените. В прокариотите (като бактериите) гените, които трябва да бъдат експресирани заедно, често са организирани в оперони. Това позволява всички гени в оперона да бъдат включени или изключени едновременно, осигурявайки координирана експресия.

Еукариотите гени, които трябва да бъдат експресирани заедно, обикновено не се намират един до друг на същата хромозома. Вместо това тези гени могат да се намират на различни

хромозоми или далеч един от друг на същата хромозома. Те се транскрибират независимо, всеки със собствен промотор и регулаторни елементи. Координираната експресия на тези гени се постига чрез действието на ТФ и други регулаторни протеини, които се свързват със специфични ДНК последователности (като енхансери и промотори) и регулират транскрипцията. Тези ТФ могат да активират или потискат транскрипцията на множество гени едновременно, дори ако са разположени далеч един от друг в генома.

Макар че генната експресия при прокариоти и еукариоти се базира на общи принципи, многоклетъчните еукариоти контролират клетъчната диференциация чрез много по-сложна и прецизна регулация на генната експресия във времето и пространството.

### 7.2.1. Основни цели на контрола на генната експресия при еукариоти

- Клетъчна диференциация и развитие  
 Специализация на клетките: По време на развитието контролът на генната експресия позволява на клетките да се диференцират в различни клетъчни типове, всеки с различни функции. Например, мускулните клетки, невроните и кожните клетки произлизат от една и съща зигота, но експресират различни набори от гени, за да развият уникални структури и функции.  
 Образуване на тъкани и органи: Правилната генна експресия е от съществено значение за образуването на тъкани и органи. Тя гарантира, че специфични гени се активират или изключват в правилния момент и място, ръководейки развитието на сложни телесни структури.  
 Ембрионално развитие: По време на ембриогенезата контролът на генната експресия е решаващ за насочване на етапите на развитие, от формирането на основния план на тялото до растежа и диференциацията на тъкани и органи.
- Отговор на сигналите на околната среда  
 Адаптация към външни стимули: Еукариотните клетки трябва да реагират на различни външни стимули, като промени в температурата, наличността на хранителни вещества, стрес и сигнални молекули. Контролът на генната експресия позволява на клетките да се адаптират, като увеличават или намаляват експресията на гените в отговор на тези условия.  
 Имунен отговор: При многоклетъчните организми контролът на генната експресия е от съществено значение за имунната система. Той гарантира, че имунните клетки експресират подходящите рецептори, сигнални молекули и ефекторни протеини в отговор на патогени или други заплахи.
- Хомеостаза и метаболитна регулация  
 Поддържане на клетъчен баланс: Контролът на генната експресия е от решаващо значение за поддържането на клетъчната хомеостаза, включително регулирането на метаболизма, прогресията на клетъчния цикъл и апоптозата. Той гарантира, че клетъчните процеси са балансирани и че клетките могат да се адаптират към метаболитните нужди и промени.  
 Пестене на енергия: Клетките регулират генната експресия, за да пестят енергия, като експресират гени само когато е необходимо. Например, при липса на глюкоза клетките могат да активират гени, свързани с алтернативни енергийни източници, като метаболизма на липидите.
- Контрол на клетъчния цикъл и растеж  
 Регулация на клетъчното делене: Контролът на генната експресия е от съществено значение за регулирането на клетъчния цикъл и гарантирането, че клетките се делят само

когато е подходящо. Това включва контролираната експресия на циклини, циклин-зависими кинази (CDK) и техните инхибитори.

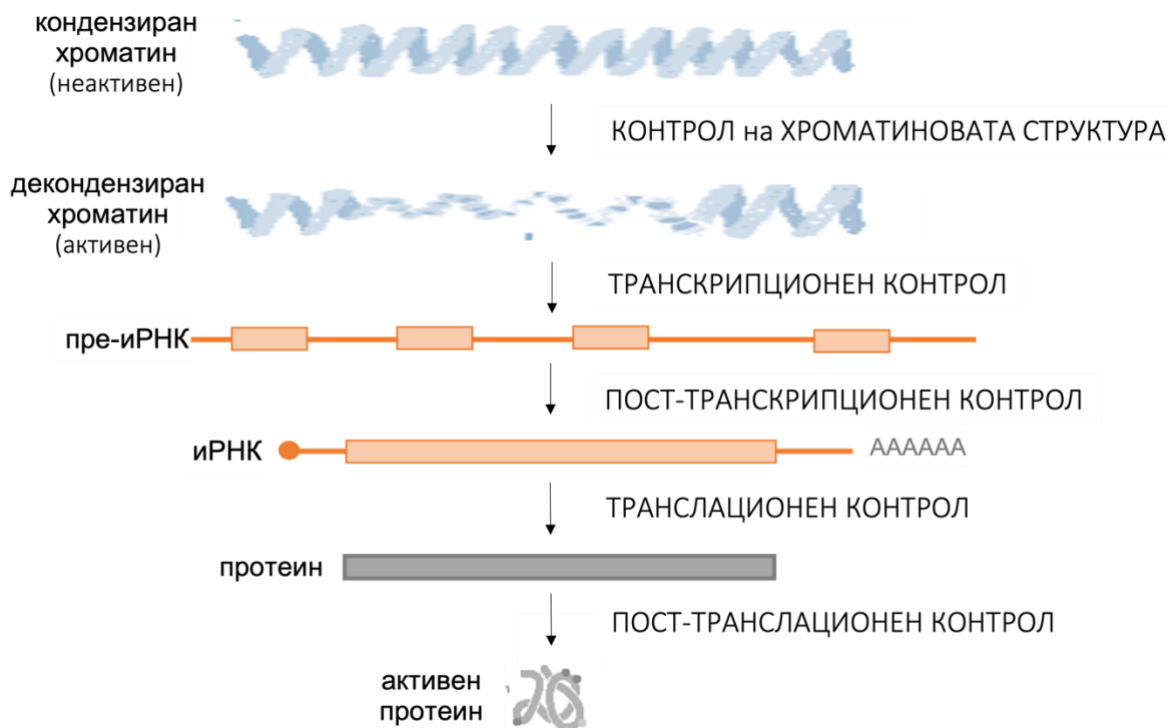
Предотвратяване на неконтролиран растеж: Чрез регулиране на гените, участващи в контролните точки на клетъчния цикъл, поправката на ДНК и апоптозата, клетките могат да предотвратят неконтролирания растеж и развитието на рак.

- Епигенетична памет и клетъчна идентичност  
Поддържане на клетъчната идентичност: Контролът на генната експресия помага за поддържане на идентичността на диференцираните клетки, като осигурява стабилни модели на експресия във времето, дори при делене на клетките. Епигенетичните механизми, като метилирането на ДНК и модификациите на хистоните, играят ключова роля в поддържането на тези експресионни модели.  
Трансгенерационно епигенетично наследяване: Някои модели на генна експресия могат да се наследяват през поколенията без промени в ДНК последователност, което позволява предаването на адаптивни черти.
- Контрол на генната доза и генетична стабилност  
Генна доза: Правилният контрол на генната експресия гарантира, че се поддържа правилната доза генни продукти. Свърхекспресията или недостатъчната експресия на гени може да доведе до нарушения в развитието, метаболитен дисбаланс и заболявания.  
Предотвратяване на генетична нестабилност: Чрез регулиране на експресията на гените, участващи в поправката на ДНК, организацията на хроматина и сегрегацията на хромозомите, клетките поддържат генетична стабилност и предотвратяват мутации или хромозомни аномалии.
- Предаване на сигнали и комуникация  
В многоклетъчните организми клетките си комуникират помежду си чрез сигнални молекули (напр. хормони, цитокини). Контролът на генната експресия е от съществено значение, за да се гарантира, че клетките могат правилно да произвеждат, отговарят и обработват тези сигнали.

### 7.2.2. Основни нива и механизми на регулация на генната експресия при еукариоти

В прокариотната клетка, прокариотната РНК полимераза транскрибира ДНК до иРНК, и още преди синтезата на иРНК да е завършена, рибозомите се прикрепват към растящата иРНК верига и започват да я транслират до белтък. В еукариотната клетка, транскрипцията и транслацията са разделени във времето и пространството. Синтезът на иРНК се извършва в ядрото, първичният транскрипт претърпява процесинг, който включва модификации на 5'- и 3'-краищата, сплайсинг, редактиране и др. Едва тогава зрелият транскрипт преминава в цитоплазмата, където се извършва белтъчната синтеза.

Контролът на генната експресия при еукариотите е сложен процес, който включва множество нива на регулация. Тези нива гарантират, че гените се експресират в правилния момент, в правилния тип клетка и в правилното количество. Принципите на контрол на генната експресия при еукариотите обхващат различни механизми, които регулират експресията на гените на различни етапи, от ДНК до функционалните протеини (Фигура 7.6).



Фигура 7.6 . Основни нива на регулация на генната експресия при еукариоти.

- Контрол на структурата и достъпността на хроматина

Най-важната структурна разлика между прокариотната и еукариотната ДНК е свързването на еукариотната ДНК с белтъци (хистонови и нехистонови) и образуването на хроматин в еукариотното ядро. Хроматинът обуславя напълно различно „изходно ниво“ на транскрипция в еукариотната клетка, в сравнение с прокариотната клетка.

При прокариоти, РНК полимеразата има достъп до промоторните области в ДНК веригата, поради което „изходното състояние“ на транскрипция е нерестриktivно или „включено“. Инициацията на транскрипцията се потиска, когато бъде блокирано свързването на РНК полимеразата от репресорни регулаторни белтъци.

При еукариоти, плътното пакетиране на ДНК в хроматин ограничава достъпа на РНК полимеразата до промоторните области, в следствие на което „изходното състояние“ на транскрипция е рестриktivно или „изключено“. Еукариотните РНК полимерази могат да започнат транскрипция само с помощта на белтъци, наречени ТФ. Но дори когато ТФ присъстват в клетката, транскрипция не винаги може да бъде осъществена, защото много често те нямат достъп до прицелните регулаторни елементи в ДНК. Затова необходимо условие за извършване на транскрипцията е преминаването на хроматина в състояние на по-слабо пакетиране. Основните механизми за контрол на състоянието на хроматина включват:

- Метилиране на ДНК: Добавянето на метилови групи към цитозиновите бази в ДНК обикновено води до репресия на гените. Моделите на метилиране могат да бъдат стабилно наследявани при клетъчно делене.
- Модификации на хистоните: Хистоните, около които е обвита ДНК, могат да претърпят различни пост-транслационни модификации (напр. ацетилиране, метилиране, фосфорилиране). Тези модификации могат или да отворят хроматина, правейки ДНК по-

достъпна за транскрипция (еухроматин), или да го кондензират, за да потискат транскрипцията (хетерохроматин). Например, ацетилирането на хистоните обикновено разхлабва хроматина и насърчава генната експресия, докато метилирането може или да активира, или да потисне генната експресия, в зависимост от специфичните хистонови остатъците, участващи в процеса.

- Ремоделиране на хроматина: Комплекси за ремоделиране на хроматина могат да преместват нуклеозоми, за да откриват или скриват ДНК региони от транскрипционния механизъм, влияейки на генната експресия.

Хроматиновата структура е отговорна за различни нива на сложност на генна регулация. Тя позволява едновременно регулиране на структурно или функционално свързани гени, които могат да бъдат пространствено отдалечени в еукариотната ДНК. Взаимодействието на хроматина с активиращи или репресиращи транскрипционни фактори води до формиране на „отворени“ домени (хлабаво пакетирани хроматин) или „затворени“ домени (плътно пакетирани хроматин), с различни размери и стабилност. Тези вариации позволяват извършването на феномени, наблюдавани единствено в еукариотите, като генно активиране на различни етапи от развитието и епигенетична памет при клетъчното делене. Те са отговорни също така за поддържането на диференцирано клетъчно състояние, което е от съществено значение за преживяване на многоклетъчния организъм. Епигенетичната регулация включва наследствени промени в генната експресия, без да се променя ДНК последователността.

- Транскрипционен контрол

Регулацията на генната експресия на нивото на транскрипцията е основна точка на контрол при еукариотите и включва множество компоненти и механизми.

- Транскрипционни фактори: Протеини, които се свързват с регулаторни ДНК елементи за иницират или модулират транскрипцията на гените. Главните ТФ са необходими за инициация на транскрипцията и свързването на РНК полимеразите в промоторните области. Другите ТФ могат да взаимодействат най-вече с РНК полимеразата II за да усилват (активатори) или потискат (репресори) инициацията на транскрипцията.
- Регулаторни ДНК елементи: Промотори, усилвателите или енхансери (от англ. enhance – усилвам), заглушителите или сайлънсери (от англ. silencer – заглушител) и други ДНК последователности се разпознават и свързват от ТФ. Промоторите са задължителни регулаторни елементи за всеки ген, към които се свързват главните ТФ за да иницират транскрипцията. Усилвателите и заглушителите са дистални регулаторни елементи, които могат да съответно да увеличат или потискат транскрипцията на целевите гени, когато са свързани със специфични ТФ. Те могат да функционират на големи разстояния и често участват в тъканно-специфичната генна експресия.
- Медиаторен комплекс: Мултипротеинов комплекс, който служи като мост между ТФ и РНК полимеразата II, помагайки за сглобяването на транскрипционния апарат в промоторите на гени.
- Регулаторни мрежи: Генната експресия при еукариотите често се контролира от сложни мрежи от ТФ и други регулаторни протеини, които взаимодействат с множество гени и помежду си, за да координират отговори на развиващи се сигнали и промени в околната среда.

Имайки предвид рестриктивното „изходно състояние“ на транскрипция, позитивната регулация чрез ТФ (активатори и ко-активатори) е доминиращата форма на контрол при еукариоти. Свързването на един тип транскрипционен фактор може да повлияе свързването на други ТФ. Генната експресия при еукариоти е силно вариабилна в зависимост от типа на

участващите активатори и сигналните молекули, които присъстват за да контролират свързването на активаторите.

- Пост-транскрипционен контрол

След транскрипция генната експресия може да се регулира на нивото на РНК чрез няколко механизма.

- Алтернативен сплайсинг: Предшествениците на иРНК могат да бъде обработвани чрез сплайсинг по различни начини, в резултат на което от един ген могат да се образуват множество варианти на иРНК и различни протеини. Това увеличава протеиновото разнообразие и позволява тъканно-специфична генна експресия.
- Редактиране на РНК: Пост-транскрипционни модификации на иРНК, като вмъкване, изтриване или заместване на нуклеотиди, могат да променят иРНК последователността и следователно протеиновия продукт.
- Стабилност и разграждане на иРНК: Стабилността на иРНК молекулите може да бъде регулирана от последователности в иРНК (като AU-богати елементи в 3'-нетранслирания регион) и от РНК-свързващи протеини и микроРНК, които насочват иРНК за разграждане или стабилизиране. Това влияе на времето, през което иРНК е налична за трансляция.
- Експорт на иРНК: Правилно обработената иРНК се изнасят селективно от ядрото в цитоплазмата, където могат да бъдат транслирани в протеини.

- Транслационен контрол

Генната експресия може да се регулира на нивото на протеинов синтез, като се контролира кога и колко ефективно иРНК се транслира в протеин.

- Инициация на трансляцията: Фазата на инициация на трансляцията често е стъпката, ограничаваща скоростта и основна точка на контрол. Регулаторни протеини могат да се свързват с 5'-нетранслируемия участък на иРНК или да взаимодействат с факторите на инициация на трансляцията, за да усилят или инхибират свързването с рибозомите.
- Локализация на РНК: Някои иРНК се транспортират до специфични места в клетката, преди да бъдат транслирани. Тази пространствена регулация позволява локализирана протеинова синтеза, която е от съществено значение в процеси като клетъчна поляризираност и ембрионално развитие.
- Малки регулаторни РНК молекули: Малки некодиращи РНК, от които най-добре проучени са микроРНК молекулите с дължина 21-23 нт, могат да се свързват с комплементарни последователности в целеви иРНК, водейки до деградация на иРНК или инхибиране на трансляцията, като по този начин фино регулират генната експресия след транскрипция.

- Пост-транслационен контрол

След като един протеин е синтезиран, неговата активност, стабилност и локализация могат да бъдат допълнително регулирани.

- Пост-транслационни модификации: Протеините могат да претърпят различни химични модификации, като фосфорилиране, ацетилиране, убиквитиниране и гликозилиране. Тези модификации могат да променят протеиновата активност, стабилност, взаимодействия и клетъчна локализация.
- Протеиново нагъване и контрол на качеството: Новосинтезираните протеини трябва да бъдат правилно нагънати, за да бъдат функционални.
- Разграждане на протеини: Протеините могат да бъдат насочени за разграждане от убиквитин-протеазомния път или чрез автофагия. Тази регулация гарантира, че повредените или ненужни протеини се отстраняват от клетката, позволявайки стриктен контрол на нивата на протеините.

### 7.2.3. Координирано-регулираните гени не са свързани при еукариоти

При висшите еукариоти, за разлика от прокариотните организми, организираност на гени в оперони не съществува. По-скоро, всеки ген се транскрибира в индивидуална моноцистронна иРНК, кодираща единичен полипептид. Тъй като гените, чиито белтъчни продукти са необходими едновременно, не са свързани, а са представени в генома самостоятелно, тяхната експресия трябва да бъде координирана. Нещо повече, такива координирано експресирани гени могат дори да не са разположени в близост в генома, а да бъдат представени върху различни хромозоми в еукариотното ядро.

Координираната експресия на еукариотни гени се отнася до регулаторните механизми, които позволяват на множество гени да се експресират заедно, по координиран начин, често в отговор на определени външни сигнали, етапи на развитие или специфични клетъчни процеси. Тази координация е от съществено значение за правилното функциониране на клетките, тъй като осигурява експресията на свързани гени (като тези, участващи в общи биохимични пътища или клетъчни процеси) в точното време, в правилния клетъчен тип и в подходящи количества.

Сложният процес на координирана експресия се контролира чрез взаимодействията на ТФ с регулаторни ДНК последователности (регулаторни елементи), структурата на хроматина, епигенетични модификации, сигнални пътища, както и различни пост-транскрипционни и пост-транслационни механизми.

- Споделени регулаторни елементи

Много еукариотни гени, които трябва да се експресират заедно, имат сходни регулаторни елементи в своите промоторни или енхансерни региони, към които се свързват специфични ТФ, за да активират или потиснат транскрипцията, например, на гените, които участват в един и същ метаболитен път.

Някои гени съдържат регулаторни елементи на отговора, кратки ДНК последователности, които реагират на определени сигнали (като хормони, стрес и др.). Гените, които споделят едни и същи елементи на отговора, могат да бъдат активирани едновременно от един и същи ТФ. Например, гените, които отговарят на топлинен стрес, може да имат елементи на топлинен шок, разпознавани от ТФ за топлинен шок.

- Транскрипционни фактори

Един транскрипционен фактор може да регулира експресията на множество гени, осигурявайки координирана експресия. Например, транскрипционният фактор p53 активира множество гени, свързани с отговор на увреждане на ДНК. Ярък пример са т. нар. главни регулатори (от англ. master regulators) - ТФ, които могат да контролират цели програми на развитие, като активират групи от гени, необходими за формирането на тъкани или органи. Например, транскрипционният фактор MyoD участва в развитието на мускулите и активира множество гени, свързани с мускулната диференциация.

- Епигенетична регулация

Епигенетични механизми, като метилиране на ДНК, модификации на хистоните и ремоделиране на хроматина, играят важна роля в координацията на генната експресия. Тези модификации могат да направят определени региони на генома по-достъпни или по-малко достъпни за транскрипционния апарат. Гените, разположени в едни и същи топологично асоциирани домени или в региони на активен хроматин, е по-вероятно да бъдат експресирани заедно. Промени в структурата на хроматина, като формирането на еухроматин (активен, декондензиран хроматин) или хетерохроматин (неактивен, кондензиран хроматин), могат да активират или заглушат групи от гени.

- Сигнални пътища

Клетките отговарят на външни стимули (като хормони, растежни фактори или стрес) чрез сигнални пътища. Тези пътища могат да доведат до активиране или репресия на ТФ, които регулират множество гени едновременно. Например, MAPK/ERK пътят активира гени, участващи в клетъчния растеж и делене.

Вътреклетъчните сигнални молекули, като cAMP, също могат да координират генната експресия, като активират специфични ТФ. Например, CREB (от англ. cAMP response element-binding protein) се активира от cAMP и се свързва с CRE елементи в промоторите на целеви гени, което води до тяхната координирана експресия.

- Генни клъстери

В някои случаи гените, които трябва да се експресират заедно, са физически групирани на една и съща хромозома. Това е особено често при гени, участващи в сходни биохимични пътища, като например Нох генните клъстери, отговорни за моделирането на тялото при животните.

- Некодиращи РНК

Регулаторни некодиращи РНК, като микроРНК и дълги некодиращи РНК, могат да контролират експресията на множество целеви гени. Така напр. една микроРНК може да свърже с различни иРНК и да инхибира транслацията или да насърчи тяхната деградация.

- Пост-транслационни модификации

Модификациите на протеини, като фосфорилиране, убиквитиниране и ацетилиране, позволяват бързото активиране или инактивиране на групи от протеини, участващи в един и същ път. Това може да работи в съчетание с генната експресия за фина настройка на клетъчните отговори.

- Обратни връзки

Много еукариотни генни мрежи включват обратни връзки, при които продуктите на определени гени регулират собствената си експресия или експресията на други гени. Например, в контрола на циркадния ритъм, CLOCK и BMAL1 активират гени, които произвеждат протеини (като PER и CRY), които от своя страна инхибират собствена си експресия.

## ГЛАВА 8 ГЕНОМНА ОРГАНИЗАЦИЯ

Геномът представлява пълният набор от генетичен материал (ДНК или РНК) в даден организъм. Той съдържа цялата информация, необходима за растеж, развитие, функциониране и възпроизводство. Геномите се различават значително по размер, структура и организация сред различните организми – от вируси и бактерии до растения и животни

### 8.1. Прокариотен геном (бактерии и археи)

- Основни характеристики на прокариотния геном:
  - обикновено представляват единична, кръгова ДНК молекула, разположена в нуклеоид;
  - имат малък размер;
  - имат висока гъстота на гените - повечето ДНК кодира белтъци, с много малко некодиращи региони;
  - някои бактерии съдържат плазмиди – малки, кръгови ДНК молекули, които носят допълнителни гени (например за антибиотична резистентност);
  - гените често са организирани в оперони, което позволява координирано генетично регулиране.

Прокариотният геном обикновено се състои от една кръгова хромозома, която носи генетичната информация на организма. Тази хромозома е свързана с хистоноподобни протеини, които подпомагат опаковането на ДНК, въпреки че истински хистони липсват при бактериите (с изключение на някои археи). Генетичният материал е локализиран в област, наречена нуклеоид, която не е оградена от мембрана, което я отличава от ядрото на еукариотните клетки.

Въпреки че най-често срещаната структура на прокариотния геном е единична кръгова хромозома, съществуват изключения. Някои бактериални видове имат множество хромозоми. Например, *Rhodobacter sphaeroides* има две хромозоми, а *Burkholderia cepacia* притежава три хромозоми. Някои прокариоти имат линейни хромозоми. Един забележителен пример е *Borrelia burgdorferi* – спирохета, причинител на лаймска болест. Грам-положителната бактерия *Streptomyces coelicolor*, известна с производството на антибиотици, също има линеен геном.

Прокариотните геноми обикновено са по-малки и по-компактни в сравнение с еукариотните. Техният размер най-често варира между 1 и 10 милиона бд, като съдържат приблизително 1 500 до 5 000 гена. Броят на гените при прокариотите е в силна корелация с размера на генома, тъй като те имат висока генна плътност и минимално количество некодираща ДНК. Например геномът на *Escherichia coli* (E. coli K-12 щам) е с дължина около 4,6 милиона бд и съдържа приблизително 4 300 протеин-кодиращи гена. Геномът на *Mycoplasma genitalium*, една от бактериите с най-малък известен геном, съдържа само 580 000 бд и кодира около 470 гена. За разлика от тях, по-големите бактериални геноми съществуват, като този на *Streptomyces coelicolor*, който има около 8,7 милиона бд и над 7 800 гена, поради сложния си метаболизъм. Някои почвени бактерии и цианобактерии могат да имат геноми, надвишаващи 10 милиона бд, доближавайки се по размер до малките еукариотни геноми.

Освен основната хромозома, много прокариоти съдържат плазмиди – малки, кръгови ДНК молекули, които се реплицират независимо от хромозомната ДНК. Плазмидите често носят

гени, които осигуряват адаптивни предимства, като резистентност към антибиотици, производство на токсини или метаболитни способности, които подпомагат оцеляването в различни среди.

Тази широка вариация в размерите и сложността на геномите отразява разнообразния начин на живот на прокариотните – от свободноживеещи организми до вътреклетъчни симбионти и патогени.

Прокариотните геноми еволюират бързо поради комбинация от високи мутационни нива, хоризонтален генен трансфер и силен селекционен натиск от околната среда. Тези механизми допринасят за тяхната изключителна адаптивност и еволюционен успех в различни екологични ниши.

Основни фактори, задвижващи еволюцията на прокариотния геном:

- Високи мутационни нива – прокариотите показват сравнително високи спонтанни мутационни нива в сравнение с еукариотите. Въпреки че тяхната репликационна система има механизми за корекция на грешки, мутациите все пак се натрупват, създавайки генетично разнообразие, което може да бъде изгодно, неутрално или вредно. При стресови условия някои бактерии увеличават мутационните си нива чрез използване на по-малко точни ДНК полимерази, което води до повишено генетично разнообразие (т.нар. мутагенни фенотипове).
- Хоризонтален генен трансфер - за разлика от вертикалния генен трансфер (от родител към потомство), този тип генен трансфер позволява на бактериите и археите да придобиват генетичен материал от неродствени организми, което значително ускорява тяхната еволюция. Трите основни механизма на хоризонтален генен трансфер включват: трансформация (приемане на свободна ДНК от околната среда - например *Streptococcus pneumoniae* може да придобие гени за антибиотична резистентност), трансдукция (пренос на ДНК чрез бактериофаги) и конюгация (директен трансфер на генетичен материал между бактерии чрез пили - например обмен на плазмиди при *E. coli*). Той е отговорен за разпространението на антибиотична резистентност, вирулентни фактори и нови метаболитни способности сред бактериалните популации.
- Адаптивност към околната среда - прокариотните геноми са силно динамични и могат да претърпяват придобиване на гени, загуба на гени и геномни реорганизации в отговор на екологичния натиск.

Еволюцията на прокариотния геном е бърза и динамична, което позволява на бактериите и археите да оцелеят, да се разпространяват и да се развиват в най-разнообразни екологични условия. Тяхната способност бързо да придобиват и модифицират генетична информация ги прави едни от най-адаптивните организми на Земята.

## 8.2. Еукариотен геном

- Основни характеристики на еукариотния геном:
  - ДНК е линейна и е опакована в хромозоми, намиращи се в ядрото; допълнителен генетичен материал се съдържа в органели като митохондрии (и хлоропласти при растенията);
  - размерът е много по-голям в сравнение с прокариоти;
  - ДНК е обвита около хистонови протеини, образувайки хроматин, който допълнително се компактизира в хромозоми;
  - имат по-сложна организация, включваща уникални и повтарящи се последователности;

- гъстотата на гените е по-ниска от тази при прокариотите, с голямо количество некодираща ДНК (включително интрони, регулаторни последователности и повтарящи се елементи);
- гените имат прекъснатата структура - съдържат екзони и интрони, които се изрязват при обработка на иРНК;
- съдържат разнообразие от регулаторни елементи - промотори, енхансъри и сайлънсъри, които контролират генната експресия;

Еукариотните геноми се състоят от множество линейни хромозоми, разположени в мембранно-обособено ядро, което ги отличава от прокариотните геноми, които обикновено са кръгови и нямат ядрен компартмент. Размерите на еукариотните геноми варират значително – от приблизително 10 милиона бд (Mb) при прости едноклетъчни еукариоти като дрождите до над 150 милиарда бд (Gb) при някои растения.

Размерът на генома не винаги корелира с биологичната сложност – явление, известно като С-стойностен парадокс. По-големият геном не означава непременно по-сложен организъм. Например човекът има около 3,2 милиарда бд, докато някои растения (напр. *Paris japonica*) имат около 150 милиарда бд, но не са по-сложни.

Еукариотните геноми съдържат много повече гени в сравнение с прокариотните, като броят им обикновено варира от няколко хиляди до над 50 000 при сложните многоклетъчни организми. Въпреки това, протеин-кодиращите гени съставляват само малка част от генома – например, при хората те заемат по-малко от 2%. Останалата част включва регулаторни елементи, интрони и повтарящи се ДНК последователности.

Броят на протеин-кодиращите гени не обяснява напълно сложността на организма. Човекът има около 20 000–25 000 гена, което е сравнимо с броя на гените в кръглия червей (*Caenorhabditis elegans*), но хората са значително по-сложни. Комплексността възниква чрез алтернативен сплайсинг, посттранскрипционни модификации и регулация на гените, а не само от броя на гените.

- Уникални и повтарящи се последователности на еукариотния геном

Еукариотният геном има много по-голяма сложност от бактериалния геном. Характерна особеност на еукариотните геноми е наличието както на уникални, така и на повтарящи се последователности, като и двата типа играят важни роли във функционирането, регулацията и еволюцията на генома.

Уникалните последователности се състоят основно от протеин-кодиращи гени, регулаторни елементи и некодиращи РНК. Тези гени съставляват само малка част от генома – например при хората около 1-2% от общата ДНК. Регулаторните региони контролират експресията на гените, докато некодиращите РНК, като микроРНК и дълги некодиращи РНК, играят важна роля в генната регулация. Псевдогените, които са останки от някога функционални гени, също присъстват в генома, но не произвеждат протеини.

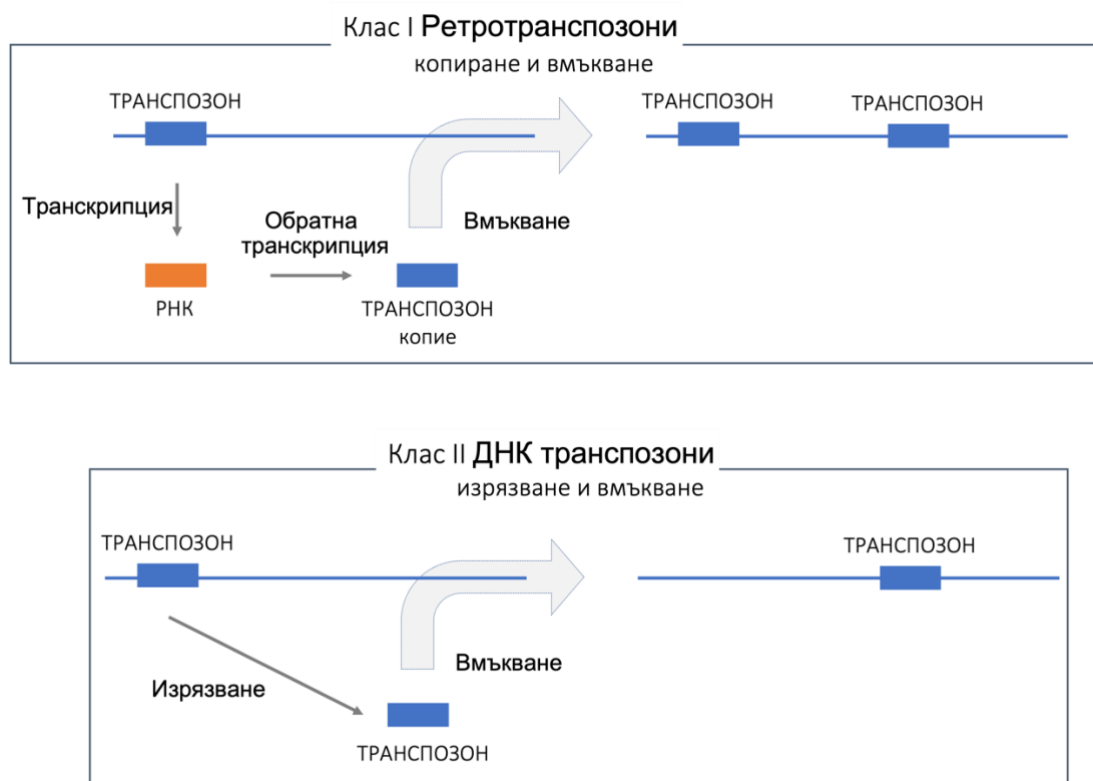
Повтарящите се последователности заемат значителна част от еукариотните геноми и могат да бъдат класифицирани като тандемни повтори и разпръснати повтори. Тандемните повтори се състоят от идентични или сходни последователности, подредени в последователен ред. Сателитната ДНК е един от тези примери и се намира главно в центромерите и хетерохроматина, като играе роля в стабилността на хромозомите. Минисателитите, известни още като повтори с променлив брой повторения (от англ. Variable Number Tandem Repeat, VNTRs), имат дължина на повторите между 10 и 60 нт и се срещат в теломерите и други геномни региони. Микросателитите, или късите тандемни повтори, се състоят от по-малки

повтори (1–6 нт) и са силно полиморфни, което ги прави полезни за генетично картографиране и съдебна експертиза.

Разпръснатите повтори, от друга страна, се срещат из целия геном и се състоят основно от подвижни елементи - транспозони. Транспозоните, известни още като скачащи гени, могат да се придвижват от едно място на друго в генома. Те са открити за първи път от Барбара МакКлинток в царевица през 40-те години на XX век, за което тя получава Нобелова награда през 1983 г.

Съществуват два основни типа транспозони според механизма им на придвижване (Фигура 8.1):

- Клас I Ретротранспозони: Придвижват се чрез образуване на РНК посредник чрез механизма „копирай и постави“ - транспозонът се транскрибира в РНК, след което обратна транскриптаза преобразува РНК обратно в ДНК. Новото ДНК копие се вмъква в ново място в генома, докато оригиналният транспозон остава на мястото си. Това води до увеличаване на броя на транспозоните с времето. Срещат се главно в еукариоти. В човешкия геном такива се дължат разпръснати ядрени елементи (от англ. Long Interspersed Nuclear Elements LINEs) и късите разпръснати ядрени елементи (от англ. Short Interspersed Nuclear Elements, SINEs).
- Клас II ДНК транспозони: Преместват се директно като ДНК в рамките на генома чрез механизма „изрежи и постави“ - транспозонът физически напуска първоначалното си място и се интегрира на друго място в генома. Те използват ензим, наречен транспозаза, който изрязва транспозона от едно място и го вмъква на друго.



Фигура 8.1. Основни класове транспозони

В бактериите транспозоните често носят гени за антибиотична резистентност и могат да ги разпространяват между различни видове чрез хоризонтален генен трансфер. Например транспозонът Tn3 носи ген за  $\beta$ -лактамаза, който осигурява резистентност към пеницилини.

Около 50% от човешкия геном е съставен от последователности, произлизащи от транспозони, повечето от които са неактивни (мутирани и неспособни да се движат). Въпреки това, някои активни транспозони, като LINE-1 (L1 елементи), все още предизвикват мутации и допринасят за генетичното разнообразие.

Повтарящите се последователности изпълняват различни биологични функции. Сателитната ДНК участва в разделянето на хромозомите, докато теломерните повтори защитават краищата на хромозомите. Транспозоните движат еволюцията на генома, като създават мутации, дублират гени и променят генната регулация. Освен това микросателитите и минисателитите допринасят за генетичното разнообразие. Въпреки това, повтарящите се последователности могат да бъдат свързани и с генетични заболявания, както се наблюдава при болести като болестта на Хънтингтън и синдрома на чупливата X-хромозома, които се причиняват от разширения на повтори.

Балансът между уникалните и повтарящите се последователности определя сложността на еукариотните геноми. Докато уникалните последователности съдържат функционалния план на организма, повтарящите се елементи влияят върху генетичното разнообразие, структурната цялост и еволюционните процеси. Еукариотните геноми обикновено еволюират по-бавно от прокариотните геноми поради по-големия си размер и сложност. Те преминават през процеси като дубликация на гени, което може да доведе до нови функции и увеличена сложност с течение на времето.

- Органелни геноми

Еукариотните организми имат допълнителни геноми в органели, като митохондрии (митохондриален геном) и, при растенията, хлоропласти (хлоропластен геном). Тези органелни геноми са кръгли и наподобяват прокариотните геноми.

Митохондриалният геном се отнася до генетичния материал, намиращ се в митохондриите – енергийните органели на еукариотните клетки. Митохондриалните геноми са много по-малки от ядрените, обикновено вариращи между 15 и 20 хиляди бд (kb) при животните и достигащи няколко хиляди бд при растенията и гъбите. Всяка митохондрия съдържа множество геномни копия. За разлика от ядрената ДНК, митохондриалната ДНК е силно компактна, с малко некодиращи или регулаторни последователности.

Митохондриалните геноми кодират ограничен брой гени, включително тези за митохондриалните рРНК, тРНК и протеини, участващи в окислително фосфорилиране – процесът, чрез който митохондриите произвеждат АТФ. Другите протеини, които се намират в митохондриите (около 1000), се кодират от ядрени гени, синтезират се в цитоплазмата върху свободни рибозоми, след което се импортират в митохондриите. Протеините, кодирани от митохондриалния геном, се синтезират вътре в митохондриите чрез транслационна система, която има сходства с бактериалните рибозоми, което отразява бактериалния произход на тези органели.

За разлика от ядрената ДНК, митохондриалната ДНК има сравнително висока мутационна честота поради излагането ѝ на реактивни кислородни видове, генерирани по време на дишането. Тази висока мутационна честота, заедно с липсата на рекомбинация, прави митохондриалните геноми полезни в популационната генетика.

При човек, митохондриалният геном кодира 22 тРНК, 2 рРНК и 13 митохондриални протеини, участващи в окислителното фосфорилиране, което е отговорно за производството на АТФ (АТФ). Протеин-кодиращите гени кодират субединици на NADH дехидрогеназа (ND1–ND6, ND4L), цитохром с оксидаза (COX1–COX3), цитохром b (CYB) и АТФ синтаза (ATP6 и ATP8). За

разлика от ядрената ДНК, която се наследява от двамата родители, митохондриалната ДНК се предава почти изцяло по майчина линия.

Митохондриалната ДНК може да натрупва мутации с времето, което може да доведе до митохондриални заболявания, засягащи енергийния метаболизъм. Тези заболявания могат да имат различни ефекти върху различни органи системи, особено тези с високи енергийни нужди, като нервната система и мускулите, както и върху стареенето, което подчертава нейната критична роля за клетъчната функция и човешкото здраве.

Хлоропластният геном е около 120–160 000 бд и съдържа около 100-120 гена. Той е кръгова молекула и често съдържа две обрънати повторения (IR), които дублират определени гени, едно голямо единично копие (LSC) и едно малко единично копие (SSC).

Хлоропластният геном кодира тРНК, рРНК и хлоропластни протеини, участващи в светлинните реакции на фотосинтезата и в цикъла на Калвин. Въпреки че хлоропластите имат собствен геном, те разчитат за пълноценно функциониране на много протеини, кодирани от ядрената ДНК.

#### ПРОЕКТЪТ за ЧОВЕШКИЯ ГЕНОМ (HGP)

Проектът за човешкия геном е международна изследователска инициатива, насочена към картографиране и секвениране на целия човешки геном, предоставяйки пълна референтна последователност на човешката ДНК. Официално стартирал през 1990 г. и завършен през 2003 г., той се превърща в едно от най-амбициозните и трансформиращи научни начинания в историята.

##### √ Основни открития и постижения:

- Определяне на пълната ДНК последователност на човешкия геном (3,2 милиарда базови двойки).
- Идентифициране и картографиране на всички гени в генома (около 20 000–21 000).
- Разкриване на важността на некодиращата ДНК, която играе ключова роля в регулацията на гените, структурата на хромозомите и клетъчните процеси.
- Анализирани генетичните вариации и тяхното въздействие върху човешкото здраве и болести. Открити бяха единични нуклеотидни полиморфизми (SNPs) и други генетични вариации, което подобри разбирането ни за човешкото разнообразие и предразположеността към заболявания
- Разработване на напреднали биоинформатични инструменти за съхранение и анализ на генетични данни.
- Подобряване на технологиите за секвениране на ДНК и геномен анализ.
- Изследване на етичните, правните и социалните последици от геномните изследвания.

##### √ Проектът за човешкия геном революционизира биомедицинските изследвания, като доведе до пробиви в:

- Изследването на генетични заболявания - чрез идентифициране на мутации, причиняващи заболявания, учените разработиха по-добри диагностични инструменти и лечения за заболявания като кистозна фиброза, болестта на Хънтингтън и рак.
- Фармакогеномика - разбирането на генетичните вариации помага на лекарите да предписват персонализирани лекарства според генетичния профил на пациента.
- Ракова геномика - проектът разкри как генетичните мутации водят до рак, което доведе до по-прецизни и ефективни лечения.
- Проучвания върху произхода и еволюцията - данните за генома предоставиха информация за произхода на човека, миграционните модели и еволюционната история.

## ПРИЛОЖЕНИЯ

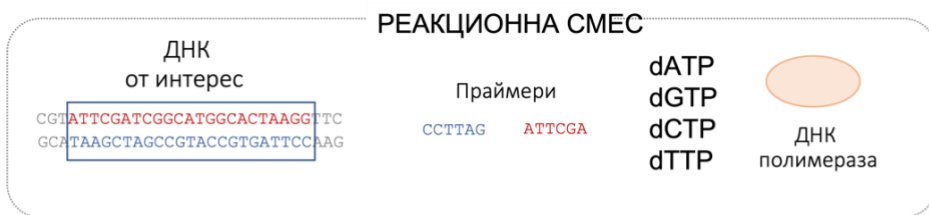
### ПРИЛОЖЕНИЕ 1 ПОЛИМЕРАЗНА ВЕРИЖНА РЕАКЦИЯ

Полимеразна верижна реакция (от англ. Polymerase Chain Reaction, PCR) е лабораторна техника за създаване *in vitro* на милиони копия на определен участък от ДНК - процесът се нарича амплификация. Първоначално е разработена през 1983 г. от американския биохимик Кари Мълис, за което той е удостоен с Нобелова награда за химия през 1993 г. PCR позволява на изследователя да произвежда голямо количество копия на ДНК, която е от интерес за него (напр. ген), и да го използва за различни експерименти.

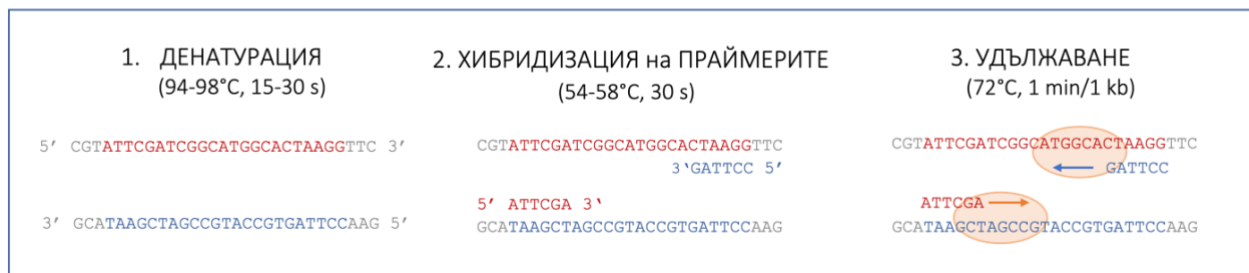
#### PCR реакционна смес

За създаването на огромен брой копия на ДНК чрез PCR са необходими следните компоненти:

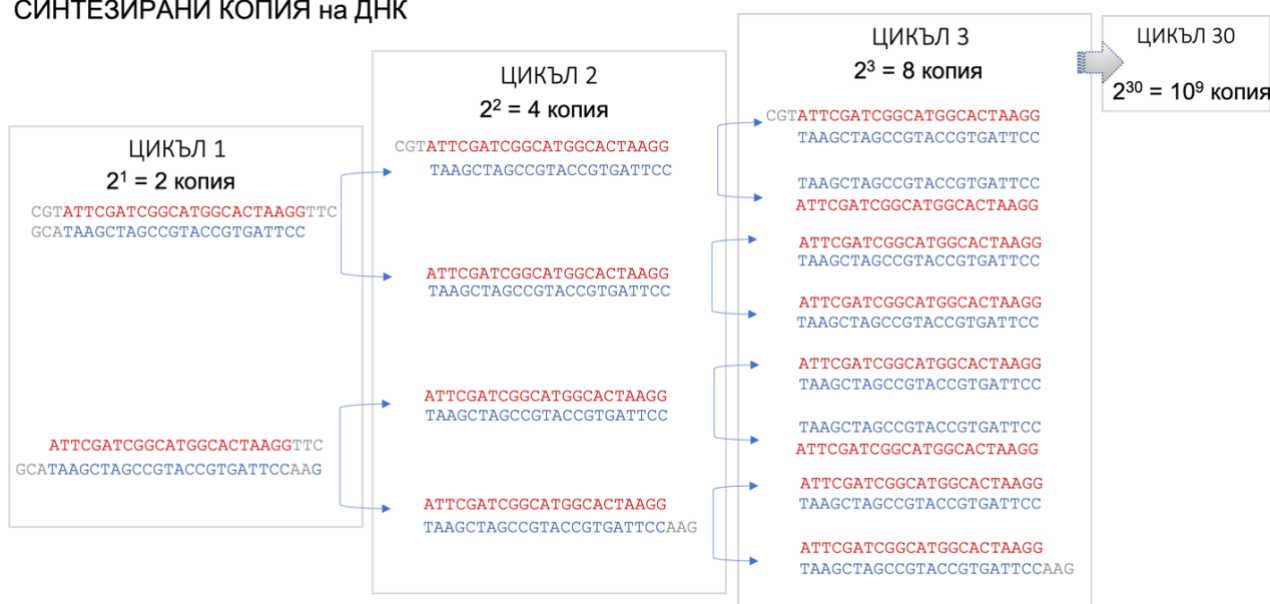
- ДНК матрица – обикновено е ДНК, която съдържа участъка, който трябва да бъде амплифициран, т. нар. ДНК от интерес или целеви участък. ДНК матрицата може да бъде получена от различни източници, напр. плазмидна ДНК, геномна ДНК или дори малко количество тъкан. Матричната ДНК обикновено присъства в много ниски концентрации в PCR реакция, 1 pg - 1 ng плазмидни или вирусни матрици, 1 ng - 1 µg геномни матрици.
- Праймери – къси олигонуклеотиди с дължина 15-25 нд (синтетична къса ДНК верига, произведена от специализиран апарат – ДНК синтезатор). PCR амплификацията изисква два праймера, които са комплементарни на последователностите в двата края на целевия участък. Те хибридизират към едно-вещната ДНК, осигурявайки отправна точка за синтеза на ДНК.
- Нуклеотиди - градивните елементи за създаване на молекулите на ДНК. При PCR реакции се добавя смес от четири типа нуклеотид трифосфати (dATP, dCTP, dGTP, dTTP; известни като dNTPs). ДНК полимеразата включва тези нуклеотиди в растящата ДНК верига, като ги добавя към 3'-края на праймера според правилата за сдвояване на базите (аденин с тимин А:Т и цитозин с гуанин С:G).
- ДНК полимераза - ДНК полимеразата, използвана в PCR, трябва да бъде термоустойчива, за да издържи на високите температури, необходими за денатурация (приблизително 95°C). Един от най-често използваните ензими е *Taq* полимеразата, която се пречиства от *Thermus aquaticus*, бактерия, която вирее в горещи извори. Този ензим функционира оптимално при около 72°C, което е температурата, използвана за удължаване по време на PCR.
- PCR буфер - той играе решаваща роля за поддържане на оптимална среда за реакцията. Той помага за стабилизиране на рН по време на циклиращия процес и осигурява основни йони, необходими за ензимната активност. Типичните PCR буфери са налични в концентрирани форми (10X или 5X) и трябва да бъдат разредени до работна концентрация от 1X преди употреба в PCR реакции.



**ЕТАПИ на ВСЕКИ ЦИКЪЛ**



**СИНТЕЗИРАНИ КОПИЯ на ДНК**



Процесът на PCR, включва повтарящи се термични цикъли (в диапазона 20-40 цикъла), осъществявани в специализирани машини – PCR апарати. Всеки цикъл включва 3 етапа с различна температура и продължителност:

- Денатурация на ДНК матрицата: Двуверижната ДНК се разделя на две единични вериги чрез разкъсване на водородните връзки. Единичните вериги са матрици за следващите стъпки на хибридикация и удължаване.

Необходим е етап на начална денатурация, преди да започне цикличността на трите етапа. Тази първоначална денатурация обикновено се осъществява в температурния диапазон 94-98°C, в зависимост от използваната ДНК полимераза и съдържанието на гуанин (G) и цитозин (C) в ДНК. За пълно денатуриране на ДНК началната денатурация може да отнеме до 2 min.

След първоначалната денатурация, всеки следващ цикъл включва етап на денатурация, обикновено извършвана в температурния диапазон 94-98°C с по-кратка продължителност, обикновено 10-30 sec.

- Хибридизация на праймерите: При охлаждането праймерите се свързват към съответните им комплементарни участъци върху едничните вериги на денатурираната ДНК – процесът се нарича, освен хибридизация, и анийлинг.

Температурата на хибридизация е критичен параметър, който значително влияе върху специфичността на свързването на праймерите. Обикновено се оптимизира да бъде 5°C под изчислената температура на топене ( $T_m$ ) на праймерите. Оптималната температура обикновено попада в диапазона 50-65°C. Времето за присъединяване обикновено е 15-60 sec, като 30 sec често са достатъчни. Ако  $T_m$  е твърде висока, праймерите може да не се свържат ефективно, което води до намалена амплификация. Обратно, ако  $T_m$  е твърде ниска, праймерите могат да се свържат неспецифично, което води до амплификация на нежелани ДНК региони. Градиентен PCR, при който се тестват едновременно множество температури на присъединяване, може да се използва за намиране на идеалната  $T_m$  за даден набор праймери.

- Удължаване на новосинтезиращата се верига: При повишаване на температурата ДНК полимеразата започва удължаване на праймерите и синтез на нова верига, комплементарна на матричната верига.

Стъпката на удължаване се извършва при температура, оптимална за активността на ДНК полимеразата, обикновено 72°C за *Taq* полимеразата. Времето за удължаване зависи от дължината на целевия ДНК регион (ампликон) и използваната полимераза. Общо правило е да се предвиди 1 min за 1 kb ДНК за *Taq* полимеразата. Някои полимерази, като *Pfu* полимераза, може да изискват по-дълго време за удължаване (напр. 2 min за 1 kb) и може да имат малко по-различни оптимални температури.

Удължаването протича при температура, която позволява оптимална полимеразна активност на свързване на нуклеотиди към нарастващата верига. Обикновено етапът на удължаване може да се извърши при 70-80°C за 1-2 min. Температурата и времето на удължаване са специфични за полимеразата. Например, *Taq* полимеразата работи най-добре при 70-80°C за 1 min за първите 2 kb и изисква още една минута за допълнителни kb; *Pfu* полимеразата функционира при 75°C и изисква 2 минути на 1kb ДНК.

Чрез повтаряне на термоцикъла 20-40 пъти, целевият ДНК регион (ампликонът) се амплифицира експоненциално. Всяка новосинтезирана ДНК молекула служи като матрица в следващите цикли, което води до многократно увеличаване на броя на копията на целевата последователност. Тази експоненциална амплификация е ключовата характеристика на PCR, позволяваща откриването и анализа на дори много малки количества изходен ДНК материал.

По-късно се разработва и техника за амплифициране на РНК. При нея се въвежда първоначална стъпка на обратна транскрипция за превръщане на РНК в комплементарна ДНК (кДНК) с помощта на ензима обратна транскриптаза. След това кДНК служи като матрица за амплификация. Тази техника е известна като RT-PCR (от англ. Reverse Transcription PCR).

- **Полимеразната верижна реакция в реално време**

За разлика от конвенционалния PCR, който обикновено предоставя качествени резултати (присъствие или отсъствие на целева ДНК), полимеразната верижна реакция в реално време, известна още като количествен PCR (qPCR) позволява прецизно количествено определяне на ДНК. Тук се измерва флуоресценцията, излъчвана по време на амплификацията, което

позволява на изследователите да определят началното количество на целевата ДНК. Количествена полимеразна верижна реакция с обратна транскрипция или RT-qPCR комбинира обратна транскрипция на РНК в комплементарна ДНК с qPCR за усилване и измерване на специфични целеви РНК.

Полимеразната верижна реакция в реално време (qPCR и RT-qPCR) предоставя надеждно количествено съотношение между началното количество целева ДНК, или РНК, и количеството на амплифицирания продукт, което я прави изключително чувствителна и точна. qPCR може точно да определи прецизния брой копия на целева нуклеинова киселина в проба. Използва се за откриване и количествено определяне на микроорганизми, които причиняват заболявания, като бактерии, вируси и гъбички. qPCR се използва за изследване на генетични варианти чрез изследване на ДНК последователността на индивида и може да открие еднонуклеотидни полиморфизми (от англ. single-nucleotide polymorphism, SNPs). Полезен за откриване, количествено определяне и генотипиране на ДНК в различни диагностични приложения, включително безопасност на храните, здравеопазване и селско стопанство.

RT-qPCR е ценен инструмент за измерване на нивата на гена експресия. Той разчита на сравнителния Cq метод, за да предостави относителна стойност на изобилието на целевите РНК молекули.

## ПРИЛОЖЕНИЕ 2 СЕКВЕНИРАНЕ НА ДНК

Секвенирането на ДНК е основна техника в молекулярната биология, която позволява на учените да определят точния ред на нуклеотидите (А, Т, С и G) в молекулата на ДНК. Тази информация е от съществено значение за разбирането на структурата, функцията и регулацията на гените.

Докато ДНК секвенирането директно чете нуклеотидната последователност на ДНК, за секвениране на РНК се използва различен подход. РНК обикновено се изследва, за да се измерят нивата на експресия на гените, да се идентифицират алтернативни събития на сплайсинг или да се разбере структурата на транскриптома на дадена клетка или организъм. Въпреки това, РНК не може да се секвенира директно чрез стандартните методи за секвениране на ДНК, тъй като тя е едноверижна и има различна химическа структура. Затова РНК първо се превръща в кДНК чрез обратна транскрипция.

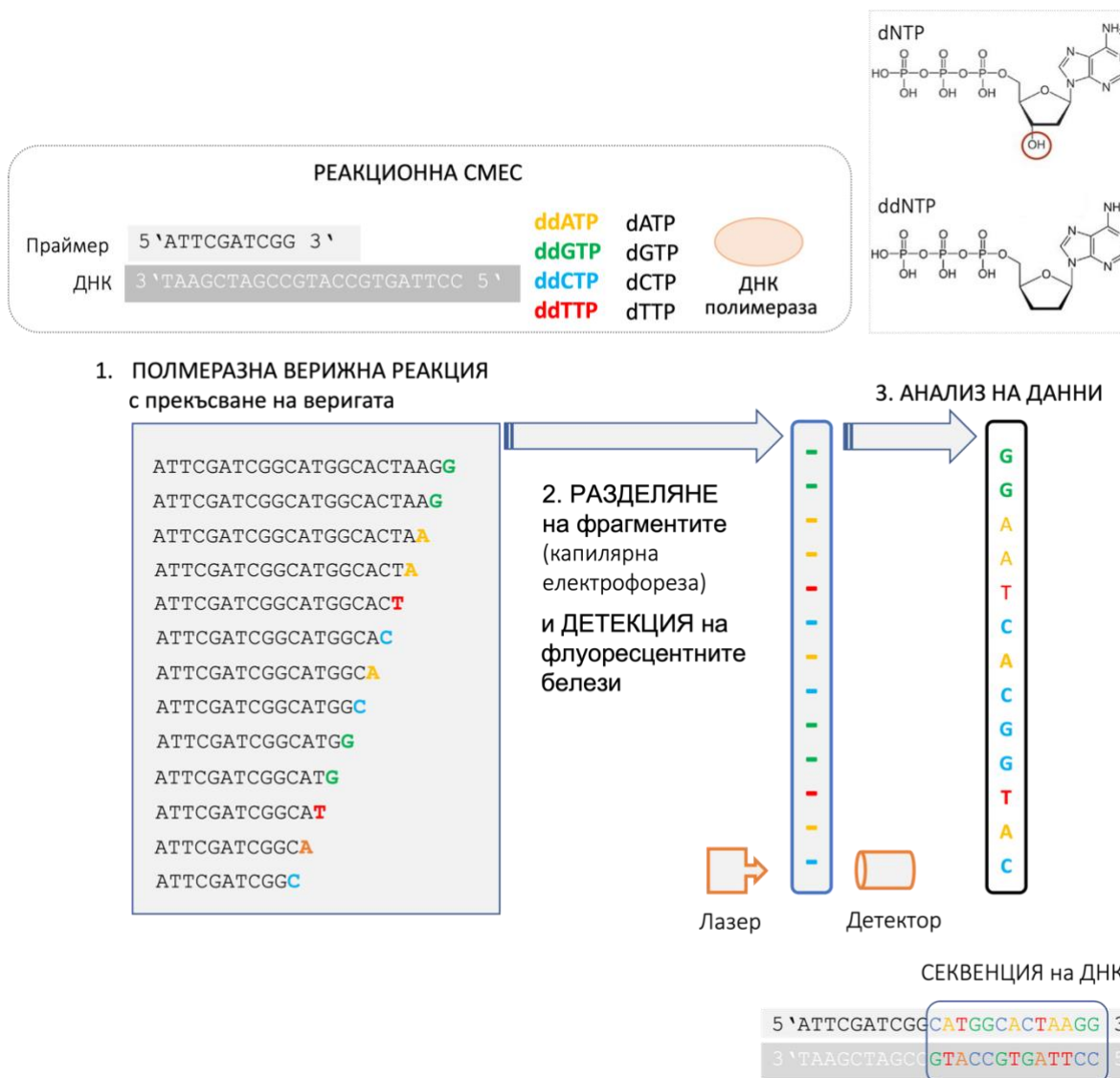
- **Първо поколение секвениране. Метод на Сенгер.**

Методите за секвениране от първо поколение, разработени в края на 70-те години на XX век, включват метода на Максам-Гилбърт и широко популярния метод на Сенгер, за който Фредерик Сенгер получава Нобелова награда. Тези методи имат различни възможности за секвениране: методът на Максам-Гилбърт може да анализира до приблизително 400 бази, докато методът на Сенгер увеличава този лимит до около 1000 бази на четене. През следващите десетилетия методът на Сенгер се утвърждава като най-широко използваният и предпочитан метод за секвениране на ДНК.

През 1987 г. компанията Applied Biosystems автоматизира и комерсиализира процеса на секвениране по метода на Сенгер, което позволява неговото широко приложение. Тази техника е от съществено значение за мащабни проекти като Проекта за човешкия геном, който през 2003 г. постига първото пълно секвениране на човешкия геном.

Методът на Сенгер разчита на включването на дидезоксинуклеотиди (ddNTPs - ddATP, ddGTP, ddCTP и ddTTP) по време на *in vitro* амплификация на ДНК чрез PCR. Дидезоксинуклеотидите са модифицирани аналози на ДНК нуклеотидите, които се отличават с липсата на 3'-хидроксилна група. Тази липсваща хидроксилна група предотвратява образуването на фосфодиестерната връзка, необходима за удължаването на веригата на ДНК, което води до прекъсване на синтеза на веригата при всяко включване на ddNTP.

В съвременните приложения на метода на Сенгер се използват флуоресцентно маркирани ddNTPs заедно със стандартни дезоксирибонуклеотидтрифосфати (dNTPs -dATP, dGTP, dCTP и dTTP). По време на реакцията на секвениране се получава смес от ДНК фрагменти с различна дължина. Тези фрагменти се разделят чрез капилярна електрофореза, а флуоресцентните маркери на ddNTPs показват терминиращия нуклеотид на всеки фрагмент. Чрез анализ на флуоресцентните сигнали се определя ДНК последователността.



Методът на Сенгер включва следните етапи:

- PCR с прекъсване на веригата: ДНК, която трябва да бъде секвенирана, служи като матрица за специализиран PCR-базиран метод, наречен PCR с прекъсване на веригата, който е основата на секвенирането по метода на Сенгер. Реакционната смес съдържа матрична ДНК, специфичен праймер, ДНК полимеразата, стандартни dNTPs и модифицирани ddNTPs. По време на етапа на удължаване ДНК полимеразата добавя dNTPs към растящата ДНК верига, като катализира образуването на фосфодиестерна връзка между свободната 3'-хидроксилна група (3'-ОН) на последния нуклеотид и 5'-фосфата на следващия нуклеотид. Когато обаче в синтеза се включи ddNTP, в който липсва 3'-ОН група, необходима за продължаване на удължаването на веригата, синтезът на ДНК спира на това място. В резултат на PCR с прекъсване на веригата се генерират милиони до милиарди ДНК фрагменти с различна дължина, като всеки фрагмент завършва на позиция, определена от включването на ddNTP. Съотношението между dNTPs и ddNTPs в реакционната смес е внимателно оптимизирано, за да се гарантира, че са представени всички възможни дължини на фрагментите и че всеки възможен терминален нуклеотид е включен на случайни позиции в секвенцията.

- Разделяне на ампликоните: В оригиналния метод на Сенгер ДНК ампликоните се разделят по размер чрез агарозна гел-електрофореза. По-късно тази техника е заменена от капилярна електрофореза, която използва като матрица силикагел за по-висока резолюция и автоматизация.

При капилярната електрофореза ДНК пробите се зареждат в ямки в единия край на колонката с матрицата, след което през гела се подава електрически ток. ДНК молекулите са отрицателно заредени поради фосфатния им скелет и мигрират към положителния електрод (анод) под влияние на електричното поле.

Тъй като ДНК фрагментите имат еднакво съотношение заряд към маса, тяхната миграция през гела зависи единствено от размера им. По-малките ДНК фрагменти срещат по-малко съпротивление, докато преминават през порите на матрицата, и затова се движат по-бързо от по-големите фрагменти. Това разделяне по размер позволява точно определяне на дължините на фрагментите.

- Визуализация и определяне на последователността на ДНК: Последователността на ДНК се определя чрез идентифициране на флуоресцентния маркер, прикрепен към терминалния ddNTP, който причинява прекъсването на веригата. ДНК полимеразата синтезира ДНК в посока 5' към 3', започвайки от известен праймер. Всеки терминален ddNTP съответства на определен нуклеотид от оригиналната ДНК последователност. Чрез анализ на реда на терминалните ddNTPs, от най-малкия до най-големия фрагмент, които съответстват на последователността, с която фрагментите излизат от капиляра, можем да реконструираме последователността на оригиналната ДНК в посока 5' към 3' края.

Когато ДНК фрагментите преминават през детектора по време на капилярната електрофореза, лазер възбужда флуоресцентните маркери, прикрепени към ddNTPs, а камера записва излъчената светлина. При автоматизираното секвениране по метода на Сенгер компютърът анализира излъчените флуоресцентни сигнали, за да идентифицира всеки терминален ddNTP. Тъй като всеки от четирите ddNTPs е маркиран с уникален флуоресцентен маркер, излъчената светлина директно идентифицира нуклеотида на точката на прекъсване.

Резултатът е хроматограма – графично изображение, което показва флуоресцентни пикове, съответстващи на всеки нуклеотид по дължината на матричната ДНК. Тази хроматограма позволява точното идентифициране на ДНК последователността.

- Второ поколение секвениране

Докато техниките от първо поколение секвенират само един ДНК фрагмент наведнъж, технологиите от второ поколение секвенират милиони до милиарди ДНК фрагменти едновременно. Тези технологии се наричат още секвениране от следващо поколение (от англ. Next-Generation Sequencing, NGS) или масивно паралелно секвениране (от англ. Massive Parallel Sequencing, MPS). В резултат на секвенирането се генерират множество прочити. Прочитът (от англ. read) е последователността от бази, съответстваща на един ДНК фрагмент, чиято дължина може да бъде в диапазона 50-500 нт.

Платформите за секвениране се различават по своите методологии, дължина на прочита, разходи и приложения, обслужващи широк спектър от изследователски нужди от клинична диагностика до екологични изследвания.

Illumina (Solexa) е най-широко използваната NGS платформа, която предоставя над 90% от глобалните данни за секвениране. Тя работи на базата на подход за секвениране чрез синтез, като открива флуоресцентни сигнали, когато нуклеотидите се добавят към нарастваща ДНК верига. Illumina предлага различни системи като MiSeq, NextSeq и NovaSeq, отговарящи на различни изисквания за дълбочина на секвениране и дължина на прочита. Тази платформа включва следните основни етапи:

- Създаване на ДНК библиотека: ДНК, или РНК, се извличат от биологичната проба. Ако се секвенира РНК, тя първо се обръща в кДНК. ДНК се накъсва на по-малки фрагменти, в двата края на които се добавят адаптори. Това позволява на фрагментите да се свържат с проточната клетка.
- Мостова амплификация на библиотеката за усилване на сигнала: След хибридизация на фрагментите върху проточната клетка, те се амплифицират (обикновено с PCR), при което се създават клъстери – групи от идентични ДНК фрагменти, с цел увеличаване сигнала по време на секвенирането.
- Секвениране на библиотеката и генериране на прочити: ДНК веригите се секвенират база по база чрез използване на флуоресцентно маркирани нуклеотиди (Illumina), промени в рН (Ion Torrent) или други методи за детекция. Секвениращата платформа „чете“ фрагментите като улавя сигнали от всеки нуклеотид (A, T, C, G) при неговото вграждане в растящата верига, при което се генерират прочитите.
- Анализ на данните: Данните от секвенирането (множествата от прочити) се анализират с помощта на специализирани софтуери за съпоставяне на прочитите, за сглобяване на прочитите в по-дълги последователности (контиги), за тяхното подравняване спрямо референтен геном и интерпретиране на резултатите, например идентифициране на генни варианти, мутации или нива на гена експресия.

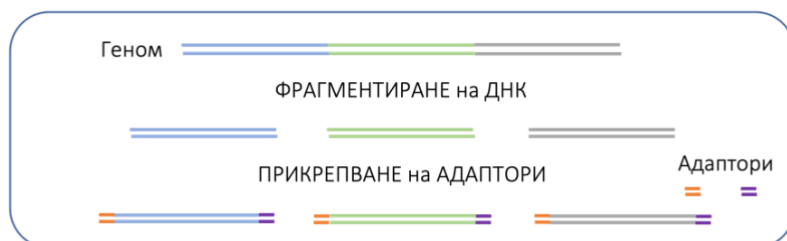
Секвенирането на ДНК (DNA-seq) позволява бързо и рентабилно да бъдат анализирани цели геноми или целеви региони.

Секвенирането на РНК (RNA-seq) се използва за анализ на транскриптома, който включва всички РНК молекули в клетката. То дава представа за нивата на гена експресия и алтернативни събития на снаждане (сплайсинг). RNA-seq може да идентифицира както кодиращи, така и не кодиращи РНК и е от решаващо значение за разбирането на клетъчните реакции и регулаторните механизми.

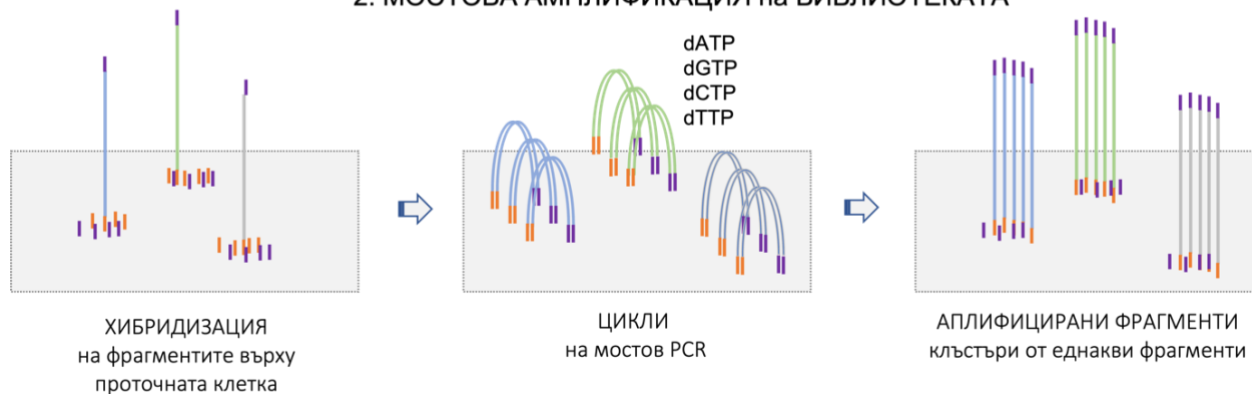
Метагеномиката включва секвениране на генетичен материал от проби от околната среда за изследване на колективните геноми на микроорганизми в дадена общност. Този подход позволява на изследователите да изследват микробното разнообразие, функционалния

потенциал и взаимодействията в екосистемите без необходимост от култивиране на организми. Има приложения в екологията, медицината и биотехнологиите.

### 1. ПОДГОТОВКА на ДНК БИБЛИОТЕКАТА



### 2. МОСТОВА АМПЛИФИКАЦИЯ на БИБЛИОТЕКАТА



### 3. СЕКВЕНИРАНЕ чрез СИНТЕЗ с флуоресцентно белязани нуклеотиди



ГЕНЕРИРАНЕ на ПРОЧИТИ

```
AGCCATGTGCACCTT
CTGATTAGCATGCCA
TCCTGAATTAGGCCG
```

### 4. АНАЛИЗ на ДАННИ

- Трето поколение секвениране

Секвенирането от трето поколение представлява значителен напредък в технологиите за секвениране на ДНК, предлагайки уникални възможности, които го отличават от предишните поколения.

- анализират се индивидуални ДНК молекули без необходимост от амплификация, позволявайки по-директно и точно секвениране на нативна ДНК.
- генерират се значително по-дълги прочити в сравнение с секвенирането от второ поколение. Например, технологии като Oxford Nanopore могат да секвенират ДНК вериги, докато преминават през нанопори, позволявайки ултра-дълги прочити, надвишаващи 1 милион бд, което е полезно за сглобяване на сложни геноми и разрешаване на повторяеми региони.
- някои платформи, като SMRT на PacBio, позволяват придобиване на данни в реално време, което дава възможност на изследователите да наблюдават процеса на секвениране, докато той протича.
- директно могат да бъдат откривани епигенетични модификации (като метилиране) на нативна ДНК, предоставяйки информация за регулацията на гените и структурата на хроматина.

РЕФЕРЕНЦИИ

- Alberts B., Johnson A., Lewis J., Raff M., Roberts K., and Walte P. (2022) *Molecular Biology of the Cell*. 7th edition. New York: Garland Science; ISBN-13. 978-0393884821.
- Baynes J. W. and Dominiczak M.H. (2022) *Medical Biochemistry*. 6th Edition, Elsevier; ISBN-9780323834506.
- Britten B.J. and Davidson E.H. (1969) Gene regulation for higher cells: a theory. *Science* 165: 349-357.
- Chang, H., Pannunzio, N., Adachi, N. *et al.* Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* 18, 495–506 (2017). <https://doi.org/10.1038/nrm.2017.48>
- Chiesa G., Sirtori C.R. Apolipoprotein A-I(Milano): current perspectives. *Curr Opin Lipidol*. 2003 Apr;14(2):159-63. doi: 10.1097/00041433-200304000-00007. PMID: 12642784
- Courey, Albert J. "Mechanisms in transcriptional regulation" (2008) Blackwell Publishing, ISBN: 978-1-4051-0370-1.
- Draz B. (2023) *Molecular Biology: Its Key Principles and Significance*. Adv Tech Biol Med. 11:418.
- Fry C.J. and Peterson C.L. (2002) Transcription. Unlocking the gates to gene expression. *Science* 295(5561):1847-8.
- Johnson, P. F., & McKnight, S. L. (1989). Eukaryotic transcriptional regulatory proteins. *Annual review of biochemistry*, 58(1), 799-839.
- Juven-Gershon T. and James T. Kadonaga (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Developmental Biology* 339: 225–229.
- Hsieh, P. (2001). Molecular mechanisms of DNA mismatch repair. *Mutation Research/DNA Repair*, 486(2), 71-87.
- Kadonaga J. T. (2012) Perspectives on the RNA Polymerase II Core Promoter. *Interdiscip Rev Dev Biol*. 1(1): 40–51.
- Krebs J. E., Goldstein E.S., Kilpatrick S.T. (2011) "Lewin's Genes XI" Jones & Bartlett Learning. ISBN-13: 978-0763766320
- Kornberg A. (1984) DNA replication. *Trends in Biochemical Sciences*, 9, 4, 122-124, [doi.org/10.1016/0968-0004\(84\)90114-2](https://doi.org/10.1016/0968-0004(84)90114-2).
- Kusakabe, M., Onishi, Y., Tada, H. *et al.* Mechanism and regulation of DNA damage recognition in nucleotide excision repair. *Genes and Environ* 41, 2 (2019). <https://doi.org/10.1186/s41021-019-0119-6>.
- Laursen, B. S., Sørensen, H. P., Mortensen, K. K., & Sperling-Petersen, H. U. (2005). Initiation of protein synthesis in bacteria. *Microbiology and molecular biology reviews*, 69(1), 101-123.
- Mohd-Sarip A. and Verrijzer C.P. (2004) Molecular biology. A higher order of silence. *Science*. 306(5701):1484-5.
- Muñoz-López M, García-Pérez JL. DNA transposons: nature and applications in genomics. *Curr Genomics*. 2010 Apr;11(2):115-28. doi: 10.2174/138920210790886871. PMID: 20885819; PMCID: PMC2874221.
- Lim, H. N., Lee, Y., & Hussein, R. (2011). Fundamental relationship between operon organization and gene expression. *Proceedings of the National Academy of Sciences*, 108(26), 10626-10631.
- López-Lastra, M., Rivas, A., & Barría, M. I. (2005). Protein synthesis in eukaryotes: the growing biological relevance of cap-independent translation initiation. *Biological research*, 38(2-3), 121-146.
- Lynch, M. (2006). The origins of eukaryotic gene structure. *Molecular biology and evolution*, 23(2), 450-468.
- Lv, X., Hueso-Gil, A., Bi, X., Wu, Y., Liu, Y., Liu, L., & Ledesma-Amaro, R. (2022). New synthetic biology tools for metabolic control. *Current Opinion in Biotechnology*, 76, 102724.

- Osbourn, A. E., & Field, B. (2009). Operons. *Cellular and Molecular Life Sciences*, 66, 3755-3775.
- Orphanides G, Lagrange T, Reinberg D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev.* 10(21):2657-83.
- Preiss, T., & W. Hentze, M. (2003). Starting the protein synthesis machine: eukaryotic translation initiation. *Bioessays*, 25(12), 1201-1211.
- Russell J. and. Zomerdijk J. C.B.M (2005) RNA-polymerase-I-directed rDNA transcription, life and works. *Trends Biochem Sci.* 30(2): 1-20.
- Sandelin A., Carninci P., Lenhard B., Ponjavic J., Hayashizaki Y., Hume D.A. (2007) Mammalian RNA polymerase II core promoters: insights from genome-wide studies. *Nat Rev Genet.* 8(6):424-36.
- Sancar, A., Lindsey-Boltz, L. A., Ünsal-Kaçmaz, K., & Linn, S. (2004). Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annual review of biochemistry*, 73(1), 39-85.
- Savić, N., & Schwank, G. (2016). Advances in therapeutic CRISPR/Cas9 genome editing. *Translational Research*, 168, 15-21.
- Sutherland, H., Bickmore, W. Transcription factories: gene expression in unions?. *Nat Rev Genet* 10, 457–466 (2009). <https://doi.org/10.1038/nrg2592>.
- Timoféeff-Ressovsky, N. W., Zimmer, K. G., & Delbruck, M. (2011). On the nature of gene mutation and gene structure. *Creating a Physical Biology: The Three-Man Paper and Early Molecular Biology*; Sloan, PR, Fogel, B., Eds.
- Thomas M. C. and Chiang C.-M. (2006) The General Transcription Machinery and General Cofactors. *Critical Reviews in Biochemistry and Molecular Biology*, 41:105–178.
- Tropp B. *Principles of Molecular Biology*, Jones and Bartlett Learning, 2014. ISBN 13: 9781449689179.
- Venkatesh, S., & Workman, J. L. (2015). Histone exchange, chromatin structure and the regulation of transcription. *Nature reviews Molecular cell biology*, 16(3), 178-189.
- Wade J.T. and Struhl K. (2008) The transition from transcriptional initiation to elongation. *Curr Opin Genet Dev.* 18(2): 130–136.
- Watson, J. D., & Crick, F. H. C. (1953) A structure for deoxyribose nucleic acid. *Nature* 171, 737–738.

Най-често срещани съкращения

бд – базова двойка

ДНК – дезоксирибонуклеинова киселина

рДНК – рибозомна ДНК

мтДНК – митохондриална ДНК

мяРНП – рибонуклеопротеинов комплекс, включващ малка ядрена РНК

нд - нуклеотид

ПИК – пре-инициаторен комплекс

РНК – рибонуклеинова киселина

РНКаза - рибонуклеаза

иРНК - информационна РНК

рРНК – рибозомна РНК

тРНК – транспортна РНК

пре-иРНК - предшественик на иРНК

пре-рРНК - предшественик на иРНК

пре-тРНК - предшественик на иРНК

СТД – С-терминален домен

ТФ – транскрипционен фактор

NMPs - Nucleoside Monophosphate (Нуклеозид монофосфати): Това са основните единици на нуклеотидите, състоящи се от азотна база, захар (рибоза или дезоксирибоза) и една фосфатна група. NMPs могат да бъдат допълнително категоризирани като dNMPs (дезоксирибонуклеозид монофосфати), които съдържат дезоксирибозна захар и са структурни единици на ДНК и rNMPs (рибонуклеозид монофосфати), които съдържат рибозна захар и са структурни единици на РНК. NMPs служат като предшественици за по-високо фосфорилирани форми като нуклеозид трифосфати (NTPs) и участват в клетъчната сигнализация, като цикличен АМР (сАМР).

AMP - Adenosine monophosphate

CMP - Cytidine monophosphate

GMP - Guanosine monophosphate

TMP - Thymidine monophosphate

UMP - Uridine monophosphate

NTPs - Nucleoside Triphosphate (Нуклеозид трифосфати): NTPs могат да бъдат допълнително категоризирани като dNTPs, които са от съществено значение за синтеза и възстановяването на ДНК, и rNTPs, които играят ключова роля в транскрипцията на РНК, преноса на енергия (АТФ) и клетъчната сигнализация (GTP).

ATP - Adenosine triphosphate

CTP - Cytidine triphosphate

GTP - Guanosine triphosphate

TTP - Thymidine triphosphate

UTP - Uridine triphosphate

PCR – Polymerase Chain Reaction, Полимеразна Верижна Реакция