

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

**А. В. Сиволоб**

# **МОЛЕКУЛЯРНА БІОЛОГІЯ**

Підручник

**Затверджено  
Міністерством освіти і науки України  
як підручник для студентів вищих навчальних закладів**



УДК 577.2(075.8)  
ББК 28.070я73  
С34

Рецензенти:

д-р біол. наук, проф., акад. НАН України Г.В. Єльська  
(директор Інституту молекулярної біології і генетики НАН України);  
д-р біол. наук, проф., чл.-кор. НАН України О.І. Корнелюк  
(Інститут молекулярної біології і генетики НАН України);  
д-р біол. наук, проф., чл.-кор. НАН України С.О. Костерін  
(Інститут біохімії імені О.В. Палладіна НАН України)

*Затверджено до друку Вченою радою  
Київського національного університету імені Тараса Шевченка  
5 березня 2007 року*

**Сиволоб, А.В.**

С34 Молекулярна біологія : підручник / А.В. Сиволоб. – К. : Видавничо-поліграфічний центр “Київський університет”, 2008. – 384 с.

ISBN 978-966-439-068-9

Подано сучасні уявлення про молекулярні механізми загальних біологічних процесів. Викладено принципи структурної організації білків і нуклеїнових кислот, а також взаємодій біологічних макромолекул. Описано молекулярні механізми транскрипції, процесингу РНК, білкового синтезу, реплікації, репарації та рекомбінації ДНК. Основну увагу приділено структурним і динамічним властивостям макромолекул і макромолекулярних комплексів, які забезпечують здійснення біологічних функцій.

Для студентів біологічних факультетів університетів, медичних і аграрних вищих навчальних закладів, а також для аспірантів і науковців, які цікавляться молекулярною біологією.

Іл. 248. Бібліогр.: 159.

**Гриф надано Міністерством освіти і науки України  
(лист №1.4/18-Г.-1431 від 27.08.07)**

УДК 577.2(075.8)  
ББК 28.070я73

ISBN 978-966-439-068-9

© А.В. Сиволоб, 2008  
© Київський національний університет імені Тараса Шевченка,  
ВПЦ “Київський університет”, 2008

## ЗМІСТ

<b>Від автора</b> .....	8
<b>Вступ</b> .....	9
<b>Розділ 1.</b> Фізико-хімічні основи молекулярної біології .....	13
Енергетика взаємодії між атомами й молекулами.....	13
Вільна енергія .....	13
Рівновага утворення комплексів молекул у розчині .....	16
Ковалентний зв'язок між атомами в молекулах .....	19
Властивості ковалентного зв'язку .....	19
Механізм гнучкості полімерного ланцюга.....	20
Перебудови системи ковалентних зв'язків у хімічних реакціях.....	22
Нековалентні взаємодії між атомами й молекулами .....	23
Вандерваальсові взаємодії.....	23
Електростатичні (іонні) взаємодії .....	25
Водневий зв'язок .....	27
Гідрофобні взаємодії.....	29
<b>Розділ 2.</b> Білки .....	33
Хімічна будова білків .....	33
Амінокислоти .....	33
Пептидний зв'язок і поліпептидний ланцюг .....	35
Вторинна структура .....	37
$\alpha$ -Спіраль .....	37
$\beta$ -Структура .....	39
Спіраль $3_{10}$ .....	40
$\beta$ -Поворот.....	40
Глобулярна структура .....	40
Роль вторинної структури в утворенні глобули.....	41
Стабілізація глобули .....	42
Структурна класифікація глобулярних білків.....	44
Структура мембранних білків.....	50
Неструктуровані білки .....	51
Конформаційна рухливість білків.....	51
Принципи функціонування білків .....	53
Механізм ферментативного каталізу .....	54
Механізми використання енергії гідролізу АТФ .....	57
<b>Розділ 3.</b> ДНК .....	65
Хімічна будова нуклеїнових кислот .....	65
Нуклеотиди .....	65

Полінуклеотидний ланцюг .....	68
Нуклеази .....	69
Подвійна спіраль.....	70
Стабілізація подвійної спіралі.....	70
Конформаційні параметри подвійної спіралі .....	74
Структурні форми ДНК .....	76
Конформація В-ДНК.....	78
Білково-нуклеїнові взаємодії.....	81
Структурна класифікація білків, які взаємодіють із ДНК.....	81
Принципи білково-нуклеїнового впізнання .....	88
Циркулярна ДНК.....	93
<b>Розділ 4.</b> Організація ДНК у клітинах: геноми та структура хроматину.....	103
Організація геномів .....	103
Генетичний код.....	103
Гени.....	105
Геноми.....	106
Молекулярна організація хроматину .....	115
Нуклеосома .....	115
Посттрансляційні модифікації гістонових хвостів .....	121
Гістон H1 .....	124
Наднуклеосомна упаковка хроматинової фібрили .....	125
Петльові домени хроматину та ядерний матрикс.....	127
<b>Розділ 5.</b> Транскрипція: прокаріоти .....	131
РНК-полімераза .....	131
Загальний сценарій транскрипції в бактерій .....	131
Структура РНК-полімерази .....	133
Ініціація транскрипції .....	136
Елонгація транскрипції .....	139
Термінація транскрипції.....	145
Регуляція транскрипції .....	147
Лактозний оперон .....	147
Антитермінація .....	149
Атенюація .....	150
Регуляція транскрипції бактеріофага $\lambda$ .....	152
Неспецифічна регуляція загального рівня транскрипції. ....	156
<b>Розділ 6.</b> Транскрипція: еукаріоти.....	159
РНК-полімераза II.....	160
Структура РНК-полімерази II.....	160
Промотор РНК-полімерази II.....	162

Базальні фактори транскрипції.....	163
Ініціація транскрипції РНК-полімеразою II .....	167
РНК-полімерази I і III .....	168
Ініціація транскрипції генів рибосомної РНК.....	168
Ініціація транскрипції РНК-полімеразою III .....	169
Механізми активації транскрипції .....	170
Транскрипційні фактори .....	171
Зовнішня регуляція активності	
транскрипційних факторів .....	173
Ацетилювання гістонів .....	176
Доступність промоторів .....	177
Комплекси ремоделювання хроматину.....	178
Роль білків НМГ в активації транскрипції .....	183
Елонгація РНК-полімерази через хроматин.....	186
Конститутивна репресія транскрипції: гетерохроматин.....	189
Деацетилювання гістонів.....	189
НР1-залежна система репресії.....	191
Метилювання ДНК.....	193
РНК-інтерференція .....	194
<b>Розділ 7.</b> Процесинг еукаріотичних мРНК .....	201
Кепування.....	201
Сплайсинг .....	204
Сплайсосома: механізм сплайсингу .....	205
Поладенілування мРНК і термінація транскрипції.....	209
Альтернативний сплайсинг.....	212
Транс-сплайсинг .....	215
Редагування мРНК .....	217
<b>Розділ 8.</b> Синтез білків .....	221
Транспортні РНК.....	221
Структура тРНК .....	221
Аміноацилювання тРНК.....	225
Рибосома .....	230
Склад рибосоми .....	230
Структура рибосоми .....	231
Елонгаційний цикл .....	239
Елонгаційний фактор EF1 .....	240
Зв'язування aa-тРНК з А-сайтом рибосоми .....	242
Транспептидація.....	247
Транслокація.....	249
Ініціація трансляції.....	253
Ініціація трансляції у прокаріотів.....	254

Ініціація трансляції в еукаріотів .....	256
Термінація трансляції .....	258
Регуляція трансляції .....	259
Формування просторової структури білка .....	262
Закономірності укладання білкової глобули .....	263
Шаперони hsp70 .....	267
Шапероніни .....	270
<b>Розділ 9. Реплікація ДНК</b> .....	275
Реплікони .....	276
ДНК-полімераза .....	279
Структура ДНК-полімерази й полімеразна реакція .....	279
Редагування помилок .....	282
Особливості ДНК-полімерази в порівнянні з РНК-полімеразою .....	283
Реплісома та інші елементи системи реплікації .....	284
Геліказа й білки SSB .....	284
Синтез ланцюга, що запізнюється .....	286
Голофермент ДНК-полімерази .....	288
Топологічні проблеми, пов'язані з реплікацією .....	290
Ініціація реплікації в бактерій .....	292
Особливості еукаріотичної системи реплікації .....	293
Еукаріотичні ДНК-полімерази .....	294
Ініціація реплікації в еукаріотів .....	295
Структурні зміни хроматину під час реплікації .....	296
Подовження кінців еукаріотичної хромосоми .....	298
Репарація ДНК .....	300
Пряма репарація .....	300
Ексцизійна репарація .....	301
Репарація некомплементарних пар основ – місметчів .....	303
Репарація без репарації .....	304
Репарація дволанцюгових розривів .....	306
<b>Розділ 10. Рекомбінація ДНК</b> .....	309
Гомологічна рекомбінація .....	310
Сайт-специфічна рекомбінація .....	316
Незаконна рекомбінація .....	318
V(D)J-рекомбінація імуноглобулінових генів .....	319
Транспозиції мобільних елементів .....	323
ДНК-транспозони .....	323
LTR-ретропозони .....	324
Мобільні елементи LINE та SINE .....	325

<b>Розділ 11. Методи молекулярної біології</b> .....	331
Клонування, ампліфікація і секвенування ДНК .....	331
Клонування ДНК .....	331
Гель-електрофорез.....	335
Створення та скринінг геномних бібліотек.....	337
Полімеразна ланцюгова реакція .....	339
Секвенування ДНК .....	341
Біоінформатика .....	343
Експресія рекомбінантних білків.....	344
Методи аналізу структури й експресії генів і геномів.....	347
Блот-гібридизація .....	347
Визначення стартових і кінцевих точок та рівня активності транскрипції.....	349
Аналіз експресії геному .....	351
Методи дослідження ДНК-білкових взаємодій.....	353
Гель-електрофорез білково-нуклеїнових комплексів.....	353
Футпринтинг .....	354
Імунопреципітація хроматину .....	355
Методи дослідження протеому.....	356
Фізичні методи дослідження структури й активності біомакромолекул.....	358
Методи безпосереднього спостереження .....	358
Рентгеноструктурний аналіз .....	360
Дослідження структури макромолекул у розчині .....	361
Методи дослідження одиночних макромолекул.....	365
Перспективи молекулярної біології .....	369
Предметний покажчик .....	377

## ВІД АВТОРА

*Мета цього видання – познайомити студентів біологічних факультетів із сучасними уявленнями про молекулярні механізми найважливіших загальних біологічних процесів. У його основу покладено курс молекулярної біології на біологічному факультеті Київського національного університету імені Тараса Шевченка. Книга передбачає попереднє ознайомлення читача з основами цитології та біохімії.*

*Якщо визначити молекулярну біологію як розділ біологічних наук, присвячений з'ясуванню молекулярних механізмів біологічних процесів, то можна сказати, що сьогодні майже вся біологія є молекулярною або, принаймні, прагне бути такою. Однак у більш вузькому сенсі під молекулярною біологією традиційно розуміють вивчення молекулярних механізмів процесів передачі та реалізації спадкової інформації – загальних процесів, що відбуваються в усіх живих системах. Цей підручник виходить саме з такого розуміння (ключовими при цьому є слова “молекулярні механізми”), яке дозволяє відокремити молекулярну біологію як навчальну дисципліну від біохімії, біофізики, генетики та фізіології. З іншого боку, оскільки загальні принципи функціонування біологічних макромолекул є універсальними, видання може розглядатися як базове для подальшого вивчення інших біологічних процесів на молекулярному рівні.*

*Усі підручники можна поділити на дві групи: видання великого обсягу, які містять вичерпно повний набір відомостей щодо певної галузі й можуть бути використані як довідники, і стислого викладу, що дають уявлення про базові принципи відповідної наукової дисципліни. Пропонована праця належить до другого типу. Основну увагу в ній приділено принципам структурної організації біологічних макромолекул та взаємодій між ними, що лежать в основі здійснення біологічних функцій.*

*Сподіваюся, що книга буде корисною не тільки для студентів, а й для аспірантів і науковців, біологів, медиків, аграріїв, а також фізиків і хіміків, які цікавляться молекулярною біологією.*

*Щиро дякую рецензентам книги Г.В. Єльській, О.І. Корнелюку й С.О. Костеріну, а також викладачам кафедри загальної та молекулярної генетики Київського національного університету імені Тараса Шевченка за обговорення рукопису й корисні зауваження.*

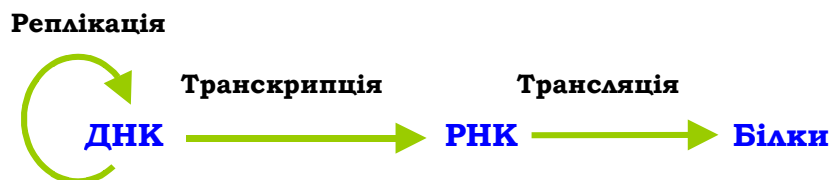
**А.В. Сиволоб**



## ВСТУП

Загальними, спільними в основних своїх рисах для всіх живих систем, процесами, що лежать в основі функціонування організмів, є процеси збереження та реалізації спадкової інформації.

Магістральний шлях передачі інформації в біологічних системах відображає схема, запропонована свого часу Френсісом Кріком під назвою *центральна догма молекулярної біології*:



Формулювання центральної догми можна вважати початком молекулярної біології як окремої науки. Саме за такою схемою, не порушуючи певної традиції, побудовано цю книгу. Слід зауважити, що крім відображених на схемі, у живих системах реалізуються також шляхи передачі інформації від РНК на РНК і на ДНК.

Вузловими точками схеми – джерелом і об'єктами передачі інформації – є біологічні *макромолекули* двох основних типів: *нуклеїнові кислоти* та *білки*. Приставка “макро-” відображає той простий факт, що молекула має великий розмір – не характерний для “звичайних” молекул неорганічного світу. Великий розмір біологічних макромолекул є принциповою відмінністю живих систем: тільки молекула, яка складається з великої кількості атомів, здатна мінімізувати перешкоджаючий вплив випадкових теплових флуктуацій і здійснювати більш-менш точні операції на основі численних взаємодій з іншими молекулами. У першому розділі книги розглянуто деякі важливі фізико-хімічні принципи, які є основою молекулярної біології.

Макромолекули – це полімерні ланцюги, що складаються з мономерних хімічних груп певного класу. Кінцевою точкою передачі біологічної інформації є білки (або протеїни) – біополімери, побудовані головним чином з 20 типів *амінокислот*. Велике розмаїття амінокислотних послідовностей створює можливості для реалізації різноманітних шляхів укладання ланцюгів у просторі – утворення специфічних просторових структур білків для виконання певних специфічних завдань. Головним

типом цих завдань є каталіз численних біохімічних реакцій (у тому числі тих, що забезпечують передачу біологічної інформації). Здійснення реакцій, які мають відбуватися у вузькому діапазоні фізіологічних умов, було б практично неможливим без такого каталізу. Крім того, білки є основним будівним матеріалом будь-якої біологічної системи й забезпечують виконання всіх інших біологічних функцій: транспорт речовин, передачу регуляторних сигналів, спрямовані рухи частин системи, захист від чужорідних молекул тощо. Принципи структурної організації білків та їхнє функціонування розглядаються в розділі 2.

Мономерами нуклеїнових кислот є *нуклеотиди*: нуклеотиди чотирьох типів утворюють полінуклеотидний ланцюг, два такі ланцюги, утримуючись разом за рахунок певних взаємодій, формують дволанцюгову спіральну молекулу *дезоксирибонуклеїнової кислоти (ДНК)*. Нуклеотиди, котрі трохи відрізняються за хімічною будовою, утворюють (як правило, одностанцюгову) полімерну молекулу *рибонуклеїнової кислоти (РНК)*. Молекула ДНК відрізняється високою стабільністю, що робить її надійним сховищем спадкової інформації. Сама інформація записана у вигляді послідовності нуклеотидів чотирьох типів, а структура дволанцюгової молекули забезпечує функціонування ДНК як носія інформації. Ключовим у цьому відношенні є *принцип комплементарності*, згідно з яким чотири нуклеотиди здатні утворити дві комплементарні пари, і ДНК є стабільною тільки за умови, що у двох ланцюгах один проти одного розташовані нуклеотиди однієї такої пари. Таким чином, кожен із ланцюгів є своєрідним дзеркальним відображенням іншого. Структурну організацію ДНК і принципи її взаємодії з білками, що забезпечують упаковку ДНК у клітині та регуляцію її активності, описано в розділі 3. Особливості різних типів РНК, які виконують різноманітні важливі допоміжні функції у процесах реалізації спадкової інформації, розглядаються окремо в розділах, присвячених механізмам відповідних процесів.

Спадкова інформація, записана в послідовності нуклеотидів на ділянці ДНК, є інформацією про послідовність амінокислот у складі білка. Окрема змістовна ділянка ДНК, у послідовності якої закодована амінокислотна послідовність білка, називається *геном*, а сукупна нуклеотидна послідовність ДНК у клітинах певного організму (яка включає не тільки гени) – *геномом*. У розділі 4 викладено загальні принципи організації геномів, а також упаковки ДНК у клітинах – у складі ДНК-білкового комплексу, який називають *хроматином*. У цьому ж розділі описано систему співвідношень між елементами послідовності

нуклеотидів і амінокислотними послідовностями білків – *генетичний код*, який використовується живими системами для перекладу нуклеотидного тексту в послідовність амінокислот у складі білка.

Наведене визначення гена є, однак, не зовсім точним. Крім білкових генів, геном кожного організму містить також гени, кінцевими продуктами яких є різноманітні молекули РНК. Але незалежно від типу гена, первинним продуктом його активності (проміжним для білкових генів) є молекула РНК. Під час *транскрипції* нуклеотидна послідовність одного з ланцюгів ДНК за принципом комплементарності переписується в нуклеотидну послідовність РНК – ДНК використовується як матриця, на якій будується комплементарна РНК-репліка. Молекула РНК, що синтезується на білковому гені, використовується далі як матриця для білкового синтезу (матрична або інформаційна РНК – мРНК). Молекулярні механізми транскрипції та її регуляції у про- та еукаріотів описано в розділах 5 і 6. Тісно пов'язаний із транскрипцією процесинг еукаріотичних мРНК (остаточне “дозрівання” функціонально-активної молекули) розглянуто в розділі 7.

Розділ 8 присвячено молекулярним механізмам кінцевого етапу експресії білкових генів – синтезу білків. Матрицею для білкового синтезу є молекула мРНК, нуклеотидна послідовність якої за правилами генетичного коду перекладається в амінокислотну послідовність білка під час *трансляції*. При цьому знов використовується принцип комплементарності: елемент послідовності мРНК упізнається комплементарною ділянкою у складі особливої молекули РНК – транспортної РНК (тРНК), яка несе на собі амінокислоту. Специфічне зв'язування певної тРНК із відповідною амінокислотою є важливою передумовою трансляції. Процес сканування матриці та синтезу амінокислотного ланцюга забезпечується *рибосомою* – мультимолекулярним комплексом, основою якого є рибосомні РНК (рРНК), зв'язані з рибосомними білками. Після закінчення синтезу на рибосомі амінокислотного ланцюга (або іноді під час синтезу) відбувається його просторове укладання у специфічну білкову структуру. Закономірності цього процесу також викладено в розділі 8.

Суттєвим моментом функціонування живої системи є не тільки реалізація (експресія) генетичної інформації, а й збереження її та подвоєння з метою передачі наступному поколінню. Подвоєння інформації – *реплікація* – це відтворення молекули ДНК у двох ідентичних дочірніх копіях. Головним механізмом реплікації є вже згадуваний кілька разів принцип комплементарності: кожен з ланцюгів ДНК використовується як матриця для синтезу комплементарної ДНК-репліки.

Спеціальні молекулярні системи забезпечують високу точність синтезу ДНК під час реплікації. Процеси виправлення помилок, які все ж таки виникають під час цього процесу, а також пошкоджень у молекулі ДНК, що з'являються під дією зовнішніх впливів, об'єднано під назвою *репарації* ДНК. Молекулярні механізми реплікації та репарації описано в розділі 9.

Розділ 10 присвячено молекулярним процесам, що приводять до надзвичайно біологічно важливого “перетасування” генетичного матеріалу – *рекомбінації* ДНК.

В останньому 11 розділі подано огляд деяких найважливіших методів дослідження, використовуваних у молекулярній біології.

З метою полегшення сприйняття інформації в тексті відсутні посилання. Перелік рекомендованої літератури наведено наприкінці кожного розділу. Перелік літератури, наведений нижче, можна вважати загальним до курсу в цілому.

Зображення структур макромолекул, якими проілюстровано книгу, створено на основі атомних координат із Банку даних білкових структур (PDB – Protein Data Bank, <http://www.rcsb.org/pdb/>, відповідні коди PDB наведено в підписах до рисунків) за допомогою програмних пакетів PyMOL (<http://www.pymol.org/>) та UCSF Chimera (<http://www.cgl.ucsf.edu/chimera>).

## **РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

*Албертс, Б., Брей, Д., Льюис, Дж. и др.* Молекулярная биология клетки : в 3 т. – М. : Мир, 1993.

*Сингер М., Берг П.* Гены и геномы : в 2 т. – М. : Мир, 1998.

*Lewin, B.* Genes VIII. Upper Saddle River. – New Jersey : Pearson Prentice Hall, 2004.

*Lodish, H., Berk, A., Zipursky, L.S. et al.* Molecular cell biology. – New York : W.H. Freeman and Company, 2000.

*Weaver, R.F.* Molecular biology. – New York : McGraw-Hill Companies, 2002.

## Розділ 1

# ФІЗИКО-ХІМІЧНІ ОСНОВИ МОЛЕКУЛЯРНОЇ БІОЛОГІЇ

All science is either physics or  
stamp collection.

*E. Rutherford*

Незважаючи на високу складність живих систем, в основі їхнього функціонування лежать такі самі закони фізики, які виконуються для неживої природи – і жодних інших. Сьогодні таке твердження може здаватися банальним, однак саме воно є базовим для сучасної парадигми молекулярної біології. Цей розділ має на меті нагадати деякі найважливіші фізичні закономірності, без розуміння яких неможливо розглядати принципи структурної організації біологічних макромолекул і взаємодії між ними.

## Енергетика взаємодії між атомами й молекулами

### *Вільна енергія*

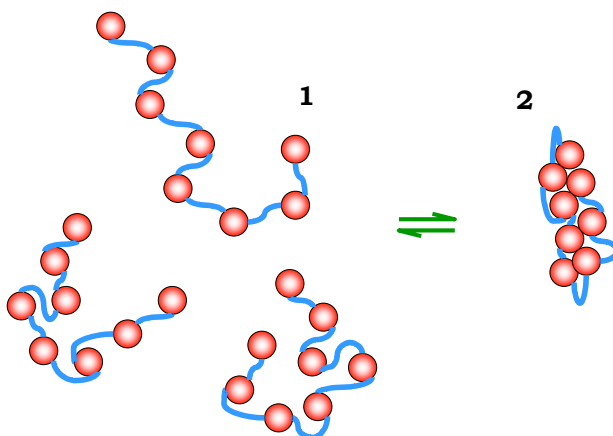
Розглянемо спрощену модель полімерного ланцюга: мономерні елементи, зображені як червоні кульки на рис. 1.1, що з'єднані гнучкими перемичками. Нехай мономери при цьому є інертними – не здатними до взаємодій між собою. Просторова форма – *конформація* – такого ланцюга буде визначатися лише конфігурацією перемичок. Унаслідок теплового руху ланцюг буде постійно змінювати свою конформацію, тобто буде реалізована так звана форма неупорядкованого клубка.

На рис. 1.1 цей стан неупорядкованого клубка зображено як стан 1 ліворуч (представлено кілька з великої кількості  $W_1$  можливих конформацій ланцюга). Певний компактний стан 2 (праворуч) може бути реалізований значно меншою кількістю  $W_2$  конформацій перемичок. Відношення ймовірності  $f_2$  знайти ланцюг у стані 2 до ймовірності  $f_1$

стану 1 (константа рівноваги  $K$  між станами 2 і 1), очевидно, визначається відношенням кількості відповідних конформацій:

$$K = \frac{f_2}{f_1} = \frac{W_2}{W_1}.$$

Оскільки  $W_2 \ll W_1$ , то рівновага зсунута ліворуч – у бік клубка. Отже, за інших однакових умов, імовірність знайти будь-яку систему в певному стані визначається кількістю мікростанів (конфігурацій системи), якими цей стан реалізується. Чим більша кількість мікростанів (у нашому прикладі  $W_1$  та  $W_2$ ), тим менш упорядкованим є даний стан системи: зрозуміло, що ступінь упорядкованості значно зростає для стану 2.



**Рис. 1.1.** Клубок (1) і компактний стан (2) полімерного ланцюга

Нехай тепер мономери здатні взаємодіяти таким чином, що при утворенні контакту між ними в навколишнє середовище виділяється тепло (знижується внутрішня енергія нашої системи). Як відомо з термодинаміки, усі процеси за інших однакових умов відбуваються в напрямку виділення тепла. Імовірність певного стану системи визначається відношенням **ентальпії**  $H$  (величина, яка за умов постійного тиску та об'єму є еквівалентною внутрішній енергії системи) до середньої кінетичної енергії теплового руху  $k_B T$  ( $k_B$  – константа Больцмана (Ludwig Boltzmann),  $T$  – абсолютна температура). А саме, імовір-

## Розділ 1. Фізико-хімічні основи молекулярної біології

ність є пропорційною до величини  $\exp\left(-\frac{H}{k_B T}\right)$ : менше значення ентальпії відповідає зростанню ймовірності (принцип Больцмана). Якщо зниження ентальпії при компактизації ланцюга є значним, то це може змінити конформаційну рівновагу на користь стану 2. Таким чином, у загальному вигляді маємо тепер записати:

$$K = \frac{f_2}{f_1} = \frac{W_2 \exp(-H_2 / k_B T)}{W_1 \exp(-H_1 / k_B T)},$$

або, позначивши  $S = k_B \ln W$  і  $G = H - TS$ ,

$$K = \frac{f_2}{f_1} = \frac{\exp\left(-\frac{H_2 - TS_2}{k_B T}\right)}{\exp\left(-\frac{H_1 - TS_1}{k_B T}\right)} = \frac{\exp\left(-\frac{G_2}{k_B T}\right)}{\exp\left(-\frac{G_1}{k_B T}\right)} = \exp\left(-\frac{\Delta G}{k_B T}\right), \quad (1.1)$$

де  $\Delta G = G_2 - G_1$ .

Останнє дуже важливе співвідношення має універсальне значення. У рівнянні (1.1) величина  $S$ , що залежить від кількості мікростанів, є мірою неупорядкованості системи й називається **ентропією**. Величина

$$G = H - TS, \quad (1.2)$$

яка була введена фундатором термодинаміки Гіббсом (Josiah Willard Gibbs), називається **вільною енергією**. Саме вона в загальному випадку визначає ймовірність перебування системи в певному стані й, відповідно, напрямку, у якому відбувається той чи інший процес. Наступні твердження, що базуються на рівняннях (1.1), (1.2), виконуються для будь-яких систем, у тому числі (і повною мірою) для біологічних та молекулярно-біологічних:

- Зміна стану системи відбувається в напрямку зниження вільної енергії, або: процес може відбуватися тільки тоді, коли він супроводжується зниженням вільної енергії.
- Рівноважний стан, що може бути стабільним скільки завгодно довго, відповідає мінімуму вільної енергії.
- Зміна вільної енергії залежить від двох складових: зміни ентальпії (внутрішньої енергії), яка супроводжується виділен-

ням / поглинанням тепла; зміни ентропії, що характеризує зростання впорядкованості / невпорядкованості в системі.

При цьому величина ентропійного внеску зростає з температурою (рівняння (1.2)), тобто за високих температур переважними завжди стають менш упорядковані стани.

Слід зробити також кілька інших важливих зауважень:

- Вільна енергія є величиною відносною: має сенс тільки *різниця* між вільними енергіями різних станів *тієї самої* системи.
- Різниця між вільними енергіями двох станів системи вказує тільки на напрямок перетворення одного стану в інший і *нічого* не говорить про те, наскільки швидко та яким шляхом буде відбуватися цей процес. Інакше кажучи, стан із меншою величиною вільної енергії неодмінно має бути реалізованим, але може виявитися, що на це потрібно дуже багато часу, і тому практично даний процес не спостерігається.
- Енергія взаємодії між елементами системи (атомами, молекулами) не обов'язково залежить тільки від ентальпійної складової, як у простому прикладі, що його ілюструє рис. 1.1. Насправді, як буде показано нижче, вільна енергія взаємодій може мати суттєву ентропійну складову або навіть залежати тільки від ентропійних ефектів.

### ***Рівновага утворення комплексів молекул у розчині***

Перш ніж перейти до розгляду типів взаємодій між атомами й молекулами, розглянемо ще один приклад рівноваги у простій, але важливій для молекулярної біології системі. У розчині певного об'єму містяться молекули двох типів (А і В), які здатні утворити між собою комплекс С (рис. 1.2):



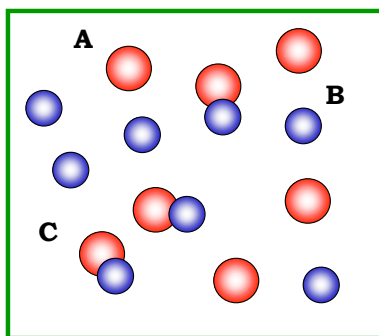
(літерою *K* тут знову позначено константу рівноваги, у цьому випадку – для процесу утворення / дисоціації комплексів). Ентропійна частина вільної енергії кожного компонента містить складову, яка залежить від концентрації. Наприклад, вільну енергію  $G_A$  компонента А (що її в цьому випадку називають також хімічним потенціалом) можна записати як

$$G_A = G_A^0 + RT \ln[A], \quad (1.4)$$



## Розділ 1. Фізико-хімічні основи молекулярної біології

де  $G_A^0$  – стандартна вільна енергія;  $R$  – універсальна газова стала, яка є константою Больцмана в перерахунку на моль речовини:  $R = k_B T N_{AV}$ ,  $N_{AV}$  – число Авогадро (Amedeo Avogadro); квадратними дужками позначено молярну концентрацію  $A$ . Зміст рівняння (1.4) дуже простий:  $G_A^0$  є вільною енергією компонента  $A$  при його концентрації 1 моль/л; зростання концентрації призводить до зниження ентропії (чим більше молекул, тим менше доступних місць розташування для кожної з них, тобто тим меншим є кількість мікростанів, див. наведене вище визначення ентропії) і, відповідно, – до зростання вільної енергії.



**Рис. 1.2.** Взаємодія між молекулами  $A$  і  $B$  з утворенням комплексу  $C$

Записавши аналогічні вирази для компонентів  $B$  і  $C$ , знайдемо різницю між правою та лівою частинами рівняння (1.3)  $\Delta G = G_C - G_A - G_B$ :

$$\Delta G = \Delta G^0 + RT \ln \frac{[C]}{[A][B]}, \quad (1.5)$$

де  $\Delta G^0$  – різниця між відповідними стандартними вільними енергіями,  $[A]$  і  $[B]$  – концентрації відповідних *вільних* компонентів у дисоційованому стані. Процес зміни концентрацій у системі буде продовжуватись до моменту досягнення рівноваги – такого співвідношення між концентраціями, коли вільна енергія комплексу дорівнюватиме сумарній енергії вільних компонентів  $A$  і  $B$ , і  $\Delta G = 0$ . (Можна показати, що загальна вільна енергія розчину буде мінімальною, коли різниця  $\Delta G$  між правою і лівою частинами рівняння (1.3) дорівнює нулю.)

Оскільки  $\Delta G^0$  – це константа, яка залежить від природи компонентів, співвідношення концентрацій під логарифмом у рівнянні (1.5) також є константним за рівноважних умов. Саме це співвідношення і є в даному прикладі константою рівноваги, яка залежить тільки від різниці стандартних вільних енергій:

$$K = \frac{[C]}{[A][B]} = \exp\left(-\frac{\Delta G^0}{RT}\right). \quad (1.6)$$

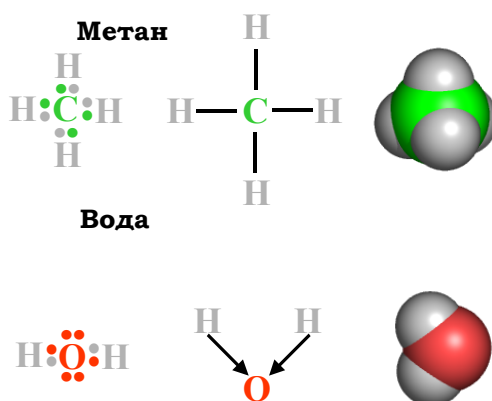
Отже, якщо при утворенні комплексу вільна енергія знижується ( $\Delta G^0 < 0$ ), це сприяє взаємодії між компонентами. У випадку, що розглядається, константу  $K$  називають також *константою асоціації*, або *константою зв'язування*, а величину  $\Delta G^0$  – стандартною *вільною енергією зв'язування*. Обидві константні величини є мірою *спорідненості* компонентів, що утворюють комплекс: чим негативнішою є енергія зв'язування (чим більше енергетично вигідних взаємодій реалізується між компонентами), тим вище значення  $K$  і тим стабільнішим є комплекс. Слід також зауважити, що ефективність утворення комплексу, під якою можна розуміти, наприклад, частку компонента А у складі комплексу –  $[C]/([C]+[A])$ , залежить від концентрації компонентів. Рівняння 1.6 є однією з форм запису відомого *закону діючих мас*, або принципу Ле Шательє (Henry Le Chatelier), згідно з яким додавання в систему одного з компонентів зсуває рівновагу в протилежний бік.

Рівняння 1.3 може означати також рівняння *хімічної реакції*, коли внаслідок перебудови *ковалентних зв'язків* (див. нижче) у молекулах А і В, утворюється нова молекула С. Тоді рівняння (1.4)–(1.6) і висновки з них також зберігають свою силу. У цьому випадку також знак зміни стандартної вільної енергії вказує на напрямок, у якому відбувається хімічна реакція. Варто підкреслити ще раз: **усі** хімічні реакції в живих системах відбуваються в напрямку зниження вільної енергії. Тут може виникнути запитання: якщо, наприклад, вільна енергія знижується при руйнуванні молекули С на А і В, але клітина має потребу саме в молекулі С, чи можливий синтез С? Відповідь є однозначною: синтез С із молекул А і В є неможливим. Але можна використати “обхідний” шлях: спочатку хімічно модифікувати одну з молекул (наприклад, утворити молекулу А') і тоді може виявитися, що утворення С із А' і В супроводжується зниженням вільної енергії, тобто є можливим. Повернемося до обговорення цього питання в розділі 2.

## Ковалентний зв'язок між атомами в молекулах

### Властивості ковалентного зв'язку

Ковалентний (або хімічний) зв'язок – це єдина сила, що підтримує цілісність певної системи атомів, яка називається молекулою. Основою утворення ковалентного зв'язку є узагальнення електронів двох атомів – ефект, що описується у квантовій механіці. Результатом узагальнення є формування так званої молекулярної орбіталі, яка “належить” двом атомам одночасно й на якій може міститися не більше двох електронів (принцип заборони Паулі (Wolfgang Pauli)).



**Рис. 1.3.** Структура молекул метану й води.

Показано пари електронів, які відповідають ковалентним зв'язкам.  
Стрілки символізують зсув електронної щільності до атома оксигену

Наприклад, на зовнішньому електронному шарі атома карбону знаходиться 4 електрони (рис. 1.3), максимально в цьому шарі може бути 8 електронів. Він “доповнюється” електронами чотирьох атомів гідрогену (по одному електрону на кожному атомі), і в результаті утворюється молекула метану. Аналогічно доповнення зовнішнього шару оксигену за рахунок двох атомів гідрогену призводить до утворення молекули води (рис. 1.3). Між двома атомами, залежно від їхньої природи, можливо також формування двох (іноді й трьох) молекулярних орбіталей (по два електрони на кожній), тобто утворення подвійного (чи потрійного) ковалентного зв'язку.

Головні властивості ковалентного зв'язку:

- Ковалентний зв'язок дуже міцний. Інакше кажучи, кожна молекула (про це говорить сам факт її існування) знаходиться в досить глибокій “потенційній ямі” – мінімумі вільної енергії.
- Кожен атом містить певну кількість електронів, які можуть бути узагальненими з іншими атомами (або “вакансій” для електронів на зовнішньому шарі). Тобто кожен атом може утворити лише певну кількість ковалентних зв'язків – має певну *валентність* (наприклад, валентність карбону дорівнює 4, оксигену – 2, водню – 1 (див. рис. 1.3)).
- Ковалентний зв'язок характеризується певною геометрією. По-перше, кожен зв'язок має конкретну довжину (зазвичай 1–2 Å). По-друге, електронні щільності різних орбіталей у молекулі взаємодіють між собою, і в результаті кожен зв'язок певним чином орієнтований у просторі відносно суміжних зв'язків. Типові значення валентних кутів (кутів між зв'язками, що йдуть від одного атома) становлять близько 109° або 120°.

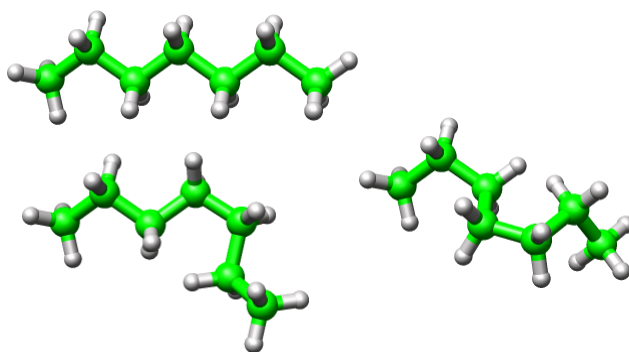
Слід також зауважити, що ступінь узагальнення електронної пари між двома атомами залежить від природи цих атомів. Наприклад, у випадку зв'язку С–Н електронна щільність розподілена приблизно рівномірно між двома атомами. Такий зв'язок називають *неполяризованим*. У випадку зв'язку О–Н, унаслідок високої *електронегативності* оксигену (здатності відтягувати на себе електронну щільність), електронна пара зсунута до цього атома (*поляризований* зв'язок показано стрілочками на рис. 1.3). У результаті в молекулі води на оксигені виникає парціальний негативний (приблизно  $-2/3$  від заряду електрона), а на кожному водні позитивний ( $+1/3$ ) заряд. Іншим важливим для біополімерів прикладом електронегативного атома є азот.

### ***Механізм гнучкості полімерного ланцюга***

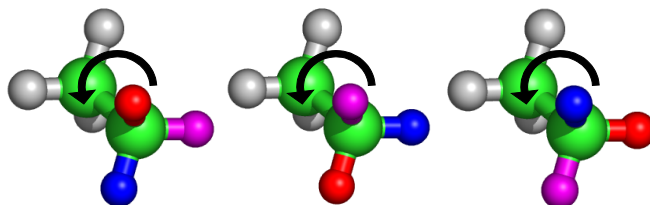
Полімерний ланцюг містить велику кількість атомів, і в його складі певні мономерні одиниці з'єднані між собою ковалентними зв'язками. У прикладі, наведеному на рис. 1.1, було припущено (і це справді так для реальних полімерів), що такий ланцюг є дуже гнучким і здатен існувати у великій кількості конформацій. Але в чому полягає механізм переходу від однієї конформації до іншої? *A priori* можна уявити три можливості: 1) зміна довжини ковалентного зв'язку; 2) зміна валентних кутів; 3) обертання навкруг ковалентного зв'язку. За умов кімнатної (чи фізіологічної) температури (20–37 °С) ковалентний зв'язок є дуже жорстким щодо повздовжніх деформацій: практично такі температури не здатні збу-

## Розділ 1. Фізико-хімічні основи молекулярної біології

дити навіть коливання довжини. Щодо коливань валентних кутів, вони є можливими за фізіологічної температури, але амплітуда цих коливань є досить невеликою (до  $5^\circ$ ). Отже, головну роль у гнучкості полімерного ланцюга відіграє обертання навкруг ковалентних зв'язків (рис. 1.4). Власне, конформація часто й описується в термінах кутів обертання.



**Рис. 1.4.** Кілька конформацій вуглеводневого ланцюжка (зелений – С, сірий – Н), отримані шляхом зміни кутів обертання навкруг зв'язків С–С



**Рис. 1.5.** Три обертальні ізомери, що є переважними для молекули етану. Атоми гідрогену при одному з карбонів забарвлені різними кольорами

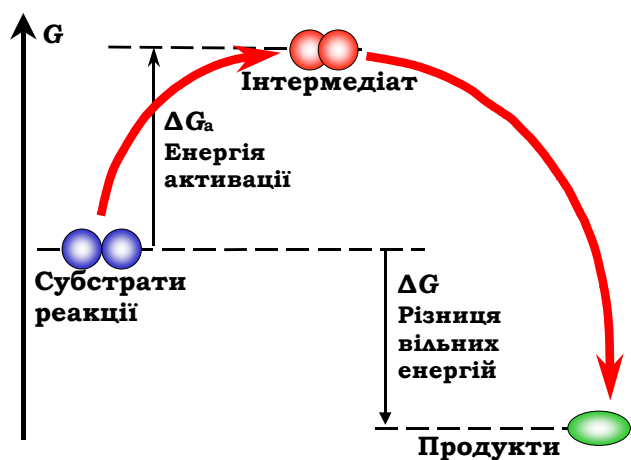
Два важливі зауваження:

- Обертання є можливим лише навкруг одинарного зв'язку. Подвійний зв'язок, унаслідок особливостей будови задіяних у його утворенні молекулярних орбіталей, допускає лише дві обертальні конформації: із кутами обертання  $0^\circ$  і  $180^\circ$ . Хімічні групи, приєднані до атомів, з'єднаних таким зв'язком, містяться тоді відповідно у *цис*- і *транс*-орієнтаціях.
- Обертання навкруг одинарного зв'язку також не є абсолютно вільним. Взаємодія між суміжними хімічними групами приво-

дить до того, що певні кути обертання є переважними (вигіднішими енергетично). Інакше кажучи, практично реалізуються лише певні діапазони значень кута обертання – різні конформації являють собою так звані обертальні ізомери (рис. 1.5).

### **Перебудови системи ковалентних зв'язків у хімічних реакціях**

Хімічну реакцію можна визначити як процес перебудови системи ковалентних зв'язків у межах однієї молекули чи групи молекул (субстрати реакції) з утворенням нових молекул-продуктів. Варто зауважити, що ковалентний зв'язок практично ніколи не руйнується в хімічних реакціях (це дуже не вигідно енергетично) – він просто замінюється на зв'язок з іншим атомом. Перший принцип, який уже було сформульовано в цьому розділі, полягає в тому, що реакція відбувається тоді й тільки тоді, коли вільна енергія продуктів реакції є нижчою, ніж вільна енергія субстратів (рис. 1.6).



**Рис. 1.6.** Енергетична схема процесу хімічної реакції

Ця умова є лише необхідною, але не достатньою. Оскільки кожна молекула вже існує в мінімумі вільної енергії (система ковалентних зв'язків є дуже стабільною), на шляху перебудови електронної щільності неодмінно має існувати перехідний стан (інтермедіат) з більш високою енергією. Це означає, що електронну систему субстратів необ-

хідно спочатку збудити, відхиливши її від рівноважного стабільного стану, після чого вона автоматично “впаде” в стан продуктів, оскільки цей стан знаходиться нижче на шкалі вільної енергії.

Отже, енергія інтермедіатного стану неодмінно є вищою за енергію субстратів: наявність інтермедіату створює на шляху кожної хімічної реакції енергетичний бар'єр, величина якого називається *енергією активації* (рис. 1.6). При цьому вільна енергія активації містить також ентропійну складову: реакція відбувається тільки тоді, коли субстрати стикаються в певній взаємній орієнтації – лише в одній із великої кількості орієнтацій при випадкових стиканнях. Тобто малоімовірна подія є передумовою утворення інтермедіату – різниця ентропії між інтермедіатом і субстратами (ентропія активації) є негативною величиною, тобто дає позитивний внесок у вільну енергію активації.

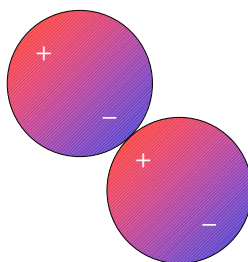
Енергія активації, згідно з принципом Больцмана (рівняння (1.1)), визначає імовірність переходу в інтермедіатний стан, тобто швидкість реакції. Великий (порівняно з тепловою енергією  $k_B T$ ) енергетичний бар'єр означає, що реакція уповільнюється настільки, що практично не відбувається. Найпростіший шлях прискорити реакцію – підвищити температуру (збільшити енергію теплового руху). Зрозуміло, що в живих системах таке рішення є неможливим. Кожна хімічна реакція, яка відбувається в клітині, потребує (для того, щоб вона могла відбуватися за розумний час) спеціального молекулярного пристрою – каталізатора, який прискорює реакцію за рахунок зниження енергії активації. Роль біокаталізаторів виконують у більшості випадків (але не завжди) білкові ферменти (ензими), механізми роботи яких розглядатимуться в розділі 2.

## **Нековалентні взаємодії між атомами й молекулами**

### ***Вандерваальсові взаємодії***

Навіть у межах неполяризованої молекули розподіл електронної щільності є рівномірним лише в разі його усереднення за часом. У кожному момент часу внаслідок флуктуацій щільності виникають локальні неоднорідності, тобто парціальні заряди різного знаку, і молекула являє собою тимчасовий *диполь*. При наближенні іншої молекули такий тимчасовий диполь першої молекули індукує диполь у межах другої молекули, унаслідок чого виникає притягання електричної природи (рис. 1.7). Квантова механіка свідчить, що енергія пари молекул, електронні сис-

теми яких поляризовані, є нижчою за енергію пари неполяризованих молекул. У результаті всі молекули та атоми на близьких відстанях притягуються один до одного. Взаємодії такого типу були описані Ван дер Ваальсом (Johannes Van der Waals) і носять його ім'я.



**Рис. 1.7.** Взаємодія між тимчасовими диполями двох молекул

Вандерваальсові взаємодії між неполяризованими молекулами називають дисперсійними; взаємодії між двома поляризованими молекулами (постійними диполями) – орієнтаційними (або диполь-дипольними); між неполяризованою молекулою та постійним диполем – індукційними (або взаємодіями типу диполь-індукований диполь).

Основні властивості вандерваальсових взаємодій:

- Вандерваальсові взаємодії, які призводять до притягання на малих відстанях, реалізуються між *будь-якими* молекулами й атомами незалежно від їхньої природи.
- Вандерваальсові взаємодії є короткодійними: енергія взаємодій швидко знижується з відстанню між молекулами чи атомами  $r$  (пропорційно до  $1/r^6$ ). Тобто вандерваальсові взаємодії є суттєвими при безпосередньому контакті між молекулами (атомами, хімічними групами).
- Слова “безпосередній контакт” означають, що існує певна оптимальна величина  $r_0$  – відстань, на якій енергія взаємодій є мінімальною. При зниженні  $r$  притягання між молекулами змінюється на відштовхування: електронні оболонки різних молекул (уже насичені валентними електронами) не можуть проникати одна в одну (відповідно до принципу заборони Паулі). Величина  $r_0$  характеризує розмір кожного атома (вандерваальсів радіус), який використовується при побудові молекулярних моделей (як, наприклад, на рис. 1.3). Непроник-



ність атомів один для одного (стеричні обмеження) має велике значення для структури біополімерів: серед великої кількості конформацій можливими є лише ті, для яких не спостерігається проникнення однієї частини ланцюга в іншу (два атоми не можуть займати одну область простору).

- Вандерваальсові взаємодії є найслабшими серед усіх нековалентних взаємодій. Для порівняння: величина середньої кінетичної енергії теплового руху  $RT$  при  $27\text{ }^\circ\text{C}$  становить  $0,6$  ккал/моль; типові значення енергії вандерваальсових взаємодій (на відстані  $r_0$ ) варіюють у межах  $0,1\text{--}0,5$  ккал/моль.

Останнє твердження потребує уточнення стосовно макромолекул. Коли взаємодіють досить великі поверхні, між якими реалізується велика кількість міжатомних контактів, вандерваальсові взаємодії стають суттєвим фактором стабілізації структури макромолекул і макромолекулярних комплексів. Для того, щоб таких контактів було справді багато (щоб не було пустот), дві поверхні повинні добре відповідати одна одній (як ключ до замка) – бути просторово *комплементарними*. Просторова комплементарність між поверхнями макромолекул є одним із важливих механізмів специфічної взаємодії між ними.

### *Електростатичні (іонні) взаємодії*

Певні хімічні групи в біологічних макромолекулах здатні іонізуватися, тобто набувати постійного негативного чи позитивного заряду, втрачаючи чи набуваючи протон. Взаємодія між такими зарядами підпорядковується добре відомому закону Кулона (Charles de Coulomb), а саме, вільна енергія взаємодії двох зарядів  $q_1$  і  $q_2$  залежить від відстані між зарядами  $r$  як

$$U = \frac{q_1 q_2}{\epsilon r},$$

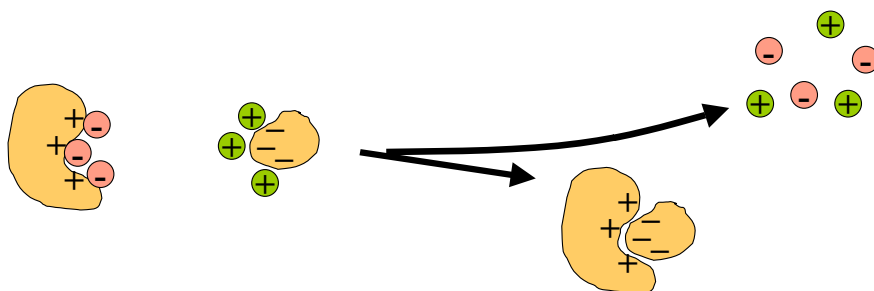
де  $\epsilon$  – діелектрична проникність середовища. При цьому слід звернути увагу на кілька досить важливих моментів.

Величина  $\epsilon$  залежить від полярності середовища. Поляризованість молекул води призводить до досить великого значення проникності ( $\epsilon \approx 80$  за кімнатної температури). У неполярному середовищі проникність є значно нижчою (наприклад, усередині молекули білка  $\epsilon \approx 3$ ) і, відповідно, електростатична енергія зростає. Занурення зарядженої

групи або навіть пари наближених різнойменних зарядів у неполярне середовище є дуже не вигідним енергетично: заряджені групи частіше залишаються на поверхні макромолекул, не занурюючись усередину.

Висока діелектрична проникність води свідчить про те, що електростатичні взаємодії у воді пов'язані в першу чергу з упорядкованістю диполів води навкруг зарядів і зміною цієї впорядкованості при наближенні чи віддаленні зарядів. Тобто електростатичні взаємодії мають ентропійну природу – при наближенні (взаємній нейтралізації) різнойменних зарядів ступінь упорядкованості диполів води знижується. Відповідно, зростає ентропія і знижується вільна енергія.

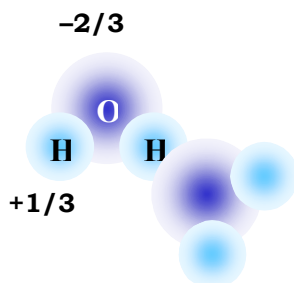
Найвагомішими електростатичні взаємодії є для макромолекул, на поверхні яких сконцентрована велика кількість заряджених груп. На перший погляд здається очевидним, що дві макромолекули з позитивно та негативно зарядженими поверхнями мають взаємодіяти за законом Кулона. Але крім води в живих системах завжди присутні неорганічні іони, які, за тим самим законом Кулона, мають зв'язуватись із зарядженими поверхнями макромолекул (рис. 1.8). Права частина рис. 1.8 енергетично краща за ліву лише одним, а саме звільненням неорганічних іонів у зовнішній розчин, тобто більш низькою впорядкованістю (вищою ентропією та нижчою вільною енергією). Таким чином, взаємодія двох заряджених макромолекул – цілком ентропійний ефект. Величина цього ефекту знижується зі зростанням концентрації неорганічних іонів (іонною силою): якщо іонів досить багато в зовнішньому розчині, ефект від їхнього звільнення є дуже незначний. Тоді взаємна спорідненість макромолекул знижується і комплекс дисоціює.



**Рис. 1.8.** Ентропійна природа електростатичних взаємодій між високзарядженими макромолекулами

### Водневий зв'язок

Розглянемо дві молекули води (рис. 1.9). Унаслідок високої електронегативності оксигену молекула води є полярною, тобто на її атомах присутні парціальні електричні заряди (приблизно  $-2/3$  заряду електрона на оксигені,  $+1/3$  на кожному гідрогені). Відповідно, дві молекули будуть взаємодіяти одна з одною електростатично. Але тут є ще дві важливі обставини: 1) атом гідрогену має лише один електрон, і цей електрон є відтягнутим на оксиген у молекулі води; 2) атом оксигену містить на зовнішньому шарі неподілену пару електронів (рис. 1.3), які не залучені до утворення ковалентного зв'язку. У результаті “слабка” електронна оболонка гідрогену “продавлюється” атомом оксигену другої молекули; його неподілена пара електронів частково узагальнюється з гідрогеном, електрон якого є відтягнутим “власним” оксигеном. Взаємодія такого типу й називається водневим зв'язком, який, являє собою дещо середнє між електростатичною взаємодією (притягання парціальних зарядів) і ковалентним зв'язком (часткове узагальнення пари електронів між двома атомами).



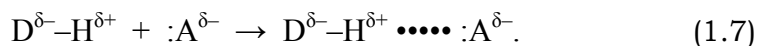
**Рис. 1.9.** Дві молекули води, зв'язані водневим зв'язком

Типові пари хімічних груп, що утворюють водневі зв'язки в макромолекулах:



Тобто водневий зв'язок утворюється між атомом гідрогену, приєднаним до електронегативного атома з частковим негативним зарядом  $\delta^-$  (ця група “віддає” протон і називається *донором* водневого зв'язку)

та електронегативним атомом іншої хімічної групи чи молекули (акцептором), який містить неподілену пару електронів:



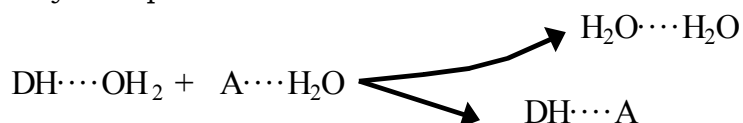
Група С–Н, наприклад, не відповідає цим принципам (зв'язок С–Н неполярний, оскільки С не є електронегативним атомом) і не може бути залученою до утворення водневого зв'язку.

Дві властивості водневого зв'язку є наслідком його природи:

- Водневий зв'язок є дуже вигідним енергетично: його енергію (утворення водневого зв'язку супроводжується виділенням тепла, тобто він має ентальпійну природу) можна оцінити приблизно в 5 ккал/моль (порівняйте з енергією теплового руху  $RT = 0,6$  ккал/моль).
- На відміну від електростатичних взаємодій, але так само, як ковалентний зв'язок, водневий зв'язок є чутливим до взаємної орієнтації донора й акцептора. Особливо це стосується напрямку донорної групи: ковалентний зв'язок D–H має бути спрямованим на акцепторний атом (припустиме відхилення зв'язку від напрямку водневого зв'язку становить 20–30°).

Друга властивість є дуже важливою. Вона означає, що водневі зв'язки між макромолекулами утворюються тільки тоді, коли між донорними та акцепторними групами є чітка просторова відповідність. Це робить водневий зв'язок найсуттєвішим фактором у визначенні *специфічності* міжмолекулярних взаємодій – міжмолекулярному впізнанні.

Перша властивість також потребує уточнень. Якщо взаємодія між донором і акцептором відбувається у водному середовищі, то рівняння (1.7) має бути переписане таким чином:

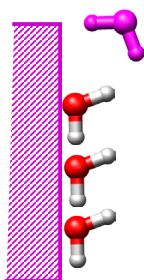


Як бачимо, кількість водневих зв'язків є однаковою ліворуч і праворуч (ентальпія системи не змінюється), різною може бути лише ентропія лівого та правого станів. Якщо донор і акцептор – маленькі молекули, то переважним буде лівий стан (велика кількість молекул води робить його менш упорядкованим): маленькі полярні молекули розчиняються у воді, тому що добре з нею взаємодіють. Якщо донор і акцептор належать макромолекулі, переважним може стати правий

стан: пара молекул води може вільніше рухатися, що збільшує ентропію системи. Отже, водневі зв'язки у водному середовищі мають ентропійну природу. Зрозуміло, що в неполярному середовищі (де немає конкуренції з водою) водневий зв'язок буде значно стабільнішим.

### **Гідрофобні взаємодії**

Сильна взаємодія молекул води за рахунок водневих зв'язків і те, що ці зв'язки можливі лише за певної взаємної орієнтації молекул, приводить до так званого гідрофобного ефекту при зануренні у воду неполярних молекул. неполярність молекули означає, що вона не здатна утворювати водневі зв'язки. А молекули води не тільки здатні, а й прагнуть обов'язково їх утворити, оскільки енергія зв'язку є високою, втрачати його дуже невигідно. Єдиний вихід для молекул води – утворити зв'язки між собою. Проте неполярна поверхня створює перешкоди: як видно з рис. 1.10, група ОН верхньої молекули “дивиться” на неполярну (гідрофобну) поверхню і не може бути задіяною в утворенні водневого зв'язку. За збереження водневих зв'язків молекулам води доводиться платити зростанням упорядкованості – вони формують поблизу гідрофобної поверхні, як показано на рис. 1.10, упорядкований кригоподібний кластер. Таким чином, занурення гідрофобної молекули у водне оточення за рахунок зниження ентропії є дуже невигідним енергетично.



**Рис. 1.10.** Кластер молекул води біля гідрофобної поверхні. Верхня молекула займає невигідну орієнтацію відносно поверхні

Якщо додати ще одну таку гідрофобну поверхню, у системи буде дуже простий вихід (рис. 1.11): злипання поверхонь супроводжується визволенням молекул води, тобто зростанням ентропії. Отже (знову!), гідрофобні взаємодії мають цілком ентропійну природу.



### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. Дайте визначення вільної енергії. Поясніть, як вона залежить від ентальпії та ентропії системи? В якому напрямку змінюється вільна енергія при переході системи до стану рівноваги?
2. Чи може при переході до стану рівноваги зрости ентальпія системи? За якої умови? Чи може зрости ступінь упорядкованості при переході до рівноважного стану?
3. Скориставшись рівнянням 1.6, розрахуйте відносну частку компонента А у складі комплексу С при концентраціях вільного компонента В, що дорівнюють  $10^{-6}$ ,  $10^{-5}$ ,  $10^{-4}$  моль/л. Вважайте, що  $K = 10^5$  л/моль.
4. Що таке ковалентний зв'язок і які основні властивості йому притаманні? За рахунок чого полімерний ланцюг може змінити свою конформацію?
5. Назвіть дві умови, за яких буде відбуватися хімічна реакція. Якою величиною визначається швидкість реакції?
6. У чому полягає фізичний механізм вандерваальсових взаємодій?
7. Поясніть ентропійну природу іонних взаємодій між двома високорядженими макромолекулами.
8. Які властивості мають бути притаманні двом хімічним групам, щоб між ними міг утворитися водневий зв'язок? Чи можливий водневий зв'язок між двома метильними групами? Між метильною та ОН-групою?
9. Чому водневий зв'язок часто буває важливим для специфічного міжмолекулярного впізнання?
10. Поясніть фізичну природу гідрофобного ефекту.

### **РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

- Гросберг, А.Ю., Хохлов, А.Р.* Фізика в мире полимеров. – М. : Наука, 1989.  
*Кантор, Ч., Шиммел, П.* Биофизическая химия : в 3 т. – М. : Мир, 1984.  
*Ленинджер, А.* Основы биохимии : в 3 т. – М. : Мир, 1985.  
*Полинг, Л. Полинг, П.* Химия. – М. : Мир, 1978.  
*Тэнфорд, Ч.* Физическая химия полимеров. – М. : Химия, 1965.  
*Финкельштейн, А.В. Птицын, О.Б.* Фізика белка. Курс лекцій. – М. : КДУ, 2005.  
*Шредингер, Э.* Что такое жизнь? Физический аспект живой клетки. – М. ; Ижевск: Регулярная и хаотическая динамика, 2002.

## Розділ 2

# БІЛКИ

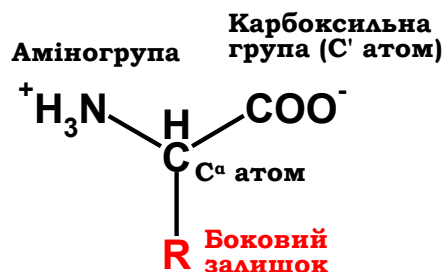
Скорлупы-то золотые,  
Ядра – чистый изумруд...

*А. Пушкин. Сказка о царе  
Салтане*

## Хімічна будова білків

### Амінокислоти

Узагальнену хімічну формулу *амінокислоти* – мономерної одиниці молекули білка (протеїну, protein) – зображено на рис. 2.1.

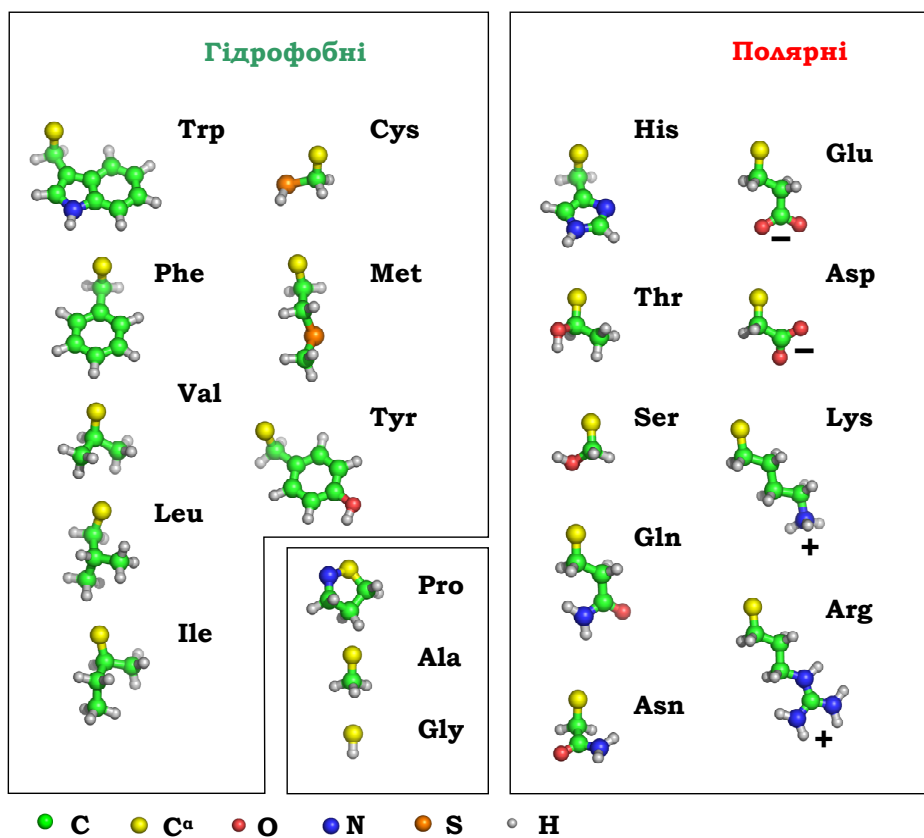


**Рис. 2.1.** Структурна формула амінокислоти

За нейтральних рН аміно- та карбоксильна групи, приєднані до спільного атома карбону ( $C^{\alpha}$ ), перебувають в іонізованому стані. Саме ці, стандартні для всіх амінокислот, групи (наявність яких і зумовила назву цих сполук) забезпечують утворення полімерної молекули білка. Розрізняються амінокислоти за типом бокового (амінокислотного) залишку. Усього у складі білків зустрічається 20 амінокислот (у дужках – загальноприйняті скорочення, три- та однолітерні): аланін (Ala, A),



аргінін (Arg, R), аспарагін (Asn, N), аспарагінова кислота (Asp, D), валін (Val, V), гістидин (His, H), гліцин (Gly, G), глютамін (Gln, Q), глютамінова кислота (Glu, E), ізолейцин (Ile, I), лейцин (Leu, L), лізин (Lys, K), метіонін (Met, M), пролін (Pro, P), серин (Ser, S), тирозин (Tyr, Y), треонін (Thr, T), триптофан (Trp, W), фенілаланін (Phe, F), цистеїн (Cys, C). Їхні бокові залишки зображено на рис. 2.2.



**Рис. 2.2.** Бокові залишки амінокислот, приєднані до C<sup>α</sup> атома, у стані іонізації, який вони мають при рН 7,0. Кольоровий код (унизу) позначення хімічних елементів зберігається на всіх подальших рисунках

Два параметри амінокислотних залишків мають особливе значення для структури білка: розмір і здатність взаємодіяти з водою. За останньою ознакою залишки можна поділити дві групи: полярні й неполярні (гідрофобні). Три гідрофобні залишки (окрема група на

рис. 2.2) – маленького розміру, тому їх можна вважати “нейтральними” щодо спорідненості до води (гідрофобний ефект залежить від розміру поверхні, див. розділ 1). У глобулярних водорозчинних білках співвідношення між гідрофобними та полярними залишками дорівнює в середньому 50 : 50. Унаслідок гідрофобного ефекту неполярні залишки прагнуть опинитися всередині структури білка, полярні залишаються на поверхні, зберігаючи взаємодії з водою. Це і є головна рушійна сила, яка змушує білковий ланцюг укладатися певним чином у просторі, формуючи компакту структуру – так звану *глобулу*.

Полярні амінокислоти тому так і називаються, що прагнуть взаємодіяти з водою за рахунок водневих зв'язків (містять донорні та акцепторні групи). Серед них є залишки, які мають постійний заряд при рН 7,0: негативний у Asp і Glu (за рахунок карбоксильних груп у складі бокового залишку), позитивний у Lys (аміногрупа) та Arg (гуанідинова група). His також може набувати позитивного заряду в структурі білка в певному мікрооточенні (у розчині він іонізується при рН < 6,5).

### ***Пептидний зв'язок і поліпептидний ланцюг***

Карбоксильна група однієї амінокислоти та аміногрупа іншої, втрачаючи молекулу води, утворюють ковалентний зв'язок, який називається *пептидним* (рис. 2.3). Група атомів O–C'–N–H називається *пептидною групою*. Оскільки в утворенні зв'язку між мономерами беруть участь різні хімічні групи, поліпептидний ланцюг має напрям: його кінець, де залишається аміногрупа, називається N-кінцем, протилежний (карбоксильна група) – C-кінцем. Синтезуються поліпептидні ланцюги (розділ 8) у напрямку N → C, у цьому ж напрямку представляються за умовчанням амінокислотні послідовності білків.

Дві властивості пептидної групи відіграють ключову роль в утворенні просторової структури білка:

1) ***Пептидна група є полярною*** – виступає одночасно донором і акцептором водневих зв'язків. Таким чином, остов поліпептидного ланцюга (пептидні групи та C<sup>α</sup>-атоми, що чергуються), від якого відгалужуються полярні та неполярні бокові залишки (див. рис. 2.4), є полярним.

2) Одна з пар електронів, що відповідає зв'язку C'–O (зазвичай позначається як подвійний, як праворуч на рис. 2.3), насправді делокалізована також між атомами C'–N: *пептидний зв'язок є “напівподвійним”*, що унеможливує обертання навкруг нього. Для більшості амінокислот реалізується більш вигідна *транс*-конформація пептидного зв'язку: при-

єднані до пептидної групи  $C^\alpha$ -атоми “дивляться” в різні боки (рис. 2.4; див. також розділ 1). Єдине виключення – пептидний зв’язок перед Pro, приблизно 10 % таких зв’язків знаходяться в *цис*-конформації.

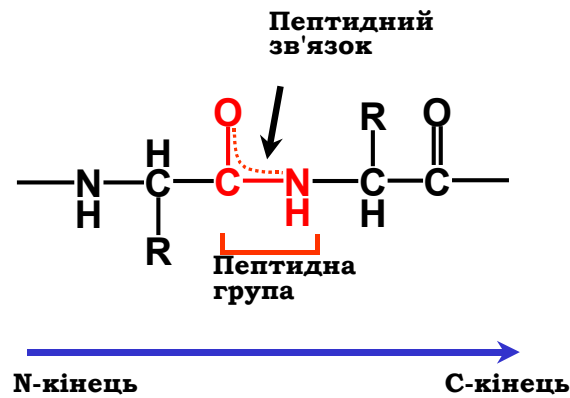


Рис. 2.3. Фрагмент поліпептидного ланцюга

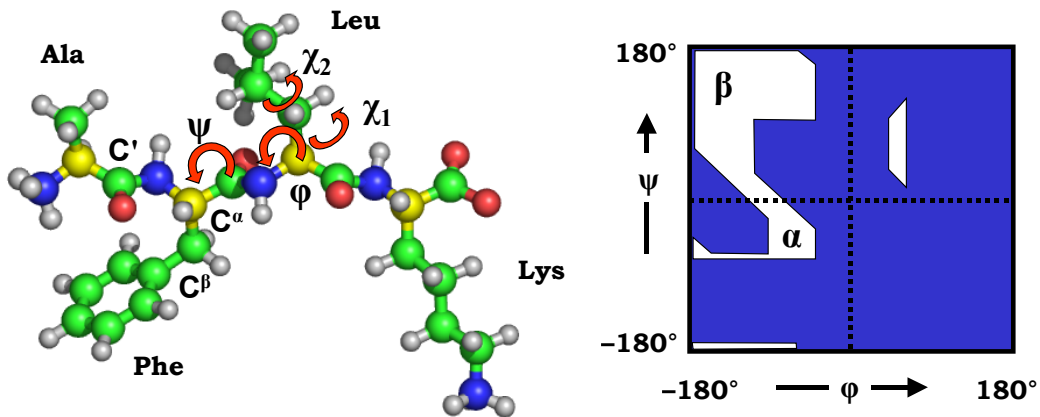


Рис. 2.4. Тетрапептид і приблизна діаграма Рамачандрана: синім кольором позначено заборонені, білим – дозволені для більшості амінокислотних залишків зони. Літери  $\alpha$  і  $\beta$  позначають зони, які відповідають певним регулярним конформаціям

Оскільки обертання навкруг пептидного зв'язку є неможливим, конформація поліпептидного ланцюга може змінюватись тільки за рахунок обертання навкруг суміжних до пептидної групи одинарних зв'язків (обертальні кути  $\phi$  і  $\psi$ , рис. 2.4), а також обертання навкруг зв'язків у межах бокових залишків ( $\chi_1, \chi_2 \dots$ ). Унаслідок стеричних обмежень не всі комбінації кутів обертання  $\phi$  і  $\psi$  є можливими – для поліпептидного ланцюга реалізується обертально-ізомерний механізм гнучкості. Дозволені для більшості залишків зони конформаційного простору в координатах  $\phi$ - $\psi$  (діаграма Рамачандрана (Gopalasamudram Narayana Ramachandran)) показано на рис. 2.4. Діаграми Рамачандрана мають свої специфічні ознаки для окремих залишків. Особливо це проявляється, якщо ввести третю координату – вільну енергію для кожної пари обертальних кутів. Інакше кажучи, переважна (найбільш імовірна) **локальна конформація ланцюга залежить від амінокислотної послідовності**.

## Вторинна структура

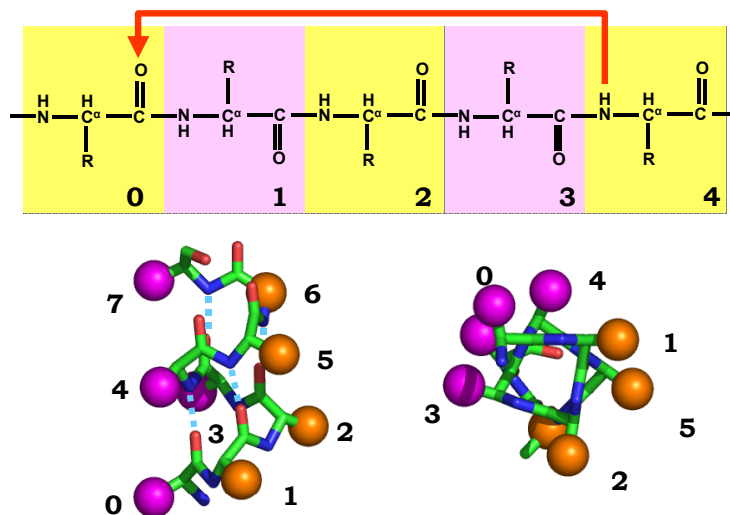
Локальну конформацію амінокислотного ланцюга називають *вторинною структурою* білка. Певна частина (іноді значна) ланцюга міститься в нерегулярній конформації (кути  $\phi$  і  $\psi$  варіюють від одного залишку до іншого). Але в глобулярних білках завжди зустрічаються також *регулярні конформації, стабілізовані водневими зв'язками між пептидними групами*, і термін “вторинна структура” іноді відносять саме до них. Загалом у глобулярних білках представлено всього два типи таких конформацій –  $\alpha$ -спіраль ( $\alpha$ -helix) і  $\beta$ -структура ( $\beta$ -structure), хоча є також інші.

### $\alpha$ -Спіраль

Одна з допустимих конформацій на карті Рамачандрана ( $\phi = -60^\circ$ ,  $\psi = -45^\circ$ , рис. 2.4) передбачає наближення оксигену пептидної групи  $i$ -го залишку до NH-групи  $i + 4$ . Перша група є акцептором водневого зв'язку, друга – донором. Система таких водневих зв'язків між кожним  $i$  та  $i + 4$  залишком стабілізує регулярну конформацію, яка й називається  $\alpha$ -спіраллю (рис. 2.5). Ця спіраль є правою (точка, що рухається вздовж ланцюга, обертається навкруг осі спіралі *за годинниковою стрілкою* й одночасно пересувається вздовж осі *вперед*), на один виток припадає 3,6 амінокислотного залишку, довжина

одного витка вздовж осі (крок спіралі) 5,4 Å. Водневий зв'язок між пептидними групами замикає в кільце 13 атомів чотирьох залишків: відповідно,  $\alpha$ -спіраль позначається також як спіраль  $4_{13}$ .  $\alpha$ -Спіраль можна приблизно уявити як циліндр діаметром  $\approx 10$  Å, на поверхні якого розташовані амінокислотні залишки.

Оскільки найімовірніша локальна конформація ланцюга залежить від послідовності, різні амінокислоти та їхні комбінації мають різний потенціал щодо включення в  $\alpha$ -спіраль (як і в будь-яку іншу вторинну структуру). Наприклад, Pro не може бути включеним до  $\alpha$ -спіралі (немає донора водневого зв'язку). Gly також уникає регулярних структур, оскільки втратив би свою високу конформаційну свободу (ентропійна ціна була б надто високою). При цьому в глобулярних білках спостерігається закономірність: гідрофобні амінокислотні залишки згруповані на поверхні спіралі в більш-менш суцільний гідрофобний кластер (рис. 2.5). Тобто  $\alpha$ -спіраль переважно утворюється на таких ділянках послідовності з високим спіральним потенціалом, які здатні сформувати гідрофобний кластер на поверхні спіралі.



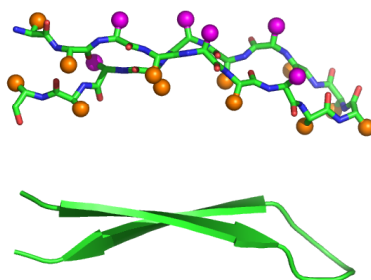
**Рис. 2.5.**  $\alpha$ -Спіраль. Червона стрілочка – водневий зв'язок.

Унизу: дві проєкції  $\alpha$ -спіралі, пунктирні лінії – водневі зв'язки між пептидними групами. Бокові залишки показані схематично, різне забарвлення символізує два типи залишків – полярні й неполярні

Така закономірність пояснюється просто. Унаслідок конкуренції з молекулами води водневі зв'язки між пептидними групами ( $\alpha$  отже, і спіраль) у водному оточенні не є дуже стабільними. Адже утворення спіралі коштує енергетичних витрат – зниження ентропії за рахунок обмеження конформаційної свободи. А водневі зв'язки у присутності води (тобто заміна зв'язку між пептидною групою та водою на зв'язок між пептидними групами) не здатні суттєво компенсувати ці витрати (див. розділ 1). Якщо ж на поверхні спіралі є гідрофобний кластер, він прагне адсорбувати іншу гідрофобну поверхню (спіралі чи інших елементів вторинної структури). Тоді частина пептидних груп і водневих зв'язків між ними опиняється в неполярному оточенні, що суттєво стабілізує і ці зв'язки, і  $\alpha$ -спіраль. У результаті у глобулі залишаються лише  $\alpha$ -спіралі з гідрофобним кластером на поверхні. Це вказує, що формування регулярної вторинної структури та укладання в глобулу – єдиний процес. Нижче при обговоренні глобулярної структури буде розглянуто питання про те, чому взагалі утворюється регулярна вторинна структура в глобулярних білках.

### $\beta$ -Структура

Інший тип регулярної вторинної структури, який часто зустрічається в білках, – це  $\beta$ -структура. Вона складається з витягнутих фрагментів поліпептидного ланцюга: кілька таких фрагментів –  $\beta$ -тяжів ( $\beta$ -strands) – утворюють  $\beta$ -структурний шар, у площині якого розташовані водневі зв'язки між пептидними групами тяжів, а від обох поверхонь відходять амінокислотні залишки (рис. 2.6).



**Рис. 2.6.** Антипаралельна  $\beta$ -структурна шпилька (амінокислотні залишки зображено кульками, як на рис. 2.5).  
Унизу: та сама шпилька в загальноприйнятому схематичному зображенні

Половина пептидних груп  $\beta$ -тяжів, які містяться на межах шару, залишаються ненасиченими водневими зв'язками. Тому чим більше тяжів входить до складу шару, тим більш стабільним він є.  $\beta$ -Шар може бути антипаралельним (сусідні тяжі спрямовані у протилежні один щодо одного напрямки, як на рис. 2.6) або паралельним (звичайно, якщо тяжі не розташовані один за одним по ланцюгу).

Як видно з рис. 2.6,  $\beta$ -шар завжди є дещо скрученим через те, що окремі тяжі не є абсолютно прямими.

Сусідні по ланцюгу амінокислотні залишки розташовані по різні боки  $\beta$ -структурного шару. І знову спостерігається вже зрозуміла закономірність: принаймні одна поверхня шару є переважно гідрофобною.

### *Спіраль $3_{10}$*

Права спіраль  $3_{10}$  (водневий зв'язок між  $i$  та  $i + 3$  залишками) зустрічається не дуже часто у вигляді коротких (три-чотири залишки) фрагментів.

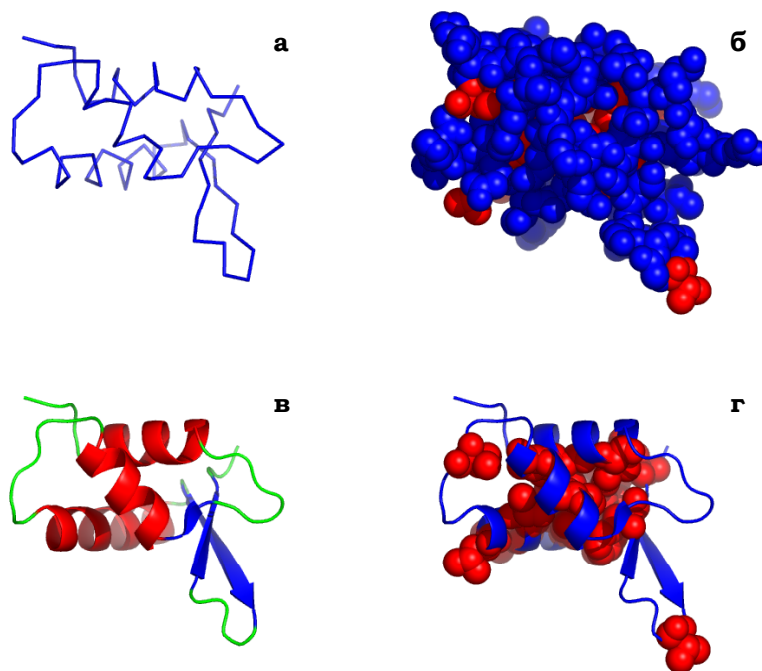
### *$\beta$ -Поворот*

$\beta$ -Поворот або  $\beta$ -вигин ( $\beta$ -turn) – невеликий елемент ланцюга із чотирьох залишків, стабілізований водневим зв'язком між пептидними групами першого та четвертого.  $\beta$ -Поворот забезпечує поворот ланцюга на  $180^\circ$  і часто знаходиться між двома сусідніми  $\beta$ -тяжами в антипаралельних  $\beta$ -шпильках.

## **Глобулярна структура**

Поліпептидний ланцюг глобулярного білка довжиною 50–150 (максимально 250) амінокислотних залишків укладається в єдину щільну глобулу розміром 25–40 Å. Усередині такої глобули утворюється **гідрофобне ядро** з неполярних залишків, заекрановане від води полярною поверхнею глобули. Довші ланцюги утворюють кілька субглобул із власним гідрофобним ядром кожна – структурних *доменів*. Просторову структуру будь-якого поліпептидного ланцюга називають *третинною структурою*. Досить часто молекула білка складається з кількох поліпептидних ланцюгів, які можуть укладатися в єдину щільну глобулу або формувати окремі глобули, що взаємодіють між собою. В обох випадках кажуть про *четвертинну* (субодиничну) структуру білка.

Різні варіанти зображення просторової структури білка показані на рис. 2.7. На частинах (б) і (г) добре видно утворення всередині глобули щільного гідрофобного ядра.



**Рис. 2.7.** Різні варіанти зображення структури глобулярного білка (глобулярний домен гістону H5, код PDB 1HST): (а) – остов ланцюга (з'єднані між собою  $C^{\alpha}$ -атоми); (б) – атомна модель, червоні – атоми неполярних, сині – полярних залишків; (в) – схема будови, де показані вторинні структури; (г) – до попередньої схеми додано атоми неполярних залишків. Усі рисунки мають однакові проекції та масштаб

### *Роль вторинної структури в утворенні глобули*

Розглядаючи структури глобулярних білків (на сьогодні їх установлено понад 40 тис.), можна сформулювати кілька загальних правил.

**Правило 1. Каркас глобули завжди складається з ділянок вторинної структури ( $\alpha$ -спіралей та/або  $\beta$ -структурних шарів), гідрофобні поверхні яких, взаємодіючи між собою, утворюють гідрофобне ядро.**

Чому саме з ділянок регулярної вторинної структури? Це стає зрозумілим, якщо брати до уваги полярність / неполярність елементів поліпептидного ланцюга. Неполярні амінокислотні залишки прагнуть опинитися



всередині глобули внаслідок гідрофобного ефекту – їхній контакт з водою є дуже не вигідним енергетично. Але разом з ними в неполярному оточенні мають опинитися *полярні* пептидні групи остова. Втрата їхніх водневих зв'язків із молекулами води є енергетично дуже не вигідною (енергія водневого зв'язку є дуже високою). Єдиний вихід для пептидних груп – утворити водневі зв'язки між собою, а це можливо тільки в складі ділянок регулярної вторинної структури. Отже, **насичення пептидних груп водневими зв'язками у складі вторинної структури є необхідною передумовою утворення всередині глобули гідрофобного ядра**. Петлі й перемички між ділянками вторинної структури знаходяться в нерегулярній конформації й залишаються на поверхні глобули, їхні пептидні групи при цьому зберігають контакт з водою. Відповідно:

**Правило 2. Сегменти вторинної структури протягнуті від одного краю глобули до іншого. Перемички між ділянками вторинної структури залишаються на поверхні й ніколи не занурюються всередину глобули.**

Обидва правила мають ясний фізичний сенс і могли б бути сформульованими, виходячи із загальних фізичних міркувань.

### *Стабілізація глобули*

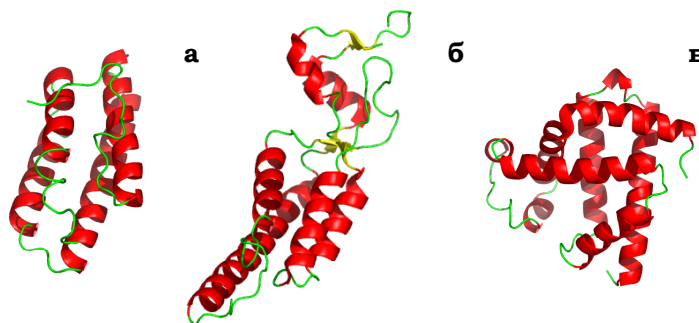
Як зрозуміло з викладеного вище, **головною рушійною силою утворення глобулярної структури є гідрофобні взаємодії**. Але, щоб здійснювати певні функції, білок повинен мати дуже стабільну просторову структуру, яку гідрофобні взаємодії не можуть забезпечити: для їхньої реалізації достатньо просто витіснити воду з інтерфейсу між гідрофобними поверхнями. При цьому поверхні, тобто сегменти вторинної структури, яким вони належать, будуть зберігати певну свободу руху. Саме така ситуація і спостерігається для проміжного стану, який реалізується при укладці білка – так званої *розплавленої глобули* (рис. 2.8). Першим етапом укладання є одночасне й досить швидке утворення вторинної структури та “злипання” сегментів вторинної структури за рахунок гідрофобних взаємодій. Загальна архітектура розплавленої глобули – така сама, як у нативного функціонального білка, але спостерігаються значні рухи ділянок вторинної структури одна відносно одної – розплавлена глобула нагадує краплю рідини, де елементами, що рухаються, є ділянки вторинної структури. На другому етапі реалізується щільна упаковка глобули за рахунок вандерваальсових взаємодій – точна взаємна підгонка амінокислотних залишків з утворенням численних контактів між ними, без порожнин усередині глобули. Результатом є **нативна глобула**, яка є вже не краплею рідини, а **твердим тілом**.



### Структурна класифікація глобулярних білків

Оскільки глобула будується з елементів регулярної вторинної структури, яких є всього два, цілком природно розділити структурні домени глобулярних білків на три групи:  $\alpha$ -спіральні,  $\beta$ -структурні та змішаного типу. Виявляється, що кількість типів взаємної упаковки сегментів вторинної структури – так званих *структурних мотивів*, або типів *укладки* (fold), які забезпечують утворення гідрофобного ядра, – досить обмежене.

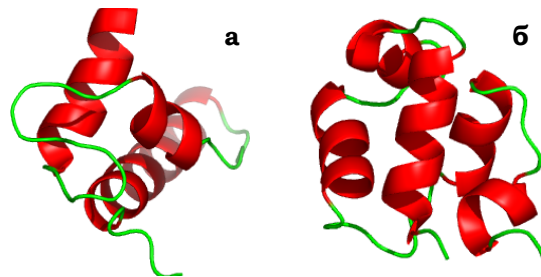
**$\alpha$ -Спіральні білки.** Серед структурних мотивів, які складаються тільки з  $\alpha$ -спіралей, в основному зустрічаються всього два типи. Довгі спіралі (довжиною до 5–6 витків) формують  $\alpha$ -спіральні пучки ( $\alpha$ -helix bundles). Дві (рідше три) спіралі у складі пучка утворюють шар з гідрофобною поверхнею, два такі шари, укладаючись паралельно (під кутом  $\sim 20^\circ$ ) або ортогонально ( $\sim 90^\circ$ ), формують гідрофобне ядро (рис. 2.9). Слід зробити наступні важливі зауваження щодо представлених структур: усі вони не мають гомології між собою щодо амінокислотної послідовності; білки подібної структури (гемеритрин і білок вірусу тютюнової мозаїки) виконують різні функції; білки різної структури (гемеритрин і міоглобін) виконують подібну функцію зв'язування кисню.



**Рис. 2.9.** Спіральні пучки з паралельною (а, б) чи ортогональною (в) упаковкою в гемеритрині (а, 1HMD), у білку оболонки вірусу тютюнової мозаїки (б, 1E17), у міоглобіні (в, 1A6M)

Другим, більш розповсюдженим, структурним мотивом є упаковка трьох або більше  $\alpha$ -спіралей у приблизно сферичну глобулу. Така упаковка реалізується для коротких спіралей, які значно частіше зустрі-

чаються в білках (близько 10–12 залишків, тобто 3–3,5 витка). Приклади наведено на рис. 2.10 (подібна упаковка реалізується також для трьох  $\alpha$ -спіралей, зображених на рис. 2.7). Короткі спіралі в цьому випадку розташовані по ребрах квазісферичного багатогранника, оточуючи приблизно сферичне гідрофобне ядро.



**Рис. 2.10.** Домени актинідину (а, 1AEC) і білка Cro бактеріофага 434 (б, 3CRO)

Указані структурні мотиви зустрічаються найчастіше, оскільки вони найкраще відповідають уже сформульованим вище двом правилам структурної організації білків. Крім того, розглядаючи ці структури можна сформулювати ще одне

**Правило 3. Перемички між сегментами вторинної структури уникають взаємного перехрестя – суміжні по ланцюгу перемички розташовані по різні боки глобули.**

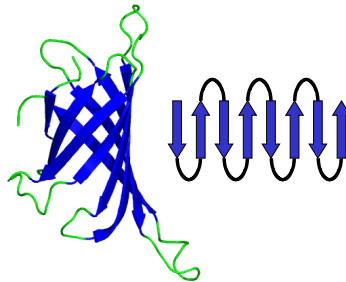
Таке перехрестя привело б або до контакту між перемичками (тобто їхньої дегідратації), або до значного вигину однієї з перемичок. Обидва наслідки є енергетично не вигідними. Інакше кажучи, у край мало можна уявити фрагментів амінокислотної послідовності, які б могли перетнутися без суттєвих енергетичних витрат.

Крім того, описані  $\alpha$ -спіральні мотиви найкраще забезпечують вимогу щільної упаковки глобули – саме за цих типів взаємної орієнтації спіралей їхні поверхні можуть здійснити найбільшу кількість вандерваальсових контактів між собою.

**$\beta$ -Структурні білки.** Основою  $\beta$ -структурних доменів є антипаралельні  $\beta$ -шари. Два шари  $\beta$ -структури, взаємодіючи своїми гідрофобними поверхнями, укладаються один на одного у своєрідний штабель,  $\beta$ -тяжі двох шарів орієнтуються при цьому паралельно або перпендикулярно (ортогональна упаковка) один до одного.

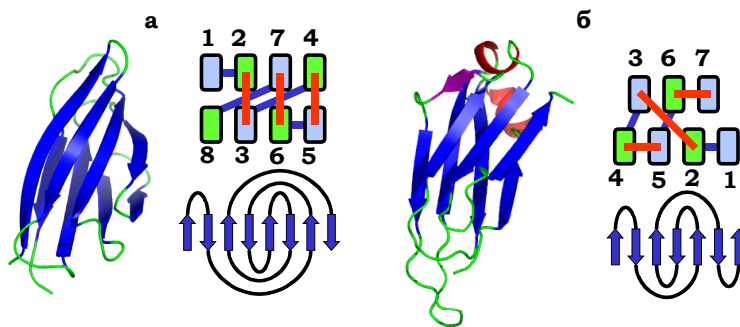
При ортогональній упаковці, унаслідок скрученості  $\beta$ -шарів, форма такого штабеля нагадує циліндр зі значним нахилом тяжів до осі (рис. 2.11). Ця укладка так і називається –  $\beta$ -циліндр, або  $\beta$ -діжка

( $\beta$ -barrel). Проекція ходу ланцюга на площину для такого структурного мотиву виглядає дуже просто: сусідні по ланцюгу ділянки є також сусідами у складі шару, і проекція еквівалентна візерунку “меандр” (часто зустрічається на античних амфорах).



**Рис. 2.11.**  $\beta$ -Циліндр у стрептовідині (1HQQ); праворуч – проекція ходу ланцюга на площину, яка відповідає візерунку “меандр”

Інший тип ходу ланцюга нагадує візерунок “грецький ключ” (рис. 2.12), відповідна укладка являє собою два паралельні шари  $\beta$ -структури ( $\beta$ -сендвіч).

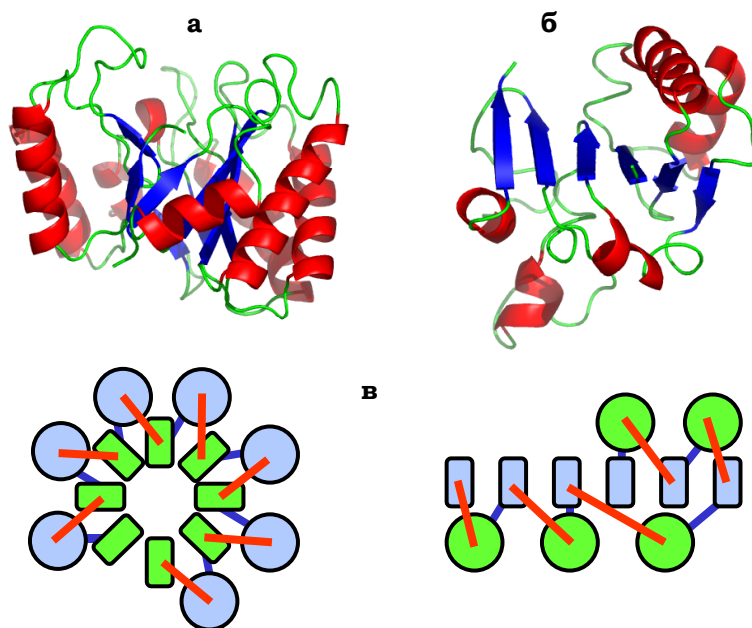


**Рис. 2.12.** (а) – Ксиланзв'язувальний домен ксиланази (1E5B); (б) – імуноглобулінова укладка в константному домені важкого ланцюга імуноглобуліну (15C8). Праворуч показано схеми упаковки двох шарів із торця ( $\beta$ -тяжі пронумеровані, блакитний – N-, зелений – С-кінець фрагмента) і проекції ходу ланцюга на площину (для структури (б) короткі  $\alpha$ -спіралі та один  $\beta$ -тяж, які прикривають торець, на схемах не показано)

Наведені приклади укладання  $\beta$ -структурних доменів з певними варіаціями зустрічаються дуже часто в глобулярних білках, оскільки

найкраще відповідають правилам структурної організації білків і вимозі щільної упаковки глобули.

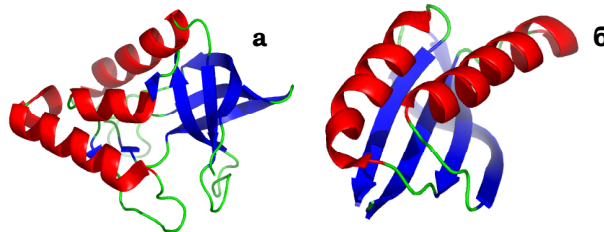
**$\alpha/\beta$ -Білки.** Так позначаються білки, в яких  $\alpha$ -спіралі та  $\beta$ -тяжі регулярно чергуються по ланцюгу (відповідно,  $\beta$ -структура є паралельною, і  $\alpha$ -спіралі також паралельні одна до одної). Для них найхарактернішими є два типи укладок (рис. 2.13): 1)  $\alpha/\beta$ -циліндр – закручений у циліндр паралельний  $\beta$ -шар, оточений циліндром з  $\alpha$ -спіралей; відповідно, утворюються два гідрофобні ядра – усередині  $\beta$ -циліндра та на інтерфейсі між  $\beta$ -структурою та  $\alpha$ -спіралями; 2) укладка Россманна (Michael Rossmann) – приблизно планарний шар паралельної  $\beta$ -структури, прикритий двома шарами  $\alpha$ -спіралей; тут також утворюються два гідрофобні ядра на інтерфейсах між  $\beta$ -структурою та  $\alpha$ -спіралями.



**Рис. 2.13.** (а) –  $\alpha/\beta$ -циліндр ендоглюканази (1A3H);  
(б) – укладка Россманна в трансгідрогеназі (1D4O).  
(в) – відповідні схеми укладок з торця  $\alpha$ -спіралей і  $\beta$ -тяжів

**$(\alpha+\beta)$ -Білки** – білки з нерегулярним чергуванням  $\alpha$ -спіралей і  $\beta$ -структурних тяжів. В їхньому складі два типи вторинної структури можуть бути рознесені: наприклад,  $\beta$ -сендвіч, торці якого прикриті

$\alpha$ -спіралями чи  $\alpha$ -спіральним субдоменом (рис. 2.14, а) – ОВ-укладка (oligonucleotide binding) із розташуванням елементів вторинної структури по ланцюгу  $\beta\beta\beta\alpha\beta\beta\alpha$ . Іншим варіантом є так звана  $\alpha\beta$ -складка ( $\alpha\beta$ -plait), зображена на рис. 2.14, б: шар антипаралельної  $\beta$ -структури, гідروفобна поверхня якого закрита  $\alpha$ -спіралями, розташування елементів вторинної структури по ланцюгу –  $\beta\alpha\beta\beta\alpha\beta$ . Отже, як і для  $\alpha/\beta$ -білків, у складі єдиної глобули, утвореної шарами сегментів вторинної структури,  $\alpha$ -спіралі та  $\beta$ -структура завжди розділяються в різні шари. Занурення  $\alpha$ -спіралі в шар  $\beta$ -структури привело б до зруйнування системи водневих зв'язків між  $\beta$ -тяжами.

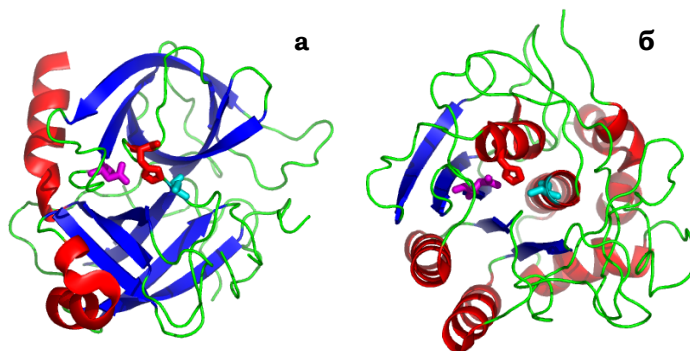


**Рис. 2.14.** (а) – стафілококова нуклеаза (1EY0);  
(б) –рибосомний білок S6 (1J5E)

Вище було розглянуто деякі найтипівіші структурні мотиви, що зустрічаються у глобулярних білках. Насправді, таких мотивів (укладок) нараховують дещо більше 100. Зрозуміло, що білки з гомологічною первинною структурою мають подібну просторову організацію. Але й десятки тисяч білків без гомології вписуються в досить обмежену кількість стандартних структурних мотивів. Причиною цього нетривіального спостереження є те, що стандартні мотиви найкраще відповідають певним правилам (частина яких сформульована вище) структурної організації, правилам, які впливають просто із загальних фізичних міркувань. У принципі, можна уявити порушення будь-якого із цих правил. Наприклад, можна так “спроектувати” амінокислотну послідовність, щоб гідروفобна перемичка між сегментами вторинної структури занурювалась у гідروفобне ядро з утворенням там водневих зв'язків її пептидними групами. Але зрозуміло, що для цього потрібен дуже обізнаний “дизайнер”. Сам факт наявності стандартних укладок свідчить про те, що **за своїм походженням молекули білків є статистичними (випадковими) сополімерами 20 амінокислот**. І чим більше послідовностей може бути вписано в певну стандартну

архітектуру глобули без порушення її стабільності, тим частіше така архітектура спостерігається у природі.

Це не означає, що *будь-який* статистичний сополімер амінокислот може сформувати стабільну глобулу. У результаті еволюційного відбору залишилися лише ті сополімери (дуже невелика кількість), які допускають утворення стабільного, щільно упакованого (з великою кількістю вандерваальсових контактів) гідрофобного ядра. Слід зауважити, що вимога утворення гідрофобного ядра не означає, що *всі* неполярні залишки мають опинитися всередині (див. рис. 2.7, (б), (г)): наприклад, у середньому для всіх білків тільки 80–85 % залишків Leu занурюються у гідрофобне ядро, решта, які залишаються на поверхні, не впливають суттєво на стабільність глобул. І це є ще одним підтвердженням випадковості білкових послідовностей.



**Рис. 2.15.** (а) –  $\alpha$ -хімотрипсин (4CHA); (б) – субтилізін (1SBC).  
Для кожного білка показані три залишки (Asp, His, Ser), які формують активний центр цих протеаз

Отже, сучасні білки є випадковими полімерами амінокислот, спочатку відібраними за критерієм стабільності їхньої глобулярної структури, а потім “відредагованими” відбором шляхом обмеженої кількості амінокислотних замін у двох напрямках: підвищення стабільності просторової структури та утворення активного центру, який має виконувати певну функцію. У результаті білки з різними функціями часто мають подібну просторову будову. І навпаки, білки, що виконують подібні функції, можуть бути зовсім різними за структурою. На рис. 2.15 зображено дві серинові протеази (однакова функція), які належать до різних структурних типів, але містять однаково організований активний центр. Таким чином, для функціонування часто важливою є не стільки просторова структура білка, скільки

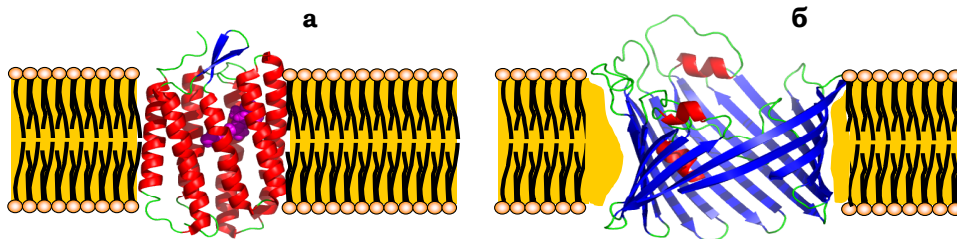


сам факт її наявності: жорсткий каркас глобули утримує активний центр у певній просторовій конфігурації, не дозволяючи йому рухатися під дією теплових флуктуацій.

### *Структура мембранних білків*

Особливості структурної організації мембранних білків або доменів, що занурюються в клітинну мембрану, пов'язані з тим, що ці білки знаходяться в неполярному оточенні. Основа мембрани являє собою подвійний шар, який утворюється за рахунок гідрофобних взаємодій між неполярними вуглеводневими хвостами молекул ліпідів (здебільшого фосфоліпідів), полярні головки ліпідів при цьому залишаються зовні подвійного шару, взаємодіючи з водою. Близько половини ваги мембрани складають мембранні білки, які виконують в основному роль каналів для пасивного чи активного транспорту через мембрану та рецепторів, які сприймають певні хімічні сигнали.

Уже описані закономірності структурної організації водорозчинних білків дозволяють легко зрозуміти основні вимоги до білків мембранних. Оскільки білок або його частина має зануритися в гідрофобне оточення, то: 1) пептидні групи мають бути насичені водневими зв'язками – у мембрану занурюються або довгі (по товщині мембрани)  $\alpha$ -спіралі, або  $\beta$ -циліндр – оскільки він з'єднує водневими зв'язками границі  $\beta$ -структурного шару (рис. 2.16); 2) зовнішній (обернений до ліпідів) бік  $\alpha$ -спіралі чи  $\beta$ -структурного тяжа має містити в основному гідрофобні залишки.



**Рис. 2.16.** Структури двох мембранних білків: бактеріородопсину (а, 1R2N) і порину (б, 2POR)

На рис. 2.16 зображено дві такі типові структури. Молекула бактеріородопсину містить пучок із семи довгих  $\alpha$ -спіралей, перемички між якими виходять з мембрани в полярне оточення. Усередині пучка

утворюється вузький канал, на внутрішній поверхні якого міститься невелика кількість полярних залишків. У каналі розміщується також кофактор небілкової природи – ретинол, за допомогою якого через канал проводиться протон. Молекула порину утворює широкий  $\beta$ -циліндр, внутрішня поверхня якого (внутрішній діаметр циліндра  $\sim 15 \text{ \AA}$ ) полярна, тобто канал є мембранною порою низької селективності.

### **Неструктуровані білки**

Далеко не всі водорозчинні білки мають упорядковану просторову структуру. Досить велика кількість білків (до 40 % у клітинах еукаріотів) не утворюють за фізіологічних умов жорсткої третинної структури, залишаючись неупорядкованими (неструктурованими), до 70 % білків мають у своєму складі окремі неупорядковані ділянки. Такі неструктуровані білки (*intrinsically disordered proteins*) або окремі ділянки глобулярних білків, часто розташовані на кінцях поліпептидного ланцюга, є, як правило, збідненими на гідрофобні амінокислоти, що й робить неупорядкований стан енергетично більш вигідним.

Досить загальною рисою неструктурованих білків є їхня здатність взаємодіяти з багатьма партнерами: це дозволяє таким білкам слугувати своєрідними платформами для збірки мультимолекулярних комплексів (у наступних розділах наведено кілька прикладів такого типу), виконуючи тим самим важливі функції в регуляції багатьох клітинних процесів. Часто зв'язування неструктурованого білка з іншими білками чи нуклеїновими кислотами приводить до утворення певної просторової структури, яка стабілізується цими новими взаємодіями. Утворення такої структури, у свою чергу, зумовлює формування робочої поверхні, що має спорідненість до наступної макромолекули: неструктурований білок (який стає структурованим унаслідок міжмолекулярних взаємодій) спрацьовує як своєрідний перемикач, необхідний для формування мультимолекулярного комплексу, задіяного до певного регуляторного акту.

### **Конформаційна рухливість білків**

У відповідь на екстремальні зовнішні впливи *in vitro* (наприклад, підвищення температури) глобулярний білок може втратити свою впорядковану структуру – стати *денатурованим*. Дослідження цього процесу є дуже важливим для з'ясування механізмів стабілізації

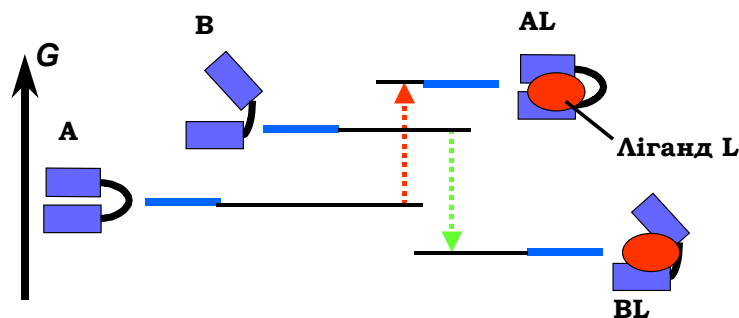
структури білка. Але ще цікавішою є можливість зміни структурної форми за фізіологічних умов – саме в цьому сенсі будемо далі розуміти конформаційну рухливість білка.

Хоча нативна молекула білка – тверде тіло, вона є досить маленькою. Тобто, на відміну від макроскопічних твердих тіл, молекула білка знаходиться під дією теплового руху як частин власної молекули, так і оточуючих молекул. Невеличкі конформаційні зміни, перш за все зміни кутів обертання основного ланцюга та бокових залишків на перемичках між елементами вторинної структури, можуть зумовити тимчасових взаємних зсувів ділянок вторинної структури. Але якщо ці структурні блоки залучені до формування спільного (і щільного) гідрофобного ядра, взаємодія між ними швидко відновить вихідну конформацію (у випадку, коли немає додаткових факторів, що можуть порушити баланс вільної енергії). Ще більша рухливість притаманна мультидоменним і субодиничним білкам: окремі домени (субглобули) можуть деформуватися і, що більш важливо, рухатися відносно один одного. Отже, в основі конформаційної рухливості білка лежить взаємний зсув великих структурних блоків (субодиниць, доменів або елементів вторинної структури) за рахунок невеликих локальних деформацій. Іноді відбуваються також локальні перебудови вторинної структури в різних структурних станах білка – руйнування або утворення спіралей тощо. Найрадикальнішою структурною зміною є утворення просторової структури у згаданих неструктурованих білках за умов їхньої взаємодії з іншими молекулами. Зрозуміло, що структурні перебудови змінюють форму робочої поверхні молекули, тобто впливають на виконання білком своїх функцій.

Припустимо, що організація білка дозволяє існування двох структурних станів – А і В. Загалом можливо, що обидва стани не розрізняються за величиною вільної енергії – тоді половину часу білок проводить в одному стані, другу – в іншому. Але, як правило, вільна енергія одного зі станів (наприклад, А) є нижчою, тоді конформаційна рівновага зсунута саме до нього (рис. 2.17, ліворуч).

Як можна перемкнути цю рівновагу на користь стану В (якщо, наприклад, саме він має виконувати якусь функцію)? Типовий шлях такого перемикавання полягає у використанні певної сполуки (ліганду). Як ліганд може виступати органічна низькомолекулярна сполука, іон, інший білок – залежно від конкретної системи. На рис. 2.17 показано ситуацію, коли ліганд має спорідненість саме до форми В і не має такої до форми А. Це означає, що вільна енергія знижується при зв'язуванні ліганду з формою В, але зростає при зв'язуванні з фор-

мою А, тобто зв'язування з А просто не відбувається. Головним є те, що у присутності ліганду здійснюється зміна балансу вільної енергії: стан В (точніше, комплекс ВL) стає переважним – його вільна енергія є нижчою, ніж вільна енергія стану А.



**Рис. 2.17.** Схема структурних переходів у молекулі білка, опосередкованих лігандом.  $G$  – шкала вільної енергії

Результатом такої фіксації одного зі структурних станів за рахунок зв'язування з лігандом є зміна робочої поверхні білка. При цьому робоча поверхня часто розташована зовсім в іншому місці відносно ділянки зв'язування ліганду. Ліганд (або його відсутність) виступає лише сигналом, регулятором, який перемикає структурні, а відповідно, і функціональні стани. Регуляція такого типу отримала назву *алостеричної регуляції* (від  $\alpha\lambda\lambda\omicron\varsigma$  – інший).

### Принципи функціонування білків

За вдалим визначенням А.В. Фінкельштейна, усі різноманітні операції, які виконують білки, можна описати трьома словами:

ЗВ'ЯЗАТИ – ТРАНСФОРМУВАТИ – ВІДПУСТИТИ

Звичайно, не обов'язково кожний білок здійснює всі три операції. Може виконуватися лише одна з них – зв'язування. Так працюють імуноглобуліни: з надзвичайно високою селективністю зв'язують антиген (точніше, антигенний детермінант – невеличку ділянку на поверхні чужорідної для організму молекули). Константа зв'язування є настільки високою, що практично взаємодія є нерівноважною (немож-

лива дисоціація комплексу). Далі цей комплекс просто руйнується певними клітинами імунної системи.

Операцію зв'язування з іншими білками й молекулами небілкової природи виконують білки, що залучені до утворення різноманітних клітинних структур. У цьому випадку зв'язування може змінюватись дисоціацією (наприклад, при утворенні / руйнуванні філаментів цитоскелета), що часто супроводжується структурними змінами в молекулі білка.

Операції зв'язати-відпустити виконують численні регуляторні білки (транспортні білки, рецептори, фактори транскрипції тощо). Для їхнього функціонування також часто велике значення мають структурні перебудови за механізмом, показаним схематично на рис. 2.17: певний ліганд, виконуючи роль хімічного сигналу, індукує або блокує спорідненість білка-регулятора до певної молекули-мішені. Конкретні приклади такого типу будуть досить часто зустрічатися в наступних розділах.

Найважливіший для живих систем функціональний клас білків – ферменти, або ензими (enzymes), які здійснюють каталіз усіх численних біохімічних реакцій, виконуючи три операції: зв'язують субстрати, сприяють їхній трансформації, відпускають продукти реакції.

### *Механізм ферментативного каталізу*

Енергетичну схему хімічної реакції (рис. 1.6) було розглянуто в розділі 1. На рис. 2.18 ту саму схему доповнено ефектами, що є наслідком присутності ферменту. Уповільнення реакції зумовлено високим значенням вільної енергії активації – у відсутності ферменту реакція практично не відбувається (червоний шлях на рис. 2.18). Відповідно, каталітична дія ферменту полягає у зниженні цього енергетичного бар'єра. Оскільки бар'єр, як будь-яка вільна енергія, має дві складові – ентальпійну та ентропійну (див. розділ 1), фермент здійснює зниження енергії активації двома шляхами (як правило, обидва мають значення).

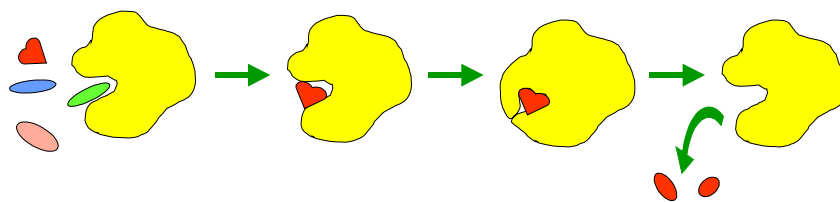
- *Ентропійний каталіз.* Перше, що робить фермент – зв'язує субстрати у своєму активному центрі за рахунок спорідненості до них (перша зелена стрілка вниз). Але головне полягає в тому, що субстрати зв'язуються не аби як, а в певній, найбільш сприятливій для реакції, взаємній орієнтації та в певній орієнтації щодо хімічних груп активного центру. Це означає, що з великої кількості орієнтацій обирається лише одна: взаємодії субстратів з активним центром компенсують ентропійні витрати на впорядкування субстратів.



продукти, які не мають спорідненості до активного центру (вільна енергія дисоційованого продукту є нижчою, ніж зв'язаного) звільняють активний центр для інших субстратів.

Для успішного виконання завдання каталізу по обом шляхам *активний центр ферменту має бути дуже жорстким* – мати чітко зафіксовану у просторі й часі структуру. Для цього, власне, і потрібна твердість глобули, яка несе на собі активний центр. І для цього зовсім не потрібна (навіпаки – шкідлива) конформаційна рухливість ферменту. Білковим ферментам деяких простих реакцій (наприклад, протеазам на рис. 2.15) конформаційна рухливість практично не притаманна.

Але ферменти часто мають вирішувати складніше завдання: дискримінувати субстрати – здійснити перетворення лише одного субстрату з певного набору аналогів. Типовий сценарій подій у цьому випадку включає в якості “дійових осіб” принаймні дві структурні форми ферменту, і, відповідно, його активного центру (рис. 2.19). У *відкритій* формі активний центр має невисоку спорідненість до субстратів певного типу (завдяки чому є можливим швидкий перебір варіантів) і не здатен каталізувати хімічне перетворення. Якщо нарешті зв'язується потрібний субстрат, він спрацьовує як ліганд на рис. 2.17 – індукує перетворення ферменту в *закриту* форму. Такий процес було названо Кошландом (Daniel E. Koshland) *індукованою відповідністю* (induced fit). Закрита форма жорстко фіксує субстрат, каталітичний центр спрацьовує, і здійснюється хімічна реакція. У результаті ліганд змінюється, – і відбувається зворотне перетворення білка у відкриту форму із звільненням продуктів.



**Рис. 2.19.** Структурні перебудови ферменту при дискримінації субстратів

Повертаючись до загальної схеми реакції, зображеної на рис. 2.18, варто ще раз зауважити: з ферментом чи без, реакція може відбуватися тільки в напрямку зниження вільної енергії. Протилежний напрямок може бути здійсненим лише за умови використання “обхідних шляхів” і особливих механізмів – також за участю білків.

### Механізми використання енергії гідролізу АТФ

Джерелом енергії для всіх процесів, що їй потребують, у живих системах є нуклеозидтрифосфати (див. розділ 3). Найчастіше використовується аденозинтрифосфат (adenosine triphosphate, АТФ). Ця сполука, яка утворюється при перенесенні протонів через мембрану хлоропластів (фотосинтез) чи мітохондрій (окислювальне фосфорилування), схематично показана на рис. 2.20.

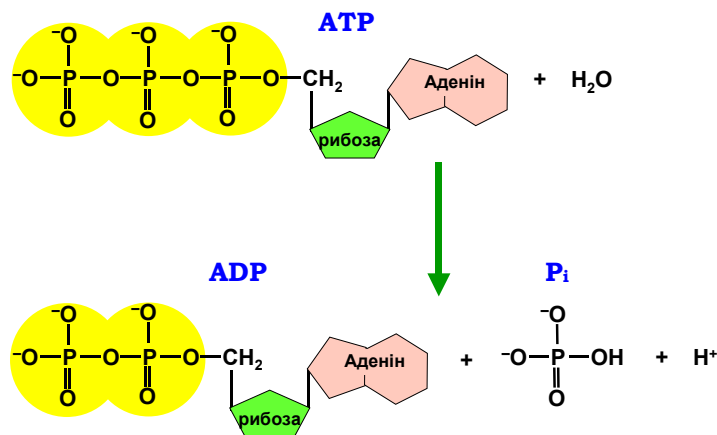


Рис. 2.20. Гідроліз АТФ

Вона складається з азотистої основи (аденіну), пентозного цукру (рибози) та трьох залишків фосфорної кислоти. Гідроліз АТФ приводить до відщеплення одного з них – залишається аденозиндифосфат (adenosine diphosphate, АДФ) і неорганічний фосфат  $P_i$  (рис. 2.20). Звичайно, як будь-яка інша реакція, гідроліз АТФ потребує каталізу (за участю досить широкого класу ферментів – АТФаз). Роль АТФ як джерела енергії зумовлена тим, що ця реакція супроводжується досить великим зниженням вільної енергії.

Зв'язок між фосфатними залишками, який руйнується при перетворенні АТФ на АДФ, іноді так і називають – макроергічним. Ця назва не має викликати помилкового уявлення, що цей зв'язок заощаджує надзвичайно велику енергію. По-перше, в цьому зв'язку немає нічого особливого. По-друге, загальна кількість ковалентних зв'язків не змінюється в ході реакції – фосфатний залишок перено-



ситься на молекулу води (рис. 2.20). Змінюється тільки *вільна* енергія набору хімічних сполук, які зображені на рис. 2.20. Причому головний внесок у цю зміну дають концентраційні (тобто ентропійні) ефекти, а не хімічні перебудови молекул.

Скориставшись рівнянням (1.5), різницю вільних енергій продуктів і субстратів реакції, зображеної на рис. 2.20, можна записати як

$$\Delta G = \Delta G^* + RT \ln \frac{[\text{ADP}][\text{P}_i][\text{H}^+]}{[\text{ATP}][\text{H}_2\text{O}]},$$

де у квадратних дужках – молярні концентрації відповідних компонентів,  $\Delta G^*$  – стандартна зміна вільної енергії,  $RT = 0,6$  ккал/моль (див. розділ 1). Оцінки показують, що величина  $\Delta G^* = +4,8$  ккал/моль (тобто є позитивною – стандартна вільна енергія зростає). Проте, ця величина не має великого сенсу: енергетичний ефект реакції  $\Delta G$  дорівнює  $\Delta G^*$  при концентраціях усіх компонентів 1 моль/л. Але концентрації води та протонів зберігаються постійними як у буферному розчині *in vitro*, так і в клітині; причому вони значно відрізняються від 1 моль/л (у розведеному розчині  $[\text{H}_2\text{O}] = 55,5$  моль/л, при рН 7,0  $[\text{H}^+] = 10^{-7}$  моль/л). Отже, відношення цих концентрацій – також постійна величина, і можна ввести іншу, більш змістовну, стандартну вільну енергію реакції

$$\Delta G^0 = \Delta G^* + RT \ln \frac{[\text{H}^+]}{[\text{H}_2\text{O}]}$$

і записати

$$\Delta G = \Delta G^0 + RT \ln \frac{[\text{ADP}][\text{P}_i]}{[\text{ATP}]}$$

Підстановка значень концентрацій води й протонів дає  $\Delta G^0 = -7,3$  ккал/моль, що вказує на енергетичну вигідність реакції гідролізу. Головний внесок дає той факт, що в реакції народжується вільний протон: за законом діючих мас, підтримання на дуже низькому рівні загальної концентрації протонів має зсувати рівновагу в бік їхнього утворення.

Але в клітинах концентрації усіх компонентів, що залишилися в останньому рівнянні, також підтримуються на певних рівнях:

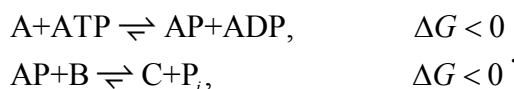
3÷8 ммоль/л АТР; 0,2÷1 ммоль/л АDP; 2÷8 ммоль/л неорганічного фосфату. Підстановка цих значень дає реальний енергетичний ефект від гідролізу АТР:  $\Delta G = -12 \div -16$  ккал/моль.

Ця енергія використовується двома шляхами: 1) для хімічної модифікації субстратів, які самі по собі не можуть вступити в певну реакцію (головним чином, маються на увазі реакції синтезу), оскільки вона супроводжувалась би зростанням вільної енергії; 2) для виконання механічної роботи молекулярними машинами.

**Модифікація субстратів.** Розглянемо хімічну реакцію



Оскільки утворення С супроводжується зростанням вільної енергії, цей процес є неможливим. Типовий спосіб, яким жива система вирішує цю проблему, виглядає так:



Тут, власне, немає гідролізу АТР: відбувається перенесення фосфатного залишку на молекулу А (фосфорилування). Реакція відбувається, оскільки супроводжується зниженням вільної енергії. Для розрахунку цієї величини можна скористатися енергіями гідролізу: щоб загальна вільна енергія реакції була негативною, енергія гідролізу АТР має бути вищою за абсолютною величиною, ніж енергія гідролізу сполуки АР, обидві величини негативні. Далі здійснюється друга енергетично вигідна реакція – утворення сполуки С з вилученням неорганічного фосфату – тут знову енергія гідролізу АР має бути вищою за абсолютною величиною, ніж (негативна) енергія гідролізу С. З певними модифікаціями (наприклад, часто переноситься не фосфат, а аденозинмонофосфат АМР) саме така схема використовується для здійснення хімічних реакцій синтезу, зокрема синтезу біополімерів.

**Молекулярні машини** – білкові молекули чи, частіше, макромолекулярні комплекси, утворені за участю білків, які здатні здійснювати спрямовані рухи. Робота таких машин забезпечує переміщення клітинних структур та органел (наприклад, розходження хромосом при клітинному поділі), зміну форми клітин, активне переміщення клітин, активний транспорт через мембрану проти градієнтів концентрації, м'язові скорочення, переміщення полімераз уздовж матриці при син-

тезі біополімерів тощо. Усе це потребує енергії, джерелом якої також є АТФ (або інші нуклеозидтрифосфати).

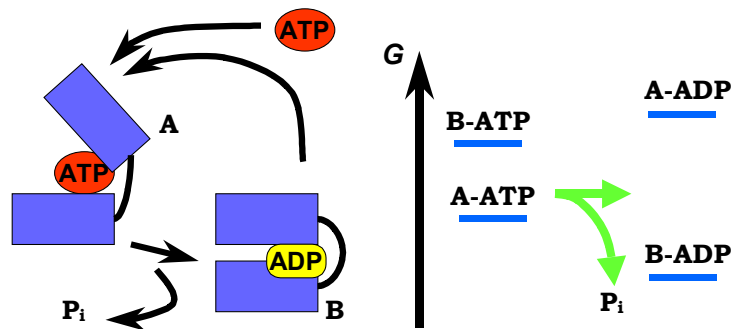
Узагалі машина – це пристрій для перетворення тієї чи іншої форми енергії в рух. Звичайна механічна машина перетворює будь-яку енергію в механічну, умовою заощадження механічної енергії при цьому є висока інерція (маса) частин машини і певна їхня жорстка взаємна орієнтація. Точніше, конструкція машини зумовлює рухи її частин тільки по певних заданих траєкторіях.

Молекулярна машина, на відміну від макроскопічної механічної машини, по-перше, маленька, тобто піддається тепловому руху своїх власних частин і зовнішніх молекул, що робить неможливим заощадження механічної енергії. По-друге, вона побудована з полімерів, а це означає, що її частинам притаманна конформаційна рухливість. Проте внаслідок того, що молекула білка має певну досить жорстку просторову організацію, конструкція молекулярної машини допускає не будь-які, а певним чином спрямовані у просторі рухи її частин. Саме ці особливості молекулярної машини й використовуються для її роботи. Основні принципи цієї роботи:

- *Конформаційна рухливість* молекулярної машини забезпечує їй можливість існувати в кількох структурних станах (двох-трьох), які розрізняються головним чином на рівні просторового розташування великих структурних блоків – доменів або субодиниць.
- Структурні стани мають різну спорідненість до певних лігандів. *Взаємодії з лігандами* (факторами) фіксують певні стани.
- *Хімічні реакції*, які каталізуються машиною, приводять до заміни лігандів, а відповідно – і до переходу в інший структурний стан.
- Рушійною силою для переміщення блоків є *тепловий рух*: блоки рухаються хаотично (хоча й у відповідності з конструкцією машини); зв'язування лігандів та заміна їх унаслідок реакцій каналізують ці рухи в певних напрямках.
- Результатом структурних перебудов є *переміщення* структурних блоків у просторі та / або зміна характеру взаємодії машини зі своїм оточенням – рух або всієї машини, або відносний рух її частин.

Загальний принцип перемикання структурних станів є точно таким, що його було зображено на рис. 2.17 – для фіксації станів використовується енергія зв'язування з лігандом. Як ліганд часто використовується саме АТФ та продукти його гідролізу. Реакція гідролізу

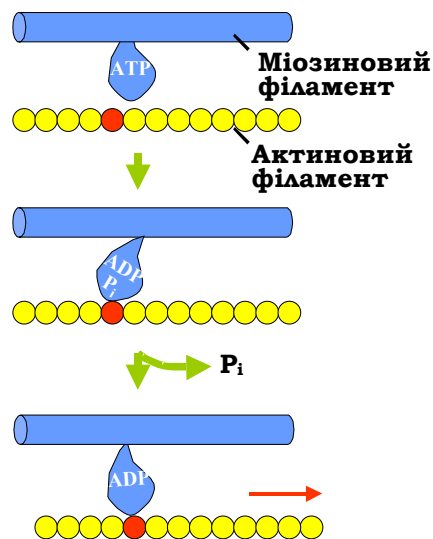
(тут відбувається саме гідроліз, як зображено на рис. 2.20) є просто засобом змінити ліганд – і перемкнути структурні стани. Вільна енергія гідролізу визначає ймовірність проходження реакції (див. рівняння (1.6)), а відповідно й ймовірність перемикавання конформацій білка. Типовий елементарний “такт” роботи молекулярної машини зображено на рис. 2.21: структурна форма А зв'язує АТФ з високою спорідненістю; спрацьовує АТФазний каталітичний центр – АДФ, що виникає внаслідок гідролізу, має підвищену спорідненість до форми В – відбувається структурна перебудова. Потім АДФ витісняється новою молекулою АТФ, і цикл повторюється.



**Рис. 2.21.** Схема роботи молекулярної машини. Праворуч: рівні вільної енергії структурних станів з різними лігандами

Приблизно за такою схемою працює, наприклад, актоміозиновий комплекс у м'язових (і не тільки) клітинах (рис. 2.22). Міозиновий філамент складається з молекул міозину – великого білка досить складної структури. Від основи цього філаменту відходять глобулярні домени міозину – головки (показано лише одну). Головка має спорідненість до АТФ і містить АТФазний каталітичний центр. Зв'язування АТФ фіксує головку в структурному стані, який не має спорідненості до філаменту, що складається з молекул іншого білка – актину. Після гідролізу АТФ нові ліганди – АДФ і неорганічний фосфат – фіксують інший стан: головка зв'язується з актином, відхиляючись ліворуч. Третій стан головки індукується окремою молекулою АДФ після дисоціації фосфату: зміщення головки праворуч без втрати зв'язку з актином. У результаті актиновий філамент зсувається разом із головкою. Далі АДФ витісняється новою молекулою АТФ, і все повторюється. Актино-

вий філамент при дисоціації головки не повертається назад, оскільки фіксується іншими головками, що перебувають на інших стадіях свого робочого циклу. Інші приклади роботи молекулярних машин розглядаються в наступних розділах.



**Рис. 2.22.** Спрощена схема роботи актоміозинового комплексу

### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. Напишіть стандартні скорочення для 20 амінокислот. На які дві групи їх можна поділити за фізико-хімічними властивостями?
2. Які властивості має пептидна група? Чи взаємодіє вона з водою?
3. Що таке вторинна структура білка? Укажіть типи й основні риси вторинних структур, що зустрічаються у глобулярних білках. Які взаємодії стабілізують регулярну вторинну структуру?
4. Які взаємодії є основною рушійною силою, що зумовлює укладання поліпептидного ланцюга в компактну глобулу?
5. Укажіть основні загальні риси структури глобулярних водорозчинних білків. Чому глобула завжди формується елементами регулярної вторинної структури?

6. Які взаємодії зумовлюють твердість білкової глобули? Яке значення має твердість білка для його функціонування?

7. Що таке структурний мотив? Укажіть основні структурні мотиви, що часто зустрічаються у глобулярних білках.

8. Як структура мембранних білків відрізняється від такої водорозчинних?

9. За рахунок чого є можливим перемикання структури білка між двома конформаційними станами?

10. У чому полягає фізична природа ферментативного каталізу?

11. За яким принципом використовується вільна енергія гідролізу АТР для здійснення реакцій синтезу?

12. Як використовується вільна енергія гідролізу АТР для виконання роботи молекулярними машинами? Сформулюйте основні принципи роботи молекулярних машин.

### **РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

*Ленинджер, А.* Основы биохимии : в 3 т. – М. : Мир, 1985.

*Фершт, Э.* Структура и механизм действия ферментов. – М. : Мир, 1980.

*Финкельштейн, А.В. Птицын, О.Б.* Физика белка. Курс лекций. – М. : КДУ, 2005.

*Чернавский, Д.С. Чернавская, Н.М.* Белок – машина. Биологические макромолекулярные конструкции. – М. : МГУ, 1999.

*Шульц, Г., Ширмер, Р.* Принципы структурной организации белков. – М. : Мир., 1982.

*Branden, C.-I., Tooze, J.* Introduction to protein structure.– New York : Garland Science, 1999.

*Creighton, T.* Proteins: structure and molecular properties. – New York: W.H.Freeman and Company.1992.

*Lesk, A.M.* Introduction to protein architecture: the structural biology of proteins. – Oxford : Oxford University Press, 2001.

*Spirin, A.A.* Ribosome as a molecular machine // FEBS Letters. – 2002. – Vol. 514. – P. 2–10.

## Розділ 3

### ДНК

They are in you and me; they created us, body and mind; and their preservation is the ultimate rationale for our existence.

*R. Dawkins. The selfish gene*

#### Хімічна будова нуклеїнових кислот

##### Нуклеотиди

Мономерна одиниця нуклеїнової кислоти – *нуклеотид* – складається з трьох елементів: азотиста основа, пентозний цукор, залишок фосфорної кислоти.

*Азотисті основи* – гетероциклічні сполуки (рис. 3.1), у кільцях яких містяться карбон і азот, а всі зв'язки мають характер частково подвійних. До складу нуклеїнових кислот входять два типи азотистих основ: пурини (purines, загальноприйняте позначення R) – аденін (A) і гуанін (G); піримідини (pyrimidines, Y) – урацил (U), тимін (T), цитозин (C). Загальне позначення для всіх основ – N.

Найсуттєвішими для структури нуклеїнових кислот та їхніх взаємодій з іншими молекулами є такі властивості азотистих основ:

- Оскільки всі зв'язки в кільці – частково подвійні, азотисті основи є **планарними** – усі атоми кільця лежать практично в одній площині.
- Атоми азоту кільця та приєднані до нього атоми кисню є акцепторами, а NH<sub>2</sub>-групи – донорами **водневого зв'язку**. Але водночас площина кільця є **гідрофобною**, тому азотисті основи погано розчиняються у воді.

Один із атомів азоту кільця (виділено червоним на рис. 3.1) приєднується у складі нуклеотиду до карбону пентозного цукру *глікозидним*

зв'язком (рис. 3.2). Цей атом пентози позначається як С1' (символ ' прийнято додавати, щоб відрізнити атоми фуранозного кільця пентози від атомів азотистої основи). Інші С'-атоми пентози нумеруються далі за порядком їхнього розташування (рис. 3.2). **До 3'-атома завжди приєднана ОН-група.** Пентоза, у складі якої ОН-група знаходиться також при 2'-атомі, називається *рибозою* (ribose). Пентоза іншого типу, яка також входить до складу нуклеїнових кислот – *дезоксирибоза* (deoxyribose) – відрізняється лише заміною цієї ОН-групи на атом гідрогену.

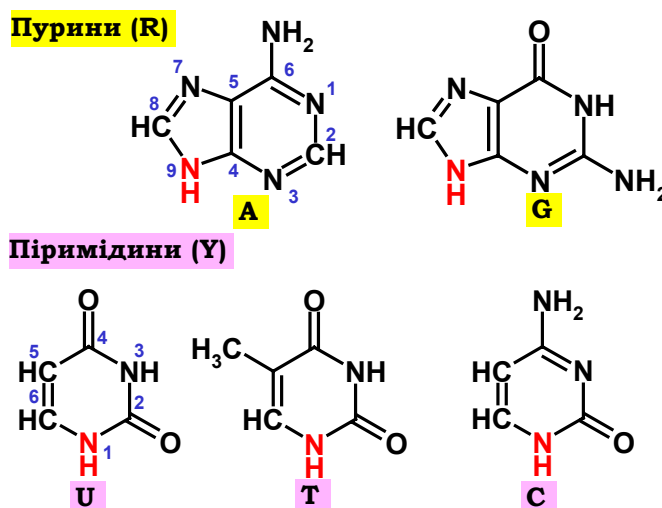


Рис. 3.1. Азотисті основи. Червоним позначено атом N, з'єднаний у нуклеотиді з С1' атомом пентози

Сполука азотистої основи й пентози називається *нуклеозидом* (залежно від типу азотистої основи: аденозин, гуанозин, цитидин, тимідин, уридин) або ж, залежно від типу пентози, рибо- чи дезоксирибонуклеозидом. Рибонуклеозиди входять до складу РНК – рибонуклеїнової кислоти (ribonucleic acid, RNA), дезоксирибонуклеозиди – до складу ДНК – дезоксирибонуклеїнової кислоти (deoxyribonucleic acid, DNA). Інша хімічна різниця між ДНК і РНК стосується складу піримідинових азотистих основ: Т у ДНК замість U в РНК.

Фосфорильовання ОН-групи при 5'-атомі пентози у складі нуклеозиду приводить до утворення нуклеотиду (рис. 3.2). Найважливішою для нуклеїнових кислот властивістю фосфатного залишку є *негативний заряд*, який він має за нейтральних рН. Таким чином, нуклеотид



є нуклеозидфосфатом або нуклеозидмонофосфатом (NMP). Зокрема, нуклеотид, зображений на рис. 3.2, є аденозинмонофосфатом (AMP); відповідний трифосфат АТР розглянуто в розділі 2 (див. рис. 2.20).

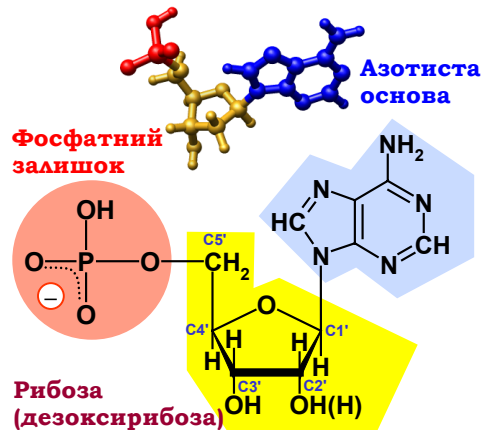


Рис. 3.2. Нуклеотид

На відміну від азотистих основ, пентози та фосфатні залишки є цілком полярними й добре взаємодіють з водою. Крім того, пентоза не містить подвійних зв'язків. Тому, на відміну від азотистої основи, у фуранозному кільці пентози є можливим обертання навкруг зв'язків, і воно не є планарним. У складі нуклеїнових кислот переважно зустрічаються дві конформації, схематично показані на рис. 3.3. Конформація *C2'-ендо* ( $C2'$ -атом виходить у бік  $C5'$ -атома з площини, заданої атомами  $C4'-O-C1'$ ) є переважною для дезоксирибози. Поява OH-групи замість гідрогену при  $C2'$ -атомі робить цю конформацію неможливою, і для рибози переважною стає конформація *C3'-ендо*. Перемикання цих двох конформацій суттєво змінює геометрію інших зв'язків у складі нуклеотиду й тому лежить в основі перетворень між структурними формами нуклеїнових кислот (див. нижче). Отже, нуклеотид складається з двох елементів, які розрізняються за полярністю та конформаційною рухливістю:

- Полярна й досить гнучка частина – пентозний цукор із фосфатним залишком, причому останній має негативний заряд.
- Жорстка планарна азотиста основа, яка в цілому є неполярною, хоча й містить донорні та акцепторні групи.

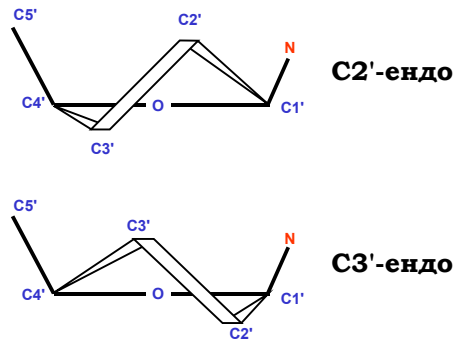


Рис. 3.3. Дві переважні конформації пентозного кільця у складі нуклеїнових кислот

### Полінуклеотидний ланцюг

Дві хімічні групи – OH-група при 3'-атомі пентози одного нуклеотиду та фосфат при 5'-атомі іншого – використовуються для утворення *фосфодієфірного зв'язку* між нуклеотидами (рис. 3.4).

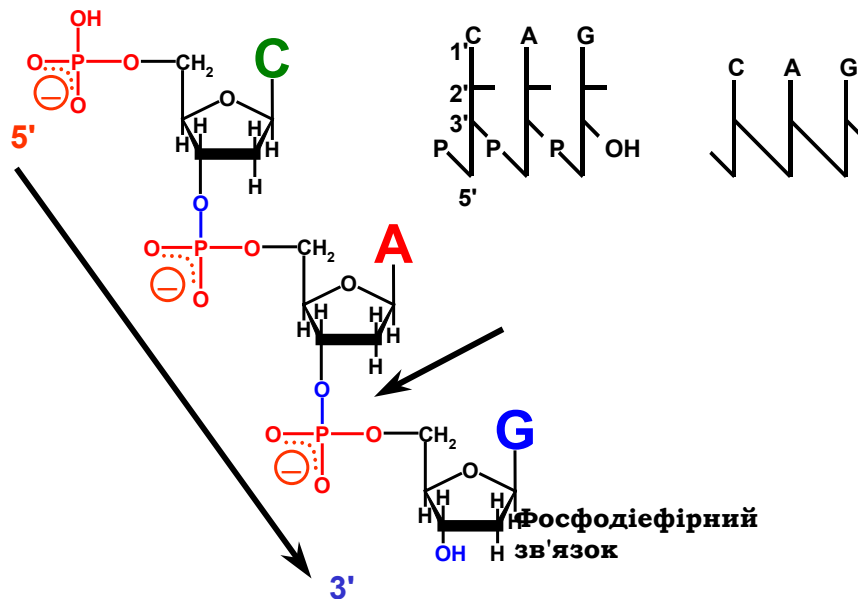


Рис. 3.4. Полінуклеотидний ланцюг. Угорі праворуч: два варіанти схематичного зображення ланцюга, вертикальна лінія позначає пентозу, P – фосфатний залишок

Таким чином, як і поліпептидний (розділ 2), полінуклеотидний ланцюг має напрямок. На одному його кінці залишається 5'-фосфат (5'-кінець), на іншому – 3' ОН-група (3'-кінець). Послідовності нуклеотидів прийнято записувати в напрямку 5' → 3', у тому ж напрямку відбувається синтез усіх нуклеїнових кислот. Слід також розуміти, що конформаційні особливості ланцюга залежать від напрямку, в якому розташовані сусідні нуклеотиди. Наприклад, послідовність СА має інші особливості, ніж послідовність АС. Загалом у складі ланцюга може бути 16 типів таких динуклеотидних контактів (16 можливих комбінацій із чотирьох по два).

Отже, полярний остов полінуклеотидного ланцюга являє собою фосфатні залишки та пентози, що чергуються – *цукрофосфатний остов*. Від цього остова відходять азотисті основи як бокові залишки.

### **Нуклеази**

Деградація нуклеїнових кислот відбувається при різноманітних функціональних процесах шляхом гідролізу фосфодієфірного зв'язку в реакції, яка каталізується спеціальними ферментами – *нуклеазами*. Нуклеази поділяють на різні типи за певними ознаками:

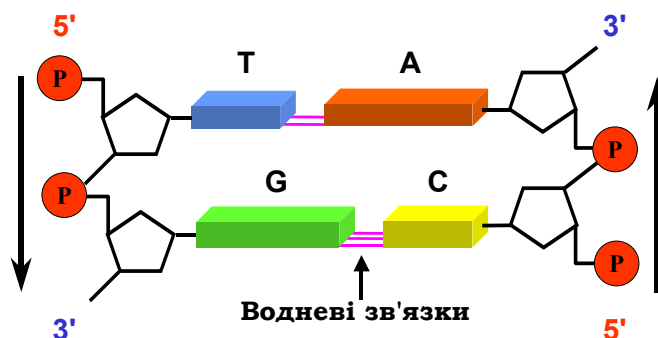
- РНКази та ДНКази є специфічними до відповідних нуклеїнових кислот.
- *Екзонуклеази* відщеплюють нуклеотиди один за одним з кінців ланцюга, відповідно розрізняють 5'- та 3'-екзонуклеази. *Ендонуклеази* гідролізують фосфодієфірний зв'язок усередині ланцюга.
- Серед ендонуклеаз, які мають своїм субстратом дволанцюгову молекулу ДНК є такі, що роблять одноланцюговий розріз – *нік* (nick), залишаючи інший ланцюг інтактним. Інші здійснюють дволанцюгові розрізи.
- Серед ендонуклеаз є неспецифічні щодо послідовності нуклеотидів. Але є також велика група бактеріальних ферментів – *рестриктаз* – які діють лише на певні невеличкі (частіше чотири або шість нуклеотидів) елементи послідовності. Рестриктази виконують у бактеріальній клітині роль захисту від чужорідної ДНК бактеріофагів і широко використовуються як інструменти дослідження в молекулярній біології (див. розділ 11).

Реакція гідролізу фосфодієфірного зв'язку є енергетично вигідною. З цього випливає, що нуклеотид (нуклеозидмонофосфат) не може бути використаним для синтезу нуклеїнової кислоти – натомість до реакцій синтезу залучаються відповідні нуклеозидтрифосфати (розділи 5, 6, 9).

## Подвійна спіраль

### Стабілізація подвійної спіралі

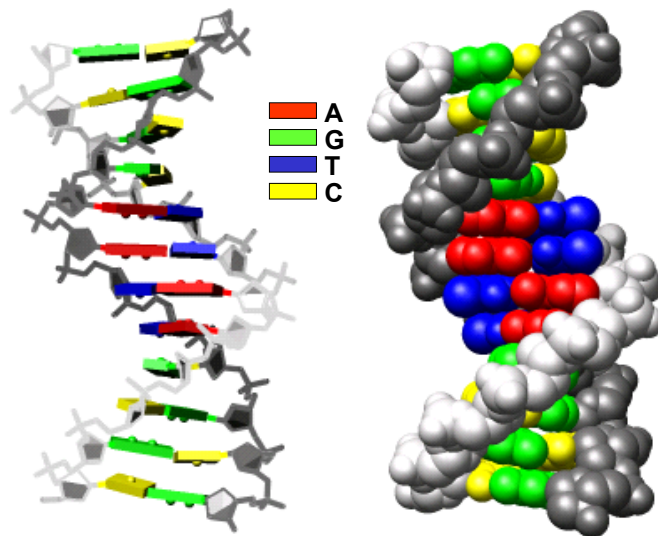
Два полінуклеотидні ланцюги (ДНК, РНК або гібридні) можуть об'єднуватись в єдину дволанцюгову структуру (*дуплекс*), схему якої зображено на рис. 3.5. Таке об'єднання відбувається за жорсткої умови: певні азотисті основи повинні стояти одна проти одної – А проти Т (чи U), G проти C. Цей *принцип комплементарності*, сформульований Уотсоном і Кріком (James D. Watson, Francis H.C. Crick), зумовлений утворенням специфічних водневих зв'язків між названими основами: два зв'язки в парі А-Т, три в парі G-С. Два ланцюги при цьому спрямовані в різні боки (є *антипаралельними*), цукрофосфатні остови розташовані зовні, пари основ – усередині цієї структури.



**Рис. 3.5.** Схема об'єднання двох динуклеотидів у дволанцюгову структуру

У складі такого дволанцюгового комплексу неполярні площини сусідніх пар основ мають наблизитися одна до одної, і відносна гнучкість остовів дозволяє це зробити простим шляхом – закрутити два ланцюги один навкруг одного у подвійну спіраль (рис. 3.6). Усередині формується *стопка* (stack) щільно укладених одна на одну пар основ, захищена на поверхні спірально закрученими один навкруг одного полярними цукрофосфатними остовами. Між остовами на поверхні спіралі утворюються два *жолобки* різного розміру – великий і маленький (*major/minor groove*), в які “дивляться” донорні та акцеп-

торні групи азотистих основ (рис. 3.7): у великому жолобку міститься по два акцептори (O, N) та одному донору (NH<sub>2</sub>) водневого зв'язку кожної пари; у маленькому – два акцептори пари А-Т, два акцептори й один донор пари G-С.

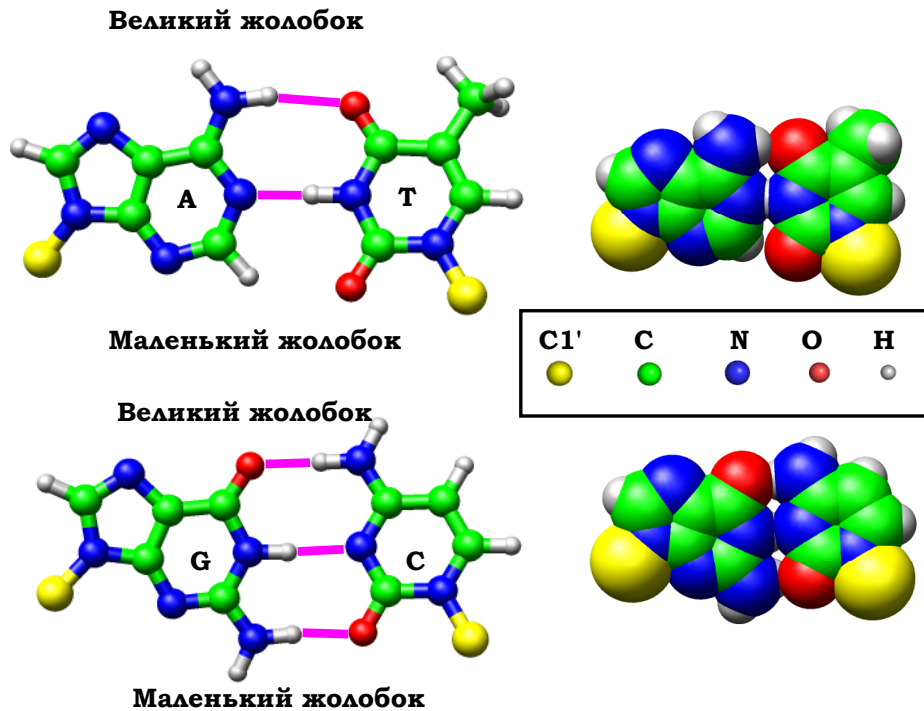


**Рис. 3.6.** Два варіанти представлення подвійної спіралі ДНК (кристалічна структура додекамеру, 355D)

Таким чином, цілком подібно до білків (розділ 2), *гідрофобні взаємодії* (розділ 1) між площинами пар азотистих основ зумовляють формування подвійної спіралі. При цьому реалізується щільна упаковка – утворюються *вандерваальсові контакти* між сусідніми парами. Велика кількість таких контактів уздовж осі спіралі робить їх суттєвим фактором стабілізації дуплекса. Вандерваальсові взаємодії між парами основ є за своїм характером орієнтаційними та, унаслідок нерівномірності розподілу електронної щільності по кільцях основ, диполь-дипольними (див. розділ 1). Такі взаємодії складного характеру (гідрофобний ефект + вандерваальсові взаємодії) у складі стопки мають спеціальну назву – *стекинг-взаємодії*.

Отже, ***стекинг-взаємодії*** – **головний фактор стабілізації подвійної спіралі**. Точніше, це практично єдиний фактор стабілізації. Роль комплементарних водневих зв'язків між основами, які самі по собі енергетично не дуже важливі (водневі зв'язки основ із водою при розхо-

дженні ланцюгів майже такі самі), полягає в тому, що вони є необхідною передумовою утворення стабільної стопки. Це знову цілком аналогічно ролі водневих зв'язків між пептидними групами вторинної структури білків у стабілізації глобули (розділ 2). Донорні й акцепторні групи азотистих основ прагнуть утворити водневі зв'язки (енергія водневого зв'язку є надто високою, щоб його втрачати). Опиняючись усередині неполярної стопки, вони стають недоступним для води і єдиний шанс зберегти водневі зв'язки – утворити їх із відповідними групами іншої основи. А це можливо лише за умови комплементарності. Відповідно, за відсутності комплементарності водневі зв'язки з водою зберігаються, що робить подвійну спіраль нестабільною.



**Рис. 3.7.** Комплементарні пари основ у складі ДНК з водневими зв'язками між основами. Праворуч: атоми тих самих пар зображені відповідно до їхніх ван-дер-ваальсових радіусів

Ефективність стекинг-взаємодій залежить від типу контактуючих пар основ, а також їхньої взаємної орієнтації та порядку розташування в ланцюгах дуплекса – типу динуклеотидного контакту вздовж осі по-

двійної спіралі. Замість 16 типів динуклеотидних контактів у одноланцюговій нуклеїновій кислоті, у подвійній спіралі залишається 10 – унаслідок комплементарності ланцюгів певні контакти є еквівалентними. Наприклад, контакт AG в одному ланцюзі автоматично означає, що в комплементарному ланцюзі реалізується контакт СТ. Слід пам'ятати, що ланцюги є антипаралельними, і що послідовність нуклеотидів читається в напрямку 5' → 3' (усі динуклеотидні контакти подвійної спіралі наведено в табл. 3.2). Для певних контактів (наприклад, пурин-піримідинових порівняно з піримідин-пуриновими) стекінг-взаємодії є більш ефективними. Відповідно, **стабільність подвійної спіралі залежить від послідовності пар основ.**

Утворення подвійної спіралі працює проти конформаційної ентропії: розходження двох ланцюгів супроводжується значним зростанням їхньої конформаційної свободи (непорядкованості). Відповідно, *стабільність подвійної спіралі залежить від температури*: при зростанні температури ентропійна компонента вільної енергії збільшується (рівняння (1.2)), і відбувається розходження ланцюгів – *плавлення* подвійної спіралі. Зрозуміло, що температура плавлення – температура за якої 50 % пар основ є розплавленими – є мірою стабільності й залежить від послідовності пар основ.

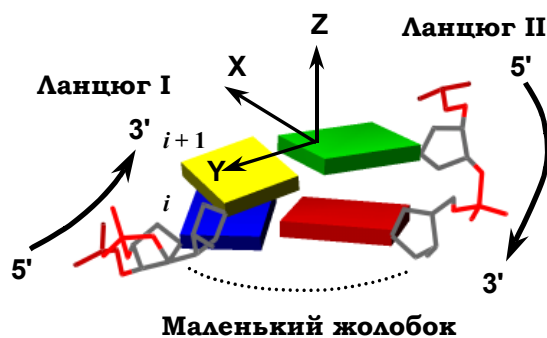
Суттєвий дестабілізуючий внесок у вільну енергію подвійної спіралі дають також електростатичні взаємодії. Оскільки кожен фосфатний залишок несе на собі негативний заряд (рис. 3.4), молекула ДНК є гігантським поліаніоном із високою поверхневою щільністю зарядів. У розчині ці заряди суттєво нейтралізовані протионами – неорганічними катіонами: навколо ДНК утворюється “іонна атмосфера” з досить високою локальною щільністю катіонів, тобто в системі виникає нерівномірний розподіл катіонів – зростає відповідна ентропійна компонента вільної енергії. Цей не вигідний ефект буде тим меншим, чим вищою є концентрація солі в розчині, тобто різниця між концентрацією катіонів поблизу від ДНК і на віддаленні від неї. Відповідно, *стабільність подвійної спіралі (і температура плавлення) залежить від іонної сили розчину* – зростає при підвищенні концентрації солі.

За фізіологічних умов подвійна спіраль є досить стабільною структурою, і саме в цій формі ДНК (і значна частина комплементарних ділянок РНК) існує в живих системах. Розходження ланцюгів, яке має тимчасово відбуватися при різноманітних функціональних процесах, потребує виконання роботи молекулярними машинами з використанням енергії гідролізу АТФ.

### Конформаційні параметри подвійної спіралі

Як для будь-якого іншого полімеру, конформаційна рухливість нуклеїнових кислот реалізується завдяки зміни кутів обертання навкруг ковалентних зв'язків. У складі подвійної спіралі ці обертальні рухи обмежені міцними стекінг-взаємодіями та комплементарними водневими зв'язками. Дозволені обертання навкруг різних зв'язків призводять в решті-решт до зміни взаємної орієнтації основ і пар основ. Саме набором параметрів, що описують такі відносні орієнтації, прийнято характеризувати конформації подвійних спіралей.

В основі номенклатури конформаційних параметрів подвійних спіралей нуклеїнових кислот, прийнятої спеціальною нарадою ЕМВО (European Molecular Biology Organization), лежить стандартна локальна декартова система координат, що ставиться у відповідність до кожної пари основ (рис. 3.8): початок координат знаходиться в геометричному центрі пари; площина  $X$ - $Y$  збігається з площиною пари, вісь  $Y$  є паралельною щодо лінії, яка з'єднує  $C1'$ -атоми, і спрямована на ланцюг, який береться за перший; вісь  $Z$  спрямована на площину наступної (у напрямку  $5'$ - $3'$  по першому ланцюгу) пари основ; напрямок осі  $X$  обирається за правилами утворення правої декартової системи координат – жолобок, на який вона спрямована, позначається як “великий”. Насправді, терміни “великий / маленький жолобок” не обов'язково мають буквально значення (див. нижче).

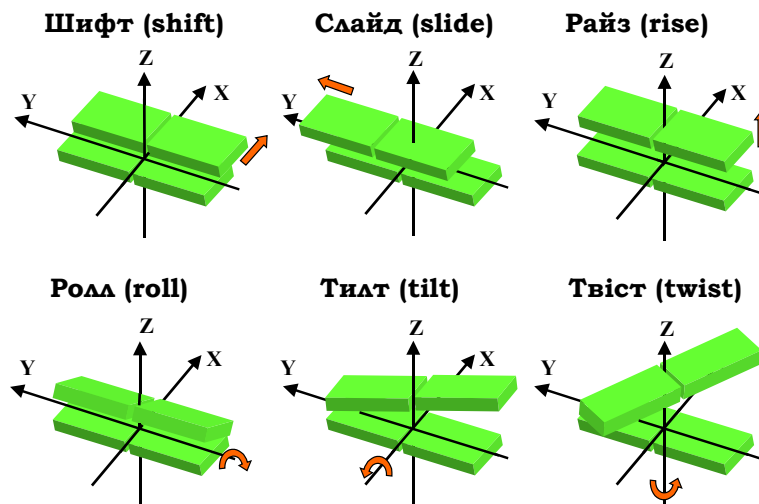


**Рис. 3.8.** Дві пари основ у молекулі ДНК і локальна декартова система координат, пов'язана з однією з них.



Для опису відносного положення двох сусідніх пар достатньо шість параметрів: три зсуви наступної пари відносно попередньої по трьох осях (будь-який зсув можна розкласти на три компоненти) і три кути обертання другої пари відносно першої навколо трьох осей (рис. 3.9):

- Шифт (shift) характеризує зсув у напрямку одного із жолобків.
- Слайд (slide) – зсув у напрямку цукрофосфатного остова.
- Райз (rise) – відстань між парами основ уздовж осі  $Z$ .
- кут рола (roll) характеризує локальний вигин подвійної спіралі в напрямку одного із жолобків.
- Кут тилта (tilt) – вигин у напрямку цукрофосфатного остова.
- Твіст (twist) є кутом обертання навкруг осі  $Z$ , який характеризує ступінь спірального закручення (для правої спіралі прийнято використовувати позитивні значення твіста).

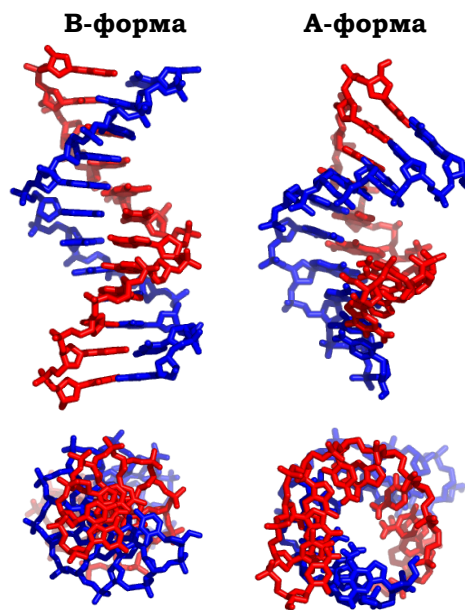


**Рис. 3.9.** Шість параметрів, які описують конформацію двох сусідніх пар основ

Аналогічно, шість параметрів характеризують відносне положення основ, що належать до однієї пари: зсув уздовж осі  $X$  і обертання навколо неї (*shear* і *buckle*), зсув і обертання відносно осі  $Y$  (*stretch* і *propeller*) та відносно осі  $Z$  (*stagger* і *opening*). Яким чином відбуваються відповідні рухи однієї основи відносно іншої легко зрозуміти з рис. 3.9.

### Структурні форми ДНК

ДНК із майже будь-якою послідовністю пар основ може існувати у двох структурних формах, що позначаються як А і В. Деякі усереднені параметри двох форм у фібрилах ДНК наведено в табл. 3.1, структури показано на рис. 3.10.



**Рис. 3.10.** В- (структура з рис. 3.6)  
та А- (декамер, 401D) форми ДНК у двох проекціях

Обидві *спіралі є правими*. Основні відмінності між двома формами ДНК зумовлені різною конформацією цукру – С2'- та С3'-ендо. Саме внаслідок цього в А-формі порівняно з В-формою:

- Скорочується відстань між фосфатними залишками.
- Зменшується ступінь правої спіральної закрутки.
- Кожна пара основ суттєво нахиляється до осі спіралі (у В-формі пари майже перпендикулярні до осі).
- Унаслідок нахилу в А-формі скорочується відстань між сусідніми парами основ уздовж осі спіралі (проекція райза на вісь спіралі).

- Кожна пара суттєво зсунута до периферії спіралі в А-формі, вісь спіралі проходить майже через центр пари в В-формі.
- “Великий” і “маленький” жолобки є справді великим і маленьким для В-форми, в А-формі їхні відносні розміри міняються місцями.

Таблиця 3.1. Порівняння між А- і В-формами подвійних спіралей ДНК

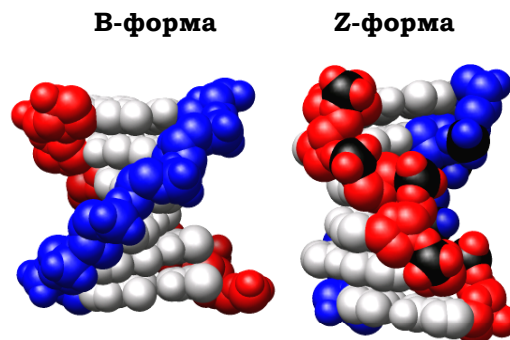
	А-форма	В-форма
Твіст $\Omega$ (°)	+32,7	+36,0
Кількість пар основ на виток спіралі ( $360^\circ/\Omega$ )	11	10
Слайд (Å)	-1,5	0
Ролл (°)	11,4	-0,2
Райз (Å)	3,3	3,3
Відстань між парами вздовж осі спіралі (Å)	2,6	3,4
Конформація дезоксирибози	C3'-ендо	C2'-ендо
Відстань між сусідніми фосфатами по ланцюгу (Å)	5,9	7,0
Великий жолобок: глибина (Å), ширина (Å)	13,5 2,7	8,5 11,7
Маленький жолобок: глибина (Å), ширина (Å)	2,8 11,0	7,5 5,7

Крім того, є різниця і в конформаційній рухливості: А-форма однорідніша за своєю структурою, В-форма допускає ширші варіації параметрів залежно від послідовності. Узагалі, частіше кажуть не про форми ДНК, а про А- і В-родинаи структурних форм.

**За фізіологічних умов у розчині ДНК існує тільки у В-формі**, перехід в А-форму здійснюється за досить екстремальних умов *in vitro* – за певних концентрацій солі та спиртів у розчині, при зниженні вологості у фібрилах. Але це не означає, що А-форма не є фізіологічною. По-перше, через те, що рибоза не може набувати С2'-ендо конформації,

усі **подвійні спіралі РНК існують в А-формі** за фізіологічних умов. З тієї ж причини набувають А-форми гібридні подвійні спіралі РНК-ДНК, які виникають тимчасово, наприклад, під час транскрипції. Крім того, ДНК може переходити в А-форму або наближену до неї в комплексах із білками.

Ще одна регулярна форма подвійної спіралі реалізується тільки для альтернуючих послідовностей poly(GC) (коли G і C чергуються по ланцюгу). Це так звана **Z-форма** (рис. 3.11). Її назва пов'язана з тим, що лінія, яка з'єднує фосфатні залишки у складі Z-форми нагадує зигзаг. Найбільш вражаюча особливість цієї форми ДНК полягає в тому, що це – **ліва спіраль** із приблизно 12 парами основ на виток. Альтернуючі послідовності G та C є в природних ДНК, але перехід у Z-форму відбувається *in vitro* за умов, які є дуже далекими від фізіологічних – при концентрації NaCl 2,5 моль/л. Біологічне значення Z-форми залишається не зовсім зрозумілим, хоча знайдені білки, які мають високу спорідненість саме до неї, тобто можуть індукувати B→Z перехід *in vivo*.



**Рис. 3.11.** В- (гексамерний фрагмент структури з рис. 3.6) і Z- (гексамер CGCGCG, 1DCG) форми ДНК. Чорним кольором на структурі праворуч позначені атоми фосфору

### Конформація В-ДНК

Структура В-форми ДНК не є абсолютно регулярною. Конформаційні особливості та особливості стекингу в різних контактах між парами основ призводять до суттєвої залежності локальної конформації подвійної спіралі від послідовності – можна сказати, що послідовність несе інформацію про структурні особливості ДНК (подібно до того, як амінокислотна послідовність зумовлює просторову структуру білка).

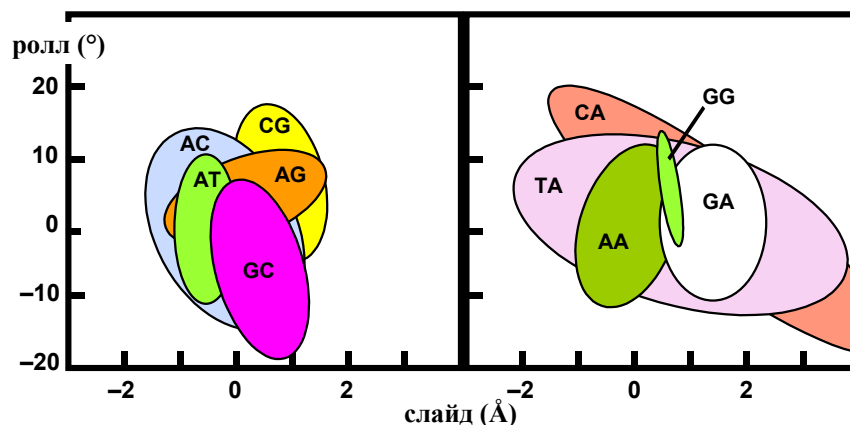
Щоб проілюструвати залежність конформації ДНК від послідовності, у табл. 3.2 наведені значення кута твіста (див. рис. 3.9) для 10 типів контактів між парами основ у кристалах вільної ДНК і білково-нуклеїнових комплексів.

Таблиця 3.2. Твіст ДНК для 10 типів контактів між парами основ (за даними Олсон та ін.)

Тип контакту	Середній кут твіста $\pm$ стандартне відхилення ( $^{\circ}$ )	
	Кристали вільної ДНК	Білково-нуклеїнові комплекси
AA/TT	35,5 $\pm$ 3,6	35,1 $\pm$ 3,9
AG/CT	30,6 $\pm$ 4,7	31,9 $\pm$ 4,5
GA/TC	39,6 $\pm$ 3,0	36,3 $\pm$ 4,4
GG/CC	35,3 $\pm$ 4,9	32,9 $\pm$ 5,2
AC/GT	33,1 $\pm$ 4,6	31,5 $\pm$ 4,2
AT	31,6 $\pm$ 3,4	29,3 $\pm$ 4,5
GC	38,4 $\pm$ 3,0	33,6 $\pm$ 4,7
CA/TG	37,7 $\pm$ 9,3	37,3 $\pm$ 6,5
CG	31,3 $\pm$ 4,7	36,1 $\pm$ 5,5
TA	43,2 $\pm$ 5,5	37,8 $\pm$ 5,5
Середнє по всіх контактах	35,6	34,2

Твіст у комплексах є дуже близьким до значення твіста вільної ДНК у розчині за фізіологічних умов – середнє значення для всіх контактів відповідає 10,1 парам основ на виток спіралі у кристалах і 10,5 парам основ у комплексах з білком. Як видно з таблиці, твіст варіює в досить широких межах – ступінь локальної спіральної закрутки ДНК суттєво змінюється залежно від послідовності пар основ. Крім того, варіюють також і стандартні відхилення від середнього значення твіста. Останній параметр показує, наскільки легко значення твіста може відхилятися від переважного середнього значення, тобто характеризує конформаційну рухливість контакту даного типу.

Інші конформаційні параметри подвійної спіралі (як середні значення, так і ширина розподілу) також залежать від послідовності пар основ. Для прикладу на рис. 3.12 показано співвідношення між ролом і слайдом (доступний сегмент конформаційного простору) для 10 контактів. Загалом конформаційна рухливість негативно корелює з ефективністю стекингу: вона є вищою для піримідин-пуринових і АТ-збагачених контактів. При цьому 6 конформаційних параметрів, наведених на рис. 3.9, не є незалежними один від одного – їхні зміни є взаємно пов'язаними.



**Рис. 3.12.** Усередині зафарбованих зон – доступні для кожного з 10 контактів значення двох конформаційних параметрів для кристалів ДНК (за даними Олсон та ін.)

Характер залежності конформації ДНК від послідовності є насправді ще складнішим: конформаційні особливості визначаються не тільки найближчими сусідами – типом контакту між парами основ, але й контекстом послідовності, де даний контакт знаходиться.

Отже, хоча конструкція молекули ДНК є простішою за білкову, на сьогодні неможливо точно передбачити деталі структури ДНК, виходячи з її послідовності. Залежні від послідовності особливості структури подвійної спіралі та потенціал щодо конформаційних змін – основа механізму специфічного впізнання послідовностей ДНК білками.

### Білково-нуклеїнові взаємодії

#### Структурна класифікація білків, які взаємодіють із ДНК

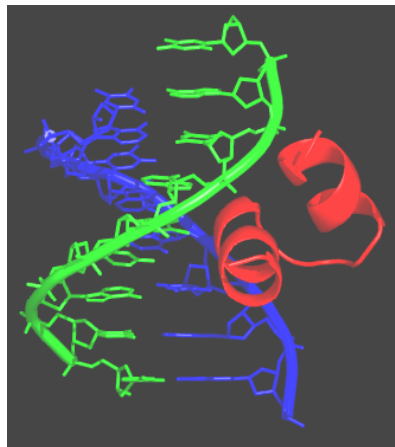
Так само, як більшість глобулярних білків можна вписати в обмежений набір стандартних укладок глобули (розділ 2), структурні елементи білків, які безпосередньо взаємодіють з ДНК, можна розділити на досить обмежену кількість стандартних типів – ДНК-зв'язувальних структурних мотивів.

**Мотив НТН** (helix-turn-helix, спіраль–поворот–спіраль), приклад якого зображено на рис. 3.13, є найбільш розповсюдженим. Мотив складається із двох  $\alpha$ -спіралей, часто майже перпендикулярних одна до одної, з'єднаних короткою перемичкою. Одна із цих спіралей (яка впізнає) виконує роль “головки, що читає”, взаємодіючи з екзоциклічними групами азотистих основ у великому жолобку подвійної спіралі. Орієнтація цієї спіралі відносно ДНК може варіювати для різних білків від приблизно перпендикулярної відносно осі подвійної спіралі до паралельної цукрофосфатному остову. Друга  $\alpha$ -спіраль часто утворює допоміжні контакти з цукрофосфатним остовом. Взаємодія  $\alpha$ -спіралі з великим жолобком часто зустрічається і для інших ДНК-зв'язувальних мотивів. Причина полягає в майже точній просторовій відповідності між цими двома елементами – можна сказати, що  $\alpha$ -спіраль є просторово комплементарною великому жолобку.

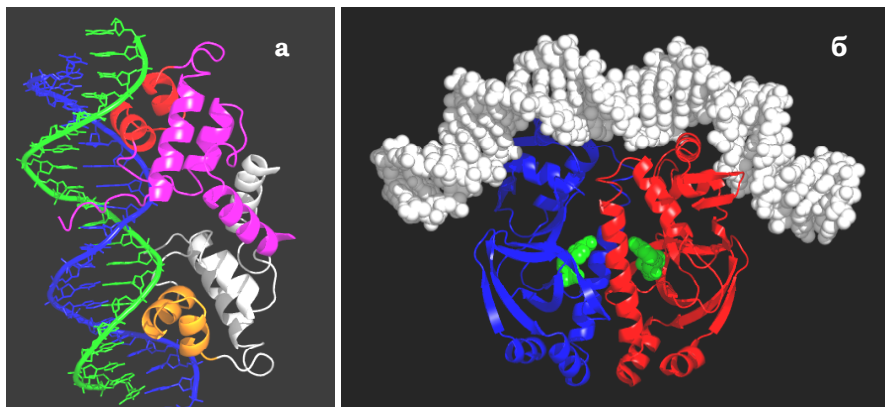
Двоспіральний НТН-мотив утримується в структурі того чи іншого структурного домену – у випадку репресора фага  $\lambda$ , у структурі  $\alpha$ -спірального пучка (рис. 3.14, а). Часто білки, що містять НТН-мотив, взаємодіють з ДНК у вигляді гомодимеру (дві однакові субодиниці) – це особливо характерно для прокаріотичних білків. Тоді дві однакові “читаючі головки” впізнають дві однакові послідовності, що симетрично розташовані в сайті взаємодії.

Інший приклад такого ж типу – катаболітний активаторний білок (CAP – Catabolite Activator Protein) *Escherichia coli* (рис. 3.14, б). При взаємодії гомодимеру CAP зі специфічним сайтом відбувається досить знач-

ний вигин ДНК у бік білка (подвійна спіраль “огортає” білок). Спорідненість білка до цього сайту залежить від ліганду – сАМР (циклічного АМР – 5'-фосфат у складі нуклеотиду утворює внутрішній ковалентний зв'язок з 3'-ОН групою рибози). Зв'язування сАМР індукує структурні зміни в молекулі білка (див. розділ 2), у результаті змінюється робоча поверхня і білок набуває здатності впізнавати специфічну послідовність.



**Рис. 3.13.** НТН-мотив репресора бактеріофага  $\lambda$  у комплексі з ДНК (1LMB)



**Рис. 3.14.** Комплекси з ДНК гомодимерів репресора бактеріофага  $\lambda$  (а, 1LMB) і білка CAP *E. coli* (б, 1CGP – зеленим кольором позначено молекули сАМР)



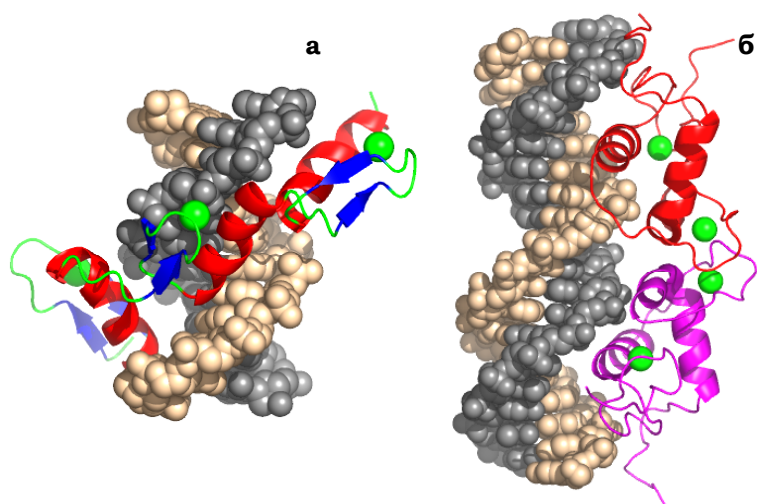
Розповсюджений клас білкових мотивів, що взаємодіють із ДНК, – **Zn-координуючі мотиви**. Це дуже невеличкі структурні елементи, які складаються з обмеженої, недостатньої для формування жорсткої глобули, кількості сегментів вторинної структури. Іон цинку надає жорсткості такому мотиву, утворюючи координаційні зв'язки з чотирма амінокислотними залишками, частіше Cys або His. Координаційний зв'язок формується неподіленою парою електронів атомів O, N або S, яка узагальнюється з атомом металу шляхом часткового “перетікання” на його низьку незаповнену орбіталь. Чотири такі зв'язки створюють жорсткий каркас, який утримує білкову поверхню, що має взаємодіяти з ДНК.

Один зі структурних мотивів такого типу, що часто зустрічається, – **цинковий палець** (Zn finger, рис. 3.15, а). Він має дуже просту будову: одна  $\alpha$ -спіраль і маленький  $\beta$ -шар із двох  $\beta$ -ділянок; одна з них і  $\alpha$ -спіраль містять залишки, що утворюють координаційні зв'язки з Zn. Як правило, кілька таких пальців (три на рис. 3.15, а), з'єднаних перемичкою, спіралью обгортають великий жолобок.

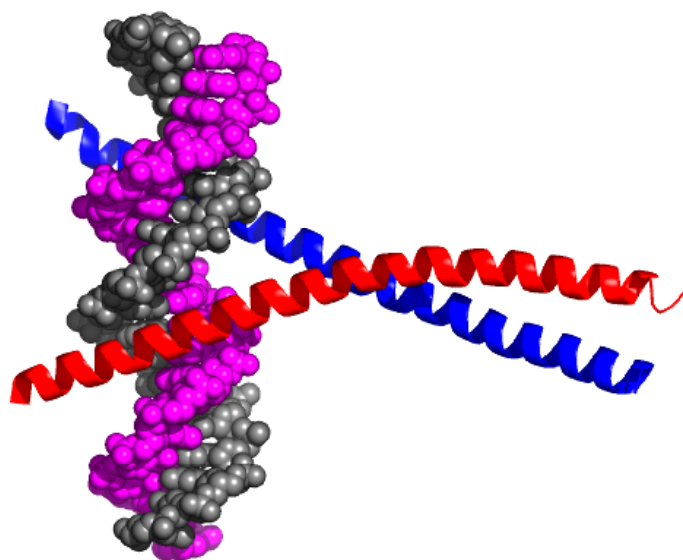
Інший приклад – ДНК зв'язувальний домен гормонового рецептора (рис. 3.15, б). Гормонові рецептори – це еукаріотичні фактори транскрипції, здебільшого гомодимери, які набувають споріденості з певними елементами послідовності ДНК після зв'язування з білком стероїдного гормону (розділ 6). ДНК-зв'язувальний домен має дві пари структурних елементів, а саме  $\alpha$ -спіраль – суміжна нерегулярна петля, кожна пара координує іон Zn. Одна пара елементів бере участь у димеризації, інша – у взаємодії з ДНК. Зустрічаються також інші Zn-координуючі мотиви, але всі вони, як і розглянуті, реалізують взаємодію  $\alpha$ -спіралі з великим жолобком.

Ще один мотив, який взаємодіє з великим жолобком ДНК за рахунок  $\alpha$ -спіралей – **лейциновий зіпер** (Leu zipper, рис. 3.16). Він являє собою дві довгі  $\alpha$ -спіралі, кожна з яких можна розділити на дві частини. Одна частина обох спіралей здійснює димеризацію – за рахунок гідрофобних взаємодій між неполярними амінокислотними залишками (часто Leu, звідки назва мотиву) утворюється трохи закручена подвійна спіраль (coiled coil). Друга частина кожної  $\alpha$ -спіралі взаємодіє з великим жолобком ДНК.

Варіацією лейцинового зіпера є мотив **спіраль – петля – спіраль** (helix – loop – helix) – різниця лише в тому, що кожна довга  $\alpha$ -спіраль розділена на дві коротші, з'єднані петлею: одна пара спіралей взаємодіє між собою, інша – з ДНК.



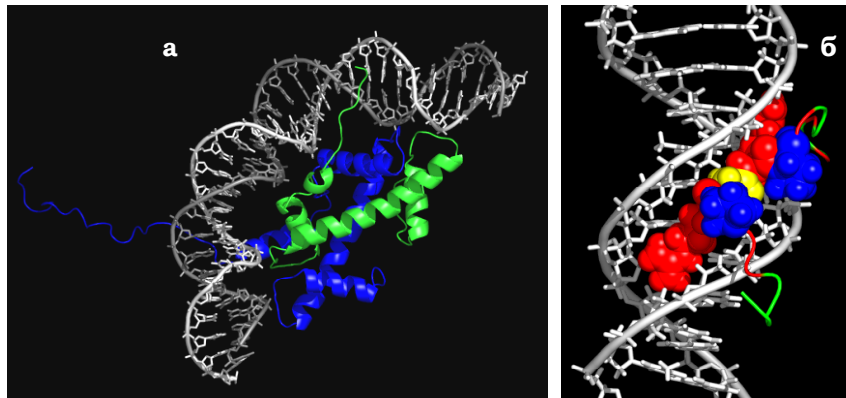
**Рис. 3.15.** Три цинкових пальця у складі ДНК-зв'язувального білка (а, 1ZAA) і ДНК-зв'язувальний домен (димер) глюкокортикоїдного рецептора (б, 1GLU). Зелені сфери – іони Zn



**Рис. 3.16.** Лейциновий зіпер (1NWQ)

Усе це не означає, що тільки  $\alpha$ -спіраль використовується для взаємодії білків з великим жолобком ДНК. У великий жолобок добре вкладається також  $\beta$ -шар із двох  $\beta$ -ділянок – саме такий шар є взаємодіючим елементом для білків типу метіонінового репресора. Численний клас білків становить родину імуноглобуліноподібних транскрипційних факторів (глобула має укладку імуноглобулінового типу, див. рис. 2.12, б), які реалізують взаємодії перемичок між ділянками  $\beta$ -структури з великим жолобком. Отже, хоча  $\alpha$ -спіраль використовується частіше, реалізуються і всі інші можливості.

Усе описане не означає також, що білки взаємодіють з ДНК тільки через великий жолобок (хоча через великий – частіше). На рис. 3.17 зображено приклади взаємодії перемичок між  $\alpha$ -спіралями (а) чи витягнутого поліпептидного ланцюга (б) з маленьким жолобком. У першому випадку – димер гістонів у складі нуклеосоми (див. розділ 4) – маємо ще один структурний мотив, який зустрічається і в інших ДНК-зв'язувальних білках, а саме гістонову укладку: одна довга  $\alpha$ -спіраль фланкована двома короткими. Димер двох таких укладок взаємодіє з  $\sim 2,5$  витками подвійної спіралі. У випадку гістонів ця взаємодія не є специфічною: утворюються електростатичні контакти позитивно заряджених амінокислотних залишків з фосфатами ДНК, але з боку маленького жолобка.

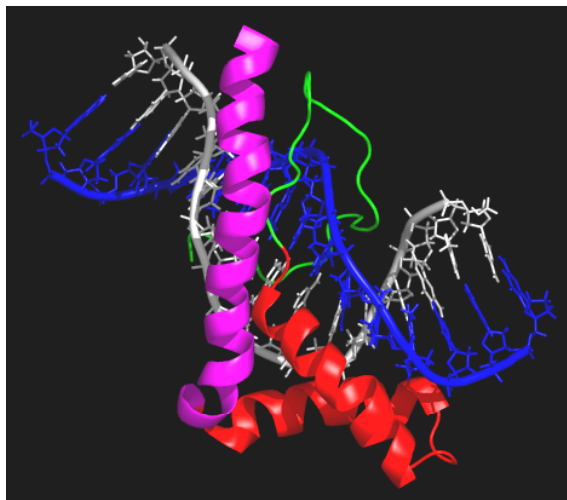


**Рис. 3.17.** Димер гістонів H3-H4 (а, 1AOI)  
і АТ-гак у складі білка HMGA (б, 2EZD)

Другий приклад (рис 3.17, б) стосується білків HMGA – білків групи високої рухливості (High Mobility Group) типу А (HMG – істо-

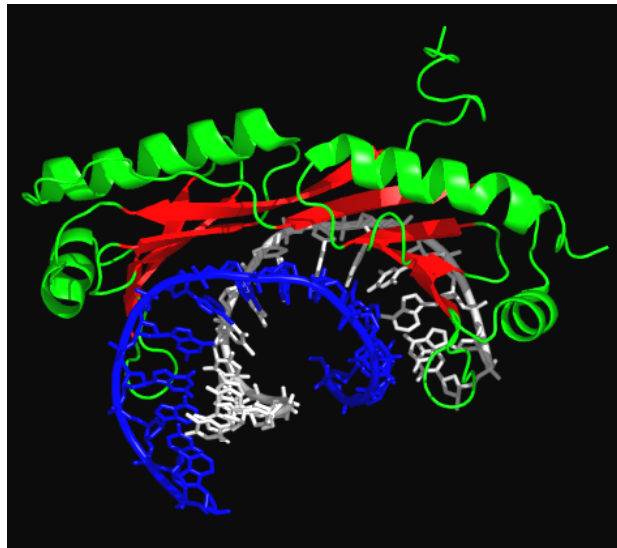
рична назва для білків різних структурних класів, мається на увазі рухливість при електрофорезі). Білки HMGА взагалі не мають глобулярної структури – є неупорядкованими. Вони містять 3 пентапептидні елементи (Pro-Arg-Gly-Arg-Pro), що називаються АТ-гаками (АТ-hook). Такий гак добре укладається у звужений маленький жолобок невеликої ділянки подвійної спіралі, збагаченої АТ-парами (тобто АТ-гак має специфічність до невеликих АТ-збагачених сайтів, завдяки чому він і отримав свою назву).

Наступні приклади стосуються дуже важливих випадків взаємодії елементів регулярної вторинної структури білків з маленьким жолобком. Оскільки  $\alpha$ -спіраль і дволанцюговий  $\beta$ -шар добре укладаються у великий жолобок, це автоматично означає, що в маленькому жолобку їм не вистачає місця. Відповідно, така взаємодія буде можливою тільки за умови суттєвої деформації подвійної спіралі ДНК (розширення жолобка) – ефект, що має важливі функціональні наслідки. На рис. 3.18 зображено досить розповсюджений мотив – НМG-бокс, який, зокрема, входить до складу білків НМGВ (ще один клас білків НМG). Мотив складається з трьох  $\alpha$ -спіралей, дві з них (позначені червоним) вбудовуються (інтеркалюють) у маленький жолобок. Це супроводжується розкрученням подвійної спіралі й вигином на  $\sim 80^\circ$  у протилежний щодо білка бік. Серед НМG-боксів є як специфічні до певних послідовностей пар основ, так і неспецифічні.



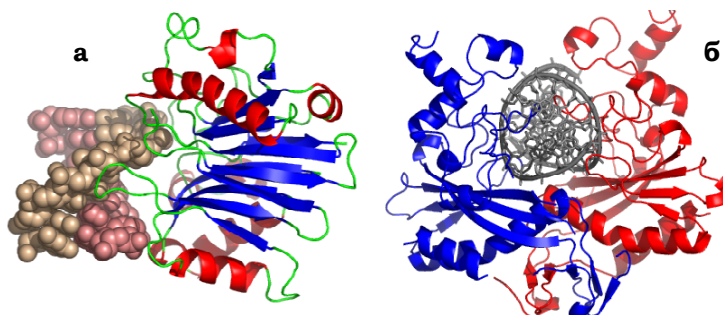
**Рис. 3.18.** НМG-бокс у складі білка НМGВ (1J5N)

У складі білка ТВР (TATA-box Binding Protein) – важливого елемента ініціації транскрипції в еукаріотів – досить широкий  $\beta$ -шар взаємодіє з маленьким жолобком подвійної спіралі в зоні ТАТА-боксу – регуляторного елемента послідовності ДНК (розділ 6). Наслідком цієї взаємодії також є значна деформація подвійної спіралі з її розкрученням і значним вигином у протилежний від білка бік (рис. 3.19). Вигин, як і у випадку НМГ-боксу, підсилюється інтеркаляцією двох гідрофобних амінокислотних залишків між парами основ. У місці інтеркаляції порушуються стекінг-взаємодії та утворюється кінк (kink) – різкий злам подвійної спіралі.



**Рис. 3.19.** ДНК-зв'язувальний домен ТВР у комплексі з ТАТА-боксом (1QNE)

На відміну від розглянутих факторів транскрипції та структурних білків, ферменти, які працюють на ДНК, досить важко описати в термінах простих структурних мотивів. Ферменти використовують складні комбінації різноманітних елементів білкової структури для впізнання ДНК і зв'язування з нею. Два приклади – мономерна ДНКаза I (неспецифічна ендонуклеаза, яка здійснює одноланцюговий розріз в ділянці маленького жолобка) і гомодимерна рестриктаза *EcoRV* – зображені на рис. 3.20.



**Рис. 3.20.** Комплекси з ДНК ДНКази I (а, 1DNK) і рестриктази *EcoRV* (б, 1AZ0)

### Принципи білково-нуклеїнового впізнання

Усі випадки взаємодії численних білків із ДНК можна поділити на дві категорії: *неспецифічне зв'язування* білка з ДНК будь-якої послідовності (характерна константа зв'язування  $K \sim 10^5\text{--}10^6$  л/моль, визначення константи зв'язування див. у розділі 1) і *специфічне впізнання* білком певної послідовності пар основ (із константою  $\sim 10^9\text{--}10^{10}$  л/моль). Білки, що здійснюють таке впізнання, як правило, взаємодіють і з будь-якою іншою ДНК неспецифічно. Розглянемо на простому прикладі, що практично означають наведені цифри.

Нехай у бактеріальній клітині радіусом 1 мкм міститься 10 молекул певного білка та 4 млн 600 тис. пар основ ДНК, де присутня одна специфічна ділянка (оператор). Виходячи з об'єму клітини, можна легко розрахувати загальні молярні концентрації білка та оператора (скориставшись числом Авогадро). Концентрація комплексу з оператором визначається рівнянням (1.6), де вільні концентрації білка та оператора дорівнюють їхнім загальним концентраціям мінус концентрація комплексу. Звідки, якщо відома  $K$ , можна розрахувати концентрацію комплексу.

Якщо білок зв'язується неспецифічно ( $K = 10^5$  л/моль), відношення концентрації комплексу до загальної концентрації оператора дорівнює  $4 \cdot 10^{-4}$  – оператор практично є вільним від білка. Та якщо замість концентрації оператора використати загальну концентрацію потенційних сайтів зв'язування на ДНК (білок зв'язується будь-де), яка є в 4 млн 600 тис. разів вищою (кожна пара основ потенційно може бути початком сайту зв'язування), то відношення кон-

центрації комплексу до загальної концентрації білка дорівнюватиме 0,995 – практично білок увесь час зв'язаний із ДНК.

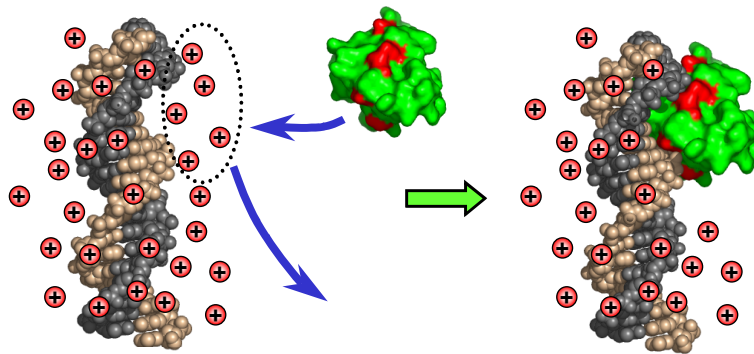
При специфічному зв'язуванні ( $K = 10^{10}$  л/моль), відношення концентрації комплексу до загальної концентрації оператора дорівнює 0,97 – десяти молекул білка виявляється достатнім, щоб більшу частину часу оператор був зв'язаним. При цьому з оператором у нашому прикладі в даний момент часу може бути зв'язана лише одна молекула білка, решта перебуває не у вільному стані, а на ДНК, взаємодіючи з нею неспецифічно.

Чим визначається висока специфічність зв'язування? *Головне правило білково-нуклеїнового впізнання – відсутність жорстких правил.* Розглянуті різні структурні мотиви білків є різними еволюційними рішеннями для специфічної взаємодії з тією чи іншою послідовністю: існує багато шляхів для того, щоб сформувати білкову поверхню для впізнання послідовності пар основ. Не існує і будь-якого коду впізнання – чіткої відповідності між амінокислотними залишками та парами основ. Але є певні загальні закономірності, частина яких уже має бути зрозумілою з розгляду ДНК-зв'язувальних структурних мотивів. Зокрема, *білково-нуклеїновий інтерфейс* частіше представлений парою  $\alpha$ -спіраль – великий жолобок завдяки хорошій просторовій відповідності між цими двома елементами. Проте використовуються також і  $\beta$ -структура, і перемички між елементами вторинної структури для впізнання пар основ у великому жолобку. Маленький жолобок також використовується для впізнання послідовностей пар основ, причому у випадку взаємодії елементів вторинної структури білка ( $\alpha$ -спіраль чи  $\beta$ -структура) із маленьким жолобком така взаємодія супроводжується розкриттям маленького жолобка. Таке розкриття стає можливим унаслідок значної деформації подвійної спіралі (розкручування, вигин, кінк у результаті порушення стекінг-взаємодій).

На поверхні ДНК розташовані фосфатні залишки та донорні / акцепторні групи азотистих основ у жолобках. Саме ці групи й залучаються до контактів з амінокислотними боковими залишками та пептидними групами на поверхні білка. **Існує декілька типів контактів (взаємодій) між ДНК і білками:**

**Електростатичні взаємодії** між позитивно зарядженими амінокислотними залишками та негативно зарядженими фосфатами присутні майже завжди в білково-нуклеїнових комплексах. Зрозуміло, що електростатичні взаємодії відповідають за неспецифічне зв'язування. Як правило, вони додатково стабілізують також і специфічні комплекси. Електростатичне зв'язування білка з нуклеїновою кислотою має цілком ентропійну природу (див. також розділ 1). Висока концентра-

ція негативних зарядів (фосфатів) на поверхні ДНК зумовлює формування навкруг ДНК “іонної атмосфери” з досить високою локальною щільністю катіонів (рис. 3.21). Зв'язування позитивно зарядженого білка призводить до визволення частки катіонів у зовнішній розчин, тобто до зростання неупорядкованості (ентропії) у системі. Енергетичний вигравш від зв'язування білка (його спорідненість до ДНК) є тим більшим, чим меншою є концентрація солі в розчині, тобто різниця між концентрацією катіонів поблизу від ДНК і на віддаленні від неї.

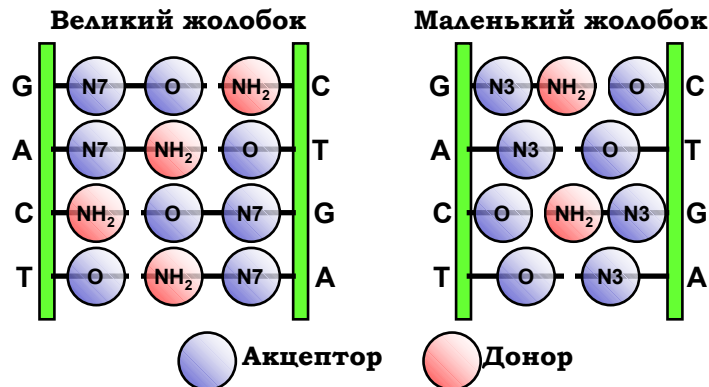


**Рис. 3.21.** Електростатичне зв'язування білка (червоним позначені позитивно заряджені амінокислотні залишки) з ДНК за рахунок визволення неорганічних катіонів

**Водневі зв'язки** між донорно-акцепторними групами білка та фосфатами й екзоциклічними групами азотистих основ. Водневі зв'язки, оскільки вони потребують чіткої взаємної орієнтації донора й акцептора, відіграють роль головного фактора специфічного впізнання. Кожна послідовність пар основ утворює в жолобках подвійної спіралі власний патерн донорних і акцепторних груп (рис. 3.22, див. також рис. 3.7), який і може впізнаватися поверхнею білка. Слід зазначити, що цей патерн більш варіабельний у великому жолобку, де легше розрізнити пари основ і динуклеотидні контакти – це ще одна причина, чому саме великий жолобок частіше використовується для впізнання. Деякі амінокислотні залишки здатні утворювати два водневі зв'язки з азотистою основою, наприклад, Arg із гуаніном. Проте Arg контактує й з усіма іншими основами. Чи буде певний залишок залученим до утворення водневого зв'язку, і якого саме, залежить від його орієнтації на поверхні. Певні закономірності (своєрідний код, коли залишок пев-



ного типу в певному місці утворює зв'язок із певною основою) спостерігаються іноді тільки в межах однієї родини структурних мотивів, але й тоді такі закономірності не мають абсолютного характеру.



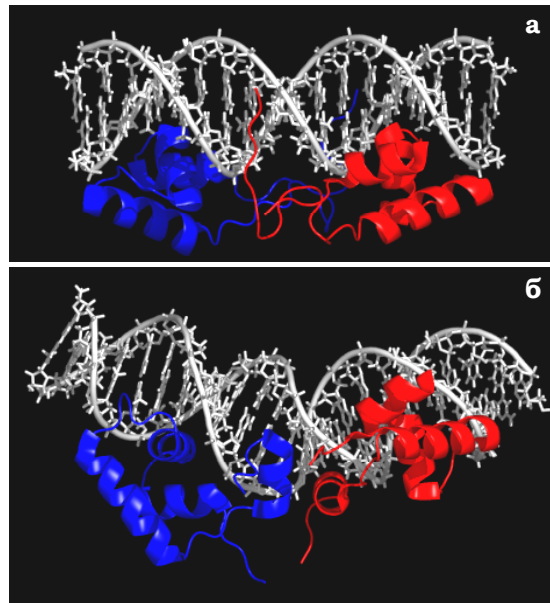
**Рис. 3.22.** Патерни донорів і акцепторів водневого зв'язку в жолобках подвійної спіралі для зазначеної послідовності чотирьох пар основ

**Водневі зв'язки, опосередковані молекулами води.** Вода взаємодіє з поверхнями як білків, так і ДНК. Визволення води з білково-нуклеїнового інтерфейсу вносить додаткову ентропійну складову у стабілізацію комплексу. Але досить часто окремі молекули води можуть залишатися в інтерфейсі, виконуючи роль біфункціональної зшивки – утворюючи водневі зв'язки з білковими групами та фосфатами або групами азотистих основ.

**Гідрофобні контакти** можуть здійснюватись за участю метильної групи тиміну у великому жолобку (рис. 3.7), але суттєвішими є гідрофобні взаємодії при інтеркаляції неполярних амінокислотних залишків між парами основ при деформації подвійної спіралі внаслідок занурення елементів вторинної структури білка в маленький жолобок (рис. 3.18, 3.19).

**Головною умовою реалізації названих контактів між ДНК і білком є заємна підгонка структури взаємодіючих елементів унаслідок відповідних конформаційних перетворень.** Так, шляхом певних конформаційних змін подвійної спіралі (вигини спіралі, зміни твіста тощо) змінюється розмір жолобків, хімічні групи "підводяться" під водневі зв'язки та міцні електростатичні контакти, збільшується загальна поверхня, що може взаємодіяти з білком. Це означає, що для ефективного впізнання необхідна певна конформація подвійної спіралі

та / або певні зміни цієї конформації. І те, й інше визначається послідовністю пар основ. Отже, основою для білково-нуклеїнового впізнання є структурно-динамічний поліморфізм нуклеотидних послідовностей – залежні від послідовності структурні особливості та конформаційна рухливість подвійної спіралі.



**Рис. 3.23.** Структура ДНК-зв'язувальних доменів димеру репресора лактозного оперона *E. coli* у комплексі з неспецифічною ДНК (а, 1OSL) та зі своїм специфічним оператором (б, 1L1M)

При цьому білок також часто потребує певної конформаційної підгонки під подвійну спіраль. Один із прикладів такої, досить значної, конформаційної зміни як у молекулі білка (формування додаткових  $\alpha$ -спіралей), так і у молекулі ДНК (вигин подвійної спіралі) при утворенні специфічного комплексу показано на рис. 3.23. Саме в процесі взаємної конформаційної підгонки білка та ДНК і реалізується специфічний патерн контактів у інтерфейсі між двома молекулами. Від того, наскільки легко потрібні конформаційні зміни відбуваються при зустрічі конкретного білка з конкретною послідовністю пар основ, залежить ефективність упізнання. Таким чином, головним механізмом упізнання послідовності ДНК білком є структурно-динамічна комплементарність двох молекул.

## Циркулярна ДНК

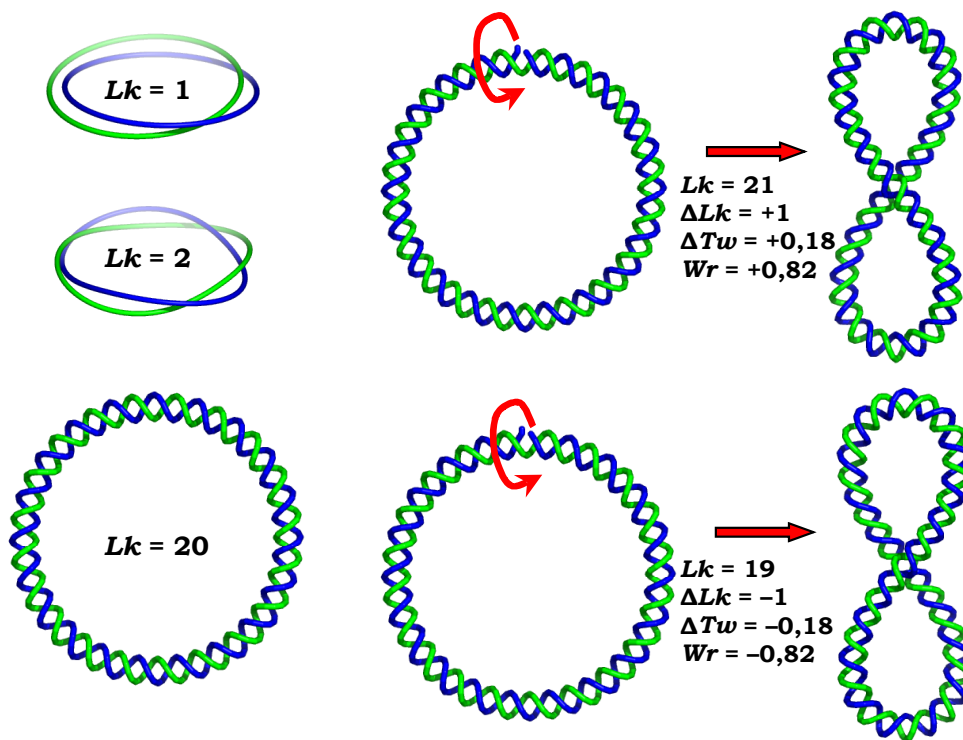
ДНК майже завжди існує *in vivo* у вигляді циркулярної ковалентно замкненої молекули (прокаріоти) або у формі, що є еквівалентною циркулярній – у вигляді петель, закріплених своїми кінцями на скелетних структурах клітинного ядра (еукаріоти, див. розділ 4). У молекулі, два кінці якої жорстко зафіксовані або з'єднані один з одним, виникають *топологічні* обмеження, і це має важливі біологічні наслідки.

Топологія – це розділ математики, присвячений вивченню таких властивостей просторових об'єктів, які не залежать від деформацій цих об'єктів. Центральним поняттям топології циркулярної ДНК є так зване число зчеплень  $Lk$  (linking number) двох полінуклеотидних ланцюгів – двох кілець, закручених одне навкруг одного в подвійну спіраль. Число зчеплень визначається як кількість перетинів одним циркулярним контуром поверхні, що натягнута на другий контур. У найпростішому випадку два кільця можна зчепити одне з одним один раз ( $Lk = 1$ , рис. 3.24), і будь-які деформації кілець (за умови збереження цілісності кожного кільця) не змінять цієї величини: кільця не можна ні розвести, ні збільшити ступінь зчеплень. Отже, і це одна з двох важливих властивостей числа зчеплень,  **$Lk$  системи двох кілець є постійною величиною** – топологічним інваріантом, – поки обидва кільця є інтактними (не містять розривів). Легко уявити два кільця, зчеплені одне з одним два (рис. 3.24) або сто разів, але не може бути двох з половиною зчеплень, тобто  **$Lk$  є цілим числом**. Саме ці дві властивості числа зчеплень і накладають топологічні обмеження на молекулу циркулярної ДНК.

На рис 3.24 показано планарну (вісь подвійної спіралі лежить в одній площині) циркулярну молекулу ДНК з  $Lk = 20$ . Зрозуміло, що для такої системи  $Lk$  дорівнює кількості витків подвійної спіралі.

Кількість витків подвійної спіралі позначається як твіст  $Tw$  (twist, дорівнює сумі кутів твіста для всіх пар основ, поділених на  $360^\circ$ , з позначкою “+” для правої спіралі та “-” для лівої), тобто для планарної циркулярної ДНК  $Lk = Tw$ . Найвигідніший твіст ДНК  $Tw_0$  (часто записується також як  $Lk_0$ ) визначається зовнішніми умовами та послідовністю пар основ (типове середнє значення за фізіологічних умов відповідає приблизно 10,5 парам основ на виток подвійної спіралі), тобто зовсім не повинен бути цілим числом. Лише тоді, коли  $Tw = Tw_0$  (припустимо, що ця умова реалізується для прикладу на

рис. 3.23, і  $T\omega_0 = 20$ ), при замиканні в кільце ДНК набуде найбільш енергетично вигідної планарної форми, у складі якої реалізується мінімальний вигин молекули.



**Рис. 3.24.** Ліворуч: два кільця з різним числом зчеплень  $Lk$ , у тому числі – планарна форма циркулярної ДНК з  $Lk = 20$ . Праворуч: збільшення чи зменшення ступеня закрутки подвійної спіралі на один оберт призводить відповідно до позитивної чи негативної надспіралізації

Якщо ж (як це частіше буває)  $T\omega_0$  не є цілим числом, замикання в кільце (точне зведення кінців обох ланцюгів) є можливим або за рахунок зміни твіста до цілого значення (треба підкрутити подвійну спіраль, щоб виставити один кінець точно напроти одного), або шляхом додаткових вигинів молекули з відхиленням від планарної форми. У будь-якому разі замикання в кільце буде можливим лише за умови деформацій молекули ДНК.

Отже, у загальному випадку, коли  $Lk$  (яке має бути цілим числом) не збігається з  $Tw_0$ , число зчеплень має дві складові згідно з рівнянням Уайта – Фуллера (James H. White, F. Brock Fuller):

$$Lk = Tw + Wr,$$

де райзинг  $Wr$  (writhing) – параметр, який є мірою відхилення осі подвійної спіралі від планарної конфігурації (для планарного кільця  $Wr = 0$ ). Розподіл величини  $Lk$  по двох складових залежить від механічних властивостей молекули ДНК: той факт, що  $Lk \neq Tw_0$ , означає наявність деформації (відхилення конформації від найвигіднішої) і ця деформація мінімізується, розподіляючись певним чином між зміною твіста  $\Delta Tw = Tw - Tw_0$  і ненульовим райзингом унаслідок зростання вигину молекули. Отже, загальною мірою деформації є величина

$$\Delta Lk = Lk - Tw_0 = \Delta Tw + Wr.$$

Оскільки зростання райзингу (при зростанні загальній деформації) пов'язане зі спіралізацією осі подвійної спіралі – надспіралізацією (supercoiling), величина  $\Delta Lk$  використовується як міра надспіралізації в циркулярній ДНК. Таким чином, надспіралізація є тим більшою, чим більше  $Lk$  відрізняється від  $Tw_0$ , і чим більша напруга, пов'язана з деформацією, накопичується в ДНК.

Описане ілюструє рис. 3.24 на прикладі ДНК із  $Tw_0 = 20$ . Якщо зробити односторонній розрив, збільшити кількість витків подвійної спіралі на 1 (закрутити подвійну спіраль торсійно) і зашити розрив, відновивши ковалентний зв'язок, то тим самим кількість зчеплень зросте на 1:  $\Delta Lk$  набуде значення +1, яке розподілиться між зміною твіста та райзингом. У цьому випадку кажуть про *позитивну надспіралізацію*, яка є топологічно еквівалентною зростанню закрутки подвійної спіралі. Відповідно, *розкручення* подвійної спіралі після розриву та наступне відновлення цілісності ланцюга призведе до *негативної надспіралізації* з  $\Delta Lk = -1$ .

Однакові циркулярні молекули ДНК, які різняться лише кількістю зчеплень (як три молекули з  $Lk = 19, 20$  та  $21$  на рис. 3.24), називаються *топоізомерами*. Зрозуміло, що практично застосувати описану процедуру для отримання різних топоізомерів досить важко. Але це можна зробити, здійснивши ензиматичне зшивання ДНК у кільце в різних умовах, що відповідають різним значенням  $Tw_0$ . Отримані топоізомери будуть відповідати найменш енергетичним

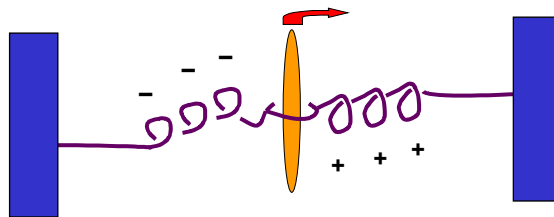
формам – таким, для яких значення  $Lk$  є максимально наближеними до значень найвигіднішого твіста. Щойно зациклення відбулося, значення  $Lk$  залишається незмінним. А зміна умов і нове значення  $T\omega_0$  приведе до зростання величини  $\Delta Lk$  у негативний чи позитивний бік і, відповідно, надспіралізації.

З іншого боку, значення  $Lk$  залишається незмінним лише за умови цілісності обох полінуклеотидних ланцюгів. Якщо внести хоча б один розрив у хоча б один із ланцюгів, два кінці ланцюга в місці розриву отримують свободу обертатися навкруг інтактного ланцюга. У результаті будь-яка надспіралізація і пов'язана з нею еластична напруга зникнуть – відбудеться *релаксація* циркулярної ДНК.

Для негативно надспіралізованої ДНК існує ще один шлях релаксації (принаймні часткової) без розривів – локальне розкручування подвійної спіралі. Припустимо, що лінійна молекула ДНК із  $T\omega_0 = 20$  (як на рис. 3.24) може існувати також у формі зі зруйнованим (розплавленим) одним витком подвійної спіралі, тобто з величиною  $T\omega_0 = 19$ . За фізіологічних умов розведення ланцюгів потребує зростання вільної енергії, тобто така форма є неможливою. Але локальне плавлення подвійної спіралі у складі негативного топоізомеру буде супроводжуватись також зниженням вільної енергії за рахунок релаксації, оскільки  $\Delta Lk = 19 - 20 = -1$  зміниться на  $\Delta Lk = 19 - 19 = 0$ . У результаті, якщо загальна вільна енергія знизиться, локальна дестабілізація подвійної спіралі стане імовірнішою. За таким самим механізмом – енергетично невигідне конформаційне перетворення стає вигідним за рахунок того, що знімає негативну надспіралізацію в кільці – у негативно надспіралізованій ДНК (за фізіологічних умов!) відбуваються й інші конформаційні перетворення (наприклад, перехід у Z-форму, яка характеризується розкрученням подвійної спіралі порівняно з B-ДНК). Оскільки більшість функціональних процесів, що відбуваються на ДНК, потребують локального руйнування подвійної спіралі, зрозуміло, що негативна надспіралізація активує ці процеси, полегшуючи таке руйнування на певних найменш стабільних ділянках.

З іншого боку, функціональні процеси (у першу чергу, транскрипція і реплікація) самі призводять до виникнення надспіралізації. Проходження такого процесу пов'язане з пересуванням (транслокацією) уздовж ДНК того чи іншого ферменту (*транслокази*), який локально руйнує подвійну спіраль. Такими транслоказами є РНК- і ДНК-полімерази, які розглядатимуться в наступних розділах, а також інші ферменти. Оскільки молекула ДНК – спіраль, пересування транслокази має супроводжуватись або її обертанням навкруг осі подвійної спіралі (на кшталт обертання гайки навкруг гвинта), або прокручуван-

ням самої подвійної спіралі (обертанням гвинта в гайці). Саме друга можливість і реалізується, оскільки транслоказа працює завжди у складі величезного мультибілкового комплексу – “гайка” є надто масивною. Крім того, транслоказа може бути закореною на скелетних структурах, що робить її обертання взагалі неможливим. Якщо (як це має місце у клітині) кінці ДНК жорстко зафіксовані на тих самих скелетних структурах і не можуть обертатися, транслокація буде створювати топологічні проблеми (рис. 3.25). У процесі пересування вздовж ДНК транслоказа руйнує подвійну спіраль попереду від себе й відновлює її позаду. Локальне розкручування подвійної спіралі попереду транслокази має бути компенсоване позитивною надспіралізацією, відновлення спіралі (закручування) – надспіралізацією негативною. У результаті попереду й позаду від транслокази виникають дві “хвилі” надспіралізації протилежного знаку (рис. 3.25).

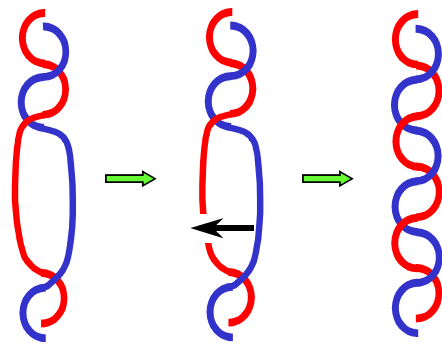


**Рис. 3.25.** Дві хвилі надспіралізації у процесі роботи транслокази

Зрозуміло, що накопичення такої надспіралізації (еластичної напруги) не може продовжуватись нескінченно: врешті-решт напруга буде блокувати процес транслокації. Отже, має бути спосіб розв'язувати цю проблему – знімати надспіралізацію, релаксуючи ДНК. Інструментом, використовуваним для цього, є спеціальні ферменти – **ДНК-топоізомерази** (topoisomerases). Оскільки зміна кількості зчеплень є можливою лише за умови порушення цілісності полінуклеотидних ланцюгів, загальний механізм роботи топоізомераз полягає у внесенні розриву (при цьому один кінець ланцюга тимчасово ковалентно пришивається до залишку Туг у активному центрі ферменту, а інший залишається вільним) і зворотному зшиванні цього розриву; зміна  $Lk$  відбувається у проміжку між цими двома подіями.

Топоізомерази поділяються на два класи. **Топоізомерази I** (мономерні білки), які поділяються на два підкласи Ia і Ib, здійснюють одноланцюговий розріз ДНК. **Топоізомерази Ia** (присутні як у про-

так і в еукаріотів) здатні релаксувати тільки негативно надспіралізовану ДНК. Фермент упізнає дестабілізовану ділянку подвійної спіралі (яка має бути присутньою в негативно надспіралізованій молекулі, див. вище), робить одноланцюговий розрив і протягує крізь нього інтактний ланцюг (рис. 3.26). У результаті число зчеплень змінюється на одиницю в напрямку зниження рівня негативно надспіралізації. Оскільки впізнається дестабілізована ділянка, топоізомераза Ia спрацьовує тільки за наявності досить високої негативно надспіралізації і не здатна релаксувати ДНК повністю – знизити  $\Delta Lk$  до 0. Основна роль топоізомераз цього типу – підтримувати певний оптимальний рівень надспіралізації в ДНК бактеріальної клітини.



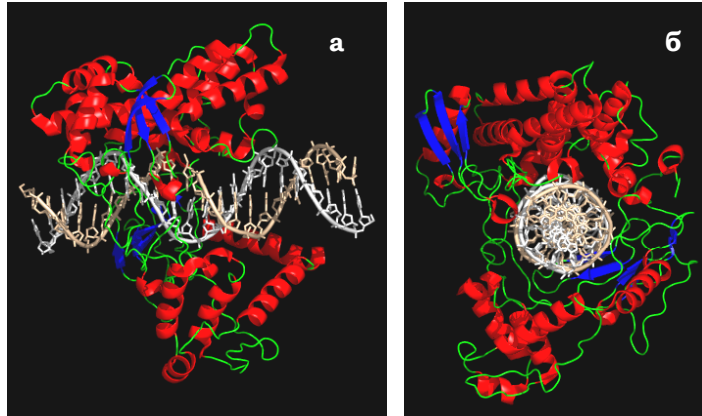
**Рис. 3.26.** Схема дії топоізомераз Ia

**Топоізомерази Ib** (тільки еукаріотичні) переводять будь-яку ДНК у максимально релаксований стан. Механізм їхньої дії є дуже простим: мультидоменний мономерний білок зв'язується з ДНК, оточуючи подвійну спіраль з усіх боків (рис. 3.27); здійснюється одноланцюговий розрив (див. рис. 3.27, a); один із кінців, які утворилися, отримує змогу вільно обертатися навколо інтактного ланцюга; після трьох-чотирьох випадкових обертів розрив зшивається, і фермент дисоціює. Оскільки оберти є вільними, вони здійснюються в напрямку зниження напруги – максимальної релаксації. Зрозуміло, що  $Lk$  в результаті роботи топоізомерази змінюється на величину, кратну одиниці, – на кількість обертів.

Таким чином, топоізомерази I обох типів змінюють число зчеплень шляхом зміни твіста подвійної спіралі. Іншою спільною ознакою топоізомераз I є незалежність їхньої активності від АТР: відновлення ціліс-



ності полінуклеотидного ланцюга відбувається після тимчасового ковалентного пришивання одного з кінців до активного центру ферменту й не потребує зовнішніх джерел вільної енергії.



**Рис. 3.27.** Еукаріотична топоізомераза I у комплексі з ДНК у двох проекціях (1K4S)

**Топоізомерази II**, на відміну від топоізомераз першого класу, змінюють число зчеплень шляхом зміни райзингу циркулярної ДНК. Є й дві інші відмінності: ці субодиничні білки мають два активні центри, в яких відбувається розрізання обох полінуклеотидних ланцюгів; фермент є активним тільки у присутності АТР, гідроліз якої використовується для здійснення конформаційних змін білка. Структурні домени двох субодиниць еукаріотичної топоізомераз II (рис. 3.28) або чотири субодиниці прокаріотичних топоізомераз цього типу утворюють своєрідні верхні та нижні “ворота”, які здатні розкриватися та закриватися при структурних перебудовах.

Схему роботи топоізомераз II зображено на рис. 3.29: 1) ділянка ДНК, яка називається G-сегментом, зв'язується з ферментом, що викликає спорідненість ферменту до АТР; 2) у відповідь на зв'язування АТР відбувається конформаційна зміна, яка супроводжується замиканням іншої зв'язаної ділянки ДНК – Т-сегмента, і дволанцюговим розривом у складі G-сегмента; 3) відбувається гідроліз АТР, що приводить до проштовхування Т-сегмента крізь розрив; 4) здійснюється зшивання розриву та звільнення обох сегментів. Шляхом описаної операції можна перетворити, наприклад, одну надспіралізовану молекулу на рис. 3.24 в іншу, звідки зрозуміло, що число зчеплень при роботі топоізомераз II змінюється на величину, кратну двом.



Рис. 3.28. Еукаріотична топоізомераза II (1BJT)

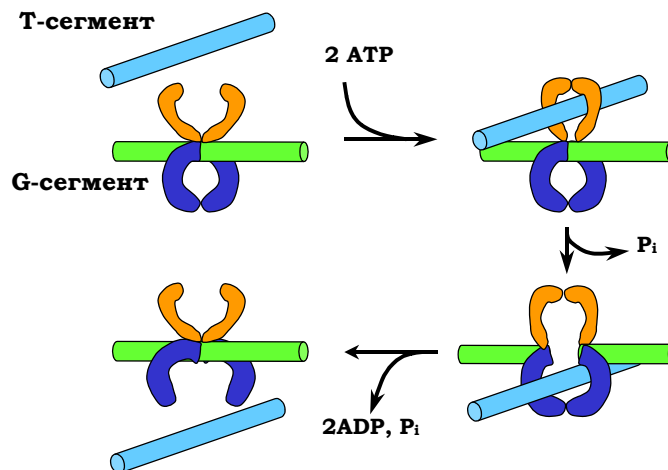


Рис. 3.29. Схема дії топоізомерази II

Напрямок зміни  $Lk$  залежить від взаємної орієнтації G- і T-сегментів. Еукаріотичні топоізомерази II дозволяють обидві орієнтації: у результаті  $Lk$  змінюється в напрямку релаксації ДНК. Але до класу топоізомераз II належить також бактеріальний фермент – *гіраза* (gyrase), який допускає тільки одну орієнтацію – її зображено на рис. 3.29. Результатом роботи гірази є зміна  $Lk$  тільки в одному напрямку – накопичення негативної надспіралізації. Це має важливі функціональні наслідки, оскільки сприяє дестабілізації подвійної спіралі ДНК на певних важливих ділянках.

### КОНТРОЛЬНІ ЗАПИТАННЯ

1. З яких трьох елементів складається нуклеотид? Назвіть основні фізико-хімічні властивості цих елементів.
2. Чим хімічно відрізняються між собою рибо- і дезоксирибонуклеїнові кислоти?
3. Між якими хімічними групами утворюється полінуклеотидний ланцюг за рахунок ковалентного зв'язку? Як позначаються кінці ланцюга й чому саме так?
4. Опишіть основні риси структури подвійної спіралі ДНК.
5. Які взаємодії стабілізують подвійну спіраль? Що лежить в основі комплементарності нуклеотидів і яке значення має комплементарність для стабілізації спіралі?
6. Які структурні форми подвійних спіралей нуклеїнових кислот ви знаєте? Чим вони відрізняються? Які з цих форм є основними формами існування подвійних спіралей ДНК і РНК *in vivo*?
7. Які конформаційні параметри характеризують вигин і ступінь спіральної закрутки молекули ДНК?
8. Напишіть усі типи динуклеотидних контактів у складі полінуклеотидного ланцюга та у складі подвійної спіралі. Як і чому різняться ці два набори контактів?
9. Чим визначається локальна конформація та динамічні властивості подвійної спіралі?
10. Назвіть основні структурні мотиви білків, котрі взаємодіють із ДНК. Які елементи структури білків залучаються до взаємодії, і з якими структурними елементами подвійної спіралі?
11. Взаємодії якого типу реалізуються між ДНК і білками?
12. До чого приводить взаємодія елементів регулярної вторинної структури білків з маленьким жолобком ДНК?
13. Сформулюйте головний принцип, за яким здійснюється специфічне впізнання білком певної послідовності пар основ ДНК.
14. Що таке число зчеплень, твіст і райзинг циркулярної ДНК? Як співвідносяться ці величини? Що таке топоізомер?
15. Як виникає надспіралізація в циркулярних ДНК? Яка різниця між негативною та позитивною надспіралізацією?
16. У чому полягають основні механізми роботи й функціональне значення ДНК-топоізомераз?

**РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

*Вологодский, А.В.* Топология и физические свойства кольцевых ДНК. – М. : Наука, 1988.

*Зенгер, В.* Принципы структурной организации нуклеиновых кислот. – М. : Мир, 1987.

*Уотсон, Дж. Д.* Двойная спираль. Воспоминания об открытии структуры ДНК. – М. : Регулярная и хаотическая динамика, 2001.

*Франк-Каменецкий, М.Д.* Век ДНК. – М. : КДУ, 2004.

*Anderson, C.F., Record, M.T.* Salt-nucleic acid interactions // *Annu. Rev. Phys. Chem.* – 1995. – Vol. 46. – P. 657–700.

*Calladine, C.R., Drew, H.R.* Principles of sequence-dependent flexure of DNA // *J. Mol. Biol.* – 1986. – Vol. 192. – P. 907–918.

*Champoux, J.J.* DNA topoisomerases: structure, function and mechanism // *Annu. Rev. Biochem.* – 2001. – Vol. 70. – P. 369–413.

*Dickerson, R.E.* DNA structure from A to Z. // *Methods Enzymol.* – 1992. – Vol. 211. – P. 67–111.

*Garvie, C.W., Wolberger, C.* Recognition of specific DNA sequences // *Mol. Cell.* – 2001. – Vol. 8. – P. 937–946.

*Olson, W.K., Gorin, A.A., Lu, X.-J. et al. V.B.* DNA sequence-dependent deformability deduced from protein-DNA crystal complexes // *Proc. Natl. Acad. Sci. USA.* – 1998. – Vol. 95. – P. 11163–11168.

*Olson, W.K., Zhurkin, V.B.* Modeling DNA deformations // *Curr. Op. Struct. Biol.* – 2000. – Vol. 10. – P. 286–297.

*Sarai, A., Kono, H.* Protein-DNA recognition patterns and predictions // *Annu. Rev. Biophys. Biomol. Struct.* – 2005. – Vol. 34. – P. 379–398.

*Vologodskii, A.V., Cozzarelli, N.R.* Conformational and thermodynamic properties of supercoiled DNA // *Annu. Rev. Biophys. Biomol. Struct.* – 1994. – Vol. 23. – P. 609–643.

*Yakovchuk, P., Protozanova, E., Frank-Kamenetskii, M.D.* Base-stacking and base-pairing contributions into thermal stability of the DNA double helix // *Nucl. Acids Res.* – 2006. – Vol. 34. – P. 564–574.

## Розділ 4

# ОРГАНІЗАЦІЯ ДНК У КЛІТИНАХ: ГЕНОМИ ТА СТРУКТУРА ХРОМАТИНУ

Il n'existe rien hors contexte.

*J. Derrida*

– Нет, – ответила Маргарита, –  
более всего меня поражает, где все  
это помещается.

*М. Булгаков. Мастер и Маргарита*

## Організація геномів

### *Генетичний код*

У послідовності нуклеотидів ДНК записана інформація про послідовність амінокислот у складі білків. Відповідність між цими двома типами текстів – між комбінаціями нуклеотидів і амінокислотами – називається *генетичним кодом*. Загальна кількість комбінацій по два нуклеотиди із чотирьох дорівнює 16, а по три – 64. Тобто мінімальна кількість нуклеотидів у одному слові нуклеотидного тексту, оскільки такими словами потрібно закодувати 20 амінокислот, має дорівнювати трьом. Саме це й спостерігається: одне слово – *кодон* – являє собою триплет нуклеотидів.

Послідовності всіх 64 кодонів (у напрямку від 5'- до 3'-кінця у складі мРНК) наведено на рис. 4.1. Серед 64 кодонів три є сигналами зупинки синтезу білка (стоп-кодони, або нонсенс-кодони), решта – 61 змістовний кодон – відповідають двадцяти амінокислотам. Співвідношення між кодонами та амінокислотами є однозначним: певний триплет кодує одну і тільки одну певну амінокислоту. Зворотне співвідношення не є однозначним: більшість амінокислот кодуються кількома триплетами – код є *виродженим*. Дві амінокислоти – Trp і Met –

невироджені й кодуються лише одним кодоном кожна. Усі інші 18 амінокислот кодуються серіями кодонів-синонімів – від двох до шести кодонів на серію.

		2й нуклеотид			
		U	C	A	G
1й нуклеотид	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } UCC } Ser UCA } UCG }	UAU } Tyr UAC } UAA } Stop UAG }	UGU } Cys UGC } UGA } Stop UGG } Trp
	C	CUU } CUC } Leu CUA } CUG }	CCU } CCC } Pro CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } CGC } Arg CGA } CGG }
	A	AUU } Ile AUC } AUA } AUG } Met	ACU } ACC } Thr ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }
	G	GUU } GUC } Val GUA } GUG }	GCU } GCC } Ala GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } GGC } Gly GGA } GGG }

Рис. 4.1. Таблиця генетичного коду

Якщо розглянути триплетні синонімичні серії, стає очевидним, що три позиції кодона нерівнозначні:

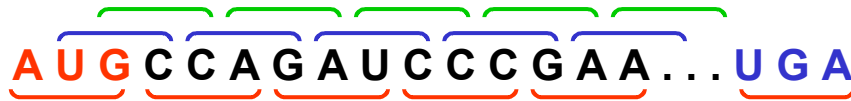
- Найважливішою для визначення змісту кодона є комбінація нуклеотидів у першій і другій позиції, при цьому друга позиція визначає зміст кодона найбільш жорстко – нуклеотидні заміни по другій позиції завжди приводять до зміни змісту (амінокислотної заміни).
- Найменш змістовною є третя позиція – нуклеотидні заміни в ній у 70 % випадків не приводять до зміни змісту кодона.

Залежно від того, який нуклеотид стоїть у другій позиції, усі кодони та відповідні амінокислоти можна розділити на чотири групи (чотири стовпчики в таблиці коду на рис. 4.1). Виявляється, що до

#### Розділ 4. Організація днк у клітинах: геноми та структура хроматину

групи U (урацил у другій позиції) належать гідрофобні амінокислоти великого розміру з повністю аліфатичними боковими залишками (див. також рис. 2.1). Тобто нуклеотидні заміни в першій та третій позиціях у межах цієї групи не викликають зміни властивостей амінокислоти, мінімізуючи вплив таких амінокислотних замін на просторову структуру білка (див. розділ 2). До групи A належать великі полярні амінокислоти (за виключенням Trp, але й він, хоча загалом гідрофобний, також здатен утворювати водневий зв'язок). Група C сформована маленькими амінокислотами. До групи G потрапляють амінокислоти з екстремальними властивостями: найбільша гідрофобна – Trp, найбільша полярна – Arg, найбільш реакційноздатна – Cys, найменша – Gly.

Генетичний код є досить універсальним – наведена таблиця коду реалізується як для бактерій, так і для ссавців. Однак у деяких випадках (окремі прокаріоти, гриби, водорості, а також автономна генетична система мітохондрій) спостерігаються невеличкі відхилення від цієї універсальної таблиці.



**Рис. 4.2.** Три можливі рамки зчитування, одна з яких (позначена червоними дужками) є відкритою

При синтезі білка (розділ 8) триплети зчитуються з нуклеотидного тексту один за одним: сусідні триплети не перекриваються, і між ними відсутні проміжки. Відповідно, будь яка послідовність нуклеотидів може бути прочитана трьома різними способами – містить три рамки зчитування (рис. 4.2). Рамка, що знаходиться між стартовим (частіше як стартовий використовується метіоніновий кодон AUG) і стоп-кодоном, називається *відкритою рамкою зчитування* (ORF, open reading frame).

#### *Гени*

Відкрита рамка зчитування (кодуюча послідовність), яка містить інформацію про амінокислотну послідовність білка, є найважливішою змістовною частиною *гена* (gene). Але для того, щоб відбулась *експресія* генетичної інформації (через синтез РНК і далі – білка), не менш важливими є регуляторні послідовності ДНК, які за рахунок спорідненості

до специфічних білків використовуються для вмикання / вимикання транскрипції (див. розділи 5, 6) як першої стадії експресії гена. Отже, визначення гена можна сформулювати так: **ген – це ділянка ДНК, яка є необхідною і достатньою для повноцінного синтезу функціональної молекули РНК**. Ділянка ДНК, яка може вважатися геном, має містити кодуєчу послідовність, де записана інформація про продукт, а також певний набір регуляторних елементів послідовності, від яких залежить запуск / блокування процесу транскрипції, шлях зчитування інформації тощо. Згідно з визначенням міжнародного консорціуму онтології послідовностей (Sequence Ontology Consortium), ген – це “певна визначена зона послідовності ДНК, яка відповідає одиниці спадковості та асоційована з регуляторними ділянками, ділянками, що транскрибуються, та / або іншими функціональними ділянками послідовності”.

Одним із найважливіших типів продуктів транскрипції генів є **mРНК** – матричні РНК (messenger RNA, mRNA), які використовуються далі як матриці для синтезу білків (піддаються трансляції, див. розділ 8). У цьому випадку білок є кінцевим продуктом гена, який, відповідно, називається білковим. Крім того, досить велика кількість генів кодує різноманітні молекули РНК, які не піддаються трансляції (є кінцевими продуктами): **pРНК** – рибосомні РНК (ribosome RNA, rRNA, розділ 8); **mРНК** – транспортні РНК (transfer RNA, tRNA, розділ 8); маленькі ядерні РНК (small nuclear RNA, **snRNA**, розділ 7); маленькі ядерцеві РНК (small nucleolar RNA, **snoRNA**, розділ 8); мікро-РНК (micro RNA, **miRNA**, розділ 6), молекули РНК, які є компонентами деяких ферментів; інші види РНК, не для всіх із яких з'ясовано їхні функції.

### *Геноми*

ДНК, що міститься у клітині, – це не тільки гени: принаймні мають бути з'єднуючі міжгенні ділянки. **Сукупність послідовностей ДНК у гаплоїдному наборі даного організму називається геномом**. На сьогодні повністю розшифровані послідовності більше 600 прокаріотичних і 80 еукаріотичних геномів. Головна відмінність між ними полягає в тому, що в прокаріотичних геномах кодуєчі послідовності становлять до 95 %, тоді як частка кодуєчих послідовностей у геномах еукаріотів не перевищує 3 %. Розміри деяких геномів і оцінка кількості білкових генів у їхньому складі наведені в табл. 4.1.



#### Розділ 4. Організація днк у клітинах: геноми та структура хроматину

Таблиця 4.1. Розміри геномів і кількість білкових генів деяких організмів

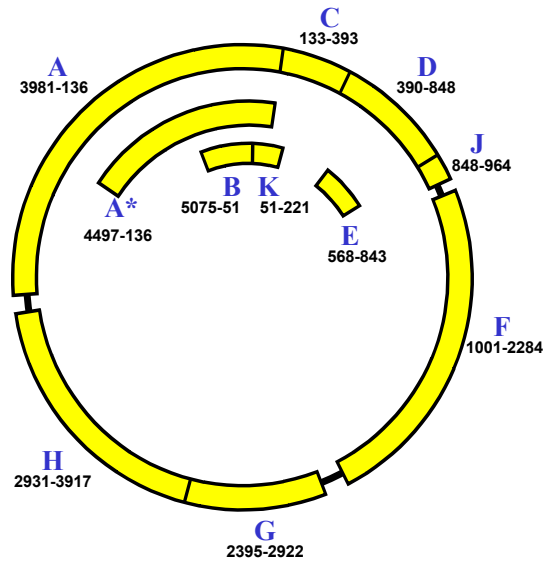
Організм	Розмір геному (пари основ)	Кількість хромосом (молекул ДНК)*	Кількість білкових генів
Бактеріофаг фХ-174	5 386	1	10
Бактерія <i>Escherichia coli</i>	$4,6 \cdot 10^6$	1	4 100
Аскоміцет <i>Saccharomyces cerevisiae</i>	$1,2 \cdot 10^7$	16	6 700
Нематода <i>Caenorhabditis elegans</i>	$10^8$	6	20 000
Плодова мушка <i>Drosophila melanogaster</i>	$1,3 \cdot 10^8$	4	14 000
Курка <i>Gallus gallus</i>	$10^9$	33	13 000
Миша <i>Mus musculus</i>	$3,3 \cdot 10^9$	20	22 000
Людина <i>Homo sapiens</i>	$3,2 \cdot 10^9$	23	21 000
Мітохондріальна ДНК людини	16 569	1	37

\* Для еукаріотів наведено кількість хромосом у гаплоїдному наборі

**Прокаріотичні й вірусні геноми.** Схему організації одного з найпростіших (і першого з вивчених) геномів – геному бактеріофага фХ-174 (слід читати “фі-десять”) – показано на рис. 4.3. Геном побудований надзвичайно “економно”: десять генів (один із них дає два різні транскрипти) займають практично всю циркулярну ДНК бактеріофага. Більше того, декілька генів перекриваються за рахунок використання різних рамок зчитування: гени А і С та С і D перекриваються своїми кінцями, гени В, К і Е повністю знаходяться у межах інших генів; три гени – А, С і К – використовують усі три можливі рамки зчитування на одній ділянці ДНК (звичайно, у даному випадку всі три рамки є відкритими). Явище перекриття генів за рахунок використання різних рамок зчитування спостерігається також для кількох інших бактеріофагів, проте іноді зустрічається в еукаріотів. Загалом геноми вірусів і бактеріофагів побудовані за подібним економним принципом: порівняно невелика кількість генів, мінімальна кількість міжгенної ДНК.

У геномі прокаріотичної клітини кількість ДНК і генів значно зростає, але зберігається принцип економності щодо використання більшості послідовностей для кодування генетичної інформації. Наприклад, геном *Escherichia coli* представлений однією циркулярною молекулою ДНК (бактеріальною хромосоною) довжиною 4,6 млн пар основ. Близько 90 % цієї ДНК припадає на кодуєчі послідовності ~4100 білкових генів і ~120 генів РНК, що не транскрибується. Середній розмір гена

складається з 950 пар основ, середня довжина міжгенної ділянки – 118 пар основ. Проте, міжгенні зони мають досить нерівномірний розподіл за довжиною, яка варіює від 0 до 1 тис. 730 пар основ.



**Рис. 4.3.** Геном бактеріофага фХ-174. Позначено початок і кінець кожного гена, загальна довжина ДНК – 5386 пар основ

Приблизно  $\frac{3}{4}$  транскрипційних одиниць *E. coli* містять один ген, решта реалізує характерний для бактерій (і тільки для них) оперонний принцип організації генетичного матеріалу. *Оперон* являє собою кластер так званих *структурних генів*, на яких синтезується одна молекула мРНК, що має кілька (на кожен структурний ген) послідовних (які не перекриваються) відкритих рамок зчитування для трансляції відповідних білків (див. приклади організації та регуляції оперонів у розділі 5). У межах оперона згруповані структурні гени, які відповідають за синтез білків, залучених до одного ланцюжка біохімічних перетворень (ферменти синтезу або деградації певної сполуки). Крім структурних генів оперон має регуляторні ділянки, за рахунок яких здійснюється регуляція транскрипції оперона як цілого. У геномі *E. coli* міститься ~650 таких одиниць транскрипції.

Характерною особливістю бактерій є наявність у клітині, поряд із бактеріальною хромосою, невеликих автономних елементів геному – *плазмід*. Плазміда є циркулярною молекулою ДНК (типовий розмір

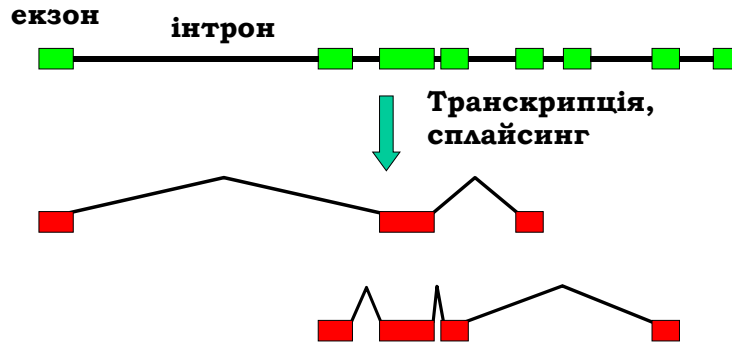
#### **Розділ 4. Організація днк у клітинах: геноми та структура хроматину**

~3 тис. пар основ), яка містить кілька генів і реплікується незалежно від бактеріальної хромосоми. Плазміди широко використовуються як зручний інструмент молекулярно-біологічних досліджень (див. розділ 11).

Циркулярна ДНК бактеріальної хромосоми існує у клітині у вигляді комплексу з білками. Ця взаємодія є досить динамічною, і практично весь бактеріальний геном (що принципово відрізняє його від еукаріотичного геному, див. нижче) перебуває в потенційно транскрипційно-активному стані: гени й оперони є об'єктами швидкої оперативної регуляції у відповідь на зміну зовнішніх умов.

**Еукаріотичні геноми** містять значно більшу кількість ДНК порівняно з геномами прокариотів (див. табл. 4.1), причому переважна частина цієї ДНК представлена послідовностями, що не є кодуєчими. У тому числі приблизно половина еукаріотичного геному – це послідовності, представлені багатьма копіями (послідовності, що повторюються). Еукаріотична ДНК існує у клітинному ядрі у складі хромосом, кожна хромосома містить одну гігантську лінійну молекулу ДНК. Послідовності, що повторюються, зосереджені на кінцях хромосом (теломери) і в зонах прикріплення хромосом до веретена поділу при мітозі (центромери).

Характерною ознакою генів еукаріотів (на відміну від прокариотів) є мозаїчний принцип будови кодуєчої частини: власне кодуєча частина – це послідовність окремих змістовних ділянок – *екзонів* (exon), розділених беззмістовними *інтронами* (intron). Часто екзони відповідають окремим структурним доменам мультидоменних білків: еволюційне збирання білка з кубиків-доменів може здійснюватись шляхом перетасування екзонів на рівні ДНК. Беззмістовними інтронами є в тому сенсі, що не містять інформації про кінцевий продукт, але в межах інтронів часто розташовані важливі регуляторні ділянки. При транскрипції молекула РНК синтезується суцільно (первинний продукт транскрипції – первинний транскрипт – містить екзони та інтрони). Отже, необхідним етапом експресії гена є процес *сплайсингу* (розділ 7) – вирізання інтронів і зшивання екзонів у кінцевий транскрипт, який може бути використаний як матриця для білкового синтезу. При цьому сплайсинг може бути спрямований різними шляхами (рис. 4.4) – *альтернативний сплайсинг*, – унаслідок чого утворюються різні кінцеві продукти – різні білки. Таким чином, альтернативний сплайсинг можна інтерпретувати як механізм “перекриття” еукаріотичних генів. Інший випадок перекриття генів у геномах еукаріотів пов'язаний з можливістю розташування гена (зі своїми інтронами та екзонами) у межах інтрона іншого гена. В окремих випадках спостерігається також перекриття відкритих рамок зчитування.



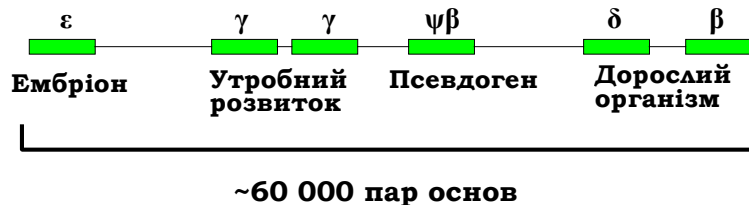
**Рис. 4.4.** Мозаїчна будова кодуєчої частини гена та схема утворення різних кінцевих продуктів унаслідок альтернативного сплайсингу

Усі гени багатоклітинного організму можна розділити на дві групи: 1) гени, від яких залежать певні універсальні функції та які активні в усіх клітинах, – “гени домашнього господарства” (housekeeping genes); 2) гени, що специфічно активуються у клітинах певного типу, – “гени розкоші” (luxury genes). Загальною ознакою генів першої групи є розташування в їхніх регуляторних зонах *СрG-острівців* (СрG-islands) – ділянок із підвищеним вмістом динуклеотидів СрG (контактів типу СG між парами основ, див. розділ 3). Вміст цих динуклеотидів в еукаріотичних геномах приблизно у п'ять разів менший за очікуваний унаслідок метилювання цитозину в складі СрG-контакту: 5mС (5-метилцитозин) спонтанно перетворюється на тимін. Метилювання цитозину в регуляторних ділянках є одним із механізмів репресії генів (див. розділ 6). Відповідно, гени, що зберігають свою активність у більшості клітин, містять неметилювані динуклеотиди СрG, вміст яких зберігається на високому рівні (для ДНК випадкової послідовності із 40 % GC-пар очікуваний вміст динуклеотидів СрG становить 4 %).

Типове значення вмісту GC-пар в еукаріотичних геномах – близько 40 % (у геномі людини – 41 %). Проте спостерігаються варіації локального GC-вмісту в широких межах – від 30 до 60 %; зустрічаються й досить довгі ділянки (до кількох мільйонів пар основ) з аномальним вмістом GC-пар. Загальна тенденція при цьому така: високий вміст GC-пар корелює з високою щільністю генів, зменшеною довжиною інтронів і низькою щільністю послідовностей, що повторюються.

#### **Розділ 4. Організація днк у клітинах: геноми та структура хроматину**

Серед еукаріотичних генів 25–50 % є унікальними (представлені в геномі єдиною копією), решта належать до родин генів, що складаються з кількох копій, як правило, не ідентичних. Відповідні (гомологічні але не ідентичні) білки складають родину білків. Кілька родин (протеїнкінази, транскрипційні фактори певного типу, імуноглобуліни) містять сотні білків, більшість родин складається з кількох (до 30) білків. Гени такої родини часто об'єднані в геномі в кластери – розташовані поряд у певній хромосомі (кластери генів теплового шоку, глобінові гени). Слід зауважити, що такий кластер не є опероном – кожен ген піддається регуляції як окрема одиниця транскрипції. Наприклад, кластер генів  $\beta$ -субодиниці гемоглобіну містить гомологічні гени, які активуються на певних стадіях індивідуального розвитку (рис. 4.5).

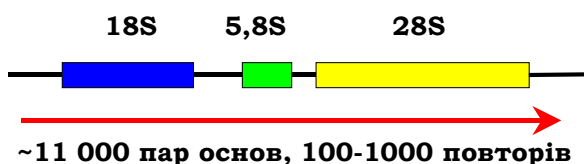


**Рис. 4.5.** Кластер  $\beta$ -глобіну в 16-й хромосомі людини (кожен ген містить інтрони). Указано стадії розвитку, на яких відповідні гени є активними

$\beta$ -Глобіновий кластер містить також неактивний *псевдоген*. Псевдогени – це послідовності, гомологічні певним генам, але не експресуються. До такої дезактивації гена може призвести мутація, яка порушує ініціацію транскрипції, сплайсинг, викликає невчасну термінацію тощо. Після дезактивації псевдоген перестає бути об'єктом відбору, і в ньому накопичується велика кількість мутацій. Зрозуміло, що в першу чергу псевдогени виникають саме в кластерах – коли є кілька копій гена, і пошкодження одного з них не приводить до фатальних наслідків.

Типовою ознакою еукаріотичних геномів є наявність кількох типів генних кластерів, що повторюються багато разів. Серед білкових генів це стосується генів гістонів – структурних білків хроматину (детальніше йтиметься у наступному підрозділі). Гени п'яти молекул гістонів завжди згруповані в кластер (кожен ген – окрема одиниця транскрипції), який повторюється до 100 разів. Іншим прикладом кластерів, що повторюються, є гени рибосомної РНК (рис. 4.6, тип рибосомних РНК позначають за їхньою константою седиментації, див. розділ 8).

У цьому випадку кластер є одиницею транскрипції, первинний транскрипт піддається далі “дозріванню” – частковій деградації з утворенням окремих молекул рРНК. Такий тандемний повтор рибосомного кластера є ядерцевим організатором – зоною, де утворюється ядерце (місце збирання рибосом). Гени рРНК ще одного типу – 5S – також тандемно повторюються в іншому місці геному. Також зібрані в тандемні кластери частина генів тРНК і маленьких ядерних РНК.

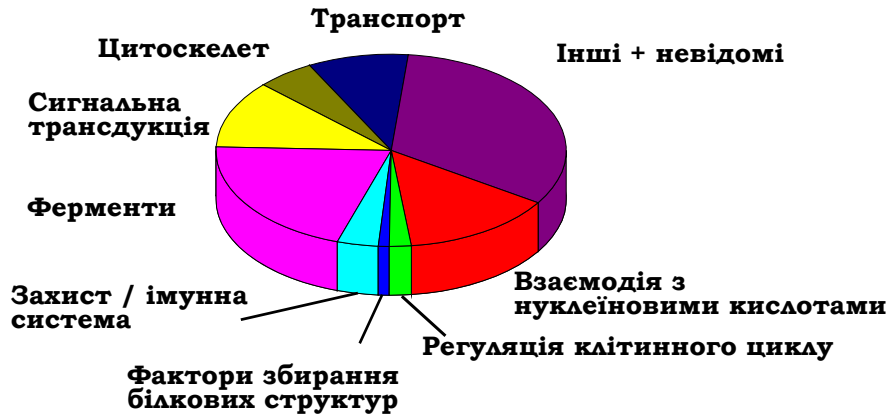


**Рис. 4.6.** Кластер генів рибосомних РНК, який повторюється в зоні ядерцевого організатора. Червона стрілка позначає первинний транскрипт

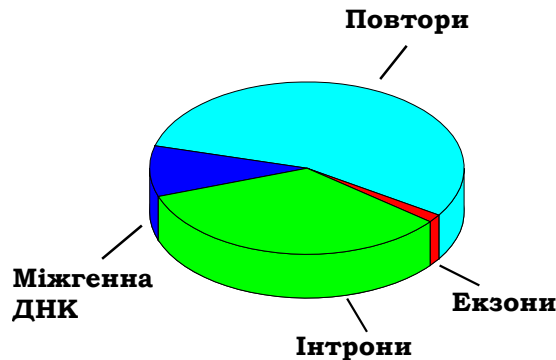
Загальна кількість генів у геномах вищих еукаріотів варіює приблизно від 20 до 30 тис. (табл. 4.1). Так, у геномі людини кількість білкових генів становить 21 тис., ще кілька тисяч генів кодують РНК, яка не піддається трансляції. Наведемо як приклад кілька загальних характеристик білкових генів людини:

- Середній розмір гена (ураховуючи інтрони) – 100 тис. пар основ (максимум – 2 млн 400·тис.).
- Середній розмір кодуючої частини – 1 тис. 400 пар основ (максимум – 80 тис.).
- Середній розмір екзона – 150 пар основ.
- Середня кількість екзонів – 9 (максимум – 178).
- Середній розмір інтрона – 5 тис. пар основ (максимум – 820 тис.).

Загальною тенденцією, що спостерігається в геномі людини, є наявність порівняно коротких екзонів, розділених довгими інтронами. Приблизний розподіл білків людини (сукупність усіх білків організму називається *протеомом*) за їхніми функціями зображено на рис. 4.7. Як видно з рис. 4.8, кодуючі послідовності цих генів займають лише ~1,5 % геному. Решта припадають на міжгенну ДНК (де містяться також регуляторні ділянки), інтрони (~34 %) і, більше половини геному, – на послідовності, що повторюються.



**Рис. 4.7.** Розподіл білків протеому людини за функціями



**Рис. 4.8.** Відносний вміст послідовностей різних типів у геномі людини

Основні типи повторів, присутні в геномі вищих еукаріотів:

1. *Псевдогени* – від 1 до кількох тисяч у геномі людини.
2. *Тандемні повтори*. До цього класу повторів можна віднести тандемні повтори генних кластерів, які вже згадувалися вище. Крім того, до тандемних повторів відносять багатократні повтори коротких послідовностей по шість-вісім пар основ у теломерах (TTAGGG у людини) і повтори  $\alpha$ -сателітної ДНК у центромерах (довжина повтору варіює від 7 пар основ у дрозофіли до 200 пар основ у ссавців, у людини – 171 пара основ). По всьому геному розподілені також так

звані прості повтори (SSR, simple sequence repeats, ~3 % у геномі людини). Зазвичай виділяють *мікросателіти* – 1–15 пар основ, що повторюються від 10 до кількох тисяч разів, і *мінісателіти* – 15–500 пар основ, які повторюються до 100 разів. У геномі людини є принаймні 30 тис. міні- та 200 тис. мікросателітних локусів.

3. *Сегментні дуплікації* – великі блоки довжиною 1–200 тис. пар основ, які характеризуються високим ступенем гомології (близько 5 % у геномі людини). Імовірно, сегментні дуплікації є продуктом порушення хромосом. Частіше зустрічаються в перицентромерних і субтеломерних зонах.

4. *Інтерсперсні (мобільні) елементи*, здатні до переміщення та розмноження в межах геному (~44 % у геномі людини). Значна частина таких послідовностей є результатом колишньої активності мобільних елементів (таких, що втратили здатність до переміщення), але деякі зберігають свою активність досі. Основні типи мобільних елементів у геномі людини (механізми їхнього переміщення детальніше розглядатимуться в розділі 10):

- ДНК-транспозони (3 %) – переміщення здійснюється шляхом вирізання ділянки ДНК із наступним вбудовуванням її в інше місце.
- LTR-ретропозони (8 %) – як і для інших двох типів мобільних елементів, переміщення відбувається через проміжну молекулу РНК: на елементі послідовності (який містить довгі кінцеві повтори – Long Terminal Repeats – і кілька генів, зокрема ген зворотної транскриптази) здійснюється транскрипція; молекула РНК прямує до цитоплазми, де відбувається *зворотна транскрипція* – синтез ДНК на РНК-матриці за допомогою зворотної транскриптази (РНК-залежна ДНК-полімераза). ДНК повертається до ядра, де вбудовується в геном.
- Мобільні елементи LINE (Long Interspersed Elements, 20 %) містять кілька генів, включаючи ген зворотної транскриптази. На відміну від LTR-ретропозонів, зворотна транскрипція здійснюється в ядрі. У геномі людини присутні три родини таких елементів, одна з них є активною і слугує основним джерелом зворотної транскриптази у клітині.
- Мобільні елементи SINE (Short Interspersed Elements, 13 %) – короткі (100–400 пар основ) беззмістовні елементи, які використовують для переміщення систему LINE. До цього класу належить і Alu-повтор (від назви відповідної рестриктази, яка здатна специфічно гідролізувати цей елемент послідовності).



#### **Розділ 4. Організація днк у клітинах: геноми та структура хроматину**

Мобільні елементи розподілені в геномі нерівномірно: є довгі ділянки, що на 90 % складаються з цих елементів, і такі, де інтерсперсні елементи відсутні. Загалом спостерігається негативна кореляція між щільністю генів і мобільних елементів. Винятком із цієї закономірності є позитивна кореляція між щільністю генів і елементів типу SINE.

**Геном мітохондрій і хлоропластів** – автономний елемент еукаріотичного геному. Циркулярна (як правило) молекула мітохондріальної ДНК містить від 6 тис. до 2 млн. пар основ і певний набір генів (рРНК, тРНК, деяких білків). Мітохондріальні гени містять інтрони (але не у ссавців). Розмір мітохондріального геному знижується в ході еволюції. Циркулярна ДНК хлоропластів більша за розміром – містить до 200 тис. пар основ, хлоропластний геном кодує до ~100 білків.

#### **Молекулярна організація хроматину**

Загальна довжина ДНК у ядрі еукаріотичної клітини становить близько 2 м. Така кількість ДНК вимагає її щільної упаковки, яка зумовлює тотальне пригнічення функціональних активностей у більшій частині геному. Але при цьому упаковка ДНК у клітинному ядрі має дозволяти вибіркову активацію певних ділянок у певні моменти часу. Ці альтернативні завдання вирішуються завдяки тому, що ДНК існує в клітинному ядрі у вигляді складного нуклеопротеїнового комплексу – *хроматину*.

На першому рівні організації хроматину ДНК формує за рахунок взаємодії з білками елементарні утворення – *нуклеосоми* – із середньою щільністю одна нуклеосома на 200 пар основ. Білковий компонент нуклеосоми (кор) складається з восьми молекул *корових гістонів* H2A, H2B, H3 і H4 – по дві молекули кожного типу. На другому рівні компактизації за участю *лінкерних гістонів H1* утворюється фібрила товщиною 30 нм. Хроматинова фібрила формує петлі розміром 20–200 тис. пар основ, кінці яких є жорстко закріпленими на скелетних структурах *ядерного матриксу*.

#### **Нуклеосома**

Білковий компонент нуклеосоми – гістони, які є одним з найбільш еволюційно консервативних класів білків. Усі корові гістони (містять від 102 до 135 амінокислотних залишків) мають спільну схему будови. У первинній структурі виділяють дві частини: глобу-

лярну та N-кінцеву невпорядковану (хвіст) довжиною від 20 (H2A) до 40 (H3) амінокислотних залишків. Гістон H2A має також помітний C-кінцевий хвіст довжиною близько 15 залишків. Невпорядковані хвости практично не містять гідрофобних залишків, збагачені позитивно зарядженими амінокислотами і є субстратами для численних посттрансляційних модифікацій (про що йтиметься нижче).

Глобулярна частина всіх корових гістонів, у свою чергу, також має спільну структуру. Вона виглядає як характерний триспіральний гістоновий мотив (histone fold), у якому одна довга  $\alpha$ -спіраль фланкована двома короткими (див. рис. 3.16, а). Гістон H3 містить також додаткову  $\alpha$ -спіраль з боку N-кінця мотиву ( $\alpha$ N-спіраль), а гістон H2B – додаткову  $\alpha$ C-спіраль. Гістоновий мотив не формує гідрофобного ядра, заекранованого від розчинника, – значна кількість гідрофобних залишків опиняється на поверхні. Унаслідок цього одна молекула корового гістону не може існувати як окремий глобулярний білок у водному середовищі. Мінімальними стабільними структурними одиницями є гетеродимери H2A-H2B та H3-H4 (мають подібну структуру, представлена на рис. 3.16, а). Два гістонові мотиви формують у складі димерів щільне гідрофобне ядро, а специфічність формування димерів залежить від наявності додаткових  $\alpha$ N- та  $\alpha$ C-спіралей у гістонах H3 та H2B відповідно.

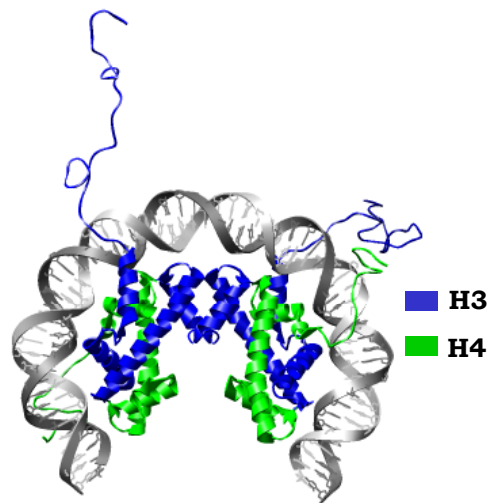
На поверхні димеру розташовані три зони скупчення позитивно заряджених амінокислотних залишків, які, відповідно до електростатичного механізму, здатні взаємодіяти з ДНК. У межах кожного такого сайту є принаймні один залишок аргініну, який відіграє найважливішу роль у взаємодії з ДНК. Разом ці три сайти створюють платформу для зв'язування ділянки ДНК довжиною 27–28 пар основ (~ 2,5 витка подвійної спіралі, рис. 3.16, а).

Два гетеродимери H3-H4 взаємодіють між собою за рахунок утворення чотириспірального пучка між гістоновими мотивами двох молекул H3. У результаті формується тетрамер  $(H3-H4)_2$  – центральний комплекс у структурі нуклеосоми (рис. 4.9). Структура тетрамеру нагадує підкову, яка характеризується хіральною – утворює елемент лівої спіралі. Вісь симетрії тетрамерного комплексу (яка водночас є і віссю симетрії всієї нуклеосоми) проходить через інтерфейс між двома молекулами H3.

Аналогічний за своєю структурою чотириспіральний пучок між гістоновими мотивами молекули H4 та H2B забезпечує взаємодію між тетрамером  $(H3-H4)_2$  і димером H2A-H2B. Таким чином, “підкова” тетрамеру симетрично продовжується двома димерами H2A-H2B в обидва боки,

#### **Розділ 4. Організація днк у клітинах: геноми та структура хроматину**

результатом чого є утворення октамеру (H2A-H2B-H3-H4)<sub>2</sub>. Отже, глобулярні частини гістонів утворюють октамерний комплекс, який і служить білковим кором нуклеосоми (рис. 4.10). На глобулярній поверхні октамеру існує своєрідний трек позитивно заряджених амінокислотних залишків, який використовується для взаємодії з нуклеосомною ДНК довжиною 145 пар основ.

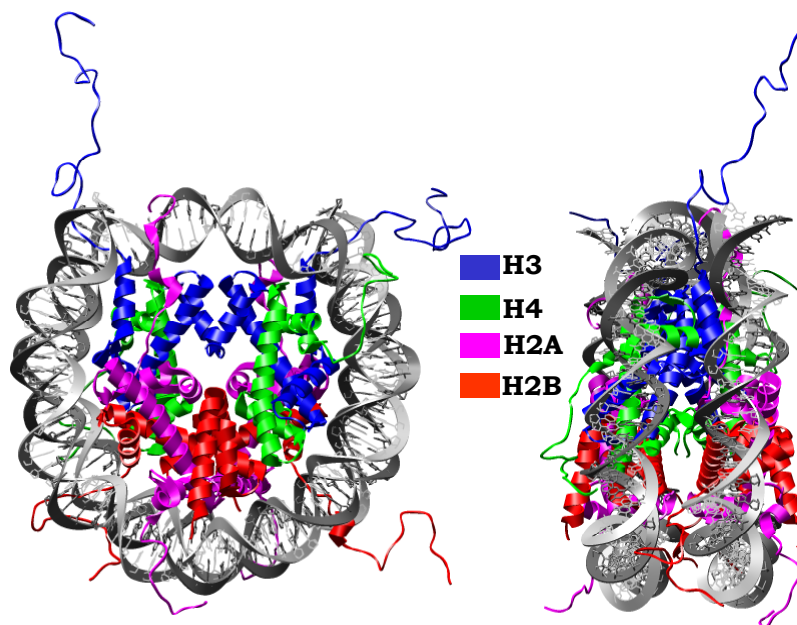


**Рис. 4.9.** Тетрамер гістонів (H3-H4)<sub>2</sub> у комплексі з ДНК у складі нуклеосоми (1KX5)

Окрема нуклеосома (*нуклеосомна кор-частинка*) – октамер гістонів + 145 пар основ ДНК – може бути вилучена з хроматину за допомогою мікрококової нуклеази: вона робить дволанцюговий розріз у ДНК, а оскільки один із ланцюгів нуклеосомної ДНК завжди взаємодіє з гістонами (рис. 4.10), доступною для нуклеази є лише ДНК за межами нуклеосоми. У хроматині вся ДНК формує нуклеосоми із середньою щільністю одна нуклеосома на 200 пар основ, сусідні нуклеосоми з'єднані міжнуклеосомними *лінкерними* (linker) ділянками. Нуклеосомна ДНК разом з лінкерною ділянкою складають так званий *нуклеосомний повтор*, довжина якого (середнє значення 200 пар основ) варіює як уздовж полінуклеосомного ланцюга, так і залежно від функціонального стану, типу клітин тощо.

Вісь симетрії нуклеосоми проходить через центральну точку нуклеосомної ДНК, у якій великий жолобок подвійної спіралі є зверненим до поверхні октамеру гістонів (рис. 4.10). ДНК-гістони взаємодіють

здійснюються в позиціях, де маленький жолобок контактує з поверхнею октамеру: саме тут реалізуються взаємодії ДНК із позитивно зарядженими сайтами на поверхні гістонових мотивів (порівн. рис. 3.16, а; 4.9, 4.10). Як видно з рис. 4.10, нуклеосомна ДНК є значно вигнутою на поверхні октамеру гістонів і формує майже ідеальну ліву суперспіраль із радіусом 41,9 Å і кроком 25,9 Å. Центральні 133 пари основ такої суперспіралі дають 1,67 витка, решта нуклеосомної ДНК (дві ділянки по сім пар основ на вході/виході) є практично прямою і просто продовжує хід нуклеосомної суперспіралі. Зрозуміло, що значний вигин нуклеосомної ДНК потребує енергетичних витрат, які компенсуються ДНК-гістоновими взаємодіями.



**Рис. 4.10.** Структура нуклеосоми у двох проекціях (1KX5)

Структура нуклеосомної ДНК зумовлена її взаємодією із глобулярною частиною октамеру гістонів. N-кінцеві хвости гістонів H3 та H2B виходять за межі нуклеосоми через канали, сформовані маленькими жолобками двох дуплексів сусідніх витків нуклеосомної суперспіралі (рис. 4.10). Ділянки хвостів, безпосередньо розташованих у каналах,

#### Розділ 4. Організація днк у клітинах: геноми та структура хроматину

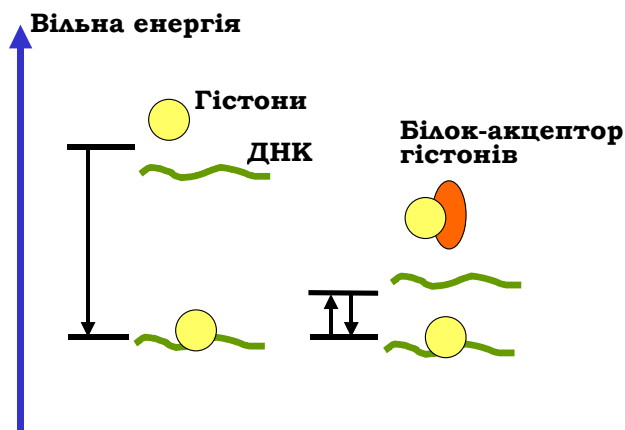
є позитивно зарядженими і, таким чином, додатково скріплюють сусідні витки суперспіралі. N-кінцеві хвости гістонів H4 і H2A також взаємодіють з маленьким жолобком зовні нуклеосомної суперспіралі. Особливе місце – входу / виходу нуклеосомної ДНК (рис. 4.10) – займає найдовший N-кінцевий хвіст гістону H3. Позитивно заряджений хвіст H3 стабілізує структуру нуклеосоми в цій зоні, де спостерігається найвища щільність негативних зарядів ДНК.

Значна частина гістонових хвостів просто виходить за межі нуклеосоми. Завдяки своїй структурній лабільності вони беруть участь в організації хроматину на наднуклеосомному рівні, а також відіграють важливу роль платформи для зв'язування різноманітних білків, що залежить від посттрансляційних модифікацій хвостів (див. нижче).

Аналіз структури нуклеосоми та молекулярних механізмів її стабілізації дозволяє сформулювати такі положення:

- Електростатичні взаємодії між ДНК і гістонами за фізіологічних умов є дуже міцними – остаточне руйнування нуклеосоми *in vitro* відбувається при концентрації солі 2 моль/л. Тобто за фізіологічних умов зв'язування гістонів з ДНК є необоротним (неможлива рівновага між зв'язаними та дисоційованими гістонами). Хоча структура нуклеосоми відповідає мінімуму вільної енергії, цього мінімуму неможливо досягти за розумний проміжок часу: гістони швидко й безладно зв'язуються з ДНК без подальшої дисоціації. Для реалізації рівноважних умов ДНК-гістонової взаємодії *in vivo* у хроматині існують *проміжні акцептори* гістонів, що виконують роль факторів збирання / руйнування нуклеосом. Ці акцептори мають спорідненість до гістонів дещо меншу, ніж спорідненість гістонів до ДНК (рис. 4.11), що й забезпечує можливість рівноважного обміну гістонами між ДНК і проміжними акцепторами із зсувом цієї рівноваги в бік ДНК-гістонових комплексів.
- Незважаючи на міцність ДНК-гістонових взаємодій, тимчасове порушення їх на кінцях нуклеосомної ДНК є можливим за фізіологічних умов: з одного боку, взаємодії на кінцевих ділянках є більш слабкими, ніж у центральній частині нуклеосомної ДНК; з іншого – висока густина негативного заряду (контакт між сусідніми витками нуклеосомної суперспіралі) дестабілізує нуклеосому ДНК на виході з нуклеосоми. Це приводить до тимчасового часткового розкручування нуклеосомної ДНК на кінцях, що є одним із головних шляхів структурної динаміки нуклеосом.

- Наявність проміжних акцепторів гістонів робить можливим обмін димерами H2A-H2B між різними нуклеосомами: тимчасове видалення димерів є іншим важливим шляхом структурної динаміки хроматину.



**Рис. 4.11.** Роль проміжних акцепторів гістонів у збиранні нуклеосоми: дисоціація гістонів є практично забороненою за фізіологічної іонної сили; у присутності акцептора взаємодія гістонів з ДНК стає рівноважною

- Незважаючи на те, що ДНК-гістонові взаємодії в нуклеосомі є неспецифічними щодо послідовності пар основ ДНК (електростатичні взаємодії між фосфатами ДНК і позитивними амінокислотними залишками), у хроматині спостерігається феномен *переважного позиціювання* нуклеосом відносно послідовності. Основною причиною позиціювання нуклеосом є залежність від послідовності здатності ДНК до деформацій, які супроводжують формування нуклеосоми (див. розділ 3). У результаті нуклеосома “обирає” таку ділянку, де вигин ДНК потребує менших енергетичних витрат. Це приводить до важливих функціональних наслідків – диференційного експонування ділянок ДНК до дії регуляторних факторів (розділ 6).
- При функціонуванні хроматину виникає необхідність у змінах згаданого вище експонування / екранування певних ділянок ДНК – репозиціюванні нуклеосом. Міцність ДНК-гістонових

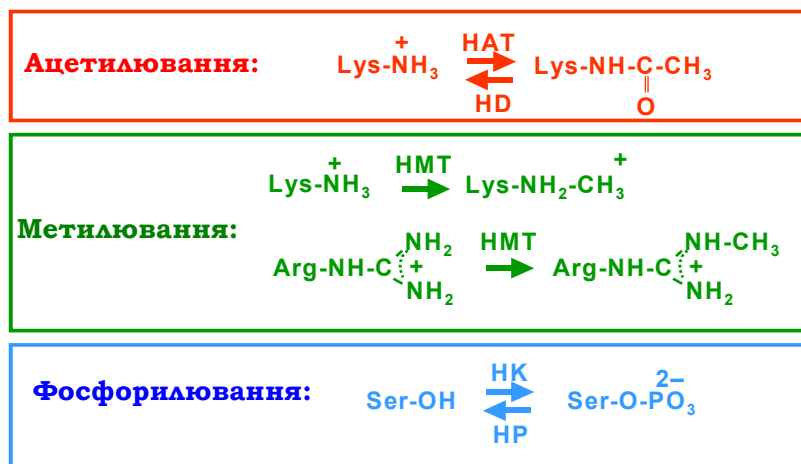
#### Розділ 4. Організація днк у клітинах: геноми та структура хроматину

взаємодій не дозволяє спонтанного зсуву нуклеосом уздовж ДНК, що зумовлює потребу в особливих АТР-залежних молекулярних пристроях, які здійснюють таке репозиціонування, – факторах ремоделювання хроматину (розділ 6), що також можуть виступати факторами збирання / руйнування нуклеосом.

#### Посттрансляційні модифікації гістонових хвостів

Невпорядковані хвости корових гістонів є субстратом для ковалентних посттрансляційних модифікацій, яким піддаються конкретні (такі, що мають конкретну позицію у складі поліпептидного ланцюга) амінокислотні залишки певного типу.

Основні типи цих модифікацій наведено на рис. 4.12:



**Рис. 4.12.** Посттрансляційні модифікації амінокислотних залишків у гістонах (показано лише хімічні групи у складі залишків, які безпосередньо модифікуються). HAT – гістон-ацетилтрансфераза, HD – гістон-деацетилаза, HMT – гістон-метилтрансфераза, HK – гістон-кіназа, HP – гістон-фосфатаза

- **Ацетилювання** залишків Lys – перенесення залишку оцтової кислоти (ацетату) на аміногрупу Lys. Реакція каталізується *гістон-ацетилтрансферазами* (HAT, Histone AcetylTransferase), різні їхні типи мають певну специфічність щодо залишків-мішеней. Результатом реакції є зникнення позитивного заряду на залишку Lys (рис. 4.12). Ацетилювання є динамічною мо-

дифікацією: гістон-деацетилази (HD, Histone Deacetylase, позначаються також як HDAC) здійснюють відщеплення ацетатних залишків. Два типи ферментів-антагоністів підтримують певний динамічний гомеостаз ацетилювання / деацетилювання, зсунутий у той чи інший бік у певних ділянках хроматину.

- **Метильовання** залишків Lys та Arg – перенесення метильної групи (однієї, двох або трьох) на аміногрупу Lys або на гуанідинову групу Arg, позитивний заряд при цьому залишається. Реакція каталізується гістон-метилтрансферазами (HMT, Histone MethylTransferase), кожна з яких є специфічною щодо залишку-мішені. На відміну від ацетилювання, метильовання є дуже стабільною у часі модифікацією. Метильовані гістони замінюються на неметильовані дуже повільно – імовірно, лише шляхом заміни на гістони, синтезовані *de novo*.
- **Фосфорильовання** залишків Ser – перенесення фосфатного залишку на ОН-групу Ser. Реакція каталізується гістон-кіназою (HK, Histone Kinase), відщеплення фосфату – фосфатазою.

Конкретні залишки, що модифікуються, показано на рис. 4.13.

Серед усіх модифікацій, лише про ацетилювання можна сказати, що воно чітко корелює з транскрипційною активністю: гіперацетилювані гістони присутні в активних ділянках хроматину, у репресованих підтримується деацетилюваний статус. Інші модифікації впливають на функціональний стан складніше: метильовання одного залишку може викликати репресію транскрипції, іншого (через кілька амінокислот від першого в тому самому хвості) – супроводжувати активацію; фосфорильовання H3-Ser10 корелює з активацією транскрипції, але також супроводжує гіперконденсацію хроматину при переході до мітозу.

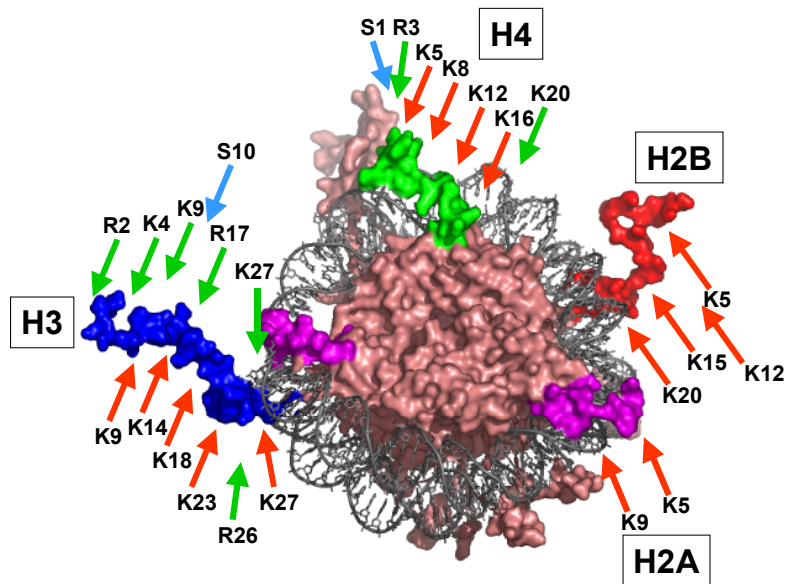
Оскільки ацетилювання численних залишків Lys (рис. 4.13) приводить до суттєвого пониження позитивного заряду хвостів, лише ацетилювання може безпосередньо впливати на взаємодію хвостів із ДНК. Це не викликає змін структури нуклеосоми (яка практично не залежить від хвостів), але змінює характер структурної динаміки як нуклеосоми, так і хроматинової фібрили на наднуклеосомному рівні її організації.

Однак головним механізмом впливу модифікованих гістонових хвостів на функціональний стан хроматину є **специфічне впізнання модифікованих хвостів іншими білками**: регуляторами, факторами транскрипції, ферментами тощо. Так, ацетилювані залишки Lys упізнаються особливим структурним блоком таких білків – *бромодоменом*.



#### Розділ 4. Організація днк у клітинах: геноми та структура хроматину

Інший блок – *хромодомен* – упізнає метильовані Lys. Гістонові хвости слугують своєрідними платформами для збирання різноманітних білкових комплексів, склад яких залежить від патерну модифікацій – розподілу певних модифікованих груп по хвостах. Співвідношення між патерном модифікацій і набором білків, які впізнають такий патерн, що у свою чергу має певні функціональні наслідки, називають **гістоновим кодом**.



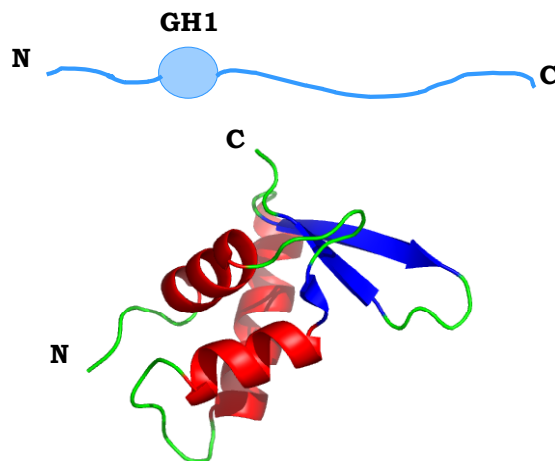
**Рис. 4.13.** Кінцеві хвости гістонів у структурі нуклеосоми (1KX5) та їхні модифікації. Кольором виділено по одному N-кінцевому хвосту кожного гістону, а також С-кінцевий – гістону H2A (поряд із N-кінцем H3). Стрілочками вказано ацетилювання (червоні), метилювання (зелені) і фосфорилювання (блакитні) залишків певного типу, які займають певні позиції в поліпептидних ланцюгах

Наприклад, фосфорилювання H3-Ser10 із одночасним ацетилюванням H3-Lys14 супроводжується активацією транскрипції, деацетилювання H3-Lys14 із одночасним метилюванням H3-Lys9 – репресією. Це лише невеличкий елемент такого коду: зрозуміло, що різноманітних комбінації модифікованих залишків на восьми гістонових хвостах (рис. 4.13) може бути велика кількість. Остаточне з'ясу-

вання гістонового коду ще далеко не завершено. Певні його закономірності та інші приклади залучення гістонового коду до регуляції транскрипції наведені у розділі 6.

### *Гістон H1*

Одна молекула п'ятого гістону – лінкерного гістону H1 – взаємодіє з кожною нуклеосомою у хроматині. Проте ця взаємодія, на відміну від корових гістонів, є динамічною: спостерігається швидкий обмін лінкерних гістонів між хроматином та їхнім пулом у ядрі.



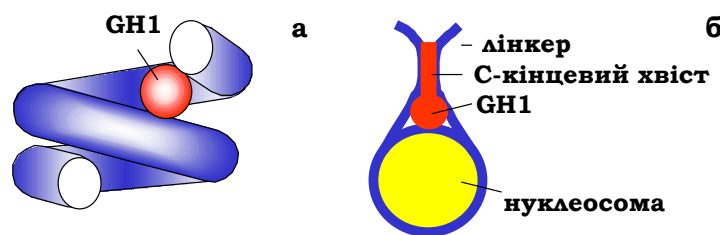
**Рис. 4.14.** Схема будови гістону H1 і структура його глобулярного домену GH1 (1HST, аналог гістону H1 з еритроцитів птахів, що позначається як H5)

За своєю структурою гістон H1 (мономерний білок) суттєво відрізняється від корових гістонів (рис. 4.14). Молекула H1 містить N-кінцевий неупорядкований хвіст, глобулярний домен GH1 і довгий, збагачений позитивно зарядженими залишками (насамперед Lys) C-кінцевий хвіст довжиною приблизно 100 амінокислотних залишків (приблизно половина молекули). Глобулярний домен (приблизно 80 амінокислот) належить до так званої родини “спіральних із крильцем” (winged helix) ДНК-зв'язувальних білків.

На поверхні домену існують два можливі сайти взаємодії з ДНК, які формуються кластерами позитивно заряджених амінокислот і локалізовані на протилежних боках глобули. Відповідно, глобулярний

#### Розділ 4. Організація днк у клітинах: геноми та структура хроматину

домен взаємодіє з нуклеосомною ДНК на виході з нуклеосоми, імовірно, між двома сусідніми витками (рис. 4.15, а), додатково стабілізуючи нуклеосомну суперспіраль. С-кінцевий хвіст Н1 взаємодіє з обома лінкерами, що виходять з нуклеосоми. У результаті дві лінкерні ділянки (довжиною по 10–30 пар основ) утворюють **стеблоподібну структуру** на виході з нуклеосоми (рис. 4.15, б).



**Рис. 4.15.** Схема взаємодії глобулярного домену GH1 (а) та цілого гістону H1 (б) із нуклеосомою

#### *Наднуклеосомна упаковка хроматинової фібрили*

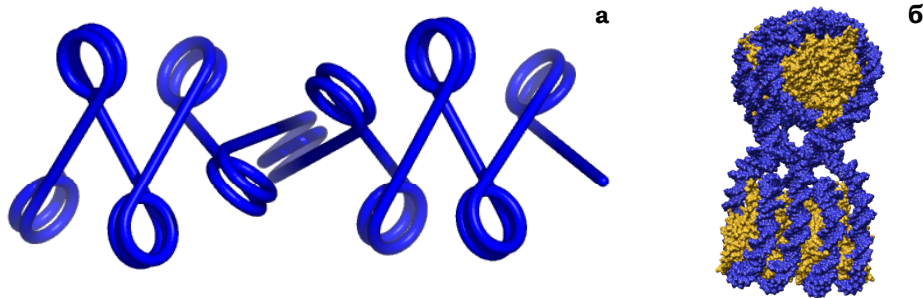
У хроматині нуклеосоми з'єднані лінкерами довжиною ~50 пар основ. Якщо лінкер (у полінуклеосомному ланцюзі без гістону H1) просто продовжує хід нуклеосомної ДНК по прямій, нуклеосоми у складі полінуклеосомної нитки мають бути розташовані зигзагом (рис. 4.16, а).

Саме такий вигляд має декомпактизована (якщо відсутній H1 і за низької іонної сили) полінуклеосомна фібрила під мікроскопом (електронним чи атомно-силовим). Товщина такого зигзагу дорівнює приблизно 30 нм. Полінуклеосомний зигзаг може конденсуватися (рис. 4.16, б).

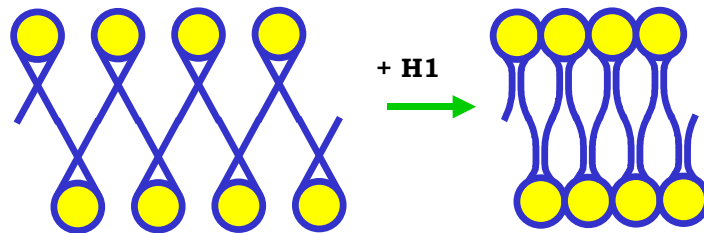
**Умова конденсації** (необхідна, але недостатня) – фізіологічна іонна сила (крім одновалентних мають бути присутніми двовалентні катіони) для зниження електростатичного розштовхування між нуклеосомами. **Фактор конденсації** (її рушійна сила) – невпорядковані хвости корових гістонів: лабільні позитивно заряджені хвости ефективно “зшивають” фібрилу, взаємодіючи з ДНК сусідніх нуклеосом.

У присутності гістону H1 компактна хроматинова фібрила товщиною 30 нм стає значно стабільнішою. Ключова роль у цій стабілізації належить стеблу, що формується унаслідок взаємодії з гістоном H1 двох лінкерних ділянок на виході з нуклеосоми: за рахунок стебла сусідні нуклеосоми значно зближуються, що сприяє конденсації (рис. 4.17). Отже, присутність H1 наближає нуклеосоми одна до одної у складі фібрили, чим сприяє “зшиванню” цих нуклеосом хвостами

корових гістонів. І навпаки, компактизація фібрили за участю хвостів корових гістонів сприяє зв'язуванню Н1 із наближеними у просторі ділянками лінкерів.



**Рис. 4.16.** Зигзагоподібна конфігурація полінуклеосомної нитки (а) і структура компактизованого елемента такої нитки у кристалах тетрануклеосом (б, 1ZBB)



**Рис. 4.17.** Роль Н1-залежного стебла на виході з нуклеосоми у стабілізації компактного стану хроматинової фібрили.

Отже, суперструктура конденсованої фібрили товщиною 30 нм являє собою тривимірний зигзаг нуклеосом, з'єднаних практично прямими – без вигинів – лінкерами, які спрямовані всередину фібрили. Усередині містяться також і гістони Н1; сумісна дія Н1 і хвостів корових гістонів підтримує компактний стан фібрили.

Фібрила товщиною 30 нм – основна форма існування інтерфазного хроматину. Але у хроматині існує значна гетерогенність за ступенем конденсації. З одного боку, передумовою активації окремих ділянок хроматину є деконденсація фібрили. Факторами такої деконденсації є зниження спорідненості до ДНК хвостів корових гістонів унаслідок їхнього ацетилювання (зниження позитивного заряду) і тимчасова дисоціація Н1. Дисоціації Н1 сприяє ацетилю-

#### **Розділ 4. Організація днк у клітинах: геноми та структура хроматину**

вання хвостів корових гістонів, а також посттрансляційні модифікації самого Н1, зокрема фосфорилування. Деконденсація фібрили та дисоціація Н1 створюють можливість взаємодії з ДНК різноманітних негістонових білків, що врешті-решт і викликає активацію транскрипції (див. розділ 6).

З іншого боку, у репресованих ділянках хроматинова фібрила може бути як додатково стабілізованою в компактному стані, так і піддаватися компактизації вищого порядку. Частина хроматину, що зберігає стан підвищеної компактизації протягом інтерфази, називається **гетерохроматином** (решта хроматину, де в принципі може відбуватися активація транскрипції, позначається як *еухроматин*). Утворення гетерохроматину здійснюватися, головним чином, у ділянках, що містять повтори – у центромерах, теломерах, зонах концентрації мобільних елементів. Деталі структури гетерохроматину залишаються нез'ясованими, механізми встановлення та підтримання неактивного конденсованого стану гетерохроматину розглядаються в розділі 6.

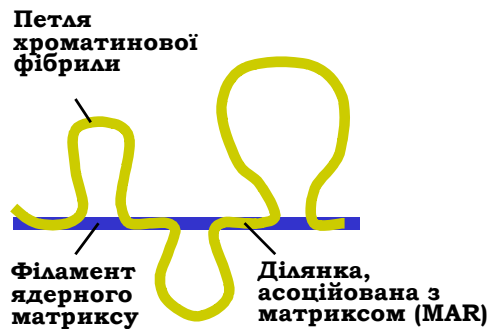
#### ***Петльові домени хроматину та ядерний матрикс***

На наступному рівні структурної організації у клітинному ядрі хроматинова фібрила формує петлі, кінці яких є жорстко закріплені на скелетних структурах клітинного ядра – *ядерному матриксі* (рис. 4.18). Оскільки основи петлі фіксуються на матриксі, ДНК у складі петлі має топологічні обмеження, тобто є еквівалентною циркулярній (див. розділ 3). Одна петля, що містить від 20 до 200 тис. пар основ ДНК (один або кілька генів), часто розглядається як важливий елемент регуляції процесів транскрипції та реплікації.

Із білками матриксу взаємодіють ділянки ДНК довжиною від 300 до 1 тис. пар основ – ділянки, асоційовані з матриксом (MAR, Matrix Associated Regions). Вони не мають між собою гомології щодо послідовності, часто є збагаченими на АТ-пари. Вважається, що більшість функціональних процесів, які відбуваються на ДНК, розпочинається саме на матриксі, тобто на MAR.

Ядерний матрикс – це система білкових філаментів, яка формує структурний каркас ядра. На периферії ядра розташована особлива частина матриксу, асоційована із внутрішньою ядерною мембраною – *ядерна ламіна*. Філаменти ламіни протягнуті від однієї ядерної пори до іншої і позиціонують порові комплекси у площині мембрани. Подвійна ядерна мембрана, ядерні пори й ламіна утворюють єдину систему – *ядерний конверт*. Від ламіни всередину ядра протягнуті філаменти

внутрішнього ядерного матриксу. З ламиною взаємодіє значна частина гетерохроматину, зокрема центромери та теломери хромосом. Еухроматинова частина хромосоми “звисяє” всередину ядра, де хроматинові петлі закріплюються на внутрішній частині матриксу. У результаті хромосома займає певну зону в об'ємі ядра – *хромосомну територію*.



**Рис. 4.18.** Схема петльової організації хроматину

Усе, про що йшлося вище, стосується організації хроматину та хромосом в інтерфазному ядрі. Коли клітина вступає в процес мітозу, після реплікації ДНК починається утворення надкомпактної мітотичної хромосоми, деталі структурної організації якої залишаються недостатньо зрозумілими. Одночасно з компактизацією хроматинової фібрили частина ламіни “розчиняється” разом із ядерною мембраною, інша частина, разом із внутрішнім матриксом, перебудовується з утворенням білкового каркасу мітотичної хромосоми – хромосомного *скефолду* (scaffold), з яким залишаються зв'язаними основи хроматинових петель.

### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. Дайте визначення кодона і відкритої рамки зчитування. Скільки існує кодонів? Які позиції нуклеотидів у складі кодона є найбільш і найменш визначальними?
2. Що таке ген?
3. Що таке геном? У чому полягає найсуттєвіша відмінність між геномами про- та еукаріотів?
4. У чому полягає перекриття генів у вірусів та еукаріотів?
5. Дайте визначення інтрона і екзона.

#### **Розділ 4. Організація днк у клітинах: геноми та структура хроматину**

6. Яка різниця між кластером генів і опероном?
7. Назвіть основні типи послідовностей ДНК, що повторюються.
8. З яких елементів складається нуклеосома? Яке місце в її структурі належить білкам? Як організована у просторі ДНК у нуклеосомі?
9. Які взаємодії стабілізують нуклеосому? За яким принципом відбувається збирання нуклеосом?
10. У чому полягає механізм позиціювання нуклеосом відносно послідовності ДНК?
11. Назвіть основні типи посттрансляційних модифікацій гістонів. Які амінокислотні залишки піддаються модифікаціям? В яких частинах молекул гістонів ці залишки розташовані?
12. Як організована хроматинова фібрила у просторі? За яким основним механізмом фібрила піддається компактизації? У чому полягає роль гістону H1?
13. Що таке ядерний матрикс і яку участь він бере в організації хроматину в клітинному ядрі?

#### **РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

- Akey, C.W., Luger K.* Histone chaperones and nucleosome assembly // *Curr. Op. Struct. Biol.* – 2003. – Vol. 13. – P. 6–14.
- Bednar, J., Horowitz, R.A., Grigoryev, S.A. et al.* Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin // *Proc. Natl. Acad. Sci. USA.* – 1998. – Vol. 95. – P. 14173–14178.
- Brown T.A.* Genomes. – New York ; London: Garland Science, 2002.
- Chromatin structure and gene expression: frontiers in molecular biology / eds. S. Elgin, J.L. Workman. – Oxford : Oxford University Press, 2002.
- Chromatin structure and dynamics: state-of-the-art. New Comprehensive Biochemistry / eds. J. Zlatanova, S.H. Leuba. Amsterdam : Elsevier, 2004. – Vol. 39.
- The ENCODE Project Consortium. Identification and analysis of functional elements in 1 % of the human genome by the ENCODE pilot project // *Nature.* – 2007. – Vol. 447. – 799–816.
- Gerstein, M.B., Bruce, C., Rozowsky, J.S. et al.* What is a gene, post-ENCODE? History and updated definition // *Genome Res.* – 2007. – Vol. 17. – P. 669–681.
- Gruenbaum, Y., Margalit, A., Goldman, R.D. et al.* The nuclear lamina comes of age // *Nat. Rev.* – 2005. – Vol. 6. – P. 21–31.

International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome // *Nature*. – 2001. – Vol. 409. – P. 860–921.

International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome // *Nature*. – 2004. – Vol. 431. – P. 931–945.

*Koonin, E.V.* How many genes can make a cell: the minimal-gene-set concept // *Annu. Rev. Genomics Hum. Genet.* – 2000. – Vol. 1. – P. 99–116.

*Lesk, A.M.* Introduction to bioinformatics. – New York : Oxford University Press, 2002.

*Luger, K., Richmond, T. J.* DNA binding within the nucleosome core // *Curr. Op. Struct. Biol.* – 1998. – Vol. 8. – P. 33–40.

*Luger, K., Richmond, T.J.* The histone tails of the nucleosome // *Curr. Opin. Genet. Dev.* – 1998. – Vol. 8 – P. 140–146.

*Pearson, H.* Genetics: what is a gene? // *Nature*. – 2006. – Vol. 441. – P. 398–401.

*Ramakrishnan, V.* Histone H1 and chromatin higher-order structure // *Crit. Rev. Eukaryot. Gene Expr.* – 1997. – Vol. 7. – P. 215–230.

*Simpson, R.T.* Nucleosome positioning: occurrence, mechanisms, and functional consequences // *Prog. Nucl. Acid Res. Mol. Biol.* – 1991. – Vol. 40. – P. 143–184.

*Sivolob, A., Prunell, A.* Nucleosome conformational flexibility and implications for chromatin dynamics // *Phil. Trans. Roy. Soc. Lond. A.* – 2004. – Vol. 362. – P. 1519–1547.

*Strahl, D.D., Allis, C.D.* The language of covalent histone modifications // *Nature*. – 2000. – Vol. 403. – P. 41–45.

*Tremethick, D.J.* Higher-order structures of chromatin: the elusive 30 nm fiber // *Cell*. – 2007. – Vol. 128. – P. 651–654.

*Wu, H.-L., Bagby, S., van den Elsen J.M.H.* Evolution of the genetic triplet code via two types of doublet codons // *J. Mol. Evol.* – 2005. – Vol. 61. – P. 54–64.



## Розділ 5

### ТРАНСКРИПЦІЯ: ПРОКАРІОТИ

Down, down, down. There was  
nothing else to do.

*Lewis Carroll.*  
Alice's adventures in Wonderland

Транскрипція (transcription) – процес синтезу РНК з використанням одного з ланцюгів ДНК як матриці, тобто “переписування” послідовності нуклеотидів ДНК у послідовність нуклеотидів РНК. Зростання ланцюга РНК відбувається в напрямку від 5'- до 3'-кінця. Субстратами реакції є 3'-кінцева ОН-група рибози зростаючого *транскрипту* (ланцюга РНК, що синтезується) і рибонуклеозидтрифосфати (rNTP). Фермент, що каталізує цю реакцію – *ДНК-залежна РНК-полімераза* (DNA dependent RNA Polymerase, RNAP). Приступаючи до розгляду прокаріотичної системи транскрипції, слід зауважити, що молекулярні механізми синтезу РНК є в основному спільними для всіх живих організмів.

#### РНК-полімераза

##### *Загальний сценарій транскрипції в бактерій*

Бактеріальна РНК-полімераза складається з кількох білкових субодиниць та існує, залежно від стадії транскрипції, у двох формах: 1) *кор-фермент* (core enzyme) у складі субодиниць, що позначаються як  $\alpha$  (дві копії),  $\beta$ ,  $\beta'$  та  $\omega$ ; 2) *голофермент* (holoenzyme) – комплекс кор-ферменту із субодиницею  $\sigma$ . Кор-фермент має досить високу неспецифічну спорідненість до ДНК, що дозволяє йому працювати на різноманітних послідовностях. Поява субодиниці  $\sigma$  у складі голоферменту приводить до *зниження* загальної неспецифічної спорідненості, але при цьому виникає *специфічна* спорідненість до особливих ділянок, з яких транскрипція має розпочинатися – *промоторів* (promoters).

За рахунок цього голофермент має змогу ефективно перебирати (шляхом зв'язування та швидкої дисоціації) різноманітні ділянки ДНК, здійснюючи пошук промотора, взаємодія з яким є міцнішою. Отже, субодиниця  $\sigma$  виконує роль *загального фактора ініціації* транскрипції. Різні промотори розрізняються за спорідненістю до голоферменту (*силою промотора*). Відповідно, упізнання слабких промоторів залежить від додаткових, специфічних для даного гена чи групи генів, факторів ініціації (див. нижче).

Робочий цикл РНК-полімерази складається з наступних стадій.

- **Ініціація** транскрипції, яка також є багатостадійним процесом (рис. 5.1):
  - зв'язування голоферменту з промотором. У результаті формується *закритий комплекс*, у складі якого ДНК зберігає форму подвійної спіралі;
  - локальне плавлення подвійної спіралі з утворенням *відкритого комплексу* – розходження ланцюгів ДНК, яке дозволяє використовувати один з них як матрицю;
  - включення перших двох нуклеотидів до молекули РНК (синтез першого фосфодіефірного зв'язку в активному центрі полімерази) – найповільніша стадія процесу;
  - зростання первинного короткого транскрипту – приєднання 8–9 нуклеотидів. Після цього є можливою абортивна ініціація (визволення короткого транскрипту), тобто невдала спроба ініціації;
  - в іншому випадку відбувається *очищення промотора* – дисоціація  $\sigma$ -фактора, яка маркує перехід до елонгації транскрипції.
- **Елонгація** транскрипції, у кожному елементарному акті якої (елонгаційному циклі) відбувається приєднання чергового нуклеотиду до 3'-кінця РНК і пересування кор-ферменту на один нуклеотид уздовж матриці (*транслокація*).
- **Термінація** транскрипції при впізнанні полімеразою спеціального сигналу термінації (особливого елемента послідовності), коли відбувається визволення транскрипту. Далі кор-фермент знову взаємодіє з  $\sigma$ -субодиницею і здійснює новий пошук промотора.

Суттєвою особливістю прокаріотичної системи транскрипції білкових генів є те, що молекула мРНК зв'язується з рибосомами безпосередньо під час транскрипції – транскрипція мРНК і білковий синтез (розділ 8) є єдиним процесом.

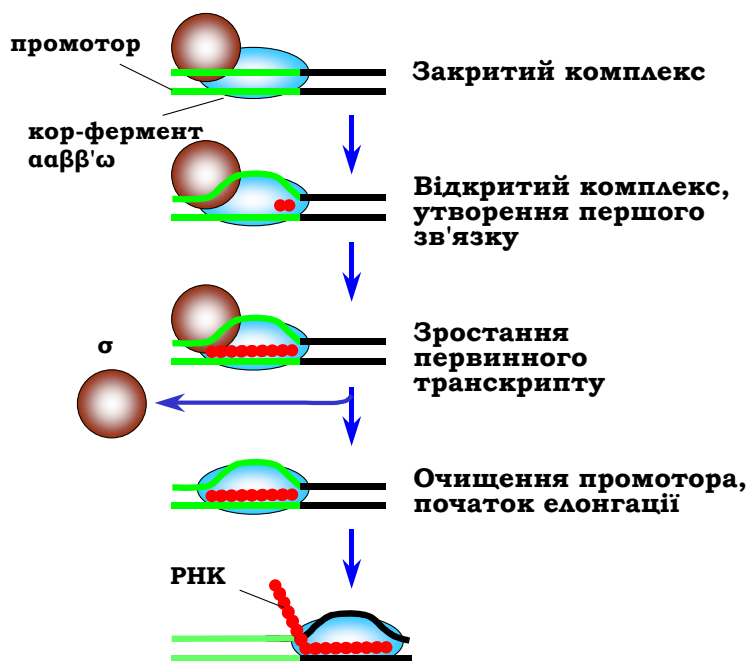


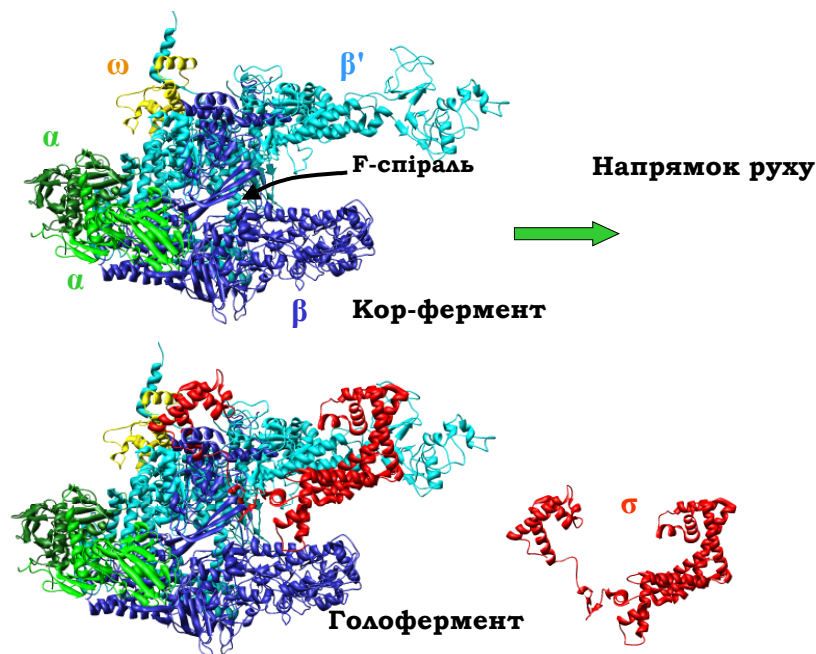
Рис. 5.1. Ініціація транскрипції бактеріальною РНК-полімеразою

### Структура РНК-полімерази

Дві субодиниці  $\alpha$  взаємодіють між собою в складі кор-ферменту бактеріальної РНК-полімерази (рис. 5.2), виконуючи структурну роль: вони сприяють утриманню разом інших частин ферменту. Дві великі субодиниці –  $\beta$  і  $\beta'$  – формують характерні “щелепи” (форма нагадує також клішню краба), у щілині між якими з ферментом взаємодіє ДНК, розташована “нижче” (downstream) у напрямку руху полімерази від точки, яка в даний момент знаходиться в активному центрі. Щелепи є рухливими елементами, здатними розмикатися /замикатися під час роботи полімерази.

Глибоко в щілині між щелепами, ліворуч від F-спіралі (F bridge helix), що зображена на рис. 5.2 (належить до найбільшої субодиниці  $\beta'$ ), розташований каталітичний активний центр: 3 залишки Asp (також належать до  $\beta'$ ), що утримують ключовий для каталізу іон  $Mg^{2+}$  (рис. 5.3). Структурні елементи, які оточують активний центр, забезпечують взаємодію полімерази з нематричним ланцюгом ДНК (підтримуючи

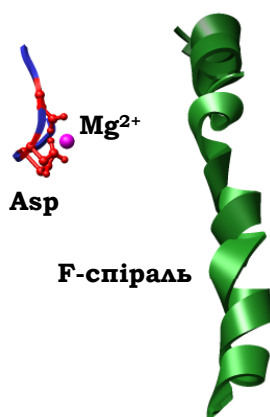
його розплавленій стан), дволанцюговим гібридом ДНК-РНК, сприяють поділу подвійної спіралі ДНК нижче від активного центру, поділу подвійної спіралі ДНК-РНК гібрида та відновленню подвійної спіралі ДНК вище (upstream) активного центру. Від активного центру, де міститься 3'-кінець РНК, що зростає, РНК виходить за межі полімеразного комплексу через спеціальний канал. Інший, так званий *вторинний канал*, дозволяє нуклеозидтрифосфатам потрапляти до точки зростання РНК. Загальну схему будови полімерази в комплексі з ДНК і РНК під час елонгації транскрипції наведено на рис. 5.4.



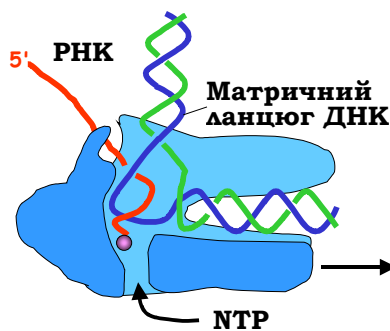
**Рис. 5.2.** Структура РНК-полімерази *Thermus thermophilus* (1IW7)

Описані риси структурної організації бактеріальної РНК-полімерази (рухливі щелепи, активний центр з іоном  $Mg^{2+}$ , канали для виходу РНК і заходу NTP) є загальними для усіх РНК-полімераз. Це стосується не тільки гомологічних еукаріотичних полімераз (див. розділ 6). Деякі бактеріофаги (зазвичай фаги використовують полімеразу клітинихазяїна) мають власні РНК-полімерази, що є мономерними білками без гомології в амінокислотній послідовності до субодиниць бактеріальної полімерази (рис. 5.5). Простота мономерних полімераз (яка робить їх

зручним об'єктом досліджень) зумовлена тим, що вони мають упізнавати лише кілька промоторів фагових генів. Але вказані вище загальні риси структурної організації РНК-полімераз та базові принципи функціонування зберігаються і для них.



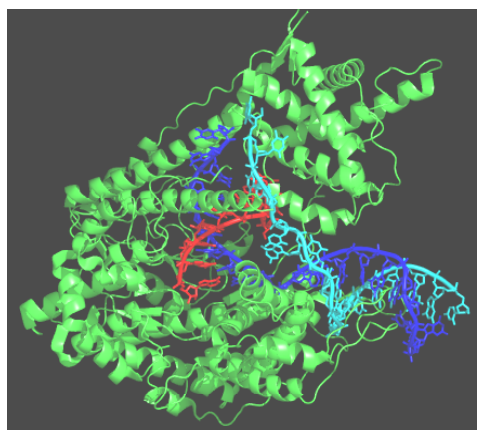
**Рис. 5.3.** Активний центр поряд із F-спіраллю у складі РНК-полімерази *Thermus aquaticus* (1HQM). Орієнтація полімерази приблизно така сама, як на рис. 5.2



**Рис. 5.4.** Схема організації полімеразного комплексу (у розрізі) під час елонгації транскрипції

У структурі голоферменту бактеріальної РНК-полімерази присутня додаткова субодиниця  $\sigma$ . Вона має мультидоменну структуру (рис. 5.2) і здійснює різноманітні взаємодії з обома великими субодиницями ( $\beta$  і  $\beta'$ ): фіксує певне положення щелеп, взаємодіє з активним центром і його оточенням. Розташування  $\sigma$ -фактора на поверхні полімеразного

комплексу дозволяє йому здійснювати впізнання промотора. Крім того, частина  $\sigma$  взаємодіє з каналом виходу РНК (порівн. рис. 5.2 і 5.4), блокуючи його: дисоціація  $\sigma$  є необхідною для виходу синтезованої РНК, тобто для вільного руху полімерази вздовж матриці під час елонгації.

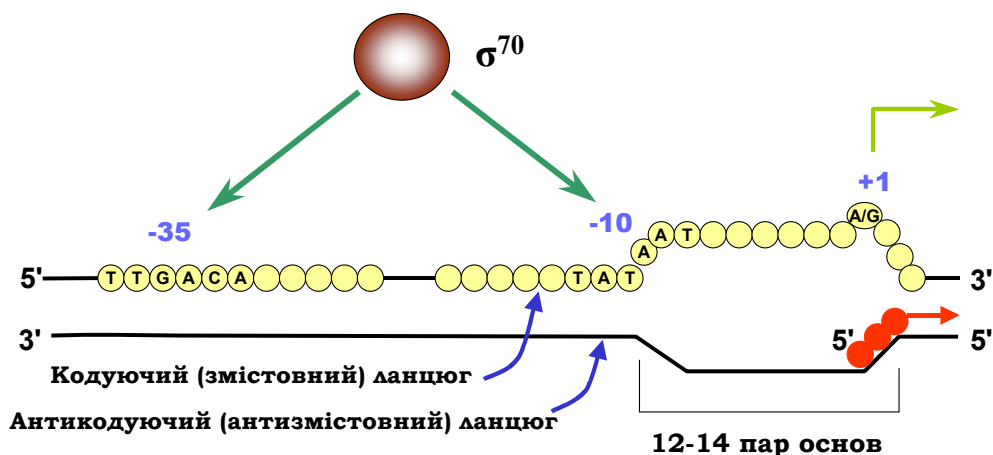


**Рис. 5.5.** Структура елонгаційного комплексу мономерної РНК-полімерази бактеріофага T7 (1MSW). Матричний ланцюг ДНК – темно-синій, РНК – червона

### Ініціація транскрипції

Зрозуміло, що РНК-полімераза може бути орієнтована відносно ДНК двома способами: кожний із цих орієнтацій буде відповідати два можливі напрямки руху ферменту при транскрипції. Відповідно, один із двох ланцюгів ДНК буде обиратися як *матричний* – той, з якого зчитується інформація в напрямку 3'-5' і на якому синтезується комплементарний ланцюг РНК у протилежному напрямку 5'-3' (рис. 5.6).

Послідовність ланцюга ДНК, який є комплементарним матричному, збігається з послідовністю РНК, що синтезується. Тому саме цей нематричний ланцюг називають *кодуючим* або *змістовним*, і саме його послідовність прийнято наводити як послідовність даного гена (матричний ланцюг називають, відповідно, антикодуючим). Нуклеотид кодуючого ланцюга, який відповідає стартовому нуклеотиду РНК (частіше пуриновий), позначається як +1, наступні нуклеотиди (чи пари основ) у напрямку руху полімерази нумерують із позначкою "+": +2, +3 і т.д. У зворотному напрямку (ліворуч на рис. 5.6) нуклеотиди нумерують із позначкою "-" (починаючи з -1).



**Рис. 5.6.** Типовий промотор *E. coli*. Зелена стрілка позначає старт транскрипції всередині розплавленої зони

Типовий бактеріальний промотор (рис 5.6) складається із двох елементів послідовності, які знаходяться на відстані приблизно  $-10$  і  $-35$  від стартової точки. Консенсусна (узагальнена для різних промоторів) послідовність елемента  $-35$  має вигляд  $T_{82}T_{84}G_{78}A_{65}C_{54}A_{45}$  (числами позначено частоту даної пари основ у відсотках). Послідовність елемента  $-10$  (називається також боксом Прибноу (David Pribnow) – Pribnow box):  $T_{80}A_{95}T_{45}A_{60}A_{50}T_{96}$ . Варіабельність послідовностей визначає різну ефективність ініціації – силу промоторів. Саме елементи  $-35$  і  $-10$  безпосередньо впізнаються  $\sigma$ -фактором при ініціації транскрипції. Наведені послідовності впізнаються варіантом  $\sigma$ -фактора з молекулярною вагою 70 кДа ( $\sigma^{70}$ ), який є досить широко вживаним, але не єдиним. Кільком іншим варіантам  $\sigma$ , під контролем яких знаходяться певні групи промоторів, відповідають інші консенсусні послідовності  $-35$  та  $-10$ . Взаємодія  $\sigma$ -фактора з промотором чітко орієнтує РНК-полімераза відносно ДНК, тобто визначає і стартову точку транскрипції, і її напрям (вибір одного з двох ланцюгів як матричного).

Вважається, що субодиниця  $\sigma$  утримує щелепи у відкритому положенні у складі голоферменту. Це сприяє підвищенню швидкості дисоціації білка від непромоторної ДНК і гарантує низьку ефективність транскрипції з випадкових сайтів на ДНК (рис. 5.7). За умови впізнання  $\sigma$ -фактором двох елементів промотора відбувається структурна

перебудова полімеразного комплексу із замиканням щелеп. У результаті ДНК у зоні пари основ +1 потрапляє всередину щілини між щелепами – утворюється закритий комплекс.

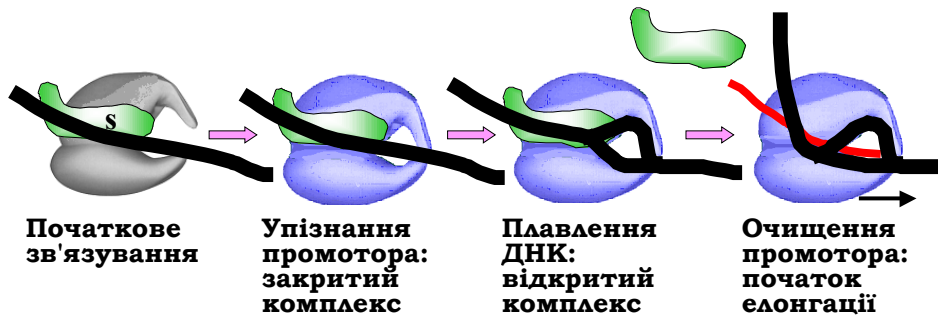


Рис. 5.7. Послідовність подій при ініціації транскрипції

Дуже швидко закритий комплекс перетворюється на відкритий: замикання ДНК у щілині приводить до того, що подвійна спіраль не може розміститися там за недостатністю простору. У результаті відбувається локальне плавлення 12–14 пар основ (формування транскрипційного міхура) у зоні від  $-7$  до  $+5$  відносно стартової точки (рис. 5.6, 5.7). Важливим фактором, що сприяє локальному плавленню подвійної спіralі у складі відкритого комплексу, є наявність у циркулярній бактеріальній ДНК певного рівня негативної надспіралізації – напружень, що полегшують розкручування дуплекса (див. розділ 3). Плавлення ДНК у щілині РНК-полімерази відбувається за рахунок взаємодії двох ланцюгів із внутрішньою поверхнею щілини: нематричний ланцюг захоплюється певними структурними елементами полімерази, матричний – опиняється в активному центрі. За активним центром ДНК робить вигин майже під прямим кутом відносно ДНК, що лежить нижче від активного центру.

Наступна, найповільніша стадія ініціації – синтез першого фосфодіефірного зв'язку, тобто ініціація власне синтезу РНК. Здатність ініціювати синтез нуклеїнової кислоти – унікальна властивість РНК-полімерази (ДНК-полімераза здатна лише продовжувати синтез уже існуючого ланцюга, див. розділ 9). Ініціація синтезу є дуже важкою, оскільки при цьому дві порівняно невеликі молекули – перші два нуклеотиди – мають бути жорстко зафіксовані в активному центрі в певній геометрії. Така фіксація залежить від взаємодій  $\sigma$ -фактора з активним центром: певною мірою субодиниця  $\sigma$  відіграє роль 3'-кінцевої ділянки РНК, яка бере участь



у створенні сайту зв'язування чергового нуклеотиду при елонгації (див. нижче), але якої ще не існує при ініціації.

Далі процес нарощування транскрипту продовжується ще до восьми-дев'яти нуклеотидів, які можуть розміститися в межах РНК-полімеразного комплексу до виходу в спеціальний канал. У процесі зростання цього первинного транскрипту можливе його визволення – абортівна ініціація. Узагалі, оточення активного центру РНК-полімерази має високу спорідненість до тимчасового РНК-ДНК-гібрида, що існує під час транскрипції (див. рис. 5.4). Але чим коротшим є гібрид, тим менше взаємодій він реалізує з полімеразою, і тим менш стабільним він є. Додаткова стабілізація короткого гібрида здійснюється завдяки взаємодії з ним  $\sigma$ -фактора.

Друга причина абортівної ініціації – конкуренція між  $\sigma$ -фактором і РНК за канал виходу. У принципі, після зростання транскрипту до 9–10 нуклеотидів  $\sigma$ -фактор більше не потрібен – “мавр зробив свою справу”. Можливий програш з боку РНК конкуренції за канал є своєрідною платою за можливість ініціювати синтез першого зв'язку. У разі такого програшу голофермент може повторити свою спробу ініціювати транскрипцію. Якщо ж конкуренцію виграє РНК,  $\sigma$ -фактор (після синтезу близько 12 нуклеотидів) нарешті витісняється з комплексу (відбувається очищення промотора) і розпочинається елонгація транскрипції.

### **Елонгація транскрипції**

При елонгації транскрипції РНК-полімераза рухається вздовж матриці разом із розплавленими 12–14 парами основ у транскрипційному міхурі (рис. 5.4, 5.8). При цьому на кожному кроці полімерази одна пара основ руйнується попереду міхура й одна відновлюється позаду. Розділення подвійної спіралі ДНК перед міхуром полегшується так званою G-петлею РНК-полімерази (G-trigger loop, рис. 5.8). У зоні міхура з полімеразою завжди зв'язаний ДНК-РНК-гібрид довжиною вісім-дев'ять пар основ: вісім-дев'ять нуклеотидів на 3'-кінці РНК залучені до утворення подвійної спіралі з матричним ланцюгом, наступні п'ять-шість нуклеотидів локалізовані в каналі виходу РНК, решта нуклеотидів на 5'-кінці виходять за межі полімеразного комплексу. Поділу гібрида, а отже, і відновленню ДНК-дуплекса, сприяє L-петля (lid loop, рис. 5.8).

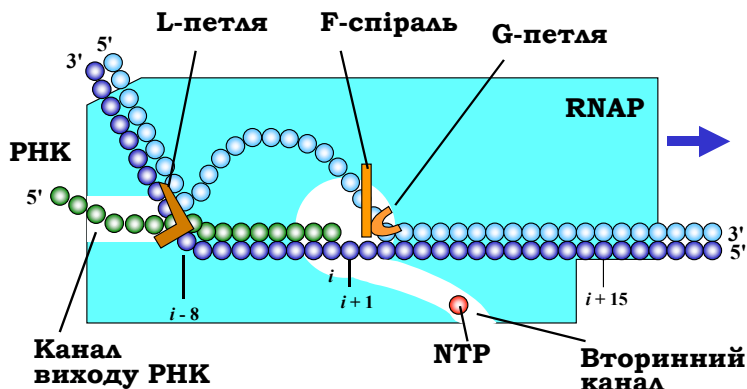


Рис. 5.8. Вихідна конфігурація системи транскрипції на початку елонгаційного циклу

Наявність ДНК-РНК гібрида при транскрипції є дуже важливою обставиною. Цей гібрид існує завдяки специфічним взаємодіям, які реалізують з ним певні елементи полімерази. Висока спорідненість РНК-полімерази до гібрида забезпечує високу **процесивність** ферменту – здатність працювати на матриці без дисоціації. Інший фактор підвищення процесивності – взаємодія транскрипту з каналом виходу РНК. Дисоціація полімерази, яка була можливою під час синтезу первинного транскрипту, поки гібрид був недостатньо довгим (абортивна ініціація, див. вище), під час елонгації стає малоімовірною подією.

На рис. 5.8 схематично зображено вихідну конфігурацію системи на початку кожного елонгаційного циклу перед зв'язуванням чергового НТР. Сайт зв'язування формується 3'-кінцевим нуклеотидом РНК, неспареним нуклеотидом матричного ланцюга ДНК (позначеним як  $i+1$ ), активним центром полімерази та його оточенням, зокрема F-спіраллю (рис. 5.3, 5.8). ОН-група на 3'-кінці РНК фіксується в активному центрі поряд з іоном  $Mg^{2+}$  (рис. 5.3, 5.9). НТР потрапляє до сайту зв'язування через вторинний канал разом з другим іоном  $Mg^{2+}$ .

**Перший етап** елонгаційного циклу – **відбір і зв'язування** нуклеозидтрифосфату, яке контролюється матрицею (нуклеотидом  $i+1$ ). Спорідненість усіх чотирьох типів НТР до РНК-полімерази є досить невисокою, що дозволяє їм здійснювати швидке зв'язування / дисоціацію. Реалізація комплементарних взаємодій НТР із відповідним

## Розділ 5. Транскрипція: Прокаріоти

нуклеотидом матриці (поява нового ліганду в сайті зв'язування – комплементарної нуклеотидної пари) викликає конформаційну зміну із замиканням NTP та жорсткою фіксацією субстратів у активному центрі (рис. 5.9, 5.10).

**Другий етап** – хімічна реакція **приєднання чергового нуклеотиду** до 3'-кінця транскрипту (рис. 5.9). За умови реалізації комплементарних взаємодій між NTP і нуклеотидом у складі матриці, NTP жорстко фіксується в активному центрі в певній “напруженій” реакційно-здатній конформації, яка може розглядатися як інтермедіат реакції.

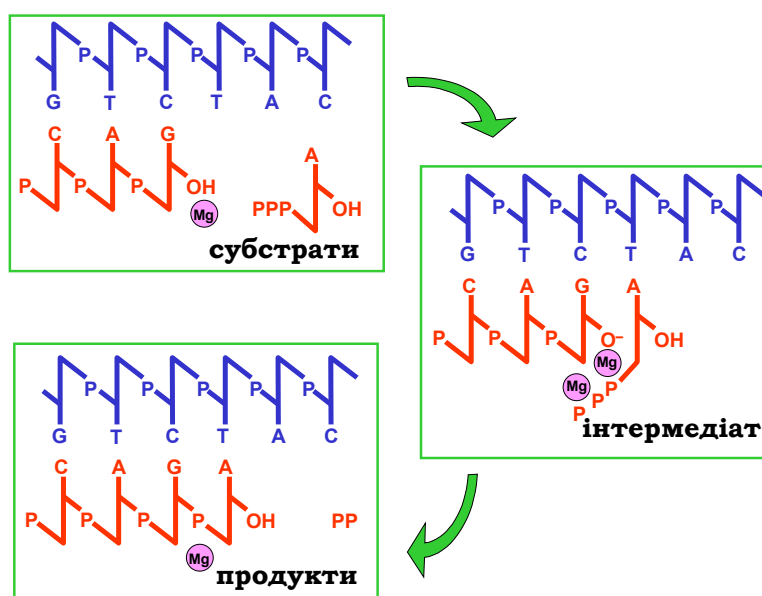
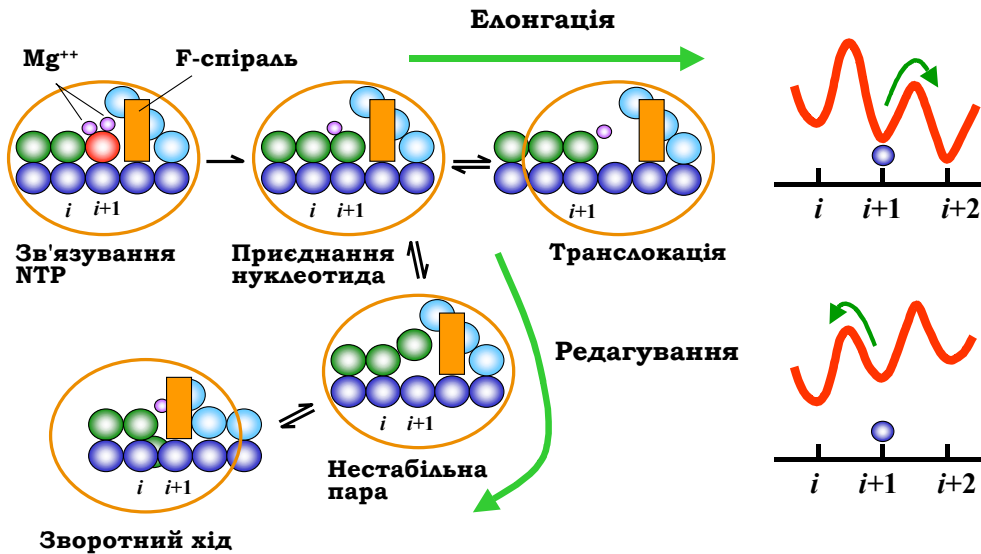


Рис. 5.9. Хімічна схема процесу приєднання нуклеотиду

Найважливішу роль у фіксації інтермедіату відіграють два іони  $Mg^{2+}$  (один утримується в активному центрі завжди, інший, який потрапив туди разом з NTP – тимчасово, через взаємодію з двома із трьох залишків Asp активного центру). Перший іон взаємодіє з  $\alpha$ -фосфатом NTP і 3'-ОН групою, забезпечуючи зокрема її іонізацію (рис. 5.9). Другий іон стабілізує певну напружену конформацію всіх трьох фосфатних залишків. Спорідненість активного центру до інтермедіату і жорстка фіксація субстратів забезпечує зниження активаційного бар'єра реакції, тобто її каталіз (див. розділ 2). Результатом реакції

є подовжений на один нуклеотид 3'-кінець РНК і відщеплений пірофосфат. Крім того, унаслідок реакції змінюється розташування ліганду (3'-кінця) відносно активного центру (рис. 5.9; 5.10).



**Рис. 5.10.** Послідовність подій в елонгаційному циклі (збільшений фрагмент схеми на рис. 5.8). Показано також альтернативний шлях при редагуванні помилок. Праворуч: енергетичні профілі, що відповідають руху полімерази вперед відносно матриці при елонгації (угорі) і назад при редагуванні (унизу)

Зміна ліганду та дисоціація пірофосфату спричиняє "розмикання" структури в зоні активного центру, яке дозволяє рух полімерази. Відбувається **третій етап – транслокація** у напрямку оптимального розташування подовженого 3'-кінця відносно активного центру (рис. 5.10). Рух полімерази при транслокації відбувається як одномір- на дифузія вздовж матриці за рахунок енергії теплових флуктуацій, тобто можливим є рух як "уперед", так і "назад". Пересування поліме- рази вперед супроводжується руйнуванням однієї нуклеотидної пари у складі РНК-ДНК гібрида та відновленням однієї пари ДНК позаду від транскрипційного міхура. Одночасно руйнується одна пара ДНК попереду від міхура. Якщо взяти до уваги, що на попередньому етапі внаслідок хімічної реакції відбулося утворення пари в гібриді, то ма- ємо практично ізоенергетичний процес – дві нуклеотидні пари утво-

рюються замість двох зруйнованих. Аналогічно, пересування полімерази назад буде супроводжуватись руйнуванням пари в гібриді на 3'-кінці РНК (3'-кінцевий нуклеотид при цьому опиняється у вторинному каналі) і пари ДНК позаду від мішура з одночасним утворенням пари ДНК попереду та гібрида позаду.

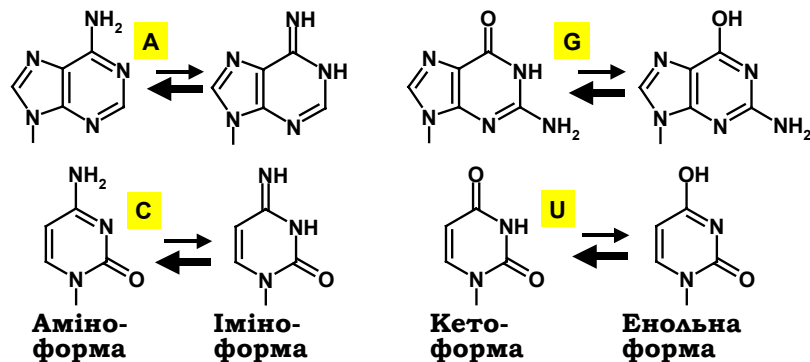
Різниця між кінцевими результатами двох рухів – в енергії взаємодій РНК-полімерази з нуклеїновими кислотами. У вихідному претранс-локаційному стані (положення  $i + 1$ , рис. 5.10) реалізуються певні нуклеопротеїнові взаємодії, що зумовлює знаходження системи в локальному мінімумі вільної енергії. Із двох інших локальних мінімумів, що відповідають положенням  $i$  та  $i + 2$ , другий є більш глибоким, оскільки для нього реалізуються найсприятливіші взаємодії між активним центром та 3'-кінцем РНК. Максимуми енергії (бар'єри) між мінімумами на рис. 5.10 відповідають проміжним станам при пересуванні полімерази (часткове порушення взаємодій, конформаційні зміни полімерази). Ці максимуми є досить невисокими, тобто легко долаються за рахунок теплової енергії. Отже, полімераза спонтанно переміщується вперед, і система опиняється в більш глибокому мінімумі вільної енергії (рис. 5.10), звідки починається наступний елонгаційний цикл.

Якщо *напрямок* переміщення полімерази визначається положенням мінімумів на шкалі вільної енергії, *швидкість* переміщення залежить від висоти максимумів вільної енергії: чим нижчим є максимум, тим легше він долається. Висота максимумів залежить від послідовності пар основ у складі гібрида, а також може змінюватись під дією додаткових факторів (фактори елонгації, шпильки в мРНК тощо).

Під час елонгації здійснюється також **редагування помилок**. Забезпечити стовідсоткову точність операцій з такими невеликими молекулами, як нуклеотиди, практично неможливо. Зокрема, джерелом (одним із основних) помилкового приєднання нуклеотидів під час транскрипції є *таутомерія* азотистих основ. Спонтанні перебудови електронних систем гетероциклів приводять до того, що кожна основа існує у вигляді двох таутомерних форм: аміно- чи іміноформи для А, С; енольної чи кетоформи для G, U, T (рис. 5.11). Рівновага зсунута в бік аміно- та кетоформ, які присутні у складі подвійних спіралей (див. рис. 3.1) і для яких реалізуються правила комплементарності А-Т, G-С. Але спарювання основ підпорядковується іншим правилам для мінорних таутомерних форм: наприклад, іміноформа А та аміноформа С утворюють між собою два водневі зв'язки, що може відбутися під час упізнання матриці нуклеозидтрифосфатом. Тоді після хімічної

реакції, у результаті швидкого повернення до мажорної таутомерної форми, у складі ДНК-РНК гібрида на 3'-кінці транскрипту залишиться некомплементарна пара нуклеотидів.

Нестабільність гібрида на 3'-кінці у випадку, коли щойно приєднаний нуклеотид є помилковим, приводить до зміни енергетичного балансу між різними позиціями РНК-полімерази відносно матриці (рис. 5.10). Найсприятливішою (такою, що відповідає найнижчому мінімуму енергії) тепер є позиція  $i$ , оскільки при цьому гібрид стає стабільним по всій довжині. Відповідно, відбувається зворотний рух полімерази з виходом 3'-кінця транскрипту у вторинний канал. При цьому ніщо більше не фіксує 3'-кінець транскрипту, і всі позиції позаду від  $i$  ( $i, i-1, i-2, \dots$ ) є майже ізоенергетичними (варіації залежать від стабільності гібрида, тобто від послідовності нуклеотидів). Отже, зворотний рух відбувається на невизначену випадкову відстань.



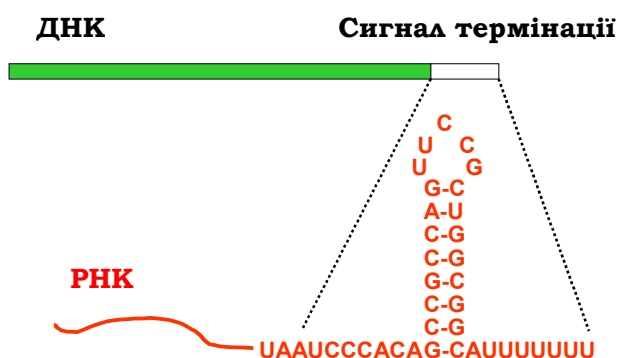
**Рис. 5.11.** Таутомерні форми азотистих основ

Коли 3'-кінцевий фрагмент транскрипту опиняється у вторинному каналі, з полімеразою в тому ж вторинному каналі взаємодіє особливий фактор елонгації транскрипції, який можна назвати фактором редагування, і який існує у двох варіантах: GreA та GreB (еукаріотичним аналогом цих факторів є фактор TFIIIS, див розділ 6). Gre-фактори, через конформаційні зміни РНК-полімерази, індують нуклеазну активність її активного центру. У результаті відбувається відщеплення 3'-кінцевого фрагмента транскрипту і в активному центрі залишається новий 3'-кінець (зі стабільною нуклеотидною парою гібрида), який використовується для відновлення елонгаційного процесу. Загальна точність процесу транскрипції дорівнює приблизно  $10^{-6}$  (частота помилково включених нуклеотидів, що залишаються у складі РНК).

З викладеного видно, що РНК-полімераза є одним із прикладів молекулярної машини (див. розділ 2). *Конформаційна рухливість* полімерази забезпечує їй можливість існувати в кількох структурних станах. Ці стани мають різну спорідненість до певних лігандів (NTP, пари основ у складі РНК-ДНК гібрида, фактори елонгації). *Взаємодії з лігандами* фіксують певні стани. *Хімічні реакції*, які каталізуються машиною, приводять до заміни лігандів, а відповідно – і до переходу в інший структурний стан. Рушійною силою для переміщення полімерази є *тепловий рух*: структурні блоки полімеразного комплексу рухаються відповідно до конструкції машини; зв'язування лігандів та їхня заміна внаслідок реакцій каналізують ці рухи в певних напрямках. Результатом структурних перебудов є переміщення полімерази вздовж матриці з одночасним синтезом транскрипту із середньою швидкістю 40 нуклеотидів/с.

### Термінація транскрипції

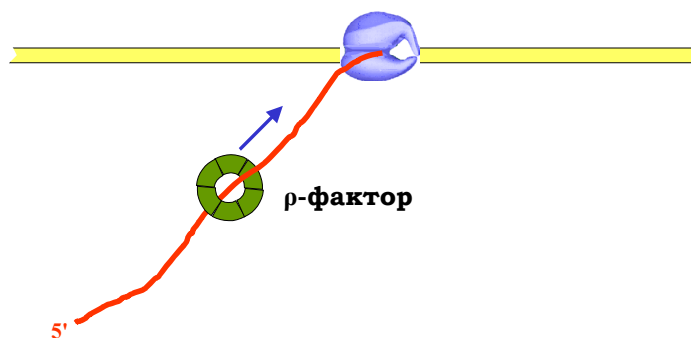
Як уже йшлося, швидкість пересування РНК-полімерази під час елонгації визначається висотою максимумів вільної енергії (див. рис. 5.10). Якщо ці максимуми (по обидва боки від мінімуму, що відповідає позиції  $i + 1$  на рис. 5.10) за певних причин виявляються досить високими, полімераза опиниться в енергетичній пасці й зупиниться на досить довгий час.



**Рис. 5.12.** Типовий сигнал термінації

Саме це й відбувається під час термінації транскрипції на певних елементах послідовності. Прокаріотичний *сигнал термінації* являє собою інвертований повтор (паліндром – послідовність ДНК, що читається однаково в обох ланцюгах у напрямках 5'-3'), безпосередньо фланкований polyT послідовністю (близько 7 нуклеотидів). Відповідно, у матричному ланцюзі розташована polyA послідовність, а у складі транскрипту – polyU. Інвертований повтор у складі транскрипту утворює дволанцюгову шпильку (рис. 5.12), а ДНК-РНК гібрид за шпилькою складається з порівняно менш стабільних А-У пар.

Загальний сценарій термінації виглядає наступним чином. Порівняно нестабільний гібрид ускладнює елонгацію транскрипції, а шпилька – зворотний рух полімерази. Під час зупинки полімерази (приблизно на 60 с) здійснюються зумовлені взаємодіями зі шпилькою структурні перебудови полімерази, руйнування гібрида, відновлення подвійної спіралі ДНК і визволення транскрипту.



**Рис. 5.13.** ρ-Залежна термінація

У кількох бактеріальних оперонах термінація залежить від так званого фактора ρ. Цей фактор (моногексамер) зв'язується з РНК-транскриптом у певних С-збагачених ділянках. Після цього починається АТР-залежне пересування ρ-фактора вздовж транскрипту в напрямку 5'-3' зі швидкістю, яка є нижчою за швидкість руху полімерази. Сигнал термінації в цьому випадку, як правило, нічим не відрізняється від описаних вище термінаторів і зумовлює зупинку полімерази. У результаті ρ-фактор дожинає полімеразу і, продовжуючи рух уздовж РНК, руйнує гібридну подвійну спіраль (працює як *геліказа* (helicase) – так називаються ферменти, що АТР-залежно руйнують подвійні спіралі нуклеїнових кислот). Результатом є колапс транскрипційного міхура (відновлення подвійної спіралі ДНК) і визволення транскрипту.



## Регуляція транскрипції

Зрозуміло, що гени та оперони (див. розділ 4) не транскрибуються постійно, а вмикаються / вимикаються в певні моменти залежно від зовнішніх умов, стадій клітинного циклу тощо. Головними елементами, взаємодія між якими приводить до активації чи репресії транскрипції, є так звані *цис-* і *транс-елементи*. *Цис*-елементи – це регуляторні елементи послідовності ДНК, які фізично зв'язані з даним геном чи опероном; у прокаріотів часто називаються *операторами* і розташовані в безпосередній близькості до промоторів. *Транс*-елементи – білкові фактори транскрипції, що вільно дифундують (*транспортуються*) у просторі клітини, шукаючи свій *цис*-елемент, з яким вони мають специфічну спорідненість. Якщо зв'язування *транс*-елемента з оператором приводить до активації транскрипції (часто за рахунок прямих білок-білкових взаємодій транскрипційного фактора з РНК-полімеразою, які підвищують її спорідненість до промотора), кажуть, що фактор є *активатором* і здійснює *позитивну* регуляцію. Якщо фактор блокує зв'язування РНК-полімерази (часто за рахунок зниження доступності промотора), його називають *репресором* і кажуть про *негативну* регуляцію.

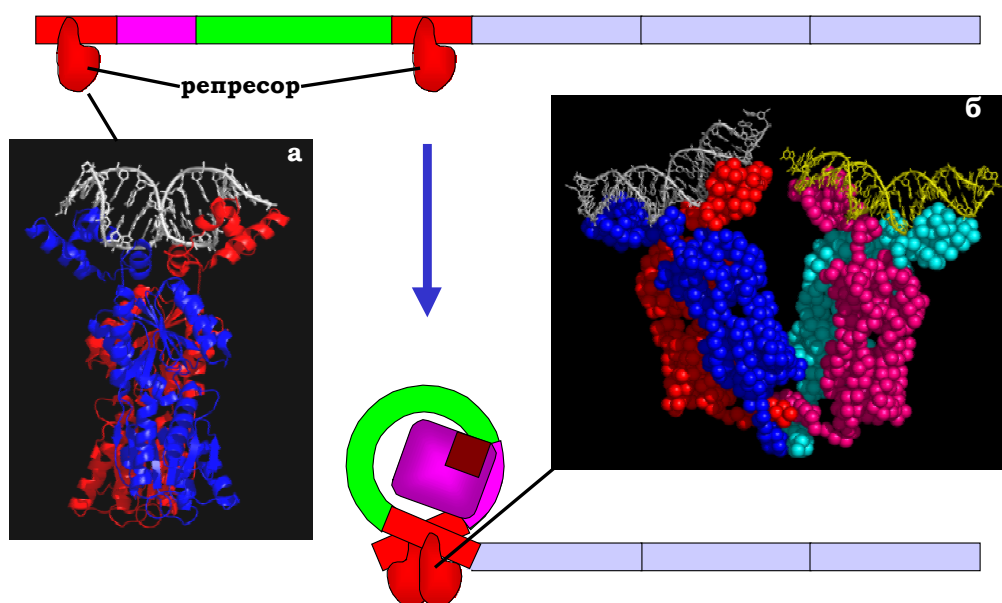
Ці загальні принципи регуляції, які, ускладнюючись, зберігаються також в еукаріотів (розділ 6), реалізуються на стадії ініціації. Крім того, для регуляції використовуються інші моменти процесу транскрипції. Найтипівіші механізми регуляції в прокаріотичних системах проілюстровано нижче на конкретних прикладах.

### *Лактозний оперон*

Лактозний оперон (*lac*-оперон) *E. coli* став у свій час, завдяки дослідженням Жакоба і Моно (François Jacob, Jacques Monod), першою детально вивченою системою регуляції транскрипції. До складу оперона (рис. 5.14) входять три структурні гени, що кодують ферменти, залучені до утилізації (катаболізму) лактози. Транскрипція всіх трьох генів здійснюється з одного промотора (синтезується єдина, так звана *поліцистронна*, молекула мРНК, яка має три послідовні відкриті рамки зчитування). Промотор оточують дві однакові операторні ділянки (*lac*-оператори), що мають спорідненість до *lac*-репресора, і сайт зв'язування CAP (Catabolite Activator Protein).



розташований промотор, і це абсолютно запобігає зв'язуванню з ним РНК-полімерази. Коли з'являється лактоза, її невелика кількість перетворюється на алолактозу, яка спрацьовує як індуктор *lac*-оперона: зв'язування алолактози з репресором індукує втрату його спорідненості до оператора. Унаслідок руйнування петлі РНК-полімераза зв'язується з промотором і оперон починає працювати.

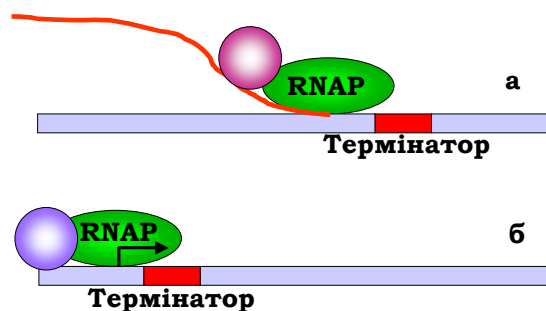


**Рис. 5.15.** Негативна регуляція *lac*-оперона *lac*-репресором.  
На вставках: структура димеру репресора в комплексі з оператором (а, 1JWL)  
і тетрамерний комплекс із двома операторами (б, 1LBG)

### Антитермінація

У системах антитермінації (кілька генів бактеріофага  $\lambda$ , гени рибосомної РНК *E. coli*) активатори транскрипції запобігають упізнанню РНК-полімеразою сигналів термінації, що знаходяться всередині кодуєчої частини гена. У відсутності факторів ген є неактивним: наявність термінуючого сигналу приводить до термінації транскрипції та визволення нефункціонального РНК-продукту. Для певних генів описано два механізми активації. У першому випадку білок-активатор упізнає специфічний сайт перед сигналом термінації у складі мРНК.

Зв'язуючись із транскриптом, білок взаємодіє також із РНК-полімеразою і впливає на неї таким чином, що термінатор стає непомітним для ферменту (рис. 5.16, а). У другому випадку транскрипційний фактор упізнає регуляторну ділянку поряд із промотором, взаємодіє з полімеразою і запобігає впізнанню термінатора, що міститься відразу за стартом транскрипції (рис. 5.16, б).



**Рис. 5.16.** Два механізми антитермінації: фактор транскрипції запобігає впізнанню термінатора РНК-полімеразою

### **Атенюація**

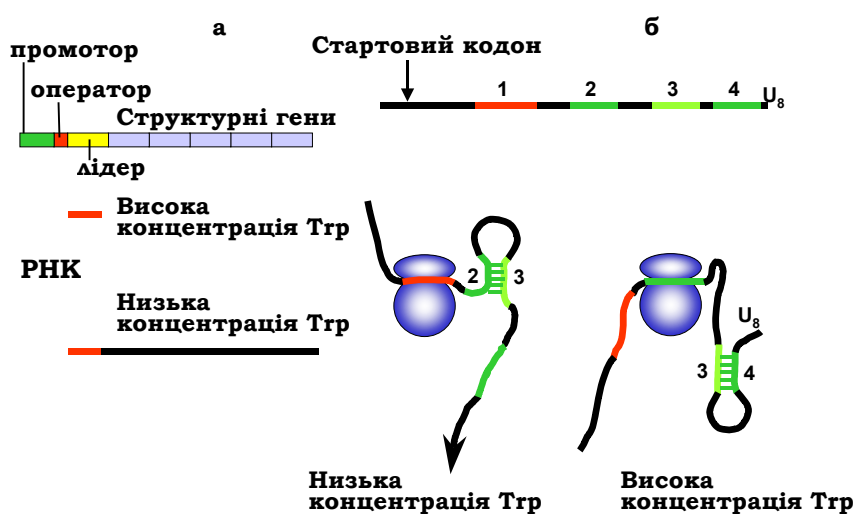
Із використанням сигналів термінації пов'язана також система *атенюації* (attenuation – послаблення), яка використовується, зокрема, для регуляції активності триптофанового оперона (*trp*-оперона) *E. coli*. Оперон містить 5 структурних генів, що відповідають за синтез амінокислоти Трп, перед ними знаходяться промотор, оператор і лідерна послідовність, з якої розпочинається транскрипція (рис. 5.17, а).

Атенюація використовує ту обставину, що прокаріотична транскрипція тісно пов'язана з трансляцією. Лідерна частина РНК містить стартовий кодон, що впізнається рибосомою, і чотири елементи послідовності: ділянка 1 містить два сусідні триптофанові кодони, ділянки 2–3 і 3–4 є попарно комплементарними, за ділянкою 4 розташована оліго-U послідовність (рис. 5.17, б). Отже, шпилька 3–4, фланкована оліго-U, є сигналом термінації транскрипції. Коли концентрація Трп є низькою (є потреба в Трп і оперон має бути активним), рибосома зупиняється на триптофанових кодонах ділянки 1 (оскільки відсутня й триптофаніл-тРНК). У цьому випадку утворюється шпилька 2–3 (ділянка 3 не залучається до утворення термінуючої шпильки), і РНК полімераза продовжує синтез повноцінної мРНК.

## Розділ 5. Транскрипція: Прокаріоти

Рибосоми зв'язуються зі стартовими кодонами, що відповідають структурним генам, і синтезуються відповідні білки.

У разі високого рівня Trp рибосома швидко проходить через ділянку 1 на ділянку 2 і зупиняється на стоп-кодоні. У результаті утворюється шпилька 3–4 – формується сигнал термінації, РНК-полімераза зупиняє транскрипцію після синтезу короткої нефункціональної лідерної РНК.

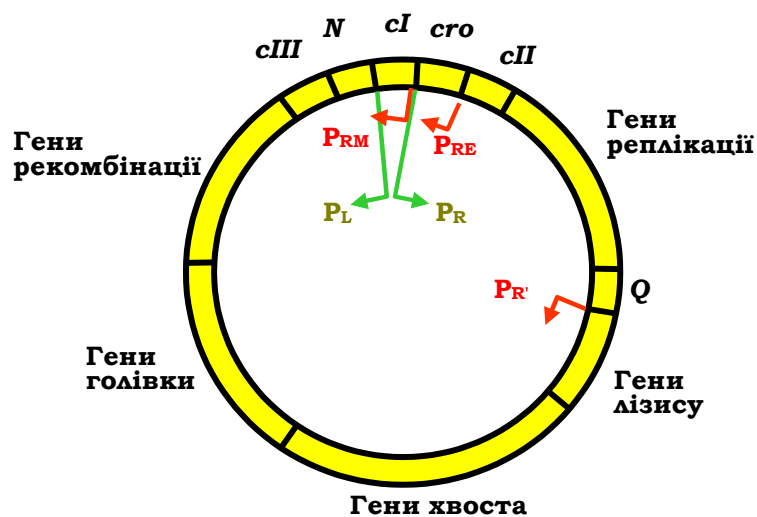


**Рис. 5.17.** (а): Схема *trp*-оперона, на якому синтезуються два РНК-продукти залежно від внутрішньоклітинної концентрації Trp. (б): Лідерна РНК і два варіанти спарювання основ у її складі залежно від розташування рибосоми

*trp*-Оперон перебуває також під контролем *trp*-репресора. Аlostеричним регулятором репресора є сам Trp: у комплексі з ним репресор набуває конформаційної форми, що має високу спорідненість до оператора. При зниженні концентрації Trp репресор дисоціює й ефективність ініціації транскрипції підвищується в ~70 разів. Атенюація є додатковим, менш ефективним механізмом регуляції: ефективність транскрипції підвищується в ~10 разів у відсутності Trp за рахунок атенюації (у присутності Trp ~10 % РНК-полімераз долають сигнал термінації й продовжують працювати, у разі відсутності Trp – практично всі). Сумісна дія атенюації та негативного контролю за рахунок репресора дозволяє змінювати активність оперона в ~700 разів залежно від внутрішньоклітинної концентрації Trp.

### Регуляція транскрипції бактеріофага $\lambda$

Після проникнення в бактеріальну клітину лінійна ДНК бактеріофага  $\lambda$  (довжиною приблизно 50 тис. пар основ) замикається в кільце бактеріальною лігазою (ligase – АТР-залежний фермент, що каталізує утворення фосфодієфірного зв'язку між 5'- та 3'-кінцевими нуклеотидами). Генوم фага містить гени, що відповідають за синтез білків головки та хвоста фагової частинки, реплікацію фагової ДНК, лізис бактерії, рекомбінацію (вбудовування фагової ДНК у бактеріальний генوم), і кілька регуляторних генів, що кодують фактори транскрипції (рис. 5.18).



**Рис. 5.18.** Циркулярна ДНК бактеріофага  $\lambda$ .

Стрілочками позначено промотори та напрямок транскрипції: зелені – сильні промотори, червоні – такі, що потребують активації

Існують два шляхи розвитку фага: *лізогенія* – лінеаризація фагової ДНК, вбудовування її в бактеріальний генوم і блокування більшості генів фага; *лізис* (після інфекції або шляхом вирізання фагової ДНК із бактеріального геному та її циркуляризації після лізогенії) – активація реплікації фагової ДНК і синтезу білків оболонки, збирання фагових частинок та руйнування клітини.

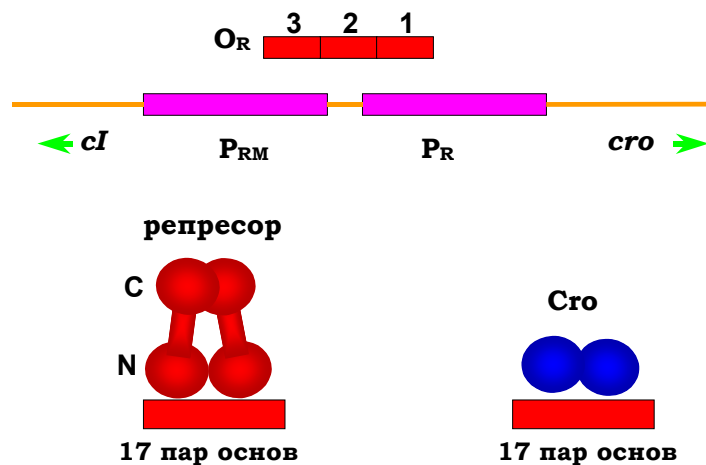
Ключову роль у виборі одного з двох шляхів виконують два фактори транскрипції: білок  $\sigma_{70}$  і  $\lambda$ -репресор – продукти відповідно генів  $\sigma_{70}$  і  $cI$ , що містяться поруч у фаговому геномі (рис. 5.18, 5.19). Відразу після інфекції РНК-полімераза зв'язується з двома сильними промоторами  $P_R$  і  $P_L$  (рис. 5.18), з яких відбувається транскрипція у протилежних напрямках на генах  $\sigma_{70}$  і  $N$  відповідно. Обидва гени закінчуються термінаторами, але щойно з'являється білок  $N$ , він зв'язується з транскриптами, рекрутує декілька бактеріальних білків, і цей комплекс запобігає впізнанню термінаторів (за механізмом антiterмінації, див. рис. 5.16, а). РНК-полімераза в цьому випадку продовжує синтезувати поліцистронну мРНК на  $cIII$  і генах рекомбінації (із промотора  $P_L$ ) і  $cII$ , генах реплікації та  $Q$  (із промотора  $P_R$ ). Після цього можливе розгалуження на два альтернативні шляхи (рис. 5.20).

Розглянемо спочатку шлях лізогенії. Продукт гена  $cII$  – активатор транскрипції, який забезпечує зв'язування РНК-полімерази зі слабким промотором  $P_{RE}$ . Із цього промотора (у напрямку, протилежному напрямку транскрипції  $\sigma_{70}$ , рис. 5.18) транскрибується ген  $cI$ , унаслідок чого з'являється  $\lambda$ -репресор. Крім того, білок  $cII$  активує ген інтегрази (міститься серед генів рекомбінації) – ферменту, що забезпечує вбудовування фагової ДНК у геном клітини-хазяїна (див. розділ 10). Продукт гена  $cIII$  захищає білок  $cII$  від бактеріальних протеаз, тобто підвищує час життя активатора. Отже, за умови високої концентрації  $cII$  виникає певна кількість  $\lambda$ -репресора, а фагова ДНК вбудовується інтегразою в бактеріальний геном.

Репресор (гомодимер, N-кінцеві частини обох субодиниць взаємодіють з ДНК (див. рис. 3.12, 3.13, а)) має спорідненість до двох наборів операторів. Один із цих наборів – оператори  $O_R 1,2,3$ , кожен довжиною 17 пар основ – частково перекриває промотори  $P_R$  і  $P_{RM}$  (рис. 5.19). Промотор  $P_{RM}$  є слабким і може використовуватись для транскрипції гена  $cI$  тільки за умови активації. Послідовності операторів є гомологічними одна одній і розрізняються за спорідненістю до репресора – найвищу спорідненість має ділянка  $O_R 1$ .

При заповненні цього оператора репресором промотор  $P_R$  гена  $\sigma_{70}$  та всіх генів праворуч від нього є заблокованим. С-кінцеві частини двох димерів репресора здатні до взаємодії: за наявності репресора в сайті  $O_R 1$  швидко заповнюється також сайт  $O_R 2$ . Інакше кажучи, репресор кооперативно взаємодіє з ділянками  $O_R 1,2$ : його зв'язування з ДНК підсилюється білок-білковими взаємодіями димерів. При заповненні оператора  $O_R 2$  репресор (який є негативним регулятором

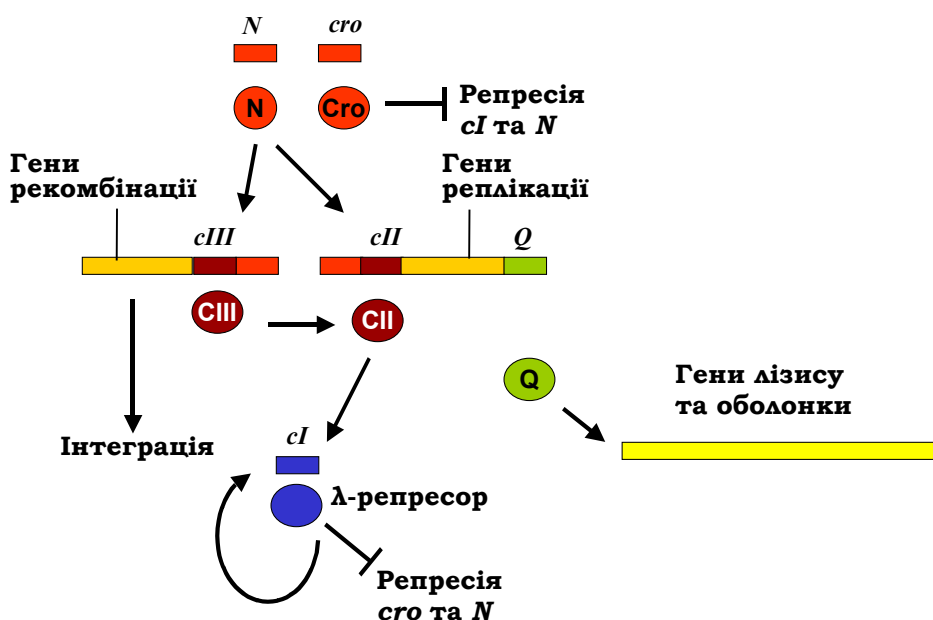
гена *cro*) спрацьовує як позитивний регулятор свого власного гена *cI* – індукує зв'язування РНК-полімерази з промотором  $P_{RM}$ . Спорідненість репресора до оператора  $O_R$  3 є найнижчою: ця ділянка некооперативно заповнюється при зростанні концентрації репресора, унаслідок чого відбувається блокування гена *cI*. Таким чином активність гена *cI* і концентрація репресора підтримуються на оптимальному рівні за принципом зворотного зв'язку. На стадії лізогенії тільки ген *cI* є активним – репресор блокує також промотор  $P_L$  (через відповідні оператори  $O_L$ ), тобто транскрипцію всіх генів ліворуч від *N*.



**Рис. 5.19.** Зона контакту між генами *cI* і *cro* фага  $\lambda$ : промотори генів частково перекривають три оператори  $O_R$  1,2,3, кожен з яких має спорідненість до репресора та білка Cro

Якщо бактерія піддається дії мутагенів, коли виникає ризик загибелі клітини (разом із вбудованою фаговою ДНК), відбувається активація певної бактеріальної протеази, яка розрізає молекулу репресора між її N- та С-кінцевими доменами. Результатом є порушення димеризації мономерів репресора і втрата спорідненості до ДНК. Унаслідок звільнення операторів  $O_R$  та  $O_L$  РНК-полімераза зв'язується з промоторами  $P_R$  та  $P_L$  і починає транскрибувати гени *cro* та *N*. Тобто знову виникає ситуація, що спостерігається відразу після інфекції – спрацьовують гени реплікації, *cII* і *Q*. Але при цьому білок *cII* швидко деградується протеазами, не встигаючи активувати синтез репресора. Така сама ситуація (висока активність протеаз) може реалізуватися відразу після інфекції за умови збагачення середовища на харчові ресурси.





**Рис. 5.20.** Мережа регуляторних взаємодій між генами фага  $\lambda$ .  
Прямокутники – транскрипти, отримані з відповідних генів,  
овали – їхні білкові продукти

Білок *Cro* (також гомодимер) некооперативно взаємодіє з операторами  $O_R 1,2,3$ , спорідненість до яких знижується в порядку 3–2–1. Заповнення  $O_R 3$  білком *Cro* остаточно вимикає синтез репресора з промотора  $P_{RM}$ . Подальше заповнення операторів при зростанні концентрації *Cro* вимикає транскрипцію *cro* і генів праворуч від нього, так само як і генів ліворуч від *N* за рахунок зв'язування *Cro* з операторами  $O_L$ . Але головний наслідок активності *cro* полягає в появі білка *Q*.

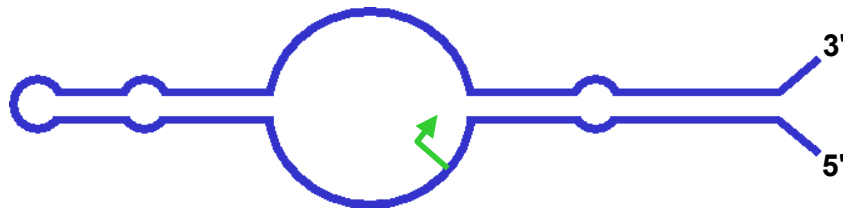
Ще один промотор  $P_R$  є насправді сильним, але відразу за ним розташований термінатор – РНК-полімераза за відсутності *Q* синтезує з цього промотора короткий нефункціональний транскрипт. Білок *Q* спрацьовує як антитермінатор (за механізмом, зображеним на рис. 5.16, б), забезпечуючи додання цього бар'єра. У результаті відбувається синтез поліцистронної мРНК на генах лізису та білків оболонки (праворуч від *Q*), і відбувається лізис клітини з виходом фагових частинок.

Розглянута досить складна система регуляції транскрипції порівняно простого геному бактеріофага  $\lambda$  дає уявлення про те, наскільки ускладнюється загальна система регуляції транскрипції у бактерій

і тим більше – в еукаріотів. Насправді ж складність загальних систем внутрішньоклітинної регуляції, які залежать від тонкого балансу великої кількості різноманітних впливів, залишається дуже далекою від остаточного розуміння.

### *Неспецифічна регуляція загального рівня транскрипції*

Поряд із специфічною активацією / репресією окремих генів, у бактеріальній клітині здійснюється неспецифічна регуляція загального рівня транскрипційної активності. Принаймні два загальні механізми використовуються для такої регуляції. Перший пов'язаний з утворенням відкритого комплексу при ініціації транскрипції – необхідністю локального плавлення подвійної спіралі, яке, у свою чергу, полегшується негативною надспіралізацією ДНК (розділ 3). У результаті активність багатьох бактеріальних промоторів суттєво залежить від рівня негативною надспіралізації, який варіює в різних фазах клітинного циклу й визначається концентраціями бактеріальних топоізомераз.



**Рис. 5.21.** Схема вторинної структури 6S РНК, яка може використовуватись як матриця для РНК-полімерази. Стрілочкою позначено початок синтезу

Другий загальний механізм регуляції використовує особливу молекулу РНК – так звану 6S РНК довжиною 184 нуклеотиди. Молекула РНК формує велику двоспіральну шпильку з кількома неспареними ділянками (рис. 5.21). Одна з неспарених ділянок містить близько 15 пар основ і нагадує транскрипційний міхур (порівн. з рис. 5.6) – саме вона сприймається РНК-полімеразою як промотор. Відбувається ефективне зв'язування, і якщо внаслідок пригнічення біохімічних синтезів (недостатності ресурсів) концентрація NTP є низькою, полімераза виявляється “заарештованою” і нездатною ініціювати транскрипцію на ДНК.

Для того, щоб звільнитися від 6S РНК, полімераза має розпочати “транскрипцію”, на цій РНК-матриці. Саме це й відбувається при підвищенні концентрації NTP (ДНК-залежна РНК-полімераза працює як

*РНК-залежна*): здійснюється елонгація транскрипції з руйнуванням подвійної спіралі РНК, синтез короткого нефункціонального транскрипту та звільнення ферменту, який далі шукає справжній промотор.

### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. Як за своїм складом і спорідненістю до ДНК розрізняються кор- і голофермент бактеріальної РНК-полімерази? Яке значення має ця різниця для ініціації транскрипції?
2. Яку будову має стандартний бактеріальний промотор?
3. Чим розрізняються закритий і відкритий комплекси РНК-полімерази з промотором?
4. У якому напрямку здійснюється зчитування інформації з ДНК під час транскрипції? Що таке змістовний і антизмістовний ланцюги? Який з них є матричним?
5. Яку роль відіграє  $\sigma$ -фактор в ініціації транскрипції?
6. Як забезпечується висока процесивність РНК-полімерази?
7. Як організовано активний центр РНК-полімерази? Які сполуки є субстратами (будівним матеріалом) синтезу РНК?
8. Опишіть основні етапи елонгаційного циклу РНК-полімерази. Чому фермент рухається вздовж ДНК під час транскрипції?
9. Яким чином здійснюється при транскрипції редагування помилок приєднання нуклеотидів?
10. Як здійснюється термінація транскрипції у бактерій?
11. Яка різниця між *цис*- і *транс*-елементами системи регуляції транскрипції?
12. Опишіть систему регуляції лактозного оперона. Яку роль у регуляції відіграють *lac*-репресор і білок CAP?
13. Що таке антитермінація?
14. Як здійснюється атенуація в триптофановому опероні?
15. За якими механізмами здійснюється регуляція загального рівня транскрипції в бактеріальній клітині?
16. Охарактеризуйте стадії життєвого циклу бактеріофага  $\lambda$ . Як відбувається перемикання між різними шляхами розвитку?
17. Репресор бактеріофага  $\lambda$  є позитивним чи негативним регулятором транскрипції гена *cro*? Гена *cI*?

**РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

*Пташине, М.* Переключение генов. Регуляция генной активности и фаг λ. – М.: Мир, 1988.

*Busby, S., Ebright, R.H.* Promoter structure, promoter recognition, and transcription activation in prokaryotes // *Cell*. – 1994. – Vol. 79. – P. 743–746.

*Ciampi, M.S.* Rho-dependent terminators and transcription termination // *Microbiology*. – 2007. – Vol. 152. – P. 2515–2528.

*Greive, S.J., von Hippel, P.H.* Thinking quantitatively about transcriptional regulation // *Nature Rev.* – 2005. – Vol. 6. – P. 221–232.

*Lane, W.J., Darst, S.A.* The structural basis for promoter – 35 element recognition by the group IV σ factors // *PLoS Biol.* – 2006. – Vol. 4(9). – P. e269 ([www.plosbiology.org](http://www.plosbiology.org)).

*Steitz, T.A.* Visualizing polynucleotide polymerase machines at work // *EMBO J.* – 2006. – Vol. 25. – P. 3458–3468.

*Wassarman, K.M., Saecker, R.M.* Synthesis-mediated release of a small RNA inhibitor of RNA polymerase // *Science*. – 2006. – Vol. 314. – P. 1601–1603.

*Willenbrock, H., Ussery, D.W.* Chromatin architecture and gene expression in *Escherichia coli*. // *Genome Biology*. – 2004. – Vol. 5. – P. 252 (<http://genomebiology.com/2004/5/12/252>).

*Wilson, C.J., Zhan, H., Swint-Kruse, L., Matthews, K.S.* The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding // *Cell Mol. Life Sci.* – 2007. – Vol. 64. – P. 3–16.

*Young, B.A., Gruber, T.M., Gross, C.A.* Views of transcription initiation // *Cell*. – 2002. – Vol. 109. – P. 417–420.

## Розділ 6

### ТРАНСКРИПЦІЯ: ЕУКАРІОТИ

С тех пор оставили его навсегда переписывать. Вне этого переписыванья, казалось, для него ничего не существовало.

*Н. Гоголь. Шинель*

В еукаріотичних клітинах функціонують РНК-полімерази трьох типів:

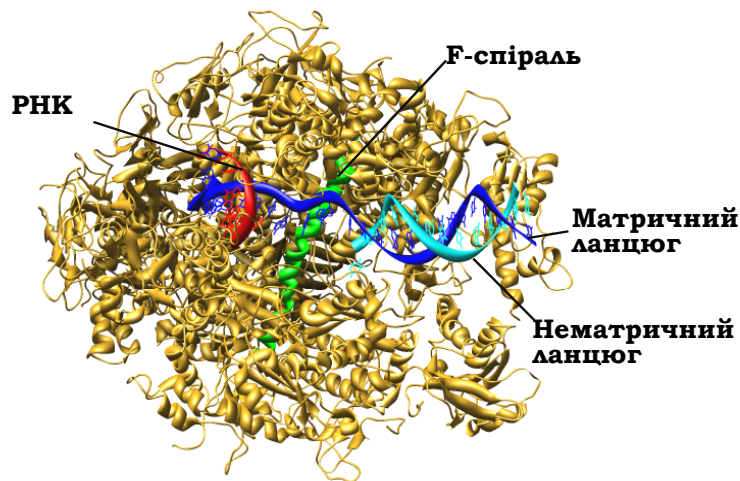
- **РНК-полімераза I** працює на кластерах генів рибосомної РНК (розділ 4) і здійснює синтез рРНК 18S, 28S та 5,8S.
- **РНК-полімераза II** транскрибує білкові гени, а також гени маленьких ядерних РНК та інших РНК, що не транслюються.
- **РНК-полімераза III** здійснює синтез тРНК, рибосомної РНК 5S і кількох інших низькомолекулярних РНК.

Кожна з цих полімераз містить чотири гомологічні корові субодиниці, які є водночас гомологами субодиниць  $\alpha$ ,  $\beta$  і  $\beta'$  прокаріотичної полімерази (розділ 5). Крім того, до складу полімераз входять п'ять спільних для всіх трьох ферментів субодиниць, а також певний набір специфічних субодиниць (у кількостях 5, 3 і 7 для РНК-полімераз I, II та III відповідно). Загальна архітектура еукаріотичних полімераз (наявність щелеп, між якими зв'язується ДНК, каналу виходу РНК, вторинного каналу для входу нуклеозидтрифосфатів) дуже схожа на таку прокаріотичної полімерази. Між про- та еукаріотичними полімеразами спостерігається також висока гомологія внутрішньої поверхні щілини щелеп та активного центру; спільними є й основні механізми роботи полімераз, розглянуті у попередньому розділі. Проте відсутня гомологія зовні: додаткові до чотирьох корових субодиниць еукаріотичних полімераз створюють специфічну поверхню для взаємодії з елементами еукаріотичної системи транскрипції.

## РНК-полімераза II

### Структура РНК-полімерази II

РНК-полімераза II, яка відповідає за синтез мРНК, складається з 12 субодиниць (рис. 6.1). Найбільша субодиниця (гомолог субодиниці  $\beta'$  прокариотичної полімерази) формує, зокрема, активний центр, структура і найближче оточення якого є практично ідентичними таким бактеріальної полімерази (розділ 5).



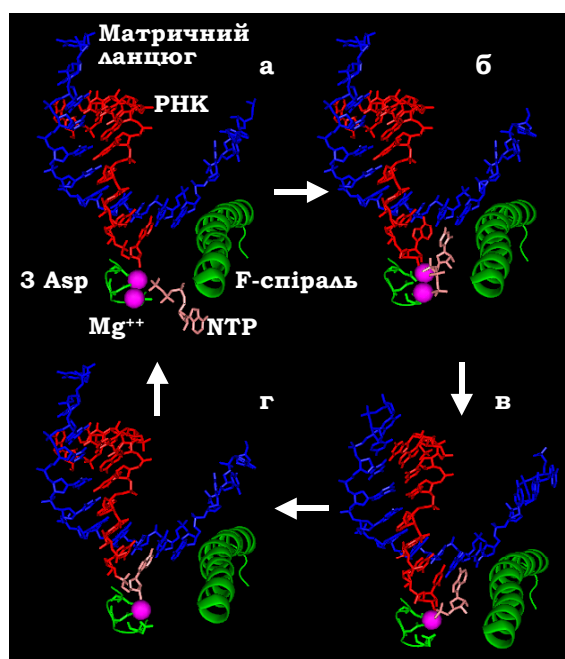
**Рис. 6.1.** Структура елонгаційного комплексу РНК-полімерази II (1R9T)

Відповідно, ідентичними є й механізми елонгації синтезу РНК. На рис. 6.2 показано структуру оточення активного центру на різних стадіях елонгаційного циклу (порівн. з рис. 5.3, 5.10): вхід NTP через вторинний канал; зв'язування NTP з матрицею та його фіксація в активному центрі завдяки двом іонам  $Mg^{2+}$ , які утримуються трьома залишками Asp; каталіз реакції приєднання нуклеотиду до 3'-кінця транскрипту; транслокація полімерази на один нуклеотид уздовж матриці, унаслідок чого система здатна вступити в новий цикл елонгації.

Під час елонгації, так само як і для бактеріальної полімерази, за умови нестабільності РНК-ДНК гібрида на 3'-кінці можливий зворотний рух полімерази з виходом 3'-кінця транскрипту у вторинний канал. У цьому випадку взаємодія з фактором транскрипції **TFIIS**

(Transcription Factor S of RNA-polymerase II, аналог бактеріальних факторів GreA/B) індукує нуклеазну активність, яка забезпечує редагування помилок при транскрипції (див. розділ 5).

Особливістю РНК-полімерази II (характерною тільки для цієї полімерази) є наявність у найбільшій субодиниці С-кінцевого домену – **CTD** (C-Terminal Domain). CTD – це довгий неупорядкований хвіст (не показаний на рис. 6.1), амінокислотна послідовність якого є гептапептидом Tyr-Ser-Pro-Thr-Ser-Pro-Ser, що тандемно повторюється 52 рази. Три залишки Ser у складі гептапептиду є субстратами фосфорилування / дефосфорилування для специфічних кіназ і фосфатаз. CTD, який відходить від полімерази поблизу від каналу виходу мРНК, є платформою для зв'язування численних білків, спорідненість яких залежить від патерна фосфорилування. Зокрема, CTD відіграє ключову роль у перемиканні між ініціацією та елонгацією транскрипції (див. нижче) та у збиранні елементів системи процесингу мРНК (розділ 7) під час елонгації.

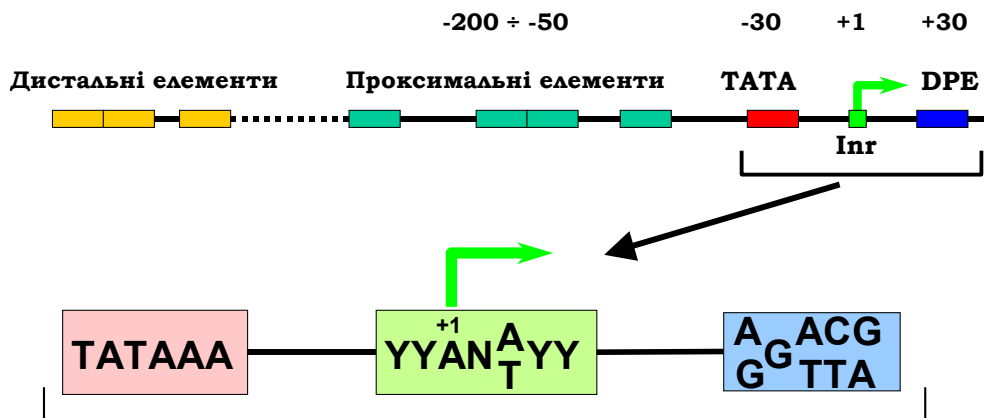


**Рис. 6.2.** Стадії елонгаційного циклу РНК-полімерази II:  
 (а) – вхід NTP до активного центру (1R9T); (б) – впізнання матриці та зв'язування NTP (1R9S); (в) – претранслокаційний стан після приєднання нуклеотиду до транскрипту (1I6H); (г) – посттранслокаційний стан (1SFO)

Для ініціації транскрипції є необхідним збирання на промоторі **преініціаторного комплексу** (PIC – Pre-Initiation Complex) за участю принаймні 12-субодиничної РНК-полімерази II та шести базальних (загальних) факторів транскрипції TFIIA, TFIIB, TFIID, TFIIЕ, TFIIF, TFIIN.

### Промотор РНК-полімерази II

Промотор РНК-полімерази II, узагальнену схему якого зображено на рис. 6.3, може складатися з наступних елементів послідовності (для конкретних промоторів спостерігаються численні варіації цієї схеми). **Проксимальні** (приблизно в зоні  $-50 \dots -200$  пар основ відносно старту транскрипції) та **дистальні** (будь-де відносно старту) регуляторні елементи мають спорідненість до специфічних факторів транскрипції, взаємодія з якими активує / блокує збирання преініціаторного комплексу. Якщо дистальний елемент підсилює ефективність ініціації, його називають *енхансером* (enhancer – підсилювач), якщо навпаки – *сайленсером* (silencer – глушник). У зоні старту знаходиться так званий **базальний промотор**, на якому, власне, і відбувається збирання PIC.



**Рис. 6.3.** Приблизна узагальнена схема організації промотора РНК-полімерази II і консенсусні послідовності основних елементів базального промотора

Базальний промотор може включати (не обов'язково) три елементи: TATA-бокс (починається приблизно з  $-30$ -ї пари основ), ініціаторний елемент (Initiator, Inr) безпосередньо в зоні старту, DPE (Downstream Promoter Element) на відстані приблизно  $+30$  пар основ від старту.



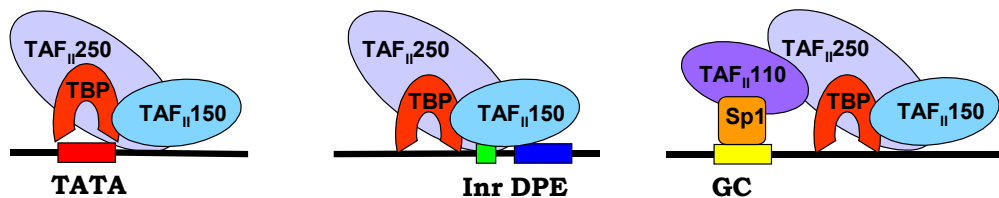
Останні два елементи, Inr та DPE, як правило, присутні (якщо присутні) одночасно. Наприклад, приблизно 14 % промоторів дрозофіли містять усі три елементи, 29 % – тільки ТАТА-бокс, 26 % – тільки DPE разом з Inr, 31 % не містять жодного з трьох елементів. У останньому випадку промотор визначається іншими, більш специфічними елементами послідовності. Зокрема, досить часто зустрічається так званий GC-бокс, що розташований, як правило, трохи вище (upstream) відносно положення ТАТА-боксу.

### *Базальні фактори транскрипції*

Послідовність збирання преініціаторного комплексу може бути різною, але всі базальні фактори мають бути присутніми у складі PIC для подальшого запуску транскрипції.

**TFIID** – один із найважливіших базальних факторів, який визначає впізнання стандартних елементів базального промотора (якщо вони присутні). Основою структури TFIID є **ТВР** (TATA-box Binding Protein) – білок зі специфічною спорідненістю до ТАТА-боксу. Досить широкий β-шар у складі ТВР взаємодіє з маленьким жолобком ДНК, наслідком чого є суттєва деформація подвійної спіралі з її розкручуванням та значним вигином у протилежний від білка бік (див. рис. 3.18). Саме з цього протилежного боку (усередині вигину) розташовуються інші білкові компоненти преініціаторного комплексу, які обгортаються промотором (див. нижче). Отже, вигин ДНК, який індукується ТВР, організує промотор і зв'язані з ним елементи в єдиний компактний комплекс і підтримує його стабільність. Досить часто (але не обов'язково) зв'язування TFIID ініціює збирання PIC.

Із ТВР у складі TFIID зв'язані так звані ТВР-асоційовані фактори (**TAFs** – TBP Associated Factors) – вісім або більше білків (конкретний склад факторів варіює для різних промоторів). TAFs взаємодіють з ТВР і між собою, деякі з них здатні також взаємодіяти з ДНК. Зокрема, TAF<sub>II</sub>150 (або його аналог, 150 – молекулярна вага в кілодальтонах) упізнає DPE (рис. 6.4), тобто забезпечує зв'язування TFIID із базальним промотором, у складі якого відсутній ТАТА-бокс. Якщо у промоторі немає також і DPE, інші TAFs здійснюють білок-білкові взаємодії з факторами транскрипції, що зв'язані з іншими елементами послідовності. Наприклад, GC-бокс упізнається транскрипційним фактором Sp1 (Specificity protein 1), який, у свою чергу, рекрутує TAF<sub>II</sub>110, що стає компонентом TFIID через взаємодію з TAF<sub>II</sub>250 (рис. 6.4).



**Рис. 6.4.** Взаємодія TFIID з базальним промотором через TBP (якщо присутній TATA-бокс), через один з TAFs (якщо присутній DPE), через інший TAF, що взаємодіє з транскрипційним фактором, зв'язаним із GC-боксом

Білок TAF<sub>II</sub>250 є практично обов'язковим компонентом TFIID (як і TBP), виконуючи роль основного фактора збирання, з яким взаємодіють інші TAFs. Крім того, TAF<sub>II</sub>250 містить у своєму складі бромодомен і гістон-ацетилтрансферазну активність (розділ 4): відповідно, здатен як впізнавати ацетильовані гістони в зоні промоторів, так і здійснювати таке ацетилювання.

**TFIIA** – гетеродимерний білок, кофактор, що стабілізує комплекс TBP з ДНК. Відповідно, цей фактор можна було б вважати ще одним TAF. У системах *in vitro* TFIIA не є необхідним для ініціації транскрипції, проте його відсутність у клітинах дріжджів є летальною.

**TFIIB** – мономерний білок із мультидоменною структурою. Два С-кінцеві домени взаємодіють із ДНК по обидва боки від сайту зв'язування TBP, із внутрішнього боку вигину, який індукується цим білком (рис. 6.5). Таким чином, два білки, які взаємодіють також між собою, зв'язуються з ДНК кооперативно, взаємно підсилюючи спорідненість до неї.

Два С-кінцеві домени TFIIB мають різну структуру і взаємодіють з різними елементами ДНК (маленьким та великим жолобками). Така асиметрія (існування двох альтернативних варіантів зв'язування з ДНК) відіграє роль у визначенні напрямку транскрипції. Це пов'язано з тим, що N-кінцева частина TFIIB (практично не має регулярної вторинної структури) взаємодіє з РНК-полімеразою II. Таким чином, саме TFIIB визначає положення полімерази на ДНК відносно стартової точки в одній із двох можливих орієнтацій ферменту.

Витягнута N-кінцева частина TFIIB взаємодіє з РНК-полімеразою в зоні активного центру, після початку транскрипції здійснює взаємодію з РНК-ДНК гібридом, а також блокує канал виходу РНК. Тобто TFIIB

певною мірою виконує функції, аналогічні до функцій бактеріального  $\sigma$ -фактора (розділ 5). Відповідно, TFIIB має бути звільненим при переході від ініціації до елонгації транскрипції (див. нижче).

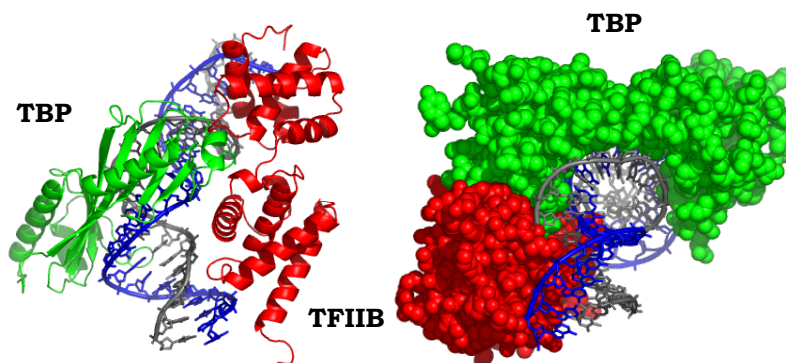


Рис. 6.5. Комплекс із ДНК ДНК-зв'язувального домену TBP і С-кінцевої частини TFIIB у двох проекціях (1D3U)

**TFIIF** – гетеродимер, взаємодіє з РНК-полімеразою II (у зоні активного центру), з ДНК, з базальними факторами TFIIE та TFIIH. Після первинного плавлення ДНК при ініціації транскрипції TFIIF здійснює взаємодії з нематричним ланцюгом, стимулює роботу TFIIH (див. нижче), запобігає абортивній ініціації. Після завершення ініціації TFIIF разом з іншими базальними факторами дисоціює від РНК-полімерази, але може реасоціювати з нею під час елонгації – у випадку зупинки ферменту. Його роль як фактора елонгації транскрипції полягає у стимуляції продовження руху полімерази.

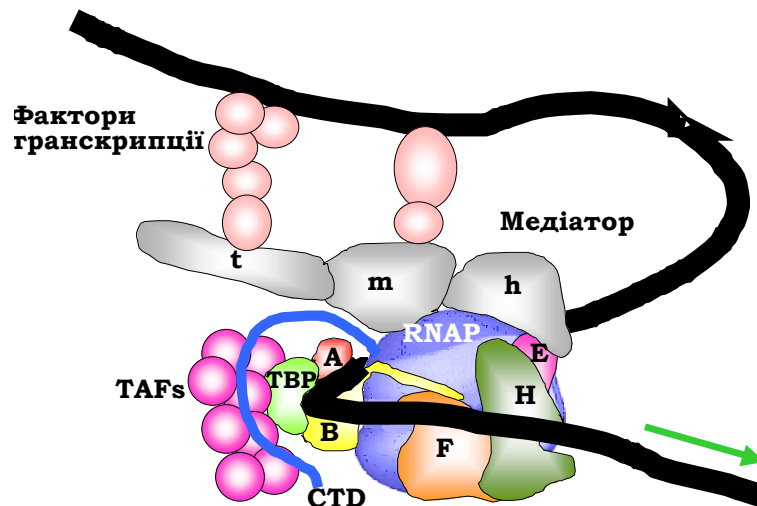
**TFIIE** – гетеродимер, стимулює роботу TFIIH, взаємодіє з транскрипційним міхуром при ініціації транскрипції.

**TFIIH** – найбільший із базальних факторів, складається з 10 субодиниць, має молекулярну вагу, порівняну з такою РНК-полімерази. TFIIH – єдиний базальний фактор, що має ферментативні активності. Перша з них – АТР-залежна ДНК-**геліказна** (гелікази – helicases – ферменти, що розкручують подвійну спіраль). Частина TFIIH взаємодіє з ДНК нижче від активного центру полімерази (по ходу транскрипції). Імовірно, конформаційні зміни білка, зумовлені гідролізом АТР, викликають обертання подвійної спіралі навкруг своєї осі, а оскільки вище від активного центру ДНК є жорстко зафіксованою, таке обертання призводить до локального плавлення дуплекса в зоні стартової точки. РНК-полімераза II є єдиною з полімераз, яка потребує такої ге-

ліказної активності при ініціації транскрипції, – інші РНК-полімерази, еукаріотичні та бактеріальні, утворюють транскрипційний міхур завдяки певним ДНК-білковим взаємодіям, яким сприяє негативна надспіралізація (див. розділ 5).

Друга ферментативна активність ТФІІІ – **кіназна**, за рахунок якої здійснюється фосфорилювання С-кінцевого домену РНК-полімерази при переході від ініціації до елонгації транскрипції. У дефосфорилюваній формі (форма А РНК-полімерази), на стадії збирання преініціаторного комплексу, СТД взаємодіє з базальними факторами транскрипції; після фосфорилювання певних залишків Ser (форма О РНК-полімерази) ця взаємодія порушується.

**Медіатор.** Ефективне збирання преініціаторного комплексу є можливою лише за участі ще одного структурного модуля – медіатора (mediator), який містить понад 20 субодиниць (склад варіює для різних промоторів – певний мінімальний набір субодиниць доповнюється більш специфічними за рахунок взаємодії з транскрипційними факторами). Цей мультибілковий комплекс має витягнуту структуру, яку можна розділити на три частини: голова (h, head), середня частина (m, middle) і хвіст (t, tail). Медіатор здійснює лише білок-білкові взаємодії: голова й середня частина взаємодіють з РНК-полімеразою, середня частина та хвіст – зі специфічними факторами транскрипції, що зв'язані на проксимальних і дистальних елементах промотора (рис. 6.6).



**Рис. 6.6.** Схема взаємного розташування елементів у складі преініціаторного комплексу РНК-полімерази II. Стрілкою позначено напрямок транскрипції

Таким чином, медіатор є засобом передачі “активаційних сигналів” з регуляторних елементів послідовності на РНК-полімеразу: збільшення кількості взаємодій підсилює ефективність збирання РІС.

### Ініціація транскрипції РНК-полімеразою II

Збирання преініціаторного комплексу є першою подією процесу ініціації. Схему будови комплексу наведено на рис. 6.6: ДНК (базальний промотор) довжиною приблизно 70 пар основ обгортає білкові компоненти РІС. При цьому промоторна ДНК зв'язана головним чином із базальними факторами транскрипції, не з полімеразою: зажим, що формується щелепами полімерази (рис. 6.1) не дозволяє розміститися дволанцюговій ДНК у щілині поблизу від активного центру.

На першому етапі ініціації утворюється *закритий* преініціаторний комплекс (рис. 6.7).

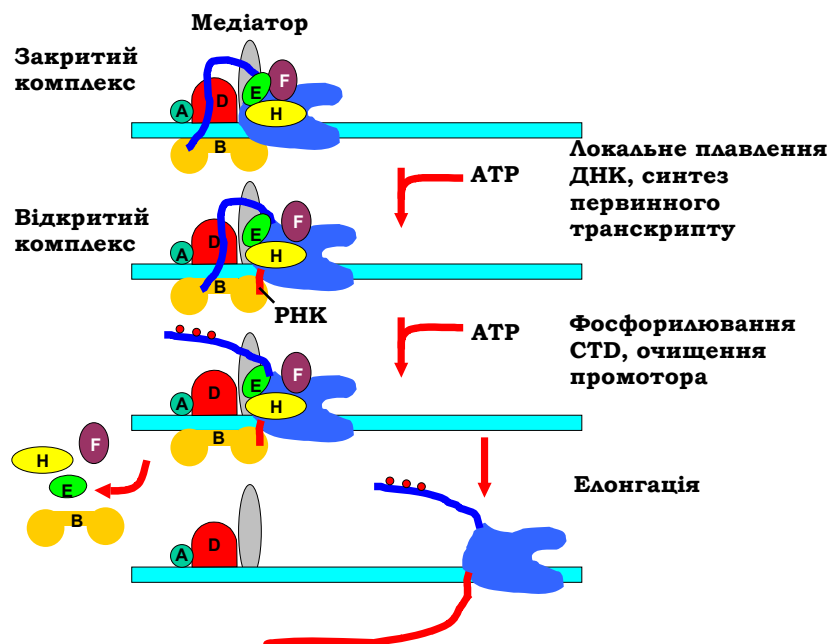


Рис. 6.7. Процес ініціації транскрипції РНК-полімеразою II

Наступна подія – утворення *відкритого* комплексу – локальне плавлення ДНК за рахунок геліказної активності ТFIІН. Нематричний ланцюг після цього захоплюється фактором ТFIІF, матричний – занурю-

ється в активний центр, де взаємодіє також із N-кінцевою частиною TFIIВ. Починається синтез короткого первинного транскрипту, коротка гібридна подвійна спіраль стабілізується тим самим N-кінцевим доменом TFIIВ, який при цьому блокує канал виходу РНК.

На наступному кроці спрацьовує кіназна активність TFIIН: відбувається фосфорилування, яке є точкою перемикання ініціації на елонгацію. STD втрачає зв'язок із базальними факторами транскрипції, які дисоціюють від полімерази – відбувається очищення промотора. РНК-полімераза продовжує синтез РНК, а частина базальних факторів може залишатися на базальному промоторі й ініціювати зв'язування іншої полімерази.

Після термінації транскрипції (розділ 7) РНК-полімераза зв'язується з TFIIF, який рекрутує специфічну фосфатазу – відбувається дефосфорилування STD, і полімераза вступає в новий робочий цикл.

## **РНК-полімерази I і III**

### *Ініціація транскрипції генів рибосомної РНК*

Система ініціації транскрипції РНК-полімеразою I виглядає значно простіше. Промотор кожного кластера генів рибосомної РНК (рис. 6.8, див. також рис. 4.6) містить два елементи послідовності: кор-елемент у зоні стартової точки та UCE (Upstream Control Element) на відстані приблизно 100 пар основ від старту. Ініціація залежить від двох транскрипційних факторів: UBF (Upstream Binding Factor) і SL1, останній позначають іноді як TFIB.

Зв'язування UBF є передумовою ініціації, але цей фактор не має високої специфічності щодо елементів промотора: він зв'язується будь-де в межах рибосомного кластера. Димери UBF містять у своїй структурі HMG-бокси (див. рис. 3.17), які сприяють вигинам ДНК. У результаті зростає імовірність наближення у просторі двох елементів промотора, з якими специфічно взаємодіє SL1. Цей фактор є аналогом фактора TFIIID РНК-полімерази II: він містить TBP і певний специфічний набір асоційованих з ним факторів, від яких і залежить спорідненість SL1 до промотора рибосомного кластера. SL1 рекрутує до промотора РНК-полімеразу I, яка починає синтез попередника рРНК.

Отже, для РНК-полімерази I виконується та сама закономірність, що й для РНК-полімерази II: незалежно від присутності чи відсутності ТАТА-боксу (який відсутній у промоторі RNAPI), TBP має бути обов'яз-

ковим компонентом преініціаторного комплексу, сприяючи організації промоторної ДНК навкруг полімерази. Ця закономірність зберігається також і для РНК-полімерази III.

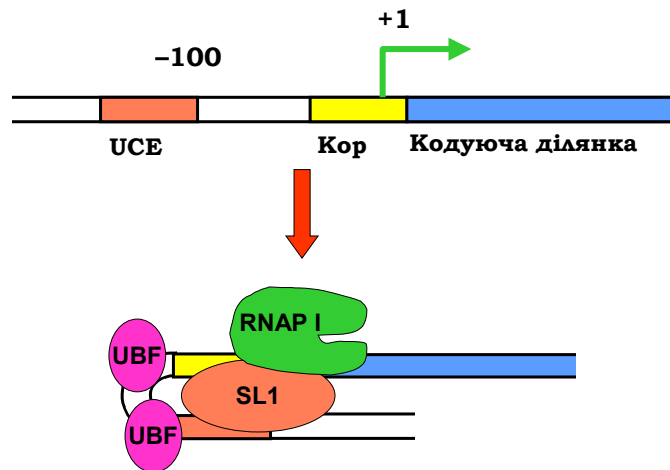


Рис. 6.8. Ініціація транскрипції РНК-полімеразою I

### Ініціація транскрипції РНК-полімеразою III

Промотори генів тРНК і 5S рРНК, які транскрибуються РНК-полімеразою III, знаходяться всередині кодуєчої частини гена (рис. 6.9). Два елементи промотора гена тРНК, так звані бокси А і В, упізнаються фактором TFIIIC, який рекрутує TFIIIB, який, у свою чергу, викликає зв'язування полімерази. Фактор TFIIIB містить у своєму складі ТВР і кілька специфічних для РНК-полімерази III TAFs.

Усередині гена 5S рРНК замість боксу В присутній інший елемент – бокс С, що впізнається фактором TFIIIA. Далі цей білок рекрутує TFIIIC і формування преініціаторного комплексу продовжується так само, як і на гені тРНК.

Описані події відбуваються однократно. Коли починається транскрипція, полімераза витісняє TFIIIC (і TFIIIA), що знаходиться попереду від неї. TFIIIB при цьому залишається на місці, рекрутуючи нову полімеразу для нового раунду транскрипції.

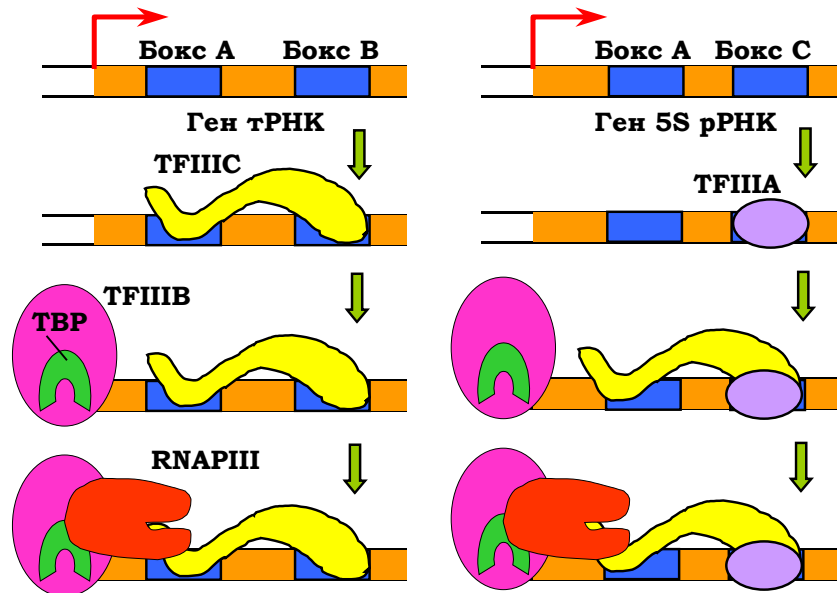


Рис. 6.9. Ініціація транскрипції РНК-полімеразою III.

### Механізми активації транскрипції

Кілька десятків тисяч еукаріотичних генів (зрозуміло, що йдеться головним чином про білкові гени, які транскрибуються РНК-полімеразою II) потребують диференційної активації / репресії в певні моменти залежно від типу клітин, стадії розвитку, зовнішніх умов тощо. Як і в прокариотів (розділ 5), ключовими елементами системи регуляції транскрипції є регуляторні *цис*-елементи послідовності (проксимальні й дистальні елементи промоторів) і транс-регулятори – білкові *транскрипційні фактори* (ТФ, під якими будемо розуміти специфічні, не базальні, фактори транскрипції). Проте кількість еукаріотичних генів є значно більшою, і вони не об'єднуються в оперони, коли кілька генів перебували б під контролем одного транскрипційного фактора. Зрозуміло, що кожен ген не може контролюватися своїм особливим фактором транскрипції: фактор також є продуктом певного гена, який також має контролюватися певним фактором.

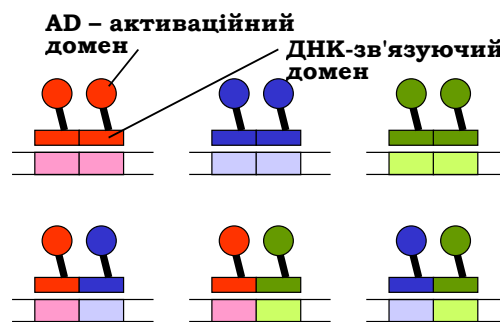
Другою особливістю еукаріотів є та обставина, що ДНК клітинного ядра організована у складні хроматинові структури (розділ 4), тобто виникає проблема доступності регуляторних елементів послідовності



(а також базальних промоторів) до транскрипційних факторів, РНК-полімерази та елементів її преініціаторного комплексу. Отже, активація транскрипції потребує перебудов структури хроматину в напрямку деконденсації хроматинової фібрили та визволення *цис*-елементів від нуклеосом. Для реалізації таких перебудов є два основні інструменти, які діють у тісній координації один з одним: система *посттрансляційних модифікацій* гістонів і АТР-залежні *фактори ремодельовання* хроматину, що здійснюють репозиціювання нуклеосом.

### Транскрипційні фактори

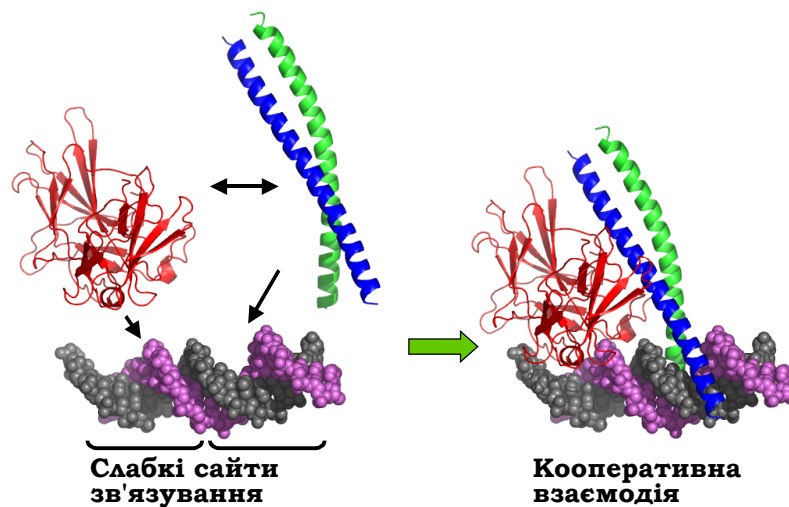
Відповіддю на необхідність регулювати окремо активність великої кількості генів лімітованим набором факторів транскрипції є принцип **модульності** будови еукаріотичних промоторів. Цей принцип ілюструє рис. 6.10: три регуляторні (наприклад, проксимальні) елементи послідовності мають спорідненість до трьох транскрипційних факторів, із трьох пар такої взаємодії можна скласти шість комбінацій. Насправді таких пар значно більше, і кількість можливих комбінацій – практично нескінченно велика. Кожен промотор може мати свій власний набір модулів, який відрізняє його від інших промоторів, і, відповідно, власний набір досить великої кількості транскрипційних факторів, потрібних для активації гена.



**Рис. 6.10.** Модульний принцип взаємодії транскрипційних факторів із промоторами

Із принципом модульності тісно пов'язаний принцип **кооперативності** взаємодії транскрипційних факторів із *цис*-елементами, що знаходяться поряд. Кожен з ТФ зазвичай має порівняно невисоку

спорідненість до відповідних елементів послідовності. Але якщо два *цис*-елементи розташовані поруч, і два ТФ здатні взаємодіяти між собою, спорідненість кожного до них підвищується (рис. 6.11): стабільність комплексу значно зростає (експоненціально, див. рівняння (1.6)) завдяки сумі невеликих знижень вільної енергії при взаємодії кожного окремого ТФ. Приклад на рис. 6.11 показує лише один “елементарний акт” кооперативної взаємодії кількох ТФ із промотором.



**Рис. 6.11.** Приклад кооперативної взаємодії двох транскрипційних факторів із ДНК (1A02)

Крім того, більшість факторів транскрипції мають у своїй структурі принаймні два домени: той, що взаємодіє з ДНК, і так званий активаційний (AD, Activation Domain), який використовується для взаємодії з іншими білками (рис. 6.10). Активаційні домени зв'язують білкові кофактори (коактиватори), у результаті на дистальних і проксимальних елементах промотора формується складний мультибілковий комплекс – **енхансосома** (enhanceosome). До складу енхансосом часто входять також гістон-ацетильтрансферази та фактори ремоделювання хроматину (див. нижче).

Різні елементи енхансосоми діють синергічно, підвищуючи загальну стабільність комплексу. З іншого боку, відсутність кількох елементів може зумовити дестабілізацію та розпад енхансосоми, де спорідненість до ДНК кожного окремого елемента є невисокою.

Це забезпечує динамізм активації: енхансосома не є фіксованою, а збирається / розбирається в певні моменти.

Активаційні домени ТФ і коактиватори мають, у свою чергу, спорідненість до медіатору та базальних факторів транскрипції. Результатом такої взаємодії є ефективне збирання преініціаторного комплексу на базальному промоторі (рис. 6.12). Слід зауважити, що певні компоненти мультибілкових комплексів, які збираються на промоторах, можуть, навпаки, блокувати ініціацію транскрипції – тоді їх називають репресорами та корепресорами.

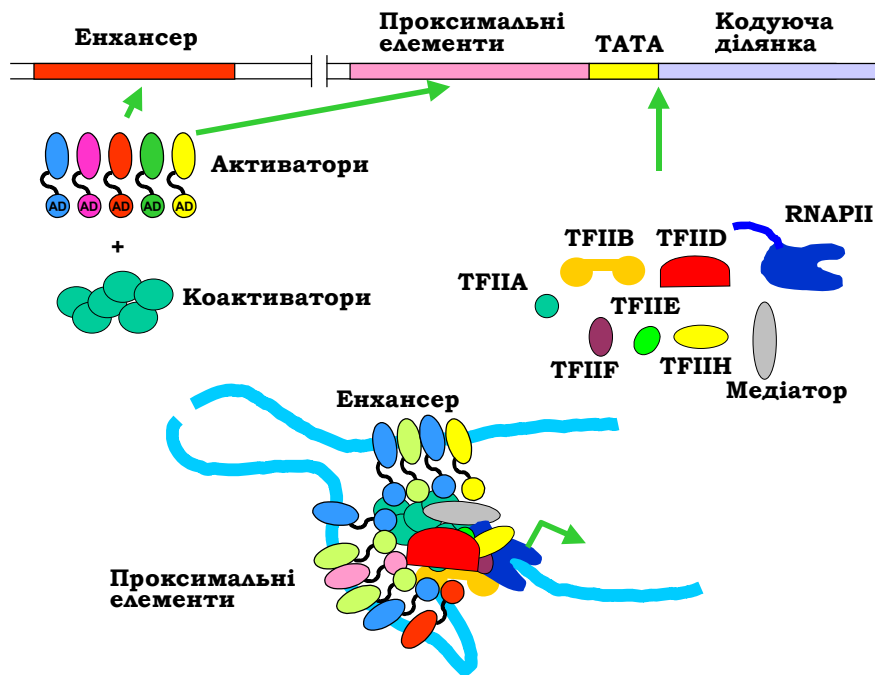


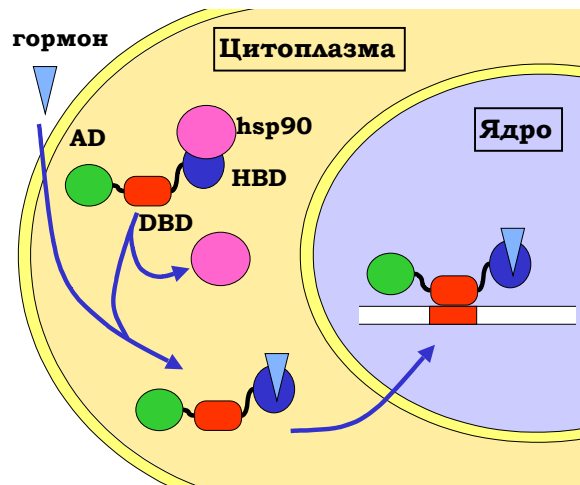
Рис. 6.12. Схема збирання комплексу активації транскрипції

### Зовнішня регуляція активності транскрипційних факторів

Активність певного гена залежить від наявності у клітині певного набору активаторів / репресорів транскрипції. Відповідно, гени самих факторів транскрипції перебувають під контролем складних систем регуляції, що працюють під час розвитку та диференціації клітин. У результаті в клітині певного типу відбувається синтез специфічного

набору ТФ, що призводить до активації специфічного набору генів. У той же час, експресія певного гена може оперативно контролюватися у відповідь на зовнішні сигнали шляхом зміни активності вже синтезованих транскрипційних факторів. Два найважливіші механізми такої регуляції – взаємодія ТФ певного типу зі стероїдними гормонами та каскади посттрансляційних модифікацій у відповідь на дію хімічних сигналів (сигнальна трансдукція).

**Гормонові рецептори** – транскрипційні фактори, активність яких залежить від стероїдних гормонів. Гормоновий рецептор (зазвичай гомодимер) складається з трьох структурних доменів: 1) ДНК-зв'язувальний домен, що має специфічну спорідненість до певних елементів послідовності ДНК (див. рис. 3.14, б); 2) активаційний домен; 3) гормон-зв'язувальний домен. У відсутності гормону білок міститься в цитоплазмі, де гормон-зв'язувальний домен взаємодіє з білком теплового шоку hsp90 (розділ 8), який підтримує недоструктурований стан домену. У результаті гормоновий рецептор є інактивованим (рис. 6.13).



**Рис. 6.13.** Активація гормонного рецептора стероїдним гормоном.  
AD – активаційний, DBD – ДНК-зв'язувальний,  
HBD – гормон-зв'язувальний домени.

Коли гормон проникає в цитоплазму, він взаємодіє з гормон-зв'язувальним доменом, витісняючи hsp90, відбувається остаточне структурування гормонного рецептора, і той стає активним – прямує до ядра, де зв'язується зі специфічним елементом послідовності. Сайт

зв'язування гормонального рецептора є досить маленьким, а чим коротшою є довжина сайту зв'язування, тим імовірнішою є його доступність. Тому взаємодія гормонального рецептора з ДНК часто запускає каскад збирання енхансосоми: рецептор рекрутує інші транскрипційні фактори за рахунок свого активаційного домену.

**Сигнальна трансдукція.** Приклад сигнальної трансдукції наведено на рис. 6.14. Білковий гормон не проникає у клітину, а зв'язується рецептором на зовнішньому боці мембрани. Рецептор є гомодимером, і у відсутності гормону дві субодиниці вільно дифундують у площині мембрани. Зв'язування гормону викликає димеризацію, унаслідок чого в цитоплазматичній частині рецептора виникає кіназна активність. Кіназа фосфорилує неактивні субодиниці транскрипційного фактора. Після фосфорилування субодиниці об'єднуються в димер, що й активує цей фактор: він проникає в ядро, де знаходить специфічну послідовність ДНК. Наведена схема є найпростішим варіантом сигнальної трансдукції: часто примембранна кіназа запускає каскад фосфорилування – фосфорилує білок, який набуває внаслідок цього кіназної активності, ця нова кіназа фосфорилує інший білок (або кілька різних білків, завдяки чому здійснюється підсилення сигналу та / або його розгалуження по кількох шляхах, спрямованих до кількох кінцевих мішеней), перетворюючи його на кіназу, і так далі – до фосфорилування та, відповідно, активації транскрипційного фактора.

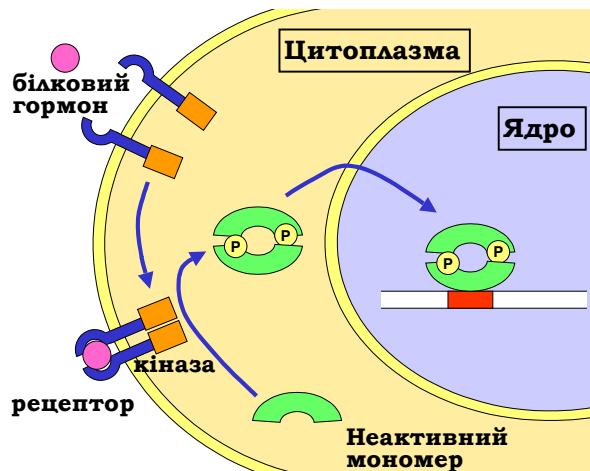


Рис. 6.14. Активація транскрипційного фактора у відповідь на дію білкового гормону

### *Ацетилювання гістонів*

Нуклеосоми та хроматинова фібрила в цілому виступають як загальний репресор генної активності. Тим самим вони допомагають забезпечити загальну інактивацию більшості генів в еукаріотичній клітині, за винятком тих, чия активація здійснюється за участю ТФ. Ключовим моментом механізмів активації є підвищення доступності регуляторних сайтів у складі промоторів, що забезпечується, зокрема, специфічними модифікаціями гістонових хвостів (див. розділ 4). Специфічна картина (патерн) модифікацій відіграє також і зворотню роль – у здійсненні гарантованої репресії гетерохроматинових ділянок.

Серед інших модифікацій ацетилювання залишків Lys (у певних консервативних позиціях, див. рис. 4.13) майже завжди корелює з активацією транскрипції – ацетилювані гістон-ацетилтрансферазами гістони акумулюються в активних промоторах, і навпаки – дія гістон-деацетилаз призводить до інактивации. Гістон-ацетилтрансферази (НАТ) входять до складу мультибілкових комплексів, які часто є компонентами енхансосою. НАТ-комплекси можуть рекрутуватися до промоторів транскрипційними факторами чи кофакторами (наприклад, активаційним доменом гормонального рецептора). Зв'язування НАТ може індукуватися іншими гістоновими модифікаціями, скажімо, через впізнання фосфорильованого Ser10 гістону H3. Крім того, часто до складу НАТ входять бромодомени – структурні модулі, які мають специфічну спорідненість до ацетилюваних лізинів. Тобто НАТ упізнають Lys, уже ацетилювані іншими НАТ, і здійснюють ацетилювання сусідніх нуклеосом, підтримуючи таким чином ацетилюваний статус певної ділянки хроматину.

Механізми головних впливів ацетилювання гістонів на структуру хроматину та активацію транскрипції полягають у наступному:

- Ацетилювання сприяє *деконденсації хроматинової фібрили* за рахунок зниження позитивного заряду головних факторів конденсації, якими є гістонові хвости. У результаті знижується спорідненість ацетилюваних нуклеосом до гістону H1, що далі сприяє *деконденсації фібрили*. Розгортання фібрили й тимчасова дисоціація H1 створює “вікно можливості” для зв'язування регуляторних факторів на виході ДНК з нуклеосоми та з міжнуклеосомною лінкерною ДНК.
- Хоча ацетилювання гістонів не змінює структуру нуклеосоми, зниження позитивного заряду гістонових хвостів при-

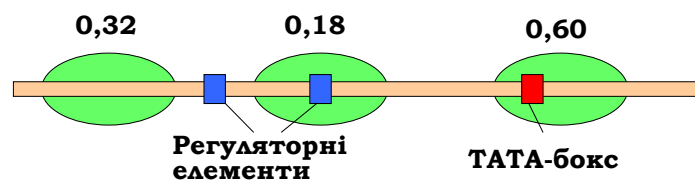
зводить до *дестабілізації нуклеосоми* на виході з неї нуклеосомної ДНК за рахунок підвищення електростатичного розштовхування між сусідніми витками нуклеосомної суперспіралі (див. розділ 4). У результаті полегшується тимчасове руйнування нуклеосом іншими факторами: перенесення гістонів на проміжні акцептори та робота комплексів ремоделювання хроматину (див. нижче).

- Ацетильовані лізинові залишки гістонів можуть безпосередньо впізнаватися факторами й кофакторами транскрипції. Так, наявність бромодому у складі TAF<sub>II</sub>250 (див. вище) сприяє підвищенню локальної концентрації TFIID у ацетильованих ділянках хроматину.

### Доступність промоторів

Залежне від послідовності ДНК переважне позиціонування нуклеосом (розділ 4) призводить до диференційного експонування ділянок ДНК до дії регуляторних факторів. З одного боку, будь-яка послідовність пар основ диктує певний “передумовлений” розподіл нуклеосом. З іншого боку, активація промоторів призводить до їхнього збіднення на нуклеосоми або за рахунок репозиціонування, або внаслідок тимчасового видалення гістонів.

Рисунок 6.15 ілюструє детально досліджений приклад: присутність нуклеосом у промоторі дріжджового гена *PHO5*, вивчену Корнбергом зі співавторами (Roger D. Kornberg).



**Рис. 6.15.** Позиції нуклеосом (овали) у промоторі гена *PHO5* дріжджів *Saccharomyces cerevisiae* та рівень їхньої присутності після активації (числа – частка часу, коли нуклеосома присутня)

У неактивному стані промотор містить 3 нуклеосоми у специфічних позиціях, у межах однієї знаходиться ТАТА-бокс, іншої – специфічна регуляторна послідовність. При активації промотора (цьому передую

гіперацетилювання гістонів) спостерігається втрата нуклеосом в усіх трьох сайтах, але ця втрата не є повною в жодному з них. У середньому втрачається 1,9 нуклеосоми з трьох; кожна з нуклеосом від 0,18 до 0,6 (специфічно для певної нуклеосоми) частини часу зберігається в активному промоторі. Це означає, що при активації здійснюється певна рівновага між видаленням та реформуванням нуклеосом.

Тимчасове видалення нуклеосом з активних промоторів є загальним правилом. Тотальний аналіз розподілу нуклеосом по ділянках усього геному дріжджів указує, що він є нерівномірним: 1) у міжгенних зонах щільність нуклеосом зменшена порівняно з відкритими рамками зчитування; 2) регуляторні області, зокрема промотори, містять менше нуклеосом, ніж інші міжгенні зони; 3) існує зворотна кореляція між щільністю нуклеосом у промоторах і рівнем транскрипційної активності відповідних генів.

### *Комплекси ремоделювання хроматину*

Підвищення доступності промоторів за їхньої активації потребує спеціальних механізмів. Адаже за фізіологічної іонної сили електростатичні взаємодії ДНК і гістонів є дуже міцними, і нуклеосома зберігає високу стабільність. Ця стабільність практично виключає навіть переміщення нуклеосоми вздовж ДНК (слайдинг): таке переміщення потребує появи високоенергетичних інтермедіатів із частковим порушенням взаємодії гістонів із ДНК. Оскільки переміщення нуклеосом є необхідним для експонування регуляторних сайтів на ДНК до дії транскрипційних факторів, у клітині існує спеціальна система, що сприяє репозиціонуванню нуклеосом за рахунок індукування проміжних структурних станів: АТР-залежні комплекси ремоделювання хроматину.

На сьогодні описано велику кількість таких комплексів у дріжджів, комах і ссавців. Усі комплекси ремоделювання (КР) – мультибілкові комплекси досить великої молекулярної ваги – містять у своєму складі субодиницю, що має АТРазну активність. Власне АТРазний домен є дуже подібним для усіх комплексів і має високу гомологію з ДНК-геліказами. Залежно від наявності додаткових структурних доменів різного типу в складі каталітичної субодиниці, КР поділяють на три основні групи:

1) **Swi/Snf** (Switch/Sucrose non-fermenting – назва походить від першого описаного КР цієї родини, який було відкрито як групу мутацій дріжджів з певними біохімічними проявами) – комплекси складаються з 10–15 субодиниць, каталітична субодиниця містить бромо-



домен, під час дії комплексу спостерігаються різноманітні проміжні структурні форми нуклеосом.

2) **ISWI** (Imitation SWItch) – 4–5 субодиниць, індукують слайдинг нуклеосом без масштабних змін їхньої структури;

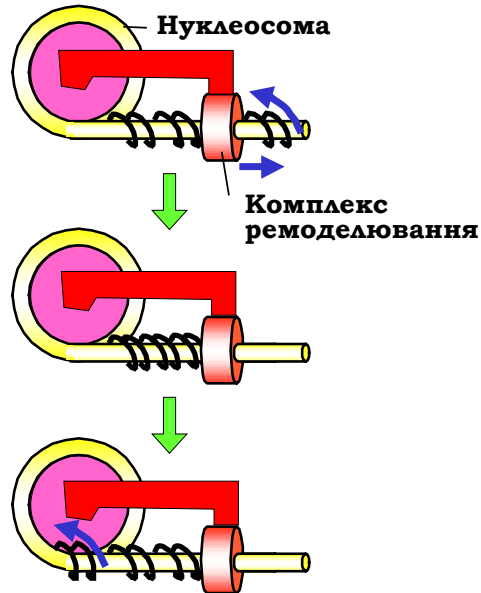
3) **CHD** (Chromodomain Helicase DNA binding) – містять хромодомен і гістон-деацетилазу, тісно пов'язані з перебудовами хроматину в репресованих ділянках.

Крім того, є низка комплексів, що не вкладаються в цю класифікацію. КР здатні здійснювати численні взаємодії з нуклеосомною ДНК, гістоновими хвостами, специфічними та загальними факторами транскрипції, гістон-ацетилтрансферазами тощо. При взаємодії з нуклеосомою КР і нуклеосома чітко орієнтуються відносно одне одного: КР родини Swi/Snf оточують нуклеосому, зв'язуючи її всередині глибокої порожнини.

*Основний результат активності всіх комплексів ремоделювання – репозиціонування нуклеосом.* Крім того, комплекси родини Swi/Snf можуть індукувати перенесення октамеру гістонів з однієї ділянки ДНК на іншу або з ДНК на проміжні акцептори гістонів (див. розділ 4).

**Механізм дії комплексів ремоделювання** базується на гомології з ДНК-геліказами. Незважаючи на це, АТРазна субодиниця КР не здатна діяти як геліказа – розводити два ланцюги. Подібно до геліказ, комплекси ремоделювання, використовуючи енергію гідролізу АТР, пересуваються вздовж ДНК, але без її руйнування. Така транслокація має бути пов'язана з одночасним обертанням ферменту навколо подвійної спіралі або навпаки – прокручуванням спіралі через фермент. Оскільки КР при цьому зафіксований на нуклеосомі, тобто взаємне обертання комплексу й нуклеосомі є неможливим, рух комплексу в бік від нуклеосомі буде генерувати торсійну напругу (зміну твіста), штовхаючи додаткові витки подвійної спіралі всередину нуклеосомної ДНК – ДНК прокручується через КР (рис. 6.16). Такий сценарій роботи комплексів ремоделювання розглядається сьогодні як найбільш імовірний.

Напруга, яка створюється за рахунок транслокації комплексу ремоделювання, може трансформуватися в зміну твіста в межах кінцевої ділянки нуклеосомної ДНК. Локальна зміна твіста дифундує всередину нуклеосомі, результатом чого є прокручування подвійної спіралі по поверхні октамеру гістонів, тобто зміна його позиції (рис. 6.17). При такому механізмі переміщення шляхом *twisting* (twisting) не повинно спостерігатися великих змін у структурі нуклеосомі під час її пересування. Саме це й характерно для комплексів ремоделювання типу ISWI.



**Рис. 6.16.** Модель дії комплексу ремоделювання шляхом створення торсійної напруги в ДНК, яка входить у нуклеосому

Комплекси ремоделювання родини Swi/Snf здатні генерувати проміжні стани нуклеосоми, в яких ДНК є більш доступною для нуклеаз, за рахунок формування петлі ДНК на поверхні октамеру гістонів. Отже, на відміну від попереднього механізму, для комплексів Swi/Snf напруга, що є результатом транслокації, викликає *випетлювання* (bulging) нуклеосомної ДНК. Така петля, імовірно, формується на кінцевій ділянці ДНК у нуклеосомі й далі пересувається по поверхні октамеру, що приводить знову до зміни позиції нуклеосоми (рис. 6.17).

Проміжні метастабільні стани нуклеосоми з випетлюванням ДНК можуть відігравати важливу функціональну роль самі по собі, оскільки вже можуть забезпечувати зростання доступності нуклеосомної ДНК: експонування петлі створює умови для її “захоплення” транскрипційними факторами. Крім того, значна дестабілізація ДНК-гістонових взаємодій у проміжному стані з випетлюванням значно підвищує імовірність руйнування нуклеосоми: у присутності Swi / Snf продемонстровано перенесення гістонів на іншу ділянку ДНК або на проміжні переносники гістонів.

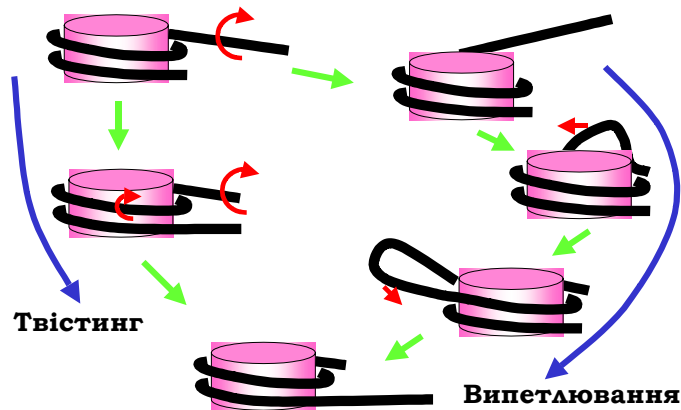


Рис. 6.17. Два механізми зміни трансляційної позиції нуклеосоми в результаті активності комплексів ремоделювання

Обидва описані механізми дії комплексів ремоделювання розпочинаються з ділянки ДНК на вході до нуклеосоми. Відповідно, ця ділянка має бути вільною від гістону H1: H1 і КР конкурують між собою за зону входу / виходу нуклеосомної ДНК.

**Шляхи рекрутування комплексів ремоделювання до промоторів.** Насправді функціональні наслідки ремоделювання хроматину певним комплексом можуть бути різноманітними. Серед відомих нині комплексів ремоделювання лише один – NuRD (Nucleosome Remodeling and Deacetylation), представник родини CHD – працює виключно як корепресор. Дія інших комплексів ремоделювання призводить або до репресії, або до активації – залежно від контексту інших функціонально важливих впливів, у кооперації з якими працює даний комплекс.

У разі активації система АТР-залежного ремоделювання хроматину завжди працює кооперативно із системою ацетилювання гістонів. Реалізуються різноманітні стратегії рекрутування КР до промоторів:

- За рахунок взаємодії КР з *транскрипційними факторами*. Активація ініціюється ТФ, який упізнає невеликий сайт зв'язування на ДНК (навіть на поверхні нуклеосоми) і безпосередньо рекрутує комплекс ремоделювання, або КР рекрутується пізніше на стадії збирання енхансоми іншими ТФ. Після цього, як правило, до енхансоми рекрутується гістон-ацетильтрансферазний комплекс.

- За рахунок *упізнання ацетильованих Lys бромодоменом*, що входить до складу КР. Один із прикладів, який поєднує цей і попередній шляхи рекрутування КР, зображено на рис. 6.18: невеликий транскрипційний фактор зв'язується у промоторі й рекрутує гістон-кіназу; кіназа здійснює фосфорильовання Ser-10 гістону H3 та дисоціює; фосфорильований Ser-10 і транскрипційний фактор упізнаються гістон-ацетилазою, відбувається ацетилювання лізинів; ацетильовані Lys та ТФ рекрутують SWI/SNF (комплекс, що належить до родини Swi/Snf), який, здійснюючи репозиціювання нуклеосоми, звільняє ТАТА-бок; відбувається зв'язування TFIID і збирання преініціаторного комплексу. Інший приклад див. на рис. 6.19.

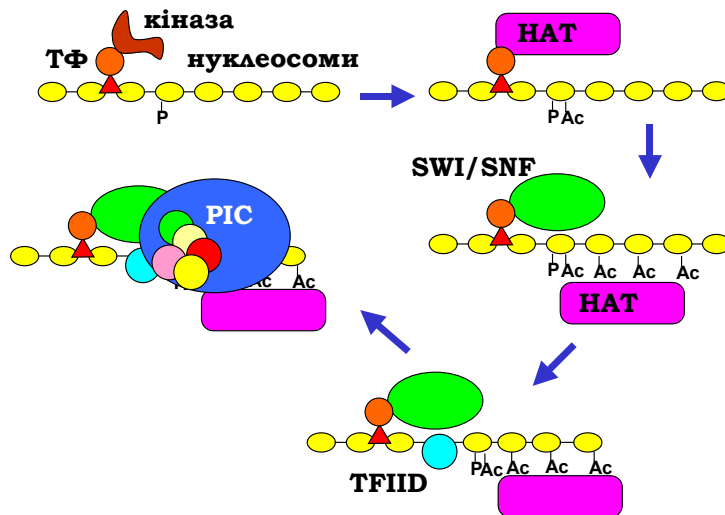


Рис. 6.18. Послідовність подій при активації дріжджового промотора *INO1*

- За рахунок *взаємодії з білками ядерного матриксу* (див. розділ 4), які можуть рекрутувати комплекс ремоделювання до основи хроматинової петлі – до ділянок, асоційованих з матриксом (MAR). Така ділянка часто використовується як своєрідні “вхідні ворота” для факторів, що забезпечують регуляцію активності в межах петлі, де міститься один або кілька генів.

- За рахунок *упізнання метильованих Lys* хромодоменом, який входить до складу КР родини СНD. Комплекси цієї родини містять також структурні модулі, здатні впізнавати *метильовані цитозини (5mC)* у ДНК.

Таким чином, комплекси ремоделювання хроматину інкорпоровані в загальну систему регуляції транскрипції. Вони працюють у тісній кооперації з транскрипційними факторами та системою посттрансляційних модифікацій гістонів.

### **Роль білків HMG в активації транскрипції**

Назва HMG (див. також розділ 3) об'єднує три, структурно абсолютно різні, родини білків. Але всі вони задіяні до системи регуляції транскрипції в еукаріотів.

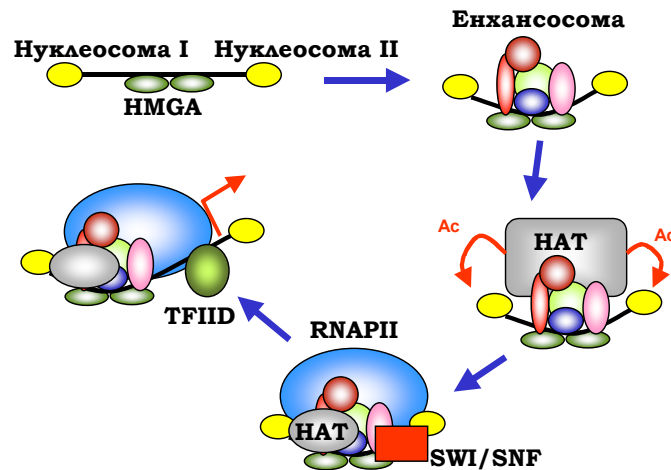
**HMGА** – невпорядковані поліпептиди, що містять у своєму складі три так звані АТ-гаки – елементи амінокислотної послідовності, здатні впізнавати короткі (чотири–п'ять пар основ), часто присутні в промоторах, АТ-збагачені ділянки ДНК (див. рис. 3.16, б). Невпорядковані частини поліпептидного ланцюга, які з'єднують АТ-гаки, мають спорідненість до певних транскрипційних факторів. Отже, зв'язування HMGА може ініціювати збирання енхансосоми. Гени HMGА активуються в результаті численних шляхів сигнальної трансдукції.

Один із прикладів участі HMGА в активації транскрипції наведено на рис. 6.19. HMGА зв'язується в лінкері між двома позиційованими нуклеосомами у промоторі β-інтерферонового гена. На цьому “зародковому” елементі відбувається збирання енхансосоми, до складу якої рекрутується гістон-ацетилтрансферазний комплекс. Ацетильовані Lys упізнаються іншою НАТ, яка продовжує процес ацетильовання, а також відбувається збирання преініціаторного комплексу – включаючи РНК-полімеразу II, але за винятком TFIID. Ацетильовані Lys та інші компоненти енхансосоми рекрутують SWI/SNF, який змінює позицію нуклеосоми II (рис. 6.19), що закривала ТАТА-бокс. Унаслідок цього нарешті зв'язується TFIID, і запускається транскрипція.

Білки HMGА не обов'язково виступають активаторами транскрипції – вони можуть також блокувати сайти зв'язування транскрипційних факторів, сприяючи репресії певних промоторів.

**HMGВ** – невеликі глобулярні білки, що містять один (дріжджі, комахи, рослини) або два (хребетні) HMG-бокси (див. рис. 3.17). У клітинних ядрах ссавців HMGВ присутні у значній кількості – до 10–20 копій на одну нуклеосому. Не дуже міцне (на невеликий проміжок часу)

зв'язування НМG-боксу з маленьким жолобком ДНК індукує значний вигин на  $\sim 80^\circ$  (див. розділ 3). Відповідно, білки НМGВ є загальними лабілізаторами структури хроматину (їх називають також архітектурними білками або хроматиновими шаперонами), що діють за принципом “торкнувся й пішов”. Після зв'язування та дисоціації НМGВ може нічого не відбутися. Але вигин, який індукується НМG-боксом може або підвищити спорідненість транскрипційних факторів, або додатково стабілізувати структуру енхансосоми (рис. 6.20). Далі НМGВ може або залишитись у складі енхансосоми, або дисоціювати, замінившись на інші транскрипційні фактори.



**Рис. 6.19.** Послідовність подій при активації  $\beta$ -інтерферонового промотора

Крім того, білки НМGВ є конкурентами гістону H1: зв'язування їх із ДНК на вході до нуклеосоми полегшує тимчасову деконденсацію хроматинової фібрили. Завдяки зв'язуванню з ДНК на вході до нуклеосоми НМGВ виступає також як кофактор комплексів ремоделювання родини Swi/Snf: вигин, що індукується НМG-боксом, полегшує випетлювання ДНК на поверхні нуклеосоми, яке мігрує всередину нуклеосомної ДНК (рис. 6.21). Ремодельованню за таким механізмом сприяє не обов'язково білок НМGВ – НМG-бокси часто містяться також у складі самих комплексів ремоделювання.

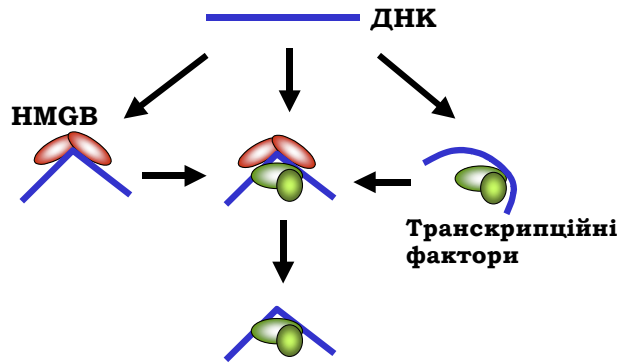


Рис. 6.20. Залучення білків HMGB до збирання енхансосоми

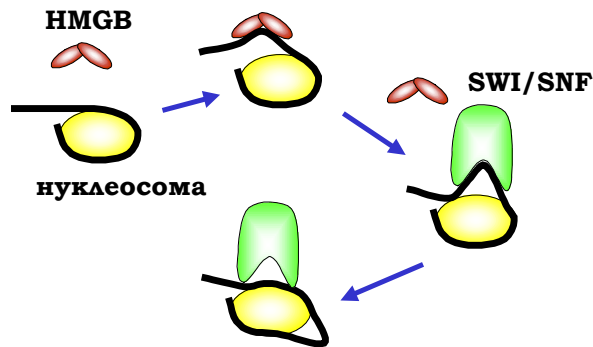


Рис. 6.21. Залучення білків HMGB до ремоделювання хроматину

**HMGN** – невпорядковані поліпептиди довжиною близько 100 амінокислот. За рахунок позитивно зарядженої частини ланцюга взаємодіють з ДНК всередині нуклеосоми. Негативно заряджена частина молекули конкурує з ДНК за гістонові хвости. У результаті HMGN, який накопичується у транскрипційно активних ділянках, сприяє деконденсації хроматинової фібрили. Взаємодія HMGN з нуклеосомою – слабка, тобто він, як і HMGB, лише на короткий час підвищує доступність ДНК для інших елементів системи активації транскрипції.

## Елонгація РНК-полімерази через хроматин

Описані вище зміни у структурі хроматину стосуються ініціації транскрипції. На стадії елонгації нуклеосома також має створювати суттєві перешкоди для проходження РНК-полімерази. При цьому хроматин, навіть у межах активних генів, у цілому зберігає нуклеосомну структуру, хоча й спостерігається збіднення на гістон H1. В експериментах *in vitro* наявність нуклеосом у складі матриці значно гальмує транскрипцію. У той же час, швидкість транскрипції *in vivo* практично така сама, як *in vitro* на вільній ДНК. Отже, мають існувати особливі механізми, котрі забезпечують “прозорість” нуклеосом для РНК-полімерази у клітинному ядрі.

На модельній системі *in vitro*, що містила мономерну РНК-полімеразу бактеріофага SP6 і коротку (227 пар основ) ДНК-матрицю з однією позиційованою нуклеосомою, було показано, що гістоновий октамер переноситься як ціле на ділянку позаду полімерази (рис. 6.22). При цьому октамер ніколи не втрачає повністю свого контакту з ДНК: у процесі перенесення спостерігаються інтермедіати, які нагадують проміжні структури з петлею ДНК на поверхні октамеру при ремоделюванні нуклеосом (порівн. рис. 6.17 та 6.22).

За аналогічним механізмом із перенесенням октамеру здійснюється проходження через нуклеосому еукаріотичної РНК-полімерази III.

Що стосується РНК-полімерази II, її проходження через хроматин супроводжується частковим руйнуванням нуклеосом попереду полімерази та відновленням їх позаду:

- В ядрах клітин спостерігається досить швидкий, транскрипційно-залежний обмін гістонів H2A і H2B між хроматином та ядерним пулом гістонів.
- Нуклеосоми, які транскрибуються РНК-полімеразою II, тимчасово збіднені на гістони H2A і H2B.
- Видалення димеру гістонів H2A-H2B РНК-полімеразою II було продемонстровано на модельній системі *in vitro*.

Узагалі, порівняно з РНК-полімеразами III і бактеріофага SP6, для РНК-полімерази II нуклеосома є значно вагомшою перешкодою. Із цього випливають два наслідки. По-перше, нуклеосомний бар'єр може використовуватися для регуляції експресії генів шляхом зниження швидкості елонгації за рахунок пауз у роботі полімерази на нуклеосомах. По-друге, наявність нуклеосомного бар'єра зумовлює необхідність спеціальних механізмів полегшення елонгації транскрипції.



Швидкість транскрипції всіх РНК-полімераз залежить від іонної сили: в умовах низької іонної сили нуклеосома гальмує проходження полімерази (SP6) або повністю зупиняє її (РНК-полімераза II); підвищення іонної сили до рівня, трохи вищого за фізіологічний (наприклад, 300 ммоль/л), призводить до зростання швидкості транскрипції. Зрозуміло, що підвищення іонної сили знижує спорідненість гістонів до ДНК і тим самим певною мірою дестабілізує нуклеосома. Роль трошки підвищеної, у порівнянні з фізіологічною, іонної сили можуть відігравати *in vivo* фактори елонгації транскрипції, які працюють як проміжні переносники гістонів (див. рис. 4.11).

Описано кілька факторів елонгації, що прискорюють роботу РНК-полімерази II:

- **Elongator** – мультибілковий комплекс, одна із субодиниць якого є гістон-ацетилтрансферазою. Комплекс взаємодіє з РНК-полімеразою II і здійснює ацетилювання гістонів під час елонгації транскрипції.
- **FACT** (FACilitates Chromatin Transcription) – складається з двох субодиниць, одна з яких є НМГВ-подібним білком. FACT має підвищену спорідненість до димерів гістонів H2A-H2B – виконує роль проміжного переносника димерів, забезпечуючи їх видалення з нуклеосом попереду полімерази та повернення до хроматину позаду.
- **Spt6** – проміжний акцептор зі спорідненістю до гістонів H3-H4.

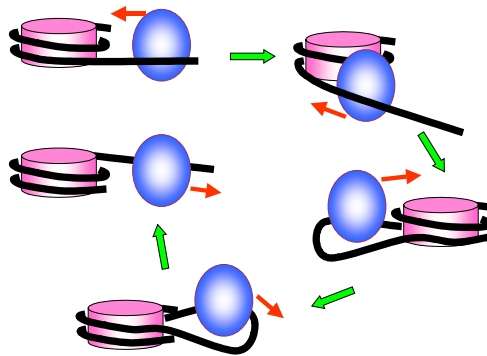
У хроматині є також інші переносники гістонів. Крім того, АТР-залежні комплекси ремоделювання хроматину виконують роль факторів, що сприяють елонгації транскрипції.

Проходження РНК-полімерази (як і будь-якого іншого ферменту, котрий здійснює транслокацію вздовж ДНК) за умов жорсткого закріплення кінців матриці (основ хроматинової петлі) має створювати позитивну надспіралізацію попереду та негативну – позаду полімеразного комплексу (див. рис. 3.24). Надспіралізація в процесі транскрипції виникає у зв'язку з тим, що обертальний рух РНК-полімеразного комплексу навкруг подвійної спіралі значно загальмований унаслідок масивності комплексу, у тому числі – присутності в його складі довгого ланцюга РНК, що синтезується. У результаті при трансляційному русі полімерази ДНК змушена обертатися навкруг своєї осі, що призводить до торсійного перекручування дуплекса попереду полімерази і, відповідно, еквівалентного розкручування – позаду.

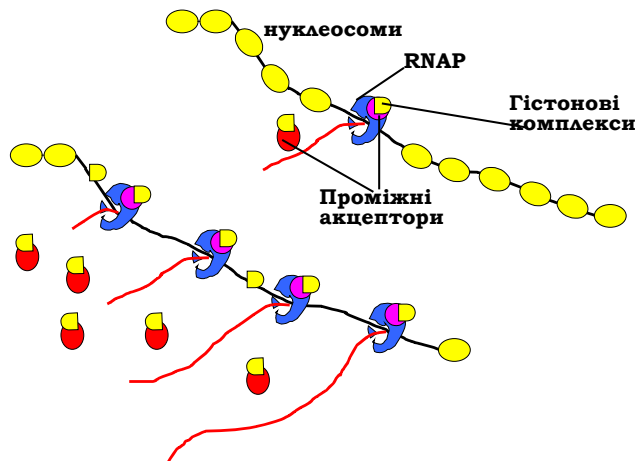
Позитивна надспіралізація дестабілізує *ліву* суперспіраль нуклеосомної ДНК попереду полімерази, що сприяє дії проміжних перенос-

ників гістонів. І навпаки – негативна надспіралізація позаду поліме-  
рази сприяє ефективному відновленню нуклеосом.

Завдяки цим механізмам (які ще далеко не з'ясовані остаточно)  
в еукаріотичних клітинах здійснюється ефективна елонгація транс-  
крипції при збереженні нуклеосомної структури навіть у транскрип-  
ційно-активних ділянках. Проте присутність нуклеосом залежить від  
рівня транскрипційної активності – кількості РНК-полімераз, що од-  
ночасно працюють на певному гені. Якщо ген є гіперактивним, ДНК  
може бути недоступною для гістонових комплексів, які залишаються  
на проміжних акцепторах (рис. 6.23).



**Рис. 6.22.** Проходження мономерної РНК-полімерази бактеріофага SP6 через нуклеосому



**Рис. 6.23.** Залежна від рівня активності транскрипції структура хроматину в транскрипційно-активних ділянках

### **Конститутивна репресія транскрипції: гетерохроматин**

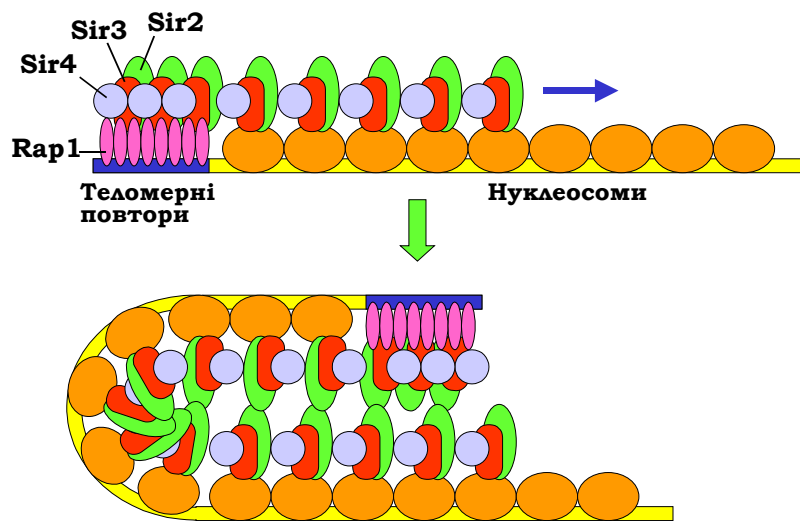
Великий розмір еукаріотичного геному (майже половина якого припадає на послідовності, що повторюються) вимагає існування систем, які б визначали гарантовану інактивіацію значної частини генетичного матеріалу. При цьому інактивованій стан як некодуючої ДНК, так і значної частини генів у клітинах певного типу має відтворюватись у дочірніх клітинах після мітозу.

Успадкування такої тканиноспецифічної системи репресії (при збереженні повної генетичної програми даного організму) називають *епігенетичною спадковістю*. В її основі лежить передача клітинам-нащадкам не просто батьківської ДНК, а хроматину разом із певними хімічними маркерами (маються на увазі патерни посттрансляційних модифікацій гістонів і метилювання ДНК). Такі маркери впізнаються відповідними білками, котрі вступають у різноманітні взаємодії з репресорами та корепресорами транскрипції, ферментами, що здійснюють посттрансляційні модифікації, і білками, які додатково компактизують хроматинову фібрилу. Подібні системи в першу чергу працюють у ділянках *гетерохроматину* – частини хроматину, що зберігає високий ступінь компактності протягом інтерфази. Прикладами конститутивного гетерохроматину (такого, що утворюється в усіх клітинах) є теломерні та центромерні зони хромосом, одна з X-хромосом самок ссавців. Інші гетерохроматинові зони є специфічними для клітин певного типу.

#### *Деацетилювання гістонів*

Якщо ацетилювання гістонів завжди корелює з підвищеною транскрипційною активністю, деацетилювання, яке здійснюється гістон-деацетилазними комплексами (HD), завжди пов'язане з репресією. Як і ацетильтрансферази, деацетилази постійно безадресно працюють у хроматині, підтримуючи певний базовий баланс ацетилювання / деацетилювання гістонів. При активації певного промотора ацетильтрансферази здійснюють адресне гіперацетилювання, а після зникнення активуючого сигналу HD повертають промотор до базового неактивного стану. Деацетилази також можуть бути адресно рекрутовані до промоторів репресорами транскрипції для підтримання гарантованого деацетилюваного статусу.

Поряд із такою динамічною регуляцією активності, деацетилювання гістонів завжди здійснюється в гетерохроматинових ділянках. Прикладом, який демонструє визначальну роль деацетилювання у підтриманні гетерохроматинового стану є теломери дріжджів (рис. 6.24).



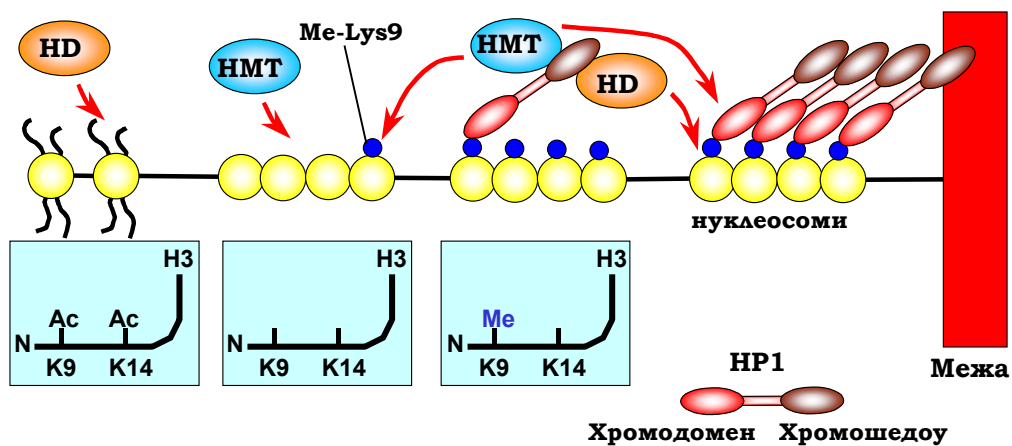
**Рис. 6.24.** Розповсюдження та самопідтримання гетерохроматинового стану в теломерах дріжджів

Теломерні повтори на кінці хромосоми впізнаються ДНК-зв'язувальним білком Rap1 (Repressor activator protein), який рекрутує білок Sir4 (Silencing information regulator). Цей останній зв'язує ще два білки – Sir3 та Sir2, останній є гістон-деацетилазою. Sir2 здійснює деацетилювання Lys16 гістону H4. Цей деацетильований залишок упізнається новим Sir3, утворюється новий Sir-комплекс, здійснюється деацетилювання нової нуклеосоми – процес деацетилювання та збирання Sir-комплексу розповсюджується вздовж фібрили. Розповсюдження гетерохроматинової зони блокуються межею, фізична природа якої є не зовсім зрозумілою.

За рахунок взаємодії між Sir-комплексами на кінцях хромосоми утворюється компактна зіпер-подібна структура (рис. 6.24). Утворення подібних структур завдяки певним білкам є досить загальною ознакою інших гетерохроматинових ділянок.

*HP1-залежна система репресії*

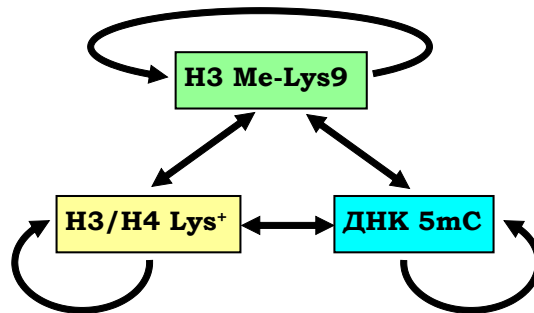
Білок HP1 (Heterochromatin Protein 1) містить два структурні домени: *хромодомен* (модуль, що впізнає метильовані лізинові залишки) і так званий *хромощедоу* (chromoshadow) – домен, який має спорідненість до певних специфічних HD і гістон-метилтрансферази (HMT), а також здатен взаємодіяти з іншою молекулою HP1.



**Рис. 6.25.** Розповсюдження та самопідтримання гетерохроматинового стану в центромерах. На вставках: стан модифікацій двох ключових залишків у N-кінцевому хвості гістону H3

Прикладом залучення HP1 до утворення гетерохроматину є центромери хромосом (рис. 6.25). Деацетилювання Lys9 і Lys14 гістону H3 сприяє зв'язуванню специфічної HMT, яка здійснює метилювання Lys9 гістону H3. Цей метильований Lys (Me-Lys9) упізнається хромодоменом HP1. Завдяки хромощедоу-домени HP1 рекрутує HD, яка підтримує деацетилюваний статус суміжних нуклеосом, і ту саму HMT, що здійснює метилювання Lys9 у складі суміжних нуклеосом: знов виникає лавиноподібний процес, що самопідтримується та розповсюджується на сусідні ділянки. Взаємодія між білками HP1 (деталі якої не з'ясовані) забезпечує додаткову компактизацію фібрили. Природа обмежувача, який не дозволяє гетерохроматину розповсюджуватися поза центромерну зону, не з'ясована, але гранична ділянка містить свій специфічний маркер – метильований Lys4 гістону H3 (і деметильований Lys9).

При реплікації білки тимчасово знімаються з ДНК у точці реплікації (розділ 9). Реплікація ДНК здійснюється нерівномірно – гетерохроматин реплікується в останню чергу. За точкою реплікації гістони батьківського хроматину (які несуть на собі гетерохроматинові маркери) повертаються на дочірні молекули ДНК разом із гістонами, синтезованими *de novo*, HP1 також повертається на той самий локус, де він був присутній, і відновлює патерни модифікацій гістонів і компактний (репресований) стан гетерохроматинової ділянки, тобто, гетерохроматиновий стан даного локусу відтворюється в дочірніх клітинах. При цьому HP1 має спорідненість до білків ламіни (розділ 4), що зумовлює розташування гетерохроматину в периферичних зонах клітинного ядра.



**Рис. 6.26.** Взаємний вплив деацетилювання гістонів, метилювання Lys9 гістону H3 і метилювання ДНК у гетерохроматині

Подібна система репресії за участю HP1 широко використовується в інших ділянках гетерохроматину, а також для гарантованого блокування генів в еухроматинових зонах. Метилювання Lys9 гістону H3 та деацетилюваний стан лізінів гістонів H3/H4 є найхарактернішою ознакою таких ділянок. Як описано вище, обидві модифікації самопідримуються та підтримують одна одну через опосередковану дію HP1. При цьому важливу роль у забезпеченні репресованого стану відіграє метилювання цитозинів ДНК – інша ковалентна модифікація, яка також відновлюється при клітинному поділі й замикає своєрідне “коло репресії” (рис. 6.26).

### Метилування ДНК

Субстратом метилування ДНК є цитозини (метильна група приєднується до п'ятого атому кільця з утворенням 5mC – 5-метилцитозину) у складі динуклеотидів CpG (рис. 6.27). Узагалі, у 70–80 % динуклеотидних контактів CpG обидва цитозини є метильованими в еукаріотичному геномі. Зони, де підтримується деметильований стан CpG (CpG-острівці, див. розділ 4), часто розташовані в промоторах генів домашнього господарства – таких, що є активними незалежно від спеціалізації клітин.

Патерн тканиноспецифічного метилування ДНК є результатом двох процесів: підтримання метильованого статусу після реплікації та метилування *de novo*.

Підтримуюча ДНК-метилтрансфераза (DNA methyltransferase, Dnmt) спрацьовує протягом 1–2 хвилини після реплікації: дві дочірні молекули ДНК містять батьківській ланцюг ДНК (з 5mC у складі CpG) і синтезований ланцюг, де С не є метильованим. Dnmt упізнає такі напівметильовані динуклеотидні контакти й відновлює симетрію щодо метилування С. За рахунок цього процесу патерн метилування відтворюється в дочірніх клітинах, що є, поряд з відновленням модифікацій гістонів, одним із важливих механізмів епігенетичного спадкування.



Рис. 6.27. Динуклеотид CpG у ДНК – субстрат метилування

Інші ДНК-метилтрансферази здійснюють метилування ДНК *de novo*. Особливо важливим цей процес є на ранніх стадіях ембріонального розвитку, коли ДНК є тотально деметильованою. У процесі диференціації здійснюється масове метилування ДНК, що визначає специфічне вимикання певних груп генів у спеціалізованих клітинах. Крім того, деметилування є можливим і в диференційованих клітинах, де Dnmt використовуються для відновлення метильованого статусу.

Отже, метилування ДНК є ознакою репресованих і гетерохроматинових ділянок. Залучення 5mC до репресії пов'язане з наявністю у складі певних білків особливих структурних модулів – MBD (Methyl

Binding Domain), які мають специфічну спорідненість до метильованих динуклеотидів CpG. Білки, що містять MBD, є компонентами різноманітних репресуючих комплексів. Зокрема, такі білки рекрутують до метильованих ділянок хроматину гістон-деацетилази. З іншого боку, деацетилюваний стан гістонових хвостів блокує деметилюючі активності, і навпаки – ацетилювання хвостів може викликати деметилювання ДНК у активних ділянках.

Аналогічно, білки, що містять MBD, рекрутують гістон-метилтрансферазу, яка здійснює метилювання Lys9 гістону H3, що призводить до репресії (рис. 6.25, 6.26). І навпаки: Me-Lys9 упізнається білком, що містить хромодомен і рекрутує ДНК-метилтрансферазу.

Хоча Me-Lys9 та 5mC є загальними маркерами конститутивно репресованих ділянок хроматину, не завжди репресія та додаткова компактизація залежить від HP1 – реалізуються також інші системи, більшість з яких є ще не достатньо вивченими. Прикладом такої системи є інактивація однієї з X-хромосом у клітинах самок ссавців. У складі X-хромосоми, яка буде інактивованою (обирається випадково на ранніх стадіях розвитку), спрацьовує ген *Xist*, що продукує велику некодуючу молекулу РНК. Ці РНК укривають собою хромосому і взаємодіють з деякими білками, серед яких – варіант гістону H2A масоH2A. Імовірно, масоH2A рекрутує гістон-метилтрансферазу (здійснюється метилювання Lys9 гістону H3) і гістон-деацетилазу. Метилювання Lys9, у свою чергу, зумовлює метилювання ДНК (що забезпечує епігенетичну спадковість). Крім того, до X-хромосоми рекрутуються структурні компактизуючі білки (але не HP1).

### ***РНК-інтерференція***

Гетерохроматин утворюється передусім на послідовностях ДНК, що повторюються. Причому має значення не послідовність як така, а саме наявність повторів. Виявляється, що у формуванні та підтриманні гетерохроматинового стану повторів (принаймні в центромерах) важливу роль відіграє процес так званої *РНК-інтерференції* (RNAi – RNA interference).

Узагалі процес РНК-інтерференції запускається будь-якою дволанцюговою РНК (dsRNA – double stranded RNA). Дволанцюгова РНК стає субстратом для РНКаз, яка отримала назву дайсер (*Dicer*) (рис. 6.28): дайсер розрізає молекулу на дволанцюгові фрагменти довжиною 19–21 пари основ (по два нуклеотиди залишаються неспареними на кінцях). Ці фрагменти позначають як siRNA (short interfering RNA),



вони можуть бути ампліфіковані РНК-залежною РНК-полімеразою (RdRP – RNA depended RNA Polymerase). Вони зв'язуються з кількома білками, утворюючи комплекс RISC (RNA Induced Silencing Complex). Один із білків має РНК-геліказну активність і розводить два ланцюги siRNA. Один із ланцюгів здійснює комплементарне спарювання з ділянкою мРНК, що синтезується у процесі транскрипції (вихідна dsRNA має збігатися за послідовністю з кодуючою частиною даного гена), спрямовуючи туди RISC. Один із компонентів комплексу – РНКаза *Slicer* – здійснює деградацію транскрипту. Крім того, RISC може індукувати активність RdRP для синтезу комплементарного ланцюга РНК на мРНК (чи її частині) як матриці. Відновлена дволанцюгова РНК буде знов підтримувати інактивацию гена шляхом інтерференції. Отже, хоча транскрипція відбувається, реалізується посттранскрипційне вимкнення гена (PTGS – posttranscriptional gene silencing).

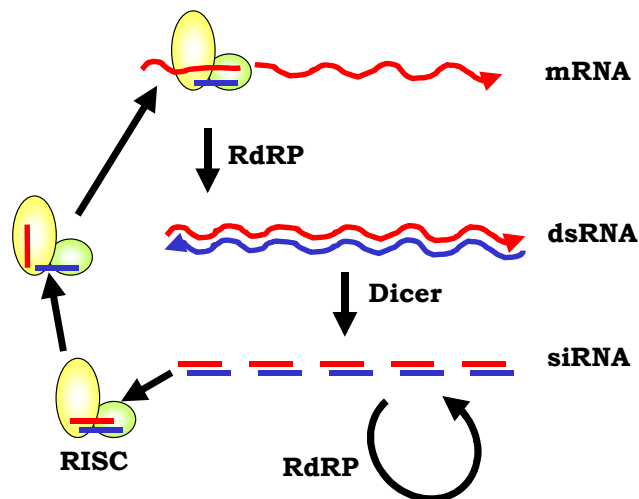
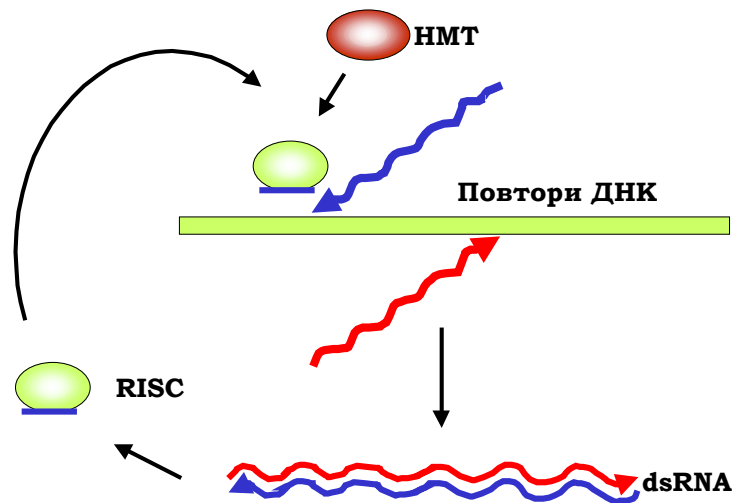


Рис. 6.28. Посттранскрипційне вимкнення гена завдяки РНК-інтерференції

РНК-інтерференція широко використовується останнім часом як дослідницький інструмент (за схемою на рис. 6.28): достатньо ввести в клітину синтетичну дволанцюгову РНК, послідовність якої ідентична послідовності певного гена, щоб запустити процес інтерференції, вимкнути ген і, таким чином, з'ясувати його функцію. Процес РНК-інтерференції розглядається також як перспективний метод лікування певних хвороб.

У клітині РНК-інтерференція є однією із систем негативної регуляції експресії генів через використання так званих мікро-РНК (miRNA). Активація генів мікро-РНК (у геномі людини такі гени, які транскрибуються РНК-полімеразою II, присутні в досить великій кількості, яку остаточно не встановлено) призводить до синтезу молекул РНК, що містять дволанцюгові шпильки. Ці шпильки (довжиною 25–35 пар основ) вирізаються нуклеазою і стають субстратом для дайсера. Результуюча дволанцюгова мікро-РНК є аналогом siRNA: один з її ланцюгів є комплементарним ділянці мРНК певного білкового гена, молекула мікро-РНК зв'язується з білками RISC і спрямовує їх до мРНК-мішені. Результатом взаємодії з мРНК може бути її посттранскрипційна деградація (більш характерно для рослин), зупинка білкового синтезу (що є більш характерним для тварин – у цьому разі взаємодія з мРНК відбувається в цитоплазмі), а також рекрутування до хроматину гістон-метилтрансфераз із наступною репресією даного гена. Дещо подібне відбувається і в гетерохроматині.



**Рис. 6.29.** РНК-інтерференція як механізм підтримання гетерохроматинового стану послідовностей ДНК, що повторюються

Участь РНК-інтерференції в підтриманні гетерохроматинового стану центромер ілюструє рис. 6.29. При порушенні компактизації на центромерних повторах може відбуватися спонтанна транскрипція в різних напрямках. Оскільки матрицею є повтори, існує висока ймовірність синтезу комплементарних молекул РНК: утвориться дволан-

цюгова РНК, яка запустить процес інтерференції. Крім деградації транскриптів, інтерференція має інший наслідок: RISC, який опиняється в зоні повторів, рекрутує до хроматину гістонметилтрансферазу, що здійснює метилювання Lys9 гістону H3. За вже відомою схемою (рис. 6.25, 6.26) відбувається зв'язування HP1 і компактизація гетерохроматину.

### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. Порівняйте структуру РНК-полімерази II з такою бактеріальної полімерази. Що є спільного між ними та чим вони відрізняються?
2. Яку будову й функціональне значення має С-кінцевий домен (CTD) РНК-полімерази II?
3. Яку будову має промотор РНК-полімерази II?
4. Охарактеризуйте базальні фактори транскрипції РНК-полімерази II та основні етапи ініціації транскрипції.
5. Яке значення має білок ТВР? Яку роль відіграють фактори TFIIВ та TFIIH у ініціації транскрипції РНК-полімеразою II?
6. Що таке медіатор? Яку роль він відіграє?
7. Яку спеціалізацію мають еукаріотичні РНК-полімерази різних типів? Порівняйте системи ініціації транскрипції РНК-полімераз I та III. Що спільного між ними та системою ініціації РНК-полімерази II?
8. Що таке енхансосома? Поясніть принципи модульності та кооперативності при взаємодії факторів транскрипції з еукаріотичними промоторами?
9. Що таке гормональний рецептор? Дайте визначення сигнальної трансдукції?
10. Яка посттрансляційна модифікація гістонів позитивно корелює з транскрипційною активністю? У чому полягають механізми такої кореляції?
11. Охарактеризуйте комплекси ремоделювання хроматину? Яким є головний результат їхні дії. У чому полягають механізми такої дії?
12. Яку роль відіграють білки HMG в активації транскрипції?
13. За якими механізмами здійснюється елонгація РНК-полімерази через нуклеосоми?
14. Опишіть функціональну роль білка HP1 у підтриманні гетерохроматинового стану центромер.

15. Яку роль відіграє метилювання ДНК у репресії транскрипції? Що називається CpG-острівцями?

16. Що таке РНК-інтерференція? Яку роль вона відіграє в підтриманні гетерохроматинового стану?

17. Дайте визначення мікро-РНК і опишіть їхню роль у регуляції транскрипції?

### РЕКОМЕНДОВАНА ЛІТЕРАТУРА

*Agresti, A., Bianchi, M.E.* HMGB proteins and gene expression // *Curr. Op. Gen. Dev.* – 2003. – Vol. 13. – P. 170–178.

*Becker, P.B., Hörz, W.* ATP-dependent nucleosome remodeling // *Annu. Rev. Biochem.* – 2002. – Vol. 71. – P. 247–273.

*Bernstein, B.E., Liu, C.L., Humphrey, E.L. et al.* Global nucleosome occupancy in yeast // *Genome Biol.* – 2004. – Vol. 5, № 9. – P. R62 (<http://genomebiology.com/2004/5/9/R62>).

*Boeger, H., Bushnell, D.A., Davis, R. et al.* Structural basis of eukaryotic gene transcription // *FEBS Letters.* – 2005. – Vol. 579. – P. 899–903.

Chromatin structure and gene expression: frontiers in molecular biology / ed. S. Elgin, J.L. Workman. – Oxford : Oxford University Press, 2002.

Chromatin structure and dynamics: state-of-the-art. // *New Comprehensive Biochemistry* / ed. J. Zlatanova, S.H. Leuba. – Amsterdam : Elsevier, 2004. – Vol. 39.

*Geiduschek, E.P., Kassavetis, G.A.* The RNA polymerase III transcription apparatus // *J. Mol. Biol.* – 2001. – Vol. 310. – P. 1–26.

*Hahn, S.* Structure and mechanism of the RNA polymerase II transcription machinery // *Nature Struct. Mol. Biol.* – 2004. – Vol. 11. – P. 394–403.

*Heintzman, N.D., Ren, B.* The gateway to transcription: identifying, characterizing and understanding promoters in the eukaryotic genome // *Cell. Mol. Life Sci.* – 2007. – Vol. 64. – P. 386–400.

*Krauss, G.* Biochemistry of signal transduction and regulation. – Weinheim : WILEI-VCH Verlag, 2003.

*Längst, G., Becker, P.B.* Nucleosome remodeling: one mechanism, many phenomena? // *Biochim. Biophys. Acta.* – 2004. – Vol. 1677. – P. 58–63.

*Maison, C., Almouzni, G.* HP1 and the dynamics of heterochromatin maintenance // *Nature Rev.* – 2004. – Vol. 5. – P. 296–304.

*Matzke, M.A., Birchler, J.A.* RNAi-mediated pathways in the nucleus // *Nature Rev.* – 2004. – Vol. 6. – P. 24–35.

## **Розділ 6. Транскрипція: Еукаріоти**

---

*Orphanides, G., Reinberg, D.* A unified theory of gene expression // *Cell*. – 2002. – Vol. 108. – P. 439–451.

*Russell, J., Zomerdijs, J.C.* The RNA polymerase I transcription machinery // *Biochem. Soc. Symp.* – 2006. – Vol. 73, P. 203–216.

*Richards, E.J., Elgin, S.C.R.* Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects // *Cell*. – 2002. – Vol. 108. – P. 489–500.

*Sims, R.J., Belotserkovskaya, R., Reinberg, D.* Elongation by RNA polymerase II: the short and long of it // *Genes Dev.* – 2004. – Vol. 18. – P. 2437–2468.

*Simpson, R.T.* Nucleosome positioning: occurrence, mechanisms, and functional consequences // *Prog. Nucl. Acid Res. Mol. Biol.* – 1991. – Vol. 40. – P. 143–184.

*Strahl, D.D., Allis, C.D.* The language of covalent histone modifications // *Nature*. – 2000. – Vol. 403. – P. 41–45.

*Studitsky, V.M., Walter, W., Kireeva, M. et al.* Chromatin remodeling by RNA polymerases // *Trends Biochem. Sci.* – 2004. – Vol. 29. – P. 127–135.

## Розділ 7

### ПРОЦЕСИНГ ЕУКАРІОТИЧНИХ мРНК

Ёжик сел на чью-то узкую  
скользкую спину и через минуту  
оказался на берегу.

С. Козлов. Ёжик в тумане

Безпосереднім продуктом еукаріотичної РНК-полімерази II є насправді пре-мРНК – попередник функціональної мРНК. Пре-мРНК не може бути використана як матриця для білкового синтезу – принаймні тому, що вона містить у своєму складі інтрони. Дозрівання пре-мРНК з утворенням функціональної матриці називають *процесингом* (processing). Він складається з трьох операцій:

- *Кепування* – модифікація 5'-кінця з утворенням так званого *кепу* (cap).
- *Сплайсинг* (splicing) – вирізання інтронів і зшивання екзонів – процес, у результаті якого мРНК стає копією лише кодуєної частини гена або її фрагментів: сплайсинг часто може відбуватися кількома альтернативними шляхами (альтернативний сплайсинг).
- *Поліаденілування* – приєднання до 3'-кінця polyA послідовності, яке тісно пов'язане з термінацією транскрипції.

**Усі операції процесингу відбуваються під час транскрипції на РНК-полімеразному комплексі**, тобто процесинг є невід'ємною частиною транскрипції. Місцем збирання машинерії процесингу служить С-кінцевий домен (CTD) РНК-полімерази II, який відходить від полімерази поряд із каналом виходу РНК (розділ 6).

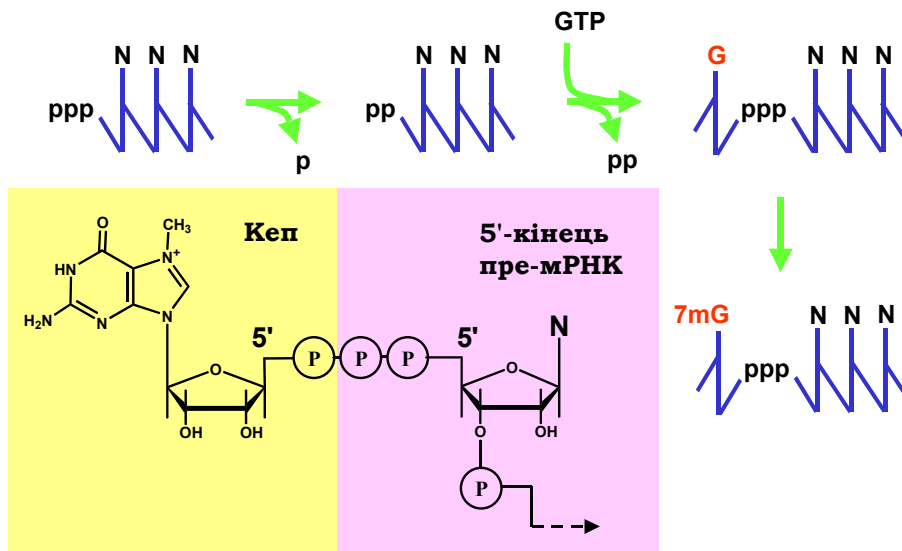
#### Кепування

Хімічну структуру кепу та процес його утворення показано на рис. 7.1. У складі пре-мРНК до 5'-атома рибози першого нуклеозидтрифосфату, який був включений до ланцюга РНК, залишаються приєднаними три фосфатні залишки.

*Перша стадія* кепування: відщеплення одного фосфату ( $\gamma$ -фосфату) із 5'-кінця, що здійснюється фосфатазою.

*Друга стадія*: перенесення гуанозинмонофосфату (GMP) з гуанозинтрифосфату (GTP) на два фосфатні залишки, що залишились на 5'-кінці (відповідний фермент – гуанілілтрансфераза). Результатом перенесення є утворення ковалентного зв'язку між 5'-фосфатом GMP та 5'-кінцевим  $\beta$ -фосфатом пре-мРНК, тобто формування нехарактерного для нуклеїнових кислот зв'язку між нуклеотидами.

*Третя стадія*: метилювання метилтрансферазою кінцевого G з утворенням 7-метилгуаніну (7mG).



**Рис. 7.1.** Стадії утворення та хімічна структура кепу

Три ферментативні активності, які здійснюють кепування, містяться в трьох (дріжджі) або двох (ссавці) білках. Ці ферменти рекрутуються на CTD і 5'-кінець пре-мРНК після синтезу перших 20–30 нуклеотидів транскрипту (рис. 7.2).

Функціонування CTD як платформи для збирання машинерії процесингу мРНК залежить від патерна фосфорильовання залишків Ser, які входять до складу гептапептиду, що повторюється (розділ 6). Наприкінці ініціації транскрипції за рахунок кіназної активності TFIIH здійснюється масоване фосфорильовання Ser5 у складі повто-

рів, що приводить до втрати зв'язку CTD із базальними факторами транскрипції та початку руху полімерази.

Під час руху, що почався, за рахунок спорідненості до фосфорильованого Ser5 із CTD зв'язується білковий фактор DSIF (DRB-Sensitivity Inducing Factor, DRB – синтетичний інгібітор транскрипції, що спрацьовує тільки за наявності фактора), який рекрутує NELF (Negative Elongation Factor). Цей останній складається з п'яти субодиниць, взаємодіє з DSIF, РНК і РНК-полімеразою, зумовлюючи зупинку руху полімерази.

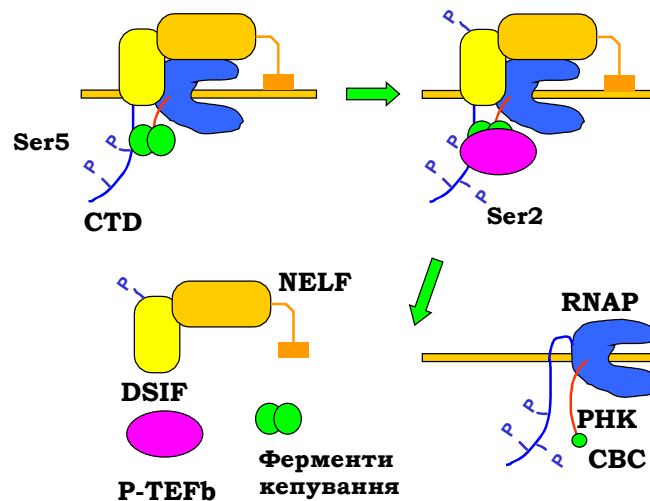


Рис. 7.2. Механізм утворення кепу в координації з транскрипцією

Під час цієї паузи (перший “кадр” на рис. 7.2) із CTD, РНК і DSIF зв'язуються ферменти кепування. Вони спрацьовують і рекрутують P-TEFb (Positive Transcription Elongation Factor b) – гетеродимер, одна із субодиниць якого має кіназну активність: відбувається фосфорильовання Ser2 у складі CTD, а також фосфорильовання DSIF, що приводить до втрати спорідненості DSIF і, відповідно, дисоціації NELF. РНК-полімераза продовжує рух, а сформований кеп відразу зв'язується з двома білками, що формують **CBC** (Cap Binding Complex). Білки CBC залишаються на кепі до моменту транспорту мРНК до цитоплазми, де замінюються на цитоплазматичні фактори.



Функціональне значення кепу є багатоплановим:

- *Захист 5'-кінця від деградації*: особливість зв'язку між першими двома нуклеотидами робить цей зв'язок непомітним для екзонуклеаз.
- *Участь в інших реакціях процесингу*: СВС стимулює сплайсинг першого інтрона та поліаденілування 3'-кінця (див. нижче).
- *Транспорт мРНК у цитоплазму* здійснюється завдяки взаємодії СВС з ядерною порою: саме своїм 5'-кінцем мРНК виштовхується в цитоплазматичний простір.
- *Ініціація трансляції*: саме кеп є первинною точкою збирання рибосоми та інших елементів ініціації білкового синтезу (розділ 8).

### Сплайсинг

Узагальнену схему інтрона зображено на рис. 7.3. На кінцях переважної більшості інтронів розташовані стандартні динуклеотиди GU та AG, що містяться у складі певних консенсусних послідовностей. Ближче до 3'-кінця, перед піримидинзбагаченим треком довжиною ~15 нуклеотидів існує ще один консенсус, до складу якого обов'язково входить аденіновий нуклеотид. Цей А є точкою розгалуження (branch point), яка виконує особливу роль.

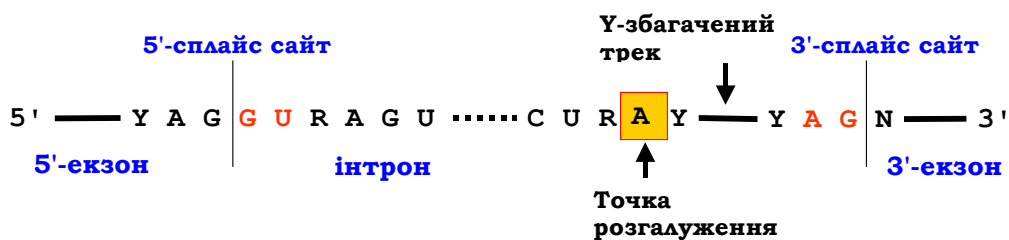


Рис. 7.3. Узагальнена схема послідовності нуклеотидів інтрона

Сплайсинг інтрона полягає у двох послідовних хімічних реакціях *трансестерифікації* (заміни одного фосфодієфірного зв'язку на інший), які показано на рис. 7.4.

1. Аденіновий нуклеотид у точці розгалуження має підвищену реакційну здатність (унаслідок особливостей просторової структури інтрона, які розглядатимуться нижче). 2'-ОН група його рибози здійснює

## Розділ 7. Процесинг еукаріотичних мРНК

нуклеофільну атаку фосфату в 5'-сплайс сайті – на межі між лівим екзоном та інтроном. У результаті зв'язок між цим фосфатом і 3'-кінцевим нуклеотидом екзона замінюється на зв'язок між фосфатом і 2'-ОН групою аденінового нуклеотиду – у 5'-кінцевій частині інтрона утворюється так зване *ласо* (lariat).

2. ОН-група, що залишилася на 3'-кінці першого екзона атакує фосфодіестерний зв'язок у 3'-сплайс-сайті. Цей зв'язок розривається, замінюючись на зв'язок між двома екзонами.

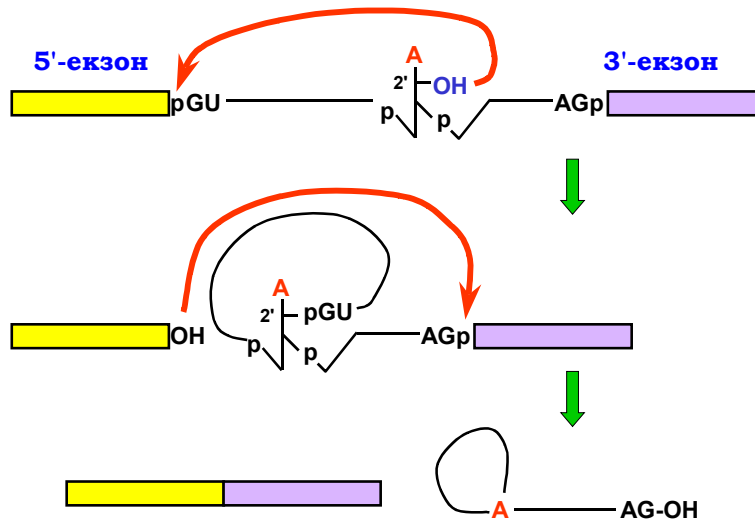


Рис. 7.4. Дві стадії сплайсингу

Оскільки кількість фосфодіестерних зв'язків не змінюється, реакція трансестерифікації є ізоенергетичною і не потребує АТФ як джерела енергії. Але безумовно, обидві реакції сплайсингу потребують каталізу, і, на відміну від інших реакцій, що розглядалися досі, **каталізаторами сплайсингу не виступають білкові ферменти.**

### Сплайсосома: механізм сплайсингу

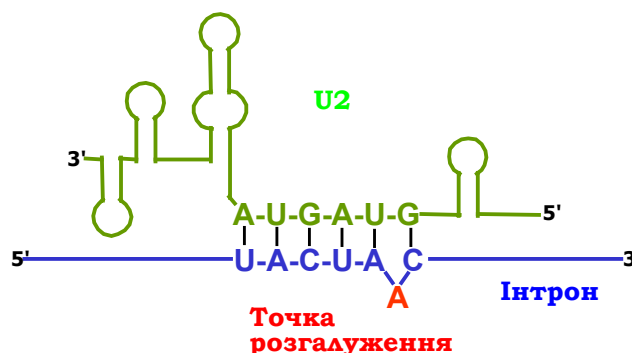
Обидві реакції сплайсингу здійснюються у складі спеціальної мультимолекулярної структури – *сплайсосоми* (spliceosome), яка зв'язана з СТД і утворюється на кожному інтроні за участю самої пре-мРНК, білків і особливих молекул *маленьких ядерних РНК* (small nuclear RNA, snRNA).

**Маленькі ядерні РНК**, які синтезуються РНК-полімеразою II на відповідних генах, згрупованих у кластери, **відіграють ключову роль у визначенні просторової структури, формуванні та функціонуванні сплайсосоми**. У сплайсингу беруть участь п'ять типів маленьких ядерних РНК (довжиною 100–200 нуклеотидів): U1, U2, U4, U5 та U6.

Маленькі ядерні РНК, які існують у клітині у вигляді комплексів зі специфічними білками, мають певну просторову структуру за рахунок утворення комплементарних подвійних спіралей – шпильок. Завдяки комплементарному спарюванню з консенсусними послідовностями інтрона відбувається їхня взаємодія з пре-мРНК та іншими маленькими ядерними РНК. Наприклад, РНК U2 впізнає консенсус у зоні розгалуження, причому згаданий *адениновий нуклеотид залишається неспареним* – “випетлююється” між сусідніми парами основ (рис. 7.5). Саме ця обставина визначає його особливий конформаційно-напружений стан і, як наслідок, підвищену реакційну здатність.

Оскільки подвійні спіралі РНК у сплайсосомі короткі, вони потребують додаткової стабілізації. Крім того, під час збирання та функціонування сплайсосоми певні спіралі мають бути зруйновані й замінені на інші. Обидві операції забезпечуються білками сплайсосоми. Отже, **роль білків сплайсосоми** зводиться до:

- розкручування подвійних спіралей РНК АТР-залежними геліказами;
- стабілізації подвійних спіралей і загальної просторової структури сплайсосоми білками зі специфічною спорідненістю до РНК;
- регуляції сплайсингу – блокування чи підсилення ефективності збирання сплайсосоми на даному інтроні білками-регуляторами сплайсингу (SR, Splicing Regulators).



**Рис. 7.5.** Схема структури маленької ядерної РНК U2 у комплексі з зоною розгалуження інтрона

Загальний сценарій **збирання та роботи сплайсосоми** схематично зображено на рис. 7.6. Після синтезу 5'-кінцевої зони інтрона, вона впізнається маленькою ядерною РНК U1 за рахунок комплементарного спарювання між U1 і 5'-сплайс сайтом (U1 при цьому зв'язана з відповідними білками). Далі в міру синтезу пре-мРНК зона розгалуження впізнається специфічним білком ВВР (Branch point Binding Protein, позначається також як SF1 у ссавців). 3'-Кінцева зона інтрона та Y-збагачений трек упізнаються допоміжним білком U2AF (U2 Auxiliary Factor), який далі сприяє заміні ВВР на РНК U2. Зв'язування U2 (яке, власне, визначає точку розгалуження, див. рис. 7.5) потребує АТР-залежного руйнування певних подвійних спіралей у складі цієї молекули. На наступних стадіях збирання та перебудов сплайсосоми (рис. 7.6) також відбувається АТР-залежне руйнування частини подвійних спіралей і заміна їх іншими.

Після зв'язування U2, на завершальному етапі збирання, з інтроном взаємодіє потрійний комплекс U4-U5-U6: U4 при цьому звільнюється, а U6 витісняє U1, взаємодіючи з 5'-кінцевою зоною інтрона. Крім того, U6 взаємодіє з U2, сприяючи наближенню 5'-сплайс-сайта до точки розгалуження. Додатково структура стабілізується маленькою ядерною РНК U5, яка взаємодіє з 3'-кінцем першого екзона та іншими елементами сплайсосоми.

Після формування сплайсосоми здійснюється її структурна перебудова (АТР-залежне розплітання частини подвійних спіралей геліказами), результатом якої є безпосереднє наближення 5'-сплайс сайта до точки розгалуження: виникають умови для першої реакції трансестерифікації з утворенням ласо (див. також рис. 7.4). Хімічна перебудова інтрона викликає нову перебудову просторової структури сплайсосоми: за рахунок одночасної взаємодії з U5 наближуються один до одного кінці екзонів, що створює умови для другої трансестерифікації. На останньому етапі від двох уже з'єднаних екзонів АТР-залежним шляхом (порушення комплементарних взаємодій) видаляється комплекс інтрона з маленькими ядерними РНК.

Отже, сплайсосома працює як АТР-залежна молекулярна машина. Структурні перебудови машини забезпечують перегрупування елементів – субстратів реакцій сплайсингу. **Каталіз обох реакцій сплайсингу здійснюється молекулами РНК.** Механізм каталізу є таким самим, як для білкових ферментів (див. розділ 2). Наприклад, для першої реакції активний центр формується зоною розгалуження інтрона й маленькими ядерними РНК U2 та U6. Просторова структура активного центру жорстко утримує субстрати – 5'-сплайс сайт і точку розгалу-

ження – у певній взаємній орієнтації та забезпечує підвищену реакційну здатність аденінового нуклеотиду, який можна розглядати одночасно і як субстрат у збудженому проміжному стані, і як компонент активного центру. У хлоропластах сплайсинг відбувається за механізмом так званого самосплайсингу (self-splicing), тобто без участі білків і маленьких РНК: мРНК сама набуває просторової структури, яка має каталітичну активність щодо власного сплайсингу.

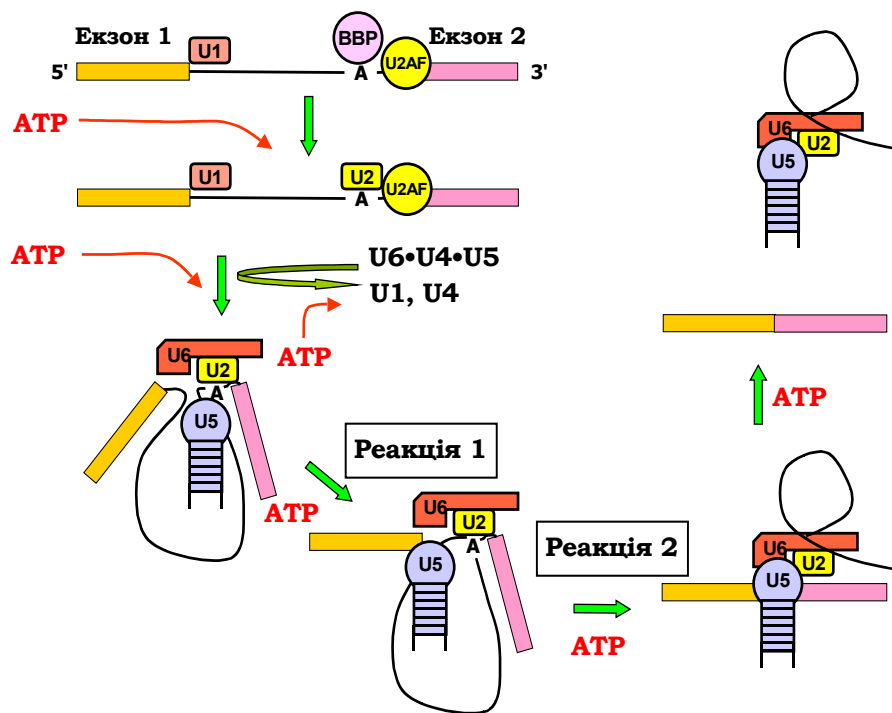


Рис. 7.6. Схема збирання сплайсоми та її функціонування

За аналогією з білковими *ензимами*, молекули РНК, які мають каталітичну активність, називають *рибозимами*. Одним із таких рибозимів є сплайсома. Зазвичай каталітична активність рибозима (як і у випадку сплайсоми) залежить від білків, які виконують допоміжні функції, стабілізуючи структуру активного центру. Вважається, що на ранніх етапах добіологічної еволюції головними (або єдиними) біологічними макромолекулами були молекули РНК (так званий РНК-світ), оскільки це єдиний тип макромолекул, що можуть одночасно виконувати роль носіїв спадкової інформації та виступати каталізаторами.

## Розділ 7. Процесинг еукаріотичних мРНК

Пізніше більш стабільні молекули ДНК перебрали на себе роль носіїв інформації, а різноманітніші за просторовою структурою білки – роль каталізаторів. Але в кількох важливих випадках (і сплайсинг – не єдиний приклад) рибозими виявилися еволюційно консервативними.

Сплайсосоми утворюються на інтронах пре-мРНК (і на STD РНК-полімерази) послідовно під час транскрипції – майже відразу після синтезу сплайс-сайтів вони впізнаються відповідними елементами (рис. 7.7). При цьому СВС, що зв'язаний з кепом, підсилює ефективність зв'язування U1 у першому 5'-сплайс-сайті – напевно, шляхом прямої взаємодії; вплив СВС на сплайсинг наступних інтронів послаблюється.

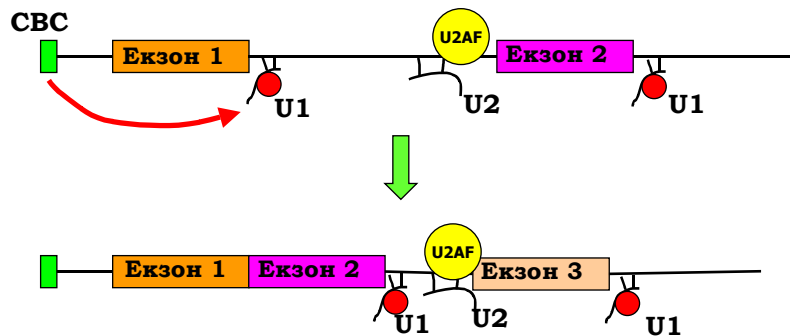


Рис. 7.7. Послідовне збирання сплайсосом на інтронах під час транскрипції

Швидкість збирання сплайсосом визначається швидкістю транскрипції. З іншого боку, зі сплайсосомою взаємодіють фактори елонгації транскрипції – наявність сплайс-сайта сприяє прискоренню руху полімерази. Такий процес синтезу РНК та її сплайсингу продовжується до того моменту, поки у складі пре-РНК не з'являється специфічна послідовність – сигнал термінації.

### Поляденілування мРНК і термінація транскрипції

Сигнал термінації транскрипції та поляденілування (так званий *polyA-сигнал*) складається з двох елементів послідовності: консенсус AAUAAA і розташована в ~20 нуклеотидах нижче від нього U- або G/U-збагачена послідовність (рис. 7.8). Перший елемент упізнається гетеротетрамерним білком CPSF (Clevage-Polyadenylation Specificity

Factor). Три із чотирьох субодиниць фактора є компонентами базального фактора транскрипції TFIID – вони переносяться на CTD після ініціації, і далі полімераза несе їх на собі, доки вони не зустрінуть сигнал. Другий елемент polyA-сигналу впізнається фактором розрізання РНК CstF (Cleavage stimulation Factor). Взаємодія обох факторів з РНК є кооперативною – вони підсилюють зв'язування одне одного.

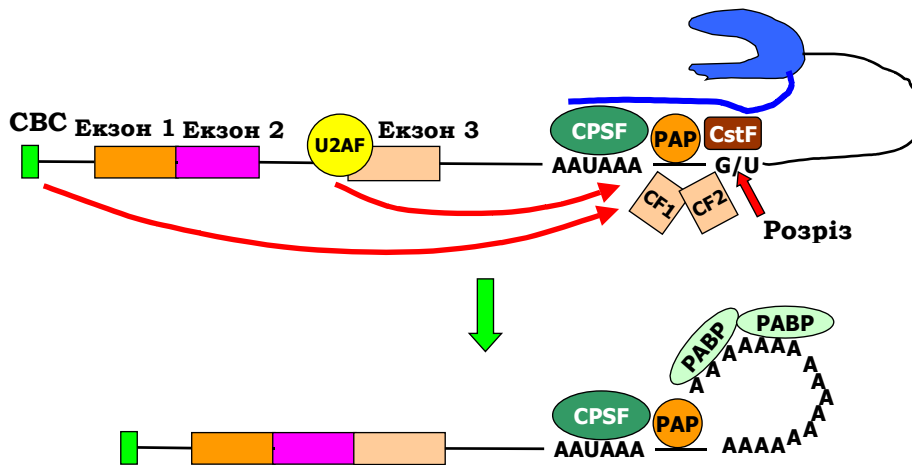
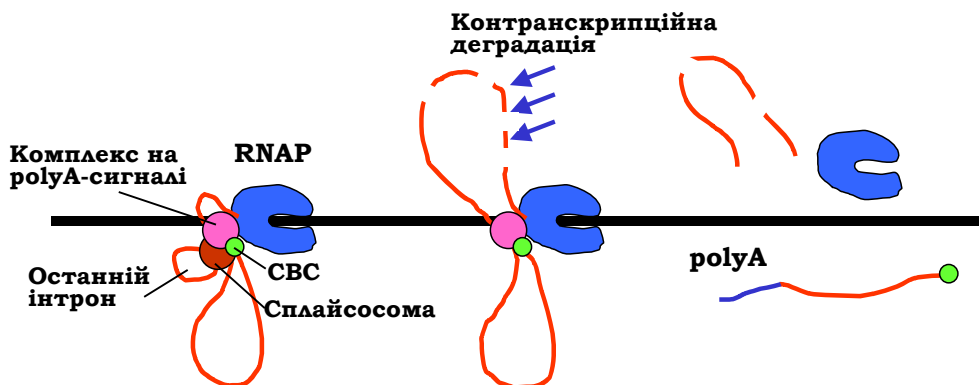


Рис. 7.8. Впізнання polyA-сигналу та поліаденілування мРНК

Після первинного впізнання polyA-сигналу, поки РНК-полімераза продовжує синтез РНК за сигналом (до 1000 нуклеотидів), до мультибілкового комплексу, що збирається на polyA-сигналі, долучаються polyA-полімераза (PAP), ще два фактори розрізання (Cleavage Factors) CF 1 і 2 та, можливо, інші білки. Збирання комплексу стимулюється зв'язаним з кепом CBC, а також сплайсосою на останньому інтроні – сплайсинг останнього інтрона та розрізання / поліаденілування РНК здійснюються одночасно і стимулюють одне одного. Зокрема, білок U2AF, який знаходиться на останньому інтроні, взаємодіє з PAP.

У межах другого елемента послідовності polyA-сигналу знаходиться консервативний динуклеотид CA, в якому й відбувається розрізання (cleavage) РНК. Від якої конкретно активності залежить це розрізання, залишається невідомим. До 3'-кінця, що виник внаслідок розрізу, PAP (за стимулюючої дії CPSF) приєднує один за одним 100–200 аденінових нуклеотидів. З polyA-хвостом, що зростає, відразу зв'язується специфічний білок PABP (PolyA Binding Protein), який підвищує процесивність PAP (рис. 7.8).

Розрізання / поліаденілування РНК є тригером термінації транскрипції (рис. 7.9). Напевно, упізнання polyA-сигналу та розрізання / поліаденілування факторами, зв'язаними з РНК-полімеразою, індукує конформаційні зміни в полімеразному комплексі, які приводять до зниження спорідненості полімерази до ДНК та / або транскрипту. Крім того, процес розрізання та поліаденілування синхронізовані з так званою котранскрипційною деградацією РНК нуклеазами нижче від polyA-сигналу. Механізми активації такої деградації не з'ясовані, але швидке руйнування транскрипту має стимулювати дисоціацію полімеразного комплексу. Термінація транскрипції іноді потребує також інших регуляторних елементів послідовності. Одні з них можуть знаходитись у межах останнього інтрона та стимулювати / блокувати впізнання polyA-сигналу, інші – нижче від polyA-сигналу й зумовлювати паузи в роботі полімерази, які сприяють термінації.



**Рис. 7.9.** Синхронізація поліаденілування і термінації транскрипції

Узагальнену схему будови зрілої еукаріотичної мРНК, яка звільняється з полімеразного комплексу, транспортується в цитоплазму і використовується як матриця для білкового синтезу, зображено на рис. 7.10. Між кепом і початком кодуючої ділянки (стартовим кодоном – найчастіше AUG) розташована 5'-кінцева зона, що не трансклюється (5' UTR – UnTranslated Region). За кодуючою ділянкою, що закінчується одним із стоп-кодонів, і перед polyA-послідовністю розташована 3'-кінцева зона, що не піддається трансляції. Обидві зони, що не трансклюються, містять важливі елементи послідовності, які використовуються для регуляції білкового синтезу (розділ 8).



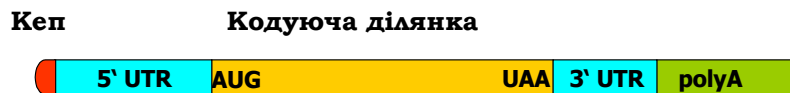


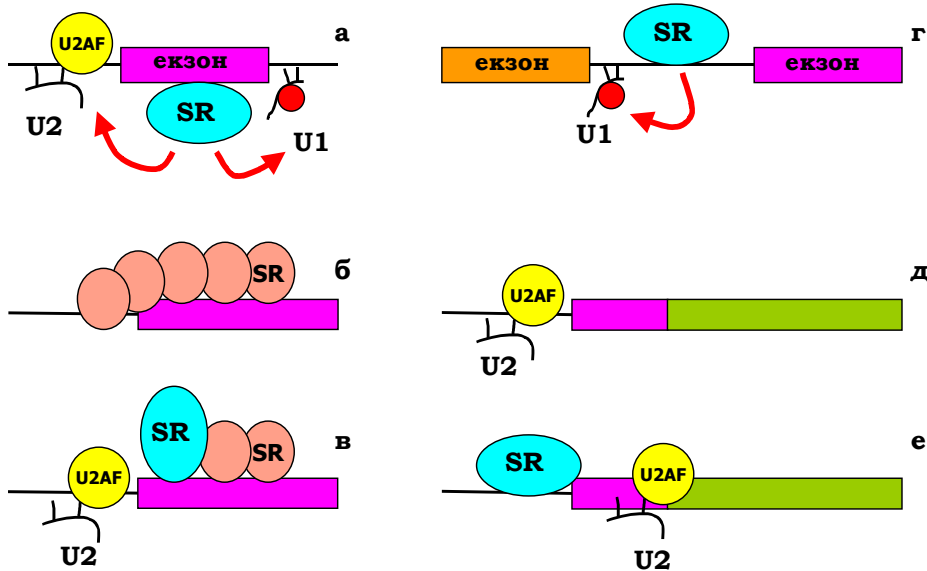
Рис. 7.10. Схема будови мРНК

Майже єдиним винятком серед інших еукаріотичних мРНК є гістонові мРНК, які не піддаються сплайсингу (гістонові гени не містять інтронів) і поліаденілуванню 3'-кінця. Термінація транскрипції гістонових генів залежить від двох сигнальних елементів послідовності РНК: один утворює шпильку (як у прокаріотів, див. розділ 5), інший спарюється з маленькою ядерною РНК U7; між цими елементами індукується розріз. Аналогічно відбувається термінація транскрипції маленьких ядерних РНК, які також не піддаються поліаденілуванню.

### Альтернативний сплайсинг

Пре-мРНК, що синтезується під час транскрипції, може піддаватися сплайсингу та поліаденілуванню різними альтернативними шляхами: кілька екзонів на початку чи всередині гена можуть вирізатися з транскрипту, останній може обрізатися та піддаватися поліаденілуванню за рахунок використання polyA-сигналу всередині одного з інтронів тощо. У результаті утворюються різні молекули мРНК, що містять різні набори екзонів і, відповідно, кодують різні білки. Багатокомпонентність (і необхідність кооперації між компонентами) сплайсосоми та системи розрізання / поліаденілування дозволяє здійснювати тонку регуляцію утворення мРНК певного типу за рахунок зміни транскрипційної активності генів, що кодують ті чи інші компоненти машинерії процесингу, зміни концентрацій компонентів, їхньої хімічної модифікації.

Ключова роль у визначенні шляху сплайсингу належить білкам-регуляторам сплайсингу (SR). Досить часто регулятор специфічно зв'язується з певною послідовністю нуклеотидів усередині ексона та стимулює впізнання сплайс-сайтів по обидва боки від нього (рис. 7.11, а). Зрозуміло, що в разі відсутності регулятора такий екзон буде вирізано разом з інтронами, що його фланкують.



**Рис. 7.11.** Кілька варіантів участі сплайсинг-регуляторів, що зв'язуються всередині екзона (а, б, в) або інтрона (г, д, е), у визначенні шляху сплайсингу

У випадку негативної регуляції зв'язування регулятора утруднює впізнання сплайс-сайтів. Наприклад, взаємодія з регулятором ініціюється всередині екзона, молекули білка кооперативно зв'язуються поруч одна з одною і в решті-решт блокують 3'-сплайс-сайт (рис. 7.11, б). Замість заблокованого сприймається наступний такий сайт, і екзон вирізається з молекули разом із двома інтронами, що його фланкують. За наявності іншого регулятора, який також упізнає специфічну послідовність в екзоні, процес зростання білкового комплексу зупиняється, і 3'-сплайс-сайт стає вільним для впізнання елементами сплайсосоми (рис. 7.11, в).

Сплайсинг-регулятори можуть також мати сайти зв'язування в інтронах, впливаючи на ефективність збирання сплайсосоми (рис. 7.11, г). Регулятор може блокувати, наприклад, 3'-сплайс-сайт: якщо така блокада відсутня, сайт розпізнається (рис. 7.11, д), за наявності регулятора сприймається інший 3'-сплайс-сайт усередині екзона – у складі мРНК залишається скорочений екзон (рис. 7.11, е).

Вибір шляху утворення кінцевого мРНК-продукту часто залежить також від альтернативного вибору роуА-сигналів, які можуть бути присутніми не лише нижче останнього екзона, а й усередині інтронів.

Приклад такого типу зображено на рис. 7.12. Узагалі, якщо polyA-сигнал міститься всередині інтрона, його впізнання блокується 5'-сплайс-сайтом – асоційованими з U1 білками (polyA-сигнал, що розташований після термінального екзона, не блокується, оскільки немає 5'-сплайс-сайта). У результаті такої блокади polyA-сигнал не впізнається і вирізається разом з інтроном. Але якщо діє специфічний сплайсинг-регулятор, який заважає впізнанню 5'-сплайс-сайта або взаємодії U1-білків з елементами системи впізнання polyA-сигналу, сигнал упізнається, і мРНК втрачає всі екзони, розташовані нижче від сигналу. Наприклад, сплайсинг мРНК важкого ланцюга імуноглобуліну IgM спрямовується першим шляхом (внутрішній polyA-сигнал ігнорується) у В-лімфоцитах, що розвиваються, – синтезується мембранна форма білка, яка містить С-кінцевий домен, зв'язаний з мембраною. У плазматичних клітинах спрацьовує внутрішній polyA-сигнал (упізнання сигналу підсилюється також завдяки зростанню концентрації CstF), і синтезується секреторна форма білка.

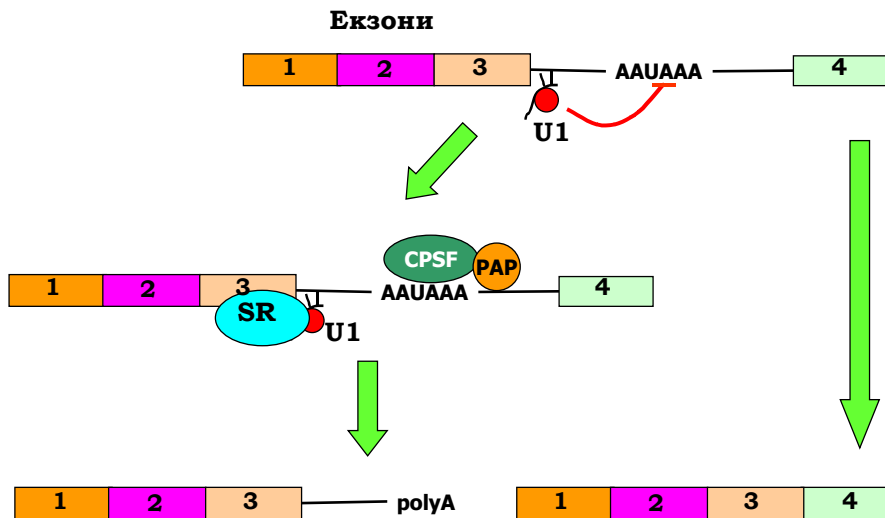
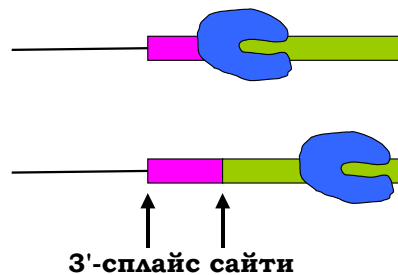


Рис. 7.12. Приклад двох шляхів процесингу, вибір яких залежить від впізнання polyA-сигналу всередині інтрона

Поряд із концентрацією елементів машинерії процесингу та сплайсинг-регуляторів, досить важливим фактором регуляції сплайсингу є швидкість транскрипції (як у розумінні швидкості руху полімерази, так і в сенсі кількості полімераз, що одночасно працюють на даному гені).

## Розділ 7. Процесинг еукаріотичних мРНК

У випадку інтрона середньої довжини 3'-сплайс-сайт синтезується РНК-полімеразою через 15–30 хвилин після 5'-сплайс-сайта. За цей час упізнання 5'-сплайс-сайта вже завершилося, і перший 3'-сплайс-сайт упізнається як такий. Якщо за досить короткий час синтезується кілька 3'-сплайс-сайтів, виникає конкуренція між ними за відповідні елементи сплайсосоми. При цьому перший із сайтів буде вигравати при низькій швидкості руху полімерази (рис. 7.13).



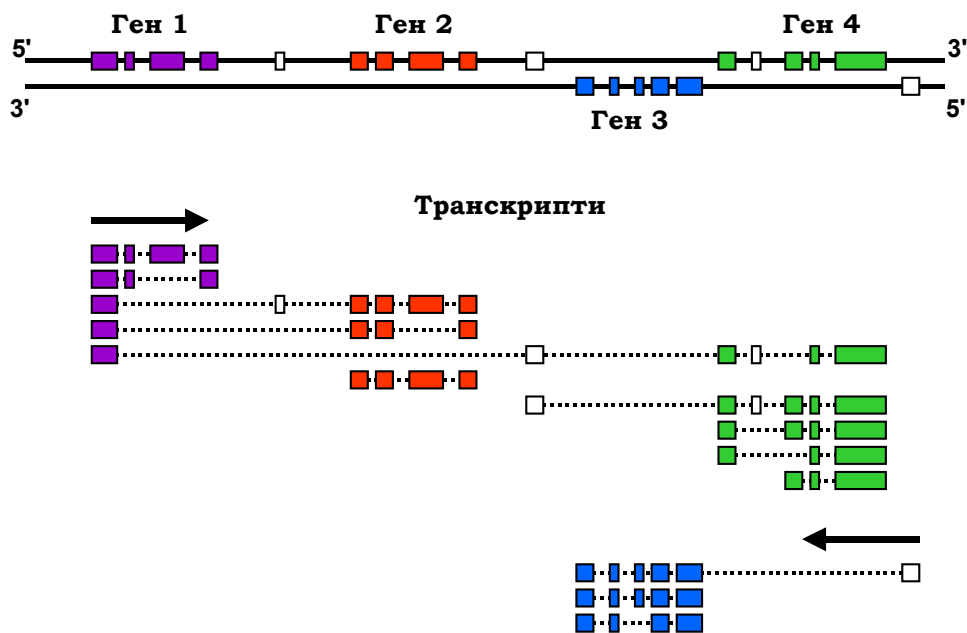
**Рис. 7.13.** Різні положення РНК-полімерази через однаковий проміжок часу внаслідок різниці у швидкості транскрипції

Швидкість транскрипції в розумінні кількості полімераз, що одночасно працюють на даному гені, також впливає на процесинг: чим нижчою є швидкість, тим більше часу “розібратися” із сигналами, на великій швидкості полімераза може “проскочити” через екзон чи *polyA*-сигнал. При цьому слід зауважити, що швидкість (частота рекрутування полімераз до промотора) визначається на стадії ініціації та контролюється транскрипційними факторами. Отже, значною мірою шлях процесингу визначається вже на етапі ініціації транскрипції. При цьому до промотора (відповідно, і до РНК-полімеразного комплексу) можуть рекрутуватися також і регулятори сплайсингу. Тому на химерних генах із чужорідним промотором часто виявляються дефекти процесингу.

### Транс-сплайсинг

Для еукаріотичних генів також досить характерним є явище *транс-сплайсингу*: об'єднання у складі мРНК екзонів різних генів. Основою для транс-сплайсингу є та обставина, що для багатьох генів існує не одна, а кілька альтернативних стартових точок транскрипції,

у тому числі такі, що розташовані досить далеко від гена й водночас є стартовими точками інших генів (рис. 7.14). У результаті транскрипція іноді здійснюється через кілька генів (така ситуація дещо нагадує прокаріотичні оперони, розділ 5), і сплайсинг відбувається на рівні таких “об’єднаних” первинних транскриптів.



**Рис. 7.14.** Зона еукаріотичного геному: показано чотири гени та відповідні екзони (прямокутники) у складі змістовних ланцюгів. Прямокутники білого кольору важко віднести до конкретного гена. Унизу: кінцеві транскрипти, синтезовані у цій зоні, стрілки вказують напрямок транскрипції

Крім того, описано також випадки транс-сплайсингу між пре-мРНК, які є продуктами різних одиниць транскрипції, іноді такими, що розташовані на різних хромосомах. У цьому випадку на 5'- і 3'-кінцях двох пре-мРНК міститься ніби “розірваний” інтрон – об’єднання двох кінців у сплайсосомі приводить до видалення цього інтрона та зшивання кінцевих екзонів. Найкраще це явище вивчено для нематод, але, імовірно, воно є досить поширеним для еукаріотів узагалі.

### **Редагування мРНК**

У кількох особливих випадках процес дозрівання мРНК включає також різноманітні модифікації, які об'єднують під назвою *редагування* мРНК (mRNA editing). Редагування, яке найчастіше відбувається під час сплайсингу або відразу після нього, не є типовим еукаріотичним процесом, хоча й зустрічається практично в усіх таксонах.

У мітохондріях простіших більшість мРНК піддається редагуванню за РНК-залежним механізмом. Так звана гідова РНК (guide RNA) спарюється з комплементарною ділянкою мРНК та ініціює збирання мультибілкової едитосоми (editosome), компоненти якої забезпечують видалення чи інсерцію уридину в певне місце ділянки мРНК.

У мітохондріях і хлоропластах рослин часто відбувається дезамінування цитидину із заміною його на уридин, що змінює кодуючі властивості багатьох мРНК.

Дезамінування аденіну в певному місці ланцюга з перетворенням його на інозин (неканонічна пуринова основа, див. розділ 8) відбувається в мРНК різних субодиниць Glu-залежного Ca<sup>2+</sup>-каналу нервових клітин ссавців. Наслідком нуклеотидної заміни є амінокислотні заміни у складі білка (усередині трансмембранної пори); комбінація різних (редагованих та не редагованих) субодиниць створює розмаїття каналів щодо їхньої чутливості до хімічного сигналу (амінокислоти Glu) і швидкості потоку Ca<sup>2+</sup> через мембрану. Редагування такого типу залежить від активності спеціальних ферментів (аденозин деаміназ). Вважається, що вони зв'язуються із дволанцюговими шпильками, які тимчасово формуються в мРНК під час синтезу, де впізнають певні елементи послідовності пар основ.

Іншим прикладом є перетворення шляхом дезамінування певного цитидину на уридин у мРНК аполіпобілка В (ApoB, apolipoprotein B) печінки ссавців. Заміна викликає перетворення глютамінового кодона на стоп-кодон. У результаті білок може існувати у двох формах з різною довжиною поліпептидного ланцюга. Обидві форми використовуються для транспорту тригліцеридів і холестерину через кров, причому довша форма пов'язана з підвищенням ризику атеросклерозу. Редагування мРНК ApoB, яке відбувається паралельно з поліаденілуванням, залежить від певної послідовності мРНК, котра впізнається комплексом білків, серед яких – специфічна деаміназа.

### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. В яких операціях полягає процесинг мРНК? Де й коли він здійснюється?
2. Опишіть загальну будову еукаріотичної мРНК.
3. Розкажіть про хімічну будову кепу. У чому полягає його функціональне значення?
4. Як і коли відбувається кепування 5'-кінця мРНК?
5. Опишіть послідовність етапів сплайсингу. Що таке точка розгалуження?
6. Дайте визначення сплайсоми. З яких елементів вона складається, де й коли формується? Опишіть основні етапи її утворення.
7. Як здійснюється каталіз хімічних реакцій сплайсингу? Що таке рибозим?
8. З якою метою при збиранні та роботі сплайсоми використовується енергія гідролізу АТФ?
9. Як відбувається термінація транскрипції РНК-полімеразою II?
10. Охарактеризуйте основні механізми альтернативного сплайсингу.

### **РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

- Colgan, D.F., Manley, J.L.* Mechanism and regulation of mRNA polyadenylation // *Genes Dev.* – 1997. – Vol. 11. – P. 2755–2766.
- The ENCODE Project Consortium. Identification and analysis of functional elements in 1 % of the human genome by the ENCODE pilot project // *Nature.* – 2007. – Vol. 447. – P. 799–816.
- Gerstein, M.B., Bruce, C., Rozowsky, J.S. et al.* What is a gene, post-ENCODE? History and updated definition // *Genome Res.* – 2007. – Vol. 17. – P. 669–681.
- In the RNA world / eds. R.F. Gesteland, J.F. Atkins.* – Cold Spring Harbor ; New York : CPHL Press, 1993.
- Jurica, M.S., Moore, M.J.* Pre-mRNA splicing: awash in a sea of proteins // *Mol. Cell.* – 2004. – Vol. 12. – P. 5–14.
- Matlin, A.J., Clark, F., Smith, C.W.J.* Understanding alternative splicing: towards a cellular code // *Nature Rev.* – 2005. – Vol. 6. – P. 386–398.

## **Розділ 7. Процесинг еукаріотичних МРНК**

---

*Palancade, B., Bensaude, O.* Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation // *Eur. J. Biochem.* – 2003. – Vol. 270. – P. 3859–3870.

*Sims, R.J., Belotserkovskaya, R., Reinberg, D.* Elongation by RNA polymerase II: the short and long of it // *Genes Dev.* – 2004. – Vol. 18. – P. 2437–2468.

*Staley, J.P., Guthrie, C.* Mechanical devices of the spliceosome: motors, clocks, springs, and things // *Cell.* – 1998. – Vol. 92. – P. 315–326.

*Stuart, K., Panigrahi, A.K.* RNA editing: complexity and complications // *Mol. Microbiology.* – 2002. – Vol. 45. – P. 591–596.

*Zorio, D.A.R., Bentley, D.L.* The link between mRNA processing and transcription: communication works both ways // *Exp. Cell Res.* – 2004. – Vol. 296. – P. 91–97.



## Розділ 8

### СИНТЕЗ БІЛКІВ

So he started to climb out of the hole... and in a little while his nose was out... and then his ears... and then his front paws... and then his shoulders...

*A. Miln.* Winnie the Pooh  
and all, all, all

Інформація щодо амінокислотної послідовності білка записана у вигляді послідовності нуклеотидів мРНК у відповідності з генетичним кодом (розділ 4). Зчитування цієї інформації та її переклад у амінокислотний текст (*трансляція*, translation) має розпочинатися зі стартового кодону, де при *ініціації трансляції* відбувається остаточне збирання головного пристрою трансляції – *рибосоми* (ribosome). Рибосома – комплекс рибосомної РНК і білків – сканує нуклеотидну послідовність мРНК, рухаючись уздовж неї кроками по три нуклеотиди від 5'- до 3'-кінця під час *елонгації трансляції* до стоп-кодону, де відбувається *термінація* процесу. Під час сканування рибосома працює як декодуєчий пристрій, забезпечуючи впізнання кодонів комплементарними триплетами (*антикодонами*) у складі *тРНК* (транспортні РНК, tRNA – transfer RNA), і як каталізатор процесу синтезу пептидного зв'язку між амінокислотами. Певний антикодон відповідає амінокислоті певного типу, яку несе на собі тРНК. Отже, тРНК виступають ключовою ланкою реалізації генетичного коду: саме вони забезпечують доставку амінокислот до рибосоми у порядку, який відповідає послідовності кодонів.

#### Транспортні РНК

##### *Структура тРНК*

Молекули тРНК містять 74–95 (найчастіше 76) нуклеотидів, серед яких зустрічаються неканонічні (утворюються внаслідок посттранскрипційних модифікацій) – тимідин, інозин (I), дигідроуридин (D), псевдоуридин ( $\psi$ ) тощо.

У складі молекули формуються комплементарні дволанцюгові стебла та шпильки за єдиною для всіх тРНК схемою, що нагадує лист конюшини (рис. 8.1). Кінцеві фрагменти ланцюга об'єднуються у дволанцюгове стебло, причому чотири нуклеотиди на 3'-кінці залишаються неспареними. 3'-Кінцевий триплет ССА є стандартним для всіх тРНК, до рибози кінцевого аденозину ковалентно приєднується амінокислота: відповідно, стебло називають **акцепторним**. 5'-Кінцева частина акцепторного стебла переходить у шпильку з петлею, яка часто містить дигідроуридин – D-стебло і D-петля. У деяких тРНК D-петля може містити на один нуклеотид менше або на один–три нуклеотиди більше відносно типового розміру петлі, зображеного на рис. 8.1. D-стебло переходить у стебло з петлею, у складі якої розташований антикодоновий триплет (у позиції 34–36), – **антикодонове стебло / антикодонова петля**. За антикодоновим стеблом розташована *варіабельна петля* (V-петля). За розміром V-петлі всі тРНК можна поділити на два структурні класи: до першого належить більшість тРНК із короткою петлею (3–5 нуклеотидів), до другого – кілька тРНК із довгою (до 16 додаткових нуклеотидів) петлею. Нумерація нуклеотидів на рис. 8.1 є стандартною, додаткові нуклеотиди D- та V-петлі, якщо вони присутні, нумерують окремо. За варіабельною петлею міститься ТψС-стебло з петлею (у складі петлі часто зустрічається консенсус ТψС), яке переходить у 3'-кінцеву частину акцепторного стебла.

тРНК конкретного типу, котра відповідає певній амінокислоті, позначають індексом, наприклад, тРНК<sup>Ala</sup>. Якщо в молекулі тРНК є амінокислота, то таку аміноацильовану тРНК позначають як Ala-тРНК<sup>Ala</sup>. Загальне позначення для аміноацильованих тРНК – aa-тРНК (aa-tRNA – aminoacyl-tRNA).

Еукаріотичні гени тРНК (близько 500 активних генів тРНК у геномі людини, частина яких зібрана в кластери) транскрибуються РНК-полімеразою III. Продуктом транскрипції генів є довші молекули-попередники, деякі містять інтрон у межах майбутньої антикодонової петлі. Процесинг цих попередників з утворенням зрілих тРНК полягає у відщепленні певними нуклеазами зайвих фрагментів на кінцях, сплайсингу інтрона (здійснюється специфічними ендонуклеазами та лігазою), приєднанні до 3'-кінця стандартного триплету ССА (у складі прокаріотичних тРНК-попередників цей триплет уже присутній), хімічній модифікації певних азотистих основ. Прокаріотичні гени тРНК (87 у геномі *E. coli*, не містять інтронів) або транскрибуються окремо (як у еукаріотів), або є частинами оперонів, і в цьому випадку

первинний транскрипт містить кілька майбутніх молекул тРНК. Крім того, деякі прокаріотичні гени тРНК знаходяться у складі оперона генів рибосомної РНК. В усіх випадках тРНК вирізаються з попередників нуклеазами, після чого піддаються хімічним модифікаціям. Одна з нуклеаз, котра бере участь у процесингу тРНК у про- та еукаріотів, – РНКаза Р – варта особливої уваги. Фермент складається з двох субодиниць, одна з яких є білковою, а друга – молекулою РНК зі складною просторовою структурою. Саме ця РНК-субодиниця має каталітичну активність, тобто РНКаза Р – рибозим.

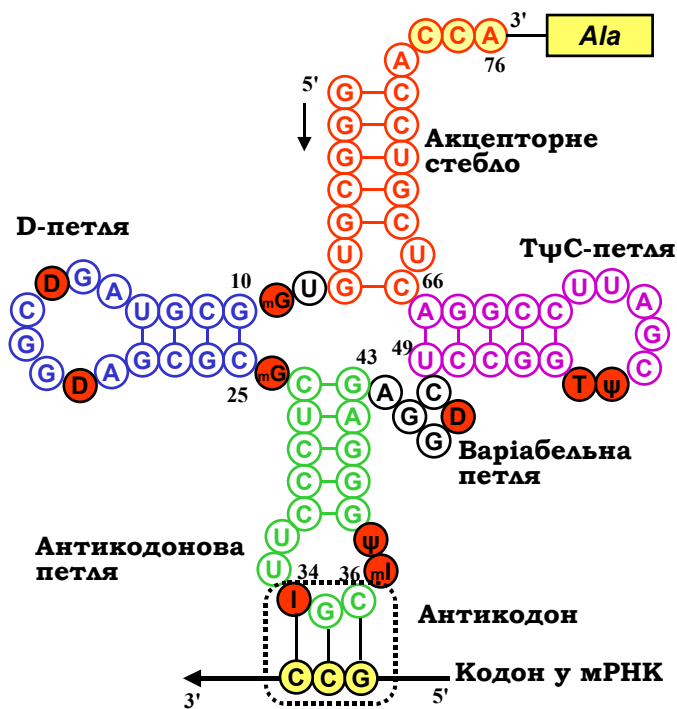
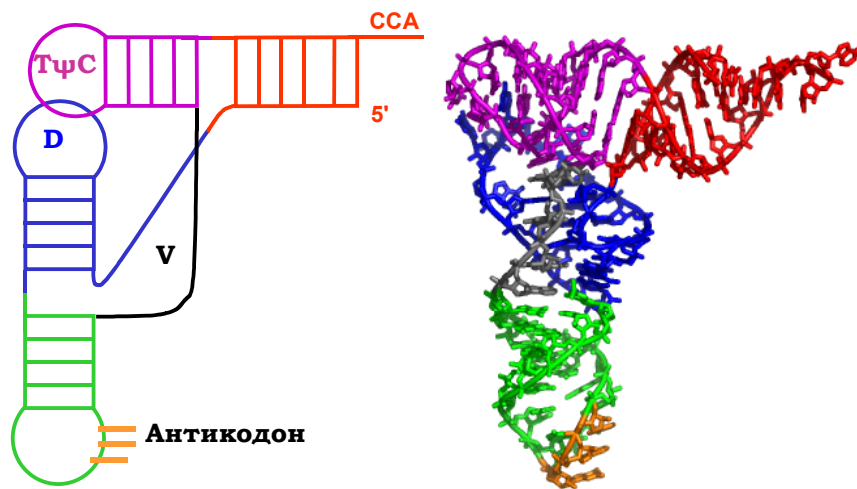


Рис. 8.1. Схема спарювання нуклеотидів у складі аланінової тРНК

Схема “листя конюшини” (рис. 8.1) не дає уяви щодо просторової структури тРНК. Насправді акцепторне та ТψС-стебло, переходячи одне в одне, утворюють єдину майже пряму подвійну спіраль (рис. 8.2), під приблизно прямим кутом до якої розташована друга подвійна спіраль, сформована D- і антикодоновим стеблом. ТψС- і D-петлі

при цьому наближаються одна до одної, між ними реалізується комплементарне спарювання основ. У результаті молекула тРНК має Г-подібну (або L-подібну) форму з двома плечами різної довжини: на кінці одного плеча акцептується амінокислота (акцепторне плече), на кінці іншого розташований антикодон (антикодонове плече).



**Рис. 8.2.** Схема і просторова структура тРНК (тРНК<sup>Phe</sup>, 6TNA). Елементи структури пофарбовані однаково ліворуч і праворуч

У складі дволанцюгових спіральних зон тРНК міститься ~55 % нуклеотидів, але ~90 % основ залучено до стекінг-взаємодій. Отже, молекула характеризується високою впорядкованістю, жорсткістю своєї структури. Це стосується також і петель, серед яких лише антикодонова петля не залучена до взаємодій з іншими елементами. Але п'ять основ петлі (включаючи антикодон) утворюють досить жорстку стопку. Додатково структура тРНК стабілізується іонами  $Mg^{2+}$ , тільки у присутності яких молекула є функціонально активною.

Загальна кількість типів тРНК, які обслуговують процес білкового синтезу, є близькою до 40 (наприклад, усі гени тРНК людини можна поділити на 49 родин за властивостями антикодонів). Оскільки типів тРНК більше, ніж амінокислот, одній амінокислоті може відповідати кілька тРНК – такі тРНК називають *ізоакцепторними*. Серед них є такі, що містять різні (але, звичайно, синонімічні) антикодони, – гетерокодонові. Є також гомокодонові тРНК, які можуть бути продуктами різних генів (розрізняються за послідовністю нуклеотидів),

а можуть бути продуктами одного гена, розрізняючись модифікаціями основ. Оскільки типів тРНК *менше, ніж кодонів*, одна тРНК здатна впізнавати кілька синонімічних кодонів, що забезпечується неоднозначністю спарювання між першою позицією антикодона і третьою (за якою, головним чином, розрізняються синонімічні кодони, див. розділ 4) – кодона. А саме, U і G здатні впізнавати по два нуклеотиди у третій позиції кодона, I (який досить часто зустрічається в першій позиції антикодона) упізнає три нуклеотиди (табл. 8.1).

**Таблиця 8.1. Відповідність між нуклеотидами в першій позиції антикодона і третій позиції кодона**

<b>Антикодон</b>	<b>Кодон</b>
C	G
A	U
U	A, G
G	U, C
I	U, C, A

### **Аміноацилювання тРНК**

Порядок залучення амінокислот до поліпептидного ланцюга, що утворюється при білковому синтезі, залежить лише від взаємодій між нуклеїновими кислотами – кодоном і антикодоном; амінокислота, яку несе тРНК, жодним чином не розпізнається рибосомою. Отже, акцептування певної амінокислоти молекулою тРНК відповідного типу (і тільки *відповідного*) є одним із найважливіших моментів білкового синтезу: від точності процесу акцептування буде залежати й точність синтезу білка в цілому.

Процес приєднання амінокислот до тРНК каталізується *аміноацил-тРНК-синтетазами* (APCаза, aaRS – aminoacyl-tRNA-Synthetase). Кожна з 20 типів (за кількістю амінокислот) цих ферментів каталізує дві хімічні реакції (рис. 8.3):

- На першій стадії відбувається так зване *активування амінокислоти* – її приєднання до АМР з утворенням аміноацил-аденілату, коли пірофосфат (pp) у складі АТР замінюється на амінокислоту (aa). Активування амінокислоти супроводжується зниженням вільної енергії, але при цьому значна частина вільної енергії, що “звільняється” при відщепленні пірофосфату

від АТР, заощаджується у формі аміноациладенілату – молекули, гідроліз якої також супроводжується великим енергетичним ефектом. Молекула АТР, що використовується на етапі активування амінокислоти, – єдине джерело енергії для майбутнього синтезу пептидного зв'язку на рибосомі.

- Аміноациладенілат утворює проміжний комплекс з активним центром ферменту й ефективно атакує ОН-групу рибози 3'-кінцевого аденозину тРНК (3'- або 2'-ОН групи залежно від класу АРСази): відбувається перенесення амінокислоти на тРНК. Після дисоціації від АРСази у складі аа-тРНК має місце спонтанний обмін амінокислоти між двома ОН-групами рибози; при зв'язуванні з рибосомою амінокислота фіксується на 3'-ОН групі. Аа-тРНК – також “макроергічна” сполука: руйнування зв'язку між амінокислотою та тРНК є енергетично вигідним, що й забезпечує утворення пептидного зв'язку на рибосомі.

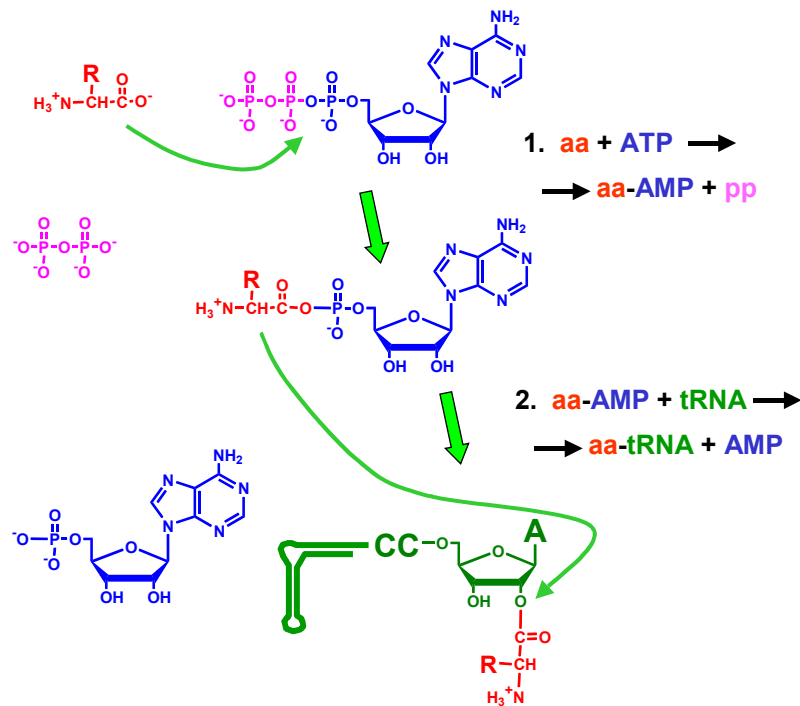


Рис. 8.3. Приєднання амінокислоти до тРНК, яке каталізується АРСазою

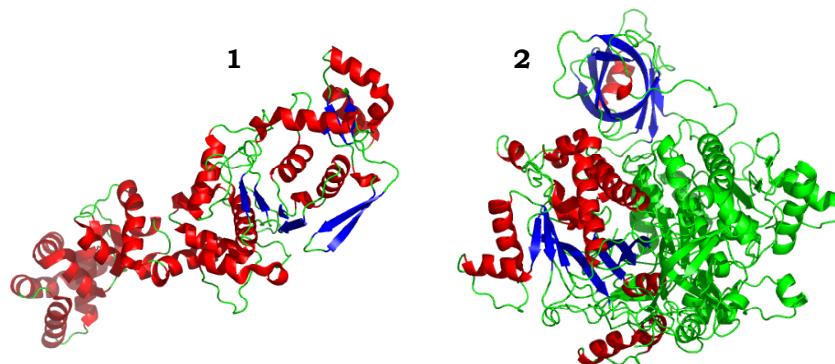
Отже, АРСази не тільки залучають амінокислоти до білкового синтезу, приєднуючи їх до тРНК, а й забезпечують заощадження вільної енергії, необхідної для приєднання амінокислоти до поліпептидного ланцюга.

Двадцять типів АРСаз, між якими немає майже нічого спільного на рівні первинної структури, можна поділити на два класи, по десять у кожному (табл. 8.2). Це завжди мультидоменні білки зі складною структурою (рис. 8.4), що зумовлено розмаїттям функцій – необхідністю специфічно зв'язати три субстрати й каталізувати дві хімічні реакції. Крім того, еукаріотичні АРСази різних типів взаємодіють між собою та з мембраною ендоплазматичного ретикулу, утворюючи так звану *кодосому*, що розташована поблизу від рибосом. Мультисубодиничні АРСази (гомодимери, гомотетрамери або гетеротетрамери) містять два ідентичні набори активних центрів та сайтів зв'язування.

Таблиця 8.2. Порівняльна характеристика двох класів АРСаз

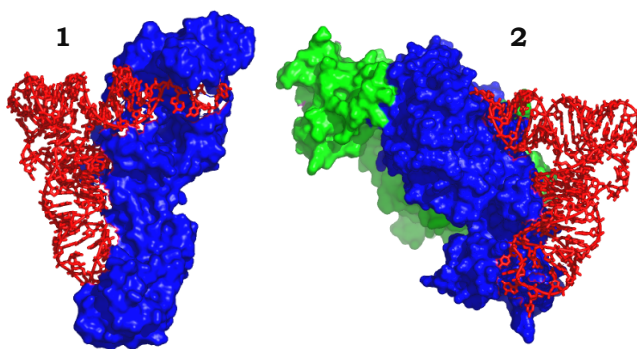
	Клас 1	Клас 2
Кількість субодиниць	1 (іноді 2)	2 або 4
Спільний структурний мотив, що оточує активний центр	Укладка Россмана	Паралельний $\beta$ -шар із 7 $\beta$ -ділянок
Амінокислоти, що акцептуються	Leu, Ile, Val Cys, Met, Glu, Gln, Arg, Tyr, Trp	His, Pro, Ser, Thr, Phe, Asp, Asn, Lys, Gly, Ala
Група рибози, що аміноацилюється	2'-ОН	3'-ОН

Центральна частина поліпептидного ланцюга АРСази формує активний центр (сайт зв'язування амінокислоти), в оточенні якого реалізується спільний для ферментів одного класу структурний мотив: для першого класу – укладка Россмана (див. розділ 2), для другого – приблизно плоский паралельний  $\beta$ -шар із семи  $\beta$ -ділянок (рис. 8.4). Як видно з табл. 8.2, амінокислоти з подібними властивостями часто є субстратами АРСаз різних класів залежно від розміру амінокислот: великі потрапляють до першого класу, маленькі – до другого (наприклад, Tyr і Phe, Glu і Asp, Arg і Lys). Відповідно, активний центр АРСаз першого класу розташований у порівняно неглибокій порожнині на поверхні ферменту, другого класу – у глибшому кармані.



**Рис. 8.4.** Структури АРСаз першого (мономерна Glu-тРНК-синтетаза, 1G59) та другого (Asp-тРНК-синтетаза, гомодимер, 1ASY) класів

Із частини молекули, яка утворює спільний структурний мотив, випетлюються інші структурні домени, які, разом із N або С-кінцевими доменами беруть участь у взаємодії з тРНК. АРСази двох класів упізнають різні елементи структури тРНК: маленький жолобок акцепторного стебла, D-петля та антикодонова петля для першого класу; великий жолобок акцепторного стебла, варіабельна та антикодонова петлі для другого (рис. 8.5). Причому специфічність упізнання тРНК залежить у першу чергу від взаємодій з акцепторним стеблом: штучна акцепторна частина (міні-тРНК) здатна специфічно зв'язатися з АРСазою та акцептувати амінокислоту.



**Рис. 8.5.** Комплекси тРНК<sup>Glu</sup> (1 клас, 1G59) і тРНК<sup>Asp</sup> (2 клас, 1ASY) з відповідними АРСазами. З димерною АРСазою зв'язано дві молекули тРНК. На даній проекції видно лише одну з них



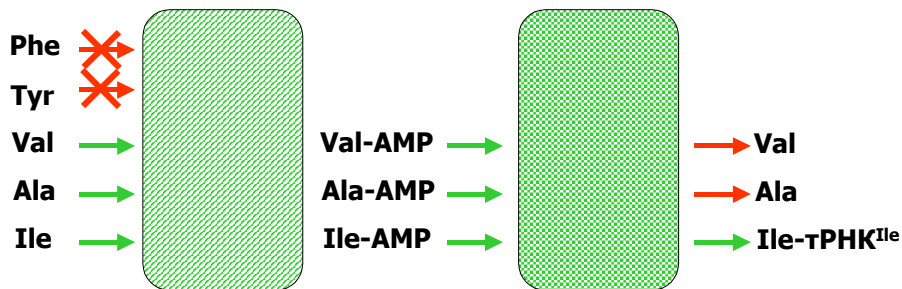


Рис. 8.6. Подвійне сито на прикладі Іле-тРНК-синтетази

Специфічність зв'язування АРСазою такого великого ліганду, як тРНК, не становить значної проблеми: велика кількість контактів дозволяє досить легко дискримінувати різні типи тРНК за рахунок різниці у вільній енергії зв'язування. Амінокислоти, навпаки, є маленькими лігандами (до того ж іноді з дуже схожою структурою), з яким неможливо реалізувати велику кількість взаємодій. Тим не менш, середня частота помилок при аміноацилюванні тРНК становить приблизно  $10^{-6}$  – такий рівень точності неможливо досягти просто за рахунок різниці у вільній енергії зв'язування, яка не може сильно відрізнятись для різних амінокислот. Відповідно, АРСази використовують певну систему корекції помилок за механізмом так званого *подвійного сита*. Наприклад, активний центр Іле-тРНК-синтетази (рис. 8.6) здійснює первинне вибракування частини амінокислот досить великого розміру на етапі зв'язування / активування, інші амінокислоти (включаючи Іле) піддаються активуванню. На другому етапі всі аміноациладенілати, що містять більш маленькі порівняно з Іле амінокислоти, гідролізуються в іншому активному центрі (центр редагування), а Іле переноситься на тРНК, оскільки його розмір не дозволяє йому зв'язатися з центром редагування. У результаті дуже близький до Іле за структурою Val акцептується тРНК<sup>Ile</sup> в одному випадку з 180 тис. Інші АРСази можуть здійснювати редагування помилок уже після акцептування амінокислоти, коли aa-тРНК упізнається як ціле, і неспоріднена амінокислота відщеплюється.

## **Рибосома**

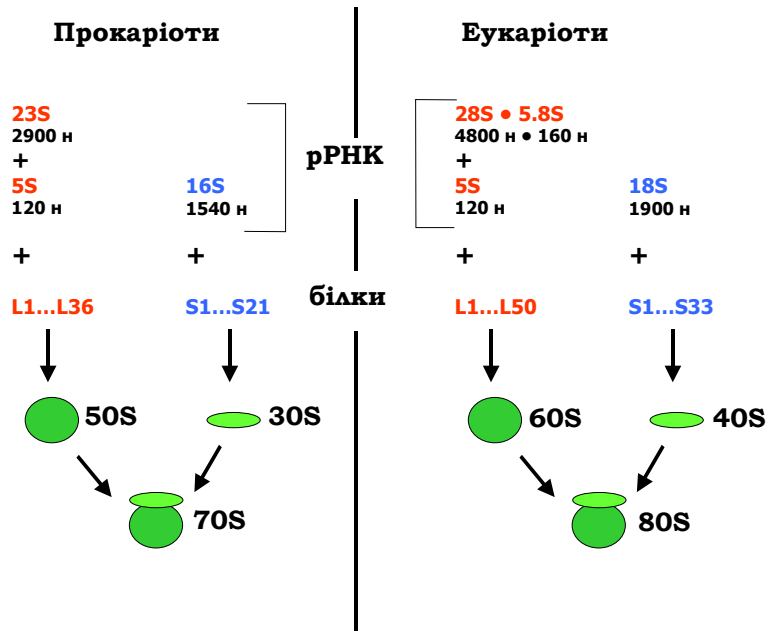
### *Склад рибосоми*

Рибосома – рибонуклеопротейдний комплекс, який складається із двох субодиниць (рис. 8.7). Компоненти рибосоми прийнято позначати *коефіцієнтами седиментації* – коефіцієнтами пропорційності, що показують, наскільки зростає швидкість руху частинки при центрифугуванні зі зростанням відцентрової сили (швидкості обертання ротора центрифуги). Коефіцієнт седиментації використовується як міра рухливості частинки, залежить від її маси, об'єму та форми, вимірюється у сведбергах (S) – позастистемних одиницях, які мають розмірність часу ( $1S = 10^{-13}$  с).

Маленька субодиниця прокариотичної рибосоми з коефіцієнтом седиментації 30S, містить одну молекулу рРНК (16S) і 21 молекулу рибосомних білків, що позначаються як S1–S21 (від Small subunit). Велика субодиниця містить дві молекули рРНК (23S і 5S) і білки L1–L36 (від Large subunit) – цей комплекс седиментує з коефіцієнтом 50S. Об'єднання субодиниць дає рибосому з коефіцієнтом седиментації 70S – оскільки коефіцієнт седиментації залежить від форми частинки, він не є адитивною величиною. Еукаріотична рибосома містить трохи більшу 18S рРНК замість 16S, дві рРНК, що міцно взаємодіють між собою (28S і 5,8S), замість 23S і більшу кількість білків. Структура обох рибосом і принципи їхньої роботи подібні.

Синтез еукаріотичних рРНК 18S, 5,8S і 28S здійснюється в ядрі, яке формується на тандемних повторях кластера відповідних генів рРНК (див. рис. 4.6). Первинний транскрипт, що синтезується РНК-полімеразою I, має константу седиментації 45S і містить три фрагменти майбутніх рРНК, розділені спейсерами. Процесинг рРНК – деградація спейсерів, а також модифікація певних основ і метилування 2'-ОН груп специфічних рибоз – здійснюється за участю близько 150 типів *маленьких ядерцевих РНК* (snoRNA – small nucleolar RNA), які, аналогічно до маленьких ядерних РНК, визначають специфічні сайти деградації та модифікацій. Частина маленьких ядерцевих РНК синтезується на певних генах РНК-полімеразами II та III. Але велика кількість цих РНК є інтронами, що визволяються в результаті сплайсингу мРНК білкових генів. 5S рРНК синтезується РНК-полімеразою III на окремих кластерах відповідних генів поза ядром. В ядрі, практично одночасно з процесин-

гом рРНК відбувається її поступова взаємодія з рибосомними білками, до ядрця ж дифундує комплекс 5S рРНК з білками: утворюються субодиниці рибосоми, які далі транспортуються до цитоплазми.

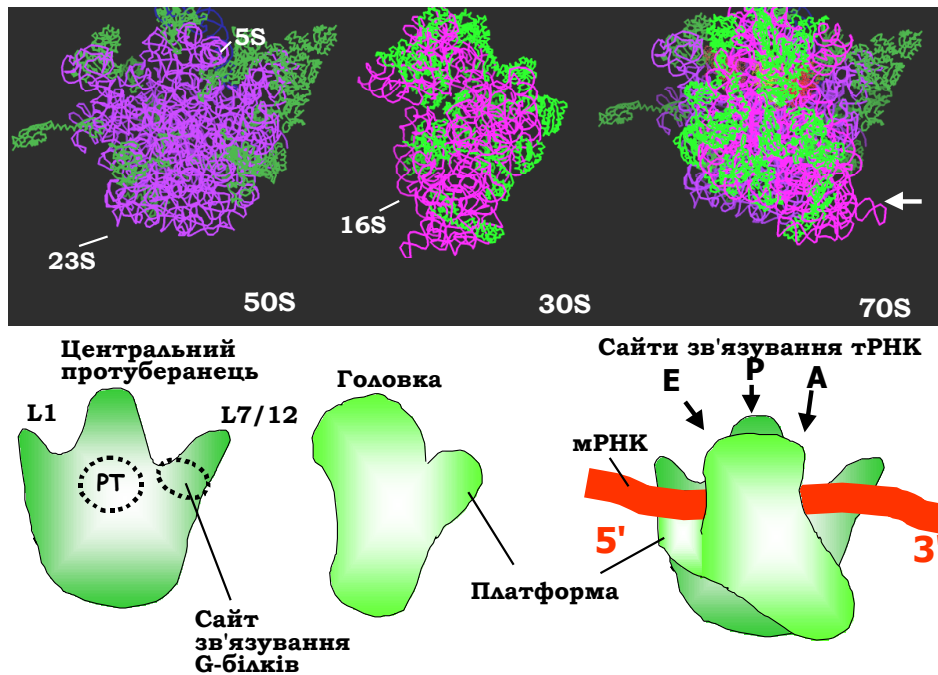


**Рис. 8.7.** Склад про- та еукаріотичної рибосом. рРНК позначено їхніми коефіцієнтами седиментації та довжиною в нуклеотидах (н)

Первинний транскрипт, що синтезується на бактеріальному рибосомному опероні, містить ділянки, які відповідають усім трьом прокаріотичним рРНК, а також кілька майбутніх тРНК. Часткова деградація транскрипту приводить до утворення зрілих молекул, які взаємодіють з рибосомними білками, формуючи дві субодиниці рибосоми. *In vivo* остаточне збирання рибосоми із двох субодиниць, як у про-, так і в еукаріотів, здійснюється при ініціації трансляції.

### Структура рибосоми

Структуру рибосоми за результатами рентгеноструктурного аналізу її кристалів показано в різних проекціях на рис. 8.8–8.10, де також схематично зображено зовнішню анатомію субодиниць та їхнє розташування у складі рибосоми відносно одна одної та деяких інших елементів системи трансляції.

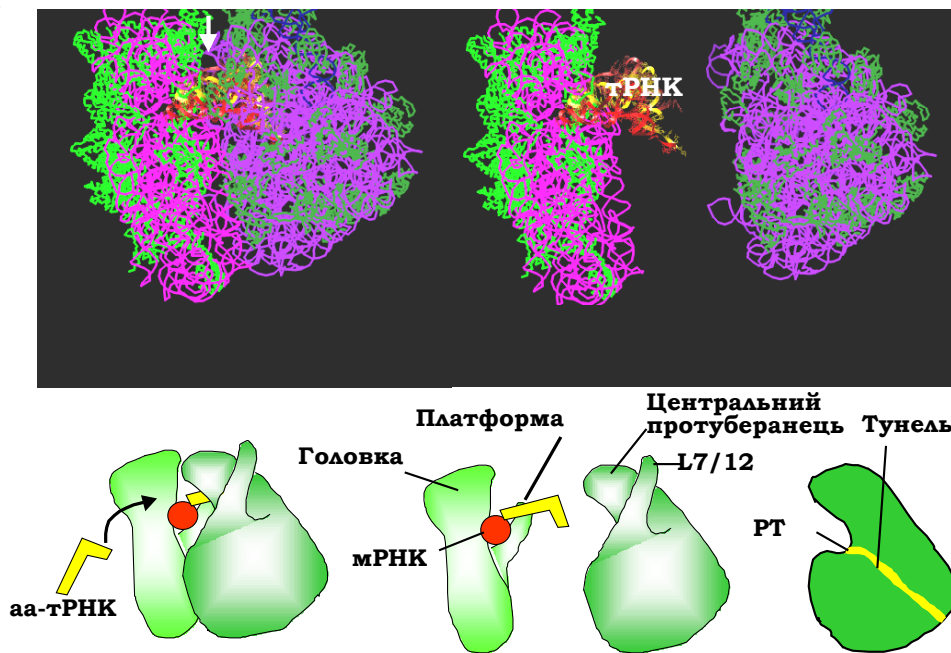


**Рис. 8.8.** Структура рибосоми *Thermus thermophilus* та її субодиниць (1GIX, 1GIY), показано лише основні ланцюги РНК і білків (білки пофарбовано зеленим). Окремі субодиниці орієнтовані інтерфейсами взаємодії між ними до глядача, у складі рибосоми орієнтація великої субодиниці збережена. Білою стрілкою позначено напрям зору, в якому структуру зображено на рис. 8.9.  
Унизу: схематичне зображення структур субодиниць та їхнього комплексу в тих самих проєкціях, РТ – пептидилтрансферазний центр

Від основного досить монолітного тіла великої субодиниці відходять три характерні відростки: виріст L1, *палець* (стебло) L7/12, сформовані відповідними рибосомними білками, і *центральный протуберанець*, утворений комплексом певних білків з рРНК 5S. Два окремі структурні домени – *головка* та *платформа* – відходять від тіла маленької субодиниці.

У складі рибосоми можна виділити три основні зони контактів між субодиницями: *головка* – *центральный протуберанець*; *платформа* – виріст L1; *центральный протуберанець* – виріст L1; *центральный протуберанець* – виріст L1; *центральный протуберанець* – виріст L1. Усі структурні елементи рухливі: можливим є переміщення *головки* й *пальця* L7/12, обертання малої субодиниці навкруг нормалі до поверхні великої субодиниці на  $\sim 6^\circ$  проти годинникової стрілки тощо. Рухи струк-

турних елементів рибосоми в зонах контактів мають важливе значення для її функціонування, оскільки саме на інтерфейсі між субодинами знаходяться всі активні центри та сайти взаємодії з елементами системи трансляції:

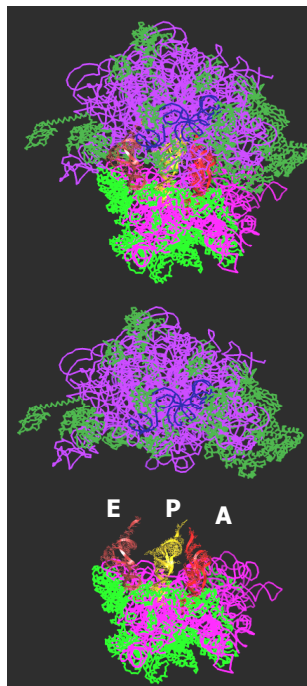


**Рис. 8.9.** Структура рибосоми та її субодинами у комплексі з трьома молекулами тРНК з рис. 8.8 у іншій проекції – з боку входу aa-тРНК до А-сайта. Білою стрілкою позначено напрям зору, у якому структуру зображено на рис. 8.10. Унизу: схематичне зображення у тій самій проекції, крайня права схема – велика субодина в розрізі

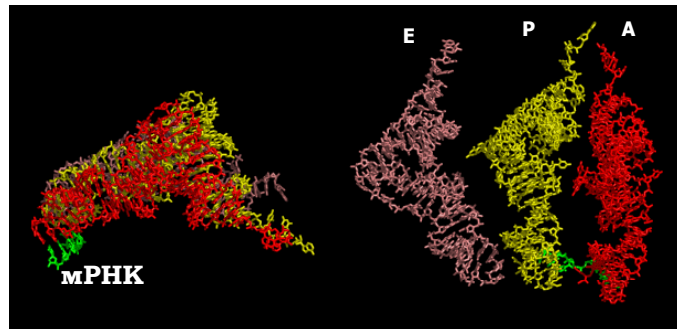
- В основі центрального протуберанця розташований пептидилтрансферазний центр (рис. 8.8, 8.9), який відповідає за каталіз реакції синтезу пептидного зв'язку. Від пептидилтрансферазного центру через тіло великої субодинаці відходить тунель – канал виходу поліпептидного ланцюга, що синтезується (рис. 8.9).
- В основі пальця L7/12 міститься сайт зв'язування G-білків (рис. 8.8) – факторів трансляції, детальне описання яких наведено нижче.

- У щілині між платформою та головкою маленької субодиниці відбувається взаємодія з мРНК (рис. 8.8, 8.9).
- Сайти зв'язування тРНК розташовані між двома субодиницями: антикодонові частини тРНК взаємодіють з мРНК і маленькою субодиницею, акцепторні частини – з великою субодиницею (рис. 8.9, 8.10). Рибосома містить три такі сайти: **А-сайт**, де відбувається зв'язування aa-тРНК; **Р-сайт**, де з рибосомою взаємодіє пептидил-тРНК (тРНК, до якої приєднаний *пептидил* – ланцюг, що синтезується); **Е-сайт** (від exit), де міститься деаміноацильована тРНК перед її звільненням з рибосоми.

Акцепторні частини тРНК, розташовані в А- і Р-сайтах, наближені одна до одної (рис. 8.10, 8.11) і взаємодіють з великою субодиницею в зоні пептидилтрансферазного центру. Кількість контактів з тРНК у Р-сайті є більшою, ніж кількість контактів, що утримують молекулу тРНК в А-сайті.



**Рис. 8.10.** Вид зверху (по стрілці на рис. 8.9) на структуру рибосоми та її субодиниць у комплексі з трьома молекулами тРНК



**Рис. 8.11.** Взаємне розташування трьох молекул тРНК у рибосомі (дві проекції приблизно збігаються з проекціями на рис. 8.9 і 8.10)

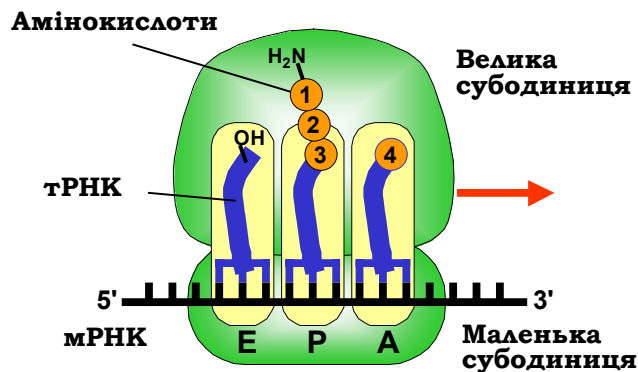
Аа-тРНК потрапляє до рибосоми через щілину між субодинаціями (рис. 8.9), розмір якої може змінюватись унаслідок рухливості структурних елементів рибосоми. Так само й розмір “воріт”, через які тРНК виходить із Е-сайта, модулюється рухливістю виросту L1.

Функціональну схему рибосоми наведено на рис. 8.12: систему зображено в момент перед додаванням четвертої амінокислоти до ланцюга, що синтезується; показано аа-тРНК, пептидил-тРНК і деаміноацильовану тРНК, що зв'язані відповідно з А-, Р- і Е-сайтами. Центральні сайти, навколо яких відбувається вся робота рибосоми – сайти зв'язування тРНК – формуються обома субодинаціями. Між субодинаціями існує своєрідне відносне *розділення праці*: маленька субодинаця, яка взаємодіє з мРНК та антикодоновими частинами тРНК, відповідає головним чином за декодуючу функцію рибосоми, а велика, яка взаємодіє з акцепторними частинами тРНК – за каталітичну функцію.

**Рибосомні РНК** становлять  $\sim 2/3$  маси рибосоми й саме вони визначають її структуру та функції. Полінуклеотидний ланцюг рРНК утворює велику кількість подвійних спіралей, з'єднаних петлями, – переважно це шпильки, проте відбувається і спарювання між віддаленими по ланцюгу ділянками. Загалом ланцюг утворює складну просторову структуру, де подвійні спіралі взаємодіють одна з одною та з одноланцюговими ділянками. Зокрема, аденозини одноланцюгових ділянок досить часто взаємодіють з маленьким жолобком подвійних спіралей, стабілізуючи структуру.

У складі 16S рРНК близько половини нуклеотидів залучено до  $\sim 60$  коротких дволанцюгових ділянок середньою довжиною вісім пар основ. Але зазвичай у складі 16S РНК виділяють 45 більш-менш

суцільних (із короткими неспареними ділянками всередині) подвійних спіралей. За своєю загальною Y-подібною формою 16S рРНК майже не відрізняється від маленької субодиниці рибосоми (рис. 8.13, порівн. рис. 8.8) – просторова структура чітко розділяється на чотири структурні домени: 5'-кінцевий формує тіло маленької субодиниці, центральний – платформу, так званий 3'-кінцевий мажорний домен – головку, 3'-кінцевий мінорний, який складається з довгої спіралі 44 (порядковий номер) і коротшої останньої спіралі 45, формує інтерфейс взаємодії з великою субодиницею. На цей РНК-скелет у складі маленької субодиниці “нарощуються” білки, однак загальна форма зберігається. При цьому сайти зв'язування мРНК і антикодонових частин тРНК формуються переважно 3'-кінцевою зоною 16S рРНК майже без участі білків; взаємодія між субодиницями також переважно забезпечується контактами РНК-РНК.



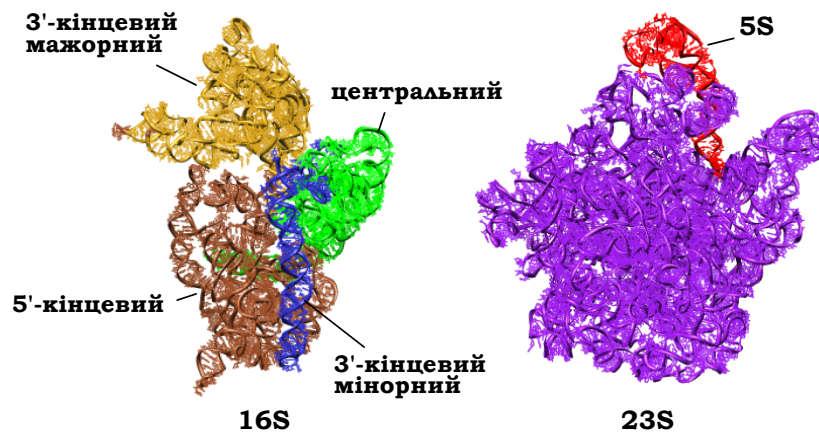
**Рис. 8.12.** Функціональна схема рибосоми: червоною стрілкою позначено напрямок руху при трансляції, амінокислоти пронумеровано відповідно до порядку включення їх до ланцюга

Еукаріотична 18S рРНК є гомологічною 16S, відрізняючись кількома вставками.

У складі 23S рРНК розрізняють шість структурних доменів, але вони тісно взаємодіють один з одним. Порівняно із 16S рРНК, ця структура є значно монолітнішою (рис. 8.13). Така монолітність зумовлена головною каталітичною функцією великої субодиниці – каталіз не передбачає конформаційної рухливості, а навпаки, вимагає жорсткості просторової організації. Каталітична активність великої субодиниці



пов'язана саме з певною зоною у складі 23S рРНК, яка також взаємодіє з акцепторними частинами тРНК. Порівняно невелика 5S рРНК у комплексі з певними рибосомними білками взаємодіє з 23S, формуючи центральний протуберанець.



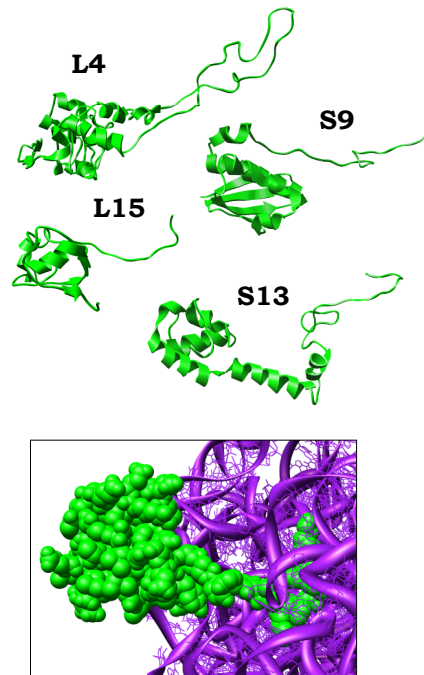
**Рис. 8.13.** Рибосомні рРНК у складі маленької (2B90) і великої (2B9P) субодиниць рибосоми *Thermus thermophilus* з боку інтерфейсу взаємодії (як на рис. 8.8). Указано структурні домени рРНК 16S, довга спіраль у складі 3'-кінцевого мінорного домену – спіраль 44

Еукаріотична 5,8S рРНК є гомологічною 5'-кінцевій зоні 23S. Узагалі еукаріотичні 28S і 5,8S рРНК, які утворюють міцний комплекс між собою, – це ніби трохи подовжена за рахунок вставок 23S рРНК, розділена на дві нерівні частини.

**Рибосомні білки** розміщуються на поверхні рРНК (відповідно, і на поверхні рибосоми). У складі маленької субодиниці платформа, щілина між платформою та головою, а також інтерфейс взаємодії з великою субодиницею майже не містять білків, котрі розміщені головним чином на зовнішньому боці субодиниці (рис. 8.8). По поверхні великої субодиниці білки розподілені рівномірніше, хоча збідненим на них є інтерфейс взаємодії з маленькою субодиницею, і навпаки, палець L7/12 утворений кількома копіями відповідних білків без участі рРНК.

Більшість рибосомних білків мають дві частини у своїй структурі: глобулу, розміщену на поверхні рибосоми, і позитивно заряджений хвіст (нерегулярна ділянка, петля або одна–дві окремі  $\alpha$ -спіралі), який занурюється в сітку рРНК і стабілізує структуру останньої (рис. 8.14). У деяких білків глобула відсутня. Отже, усі рибосомні білки (за винятком L7/12) взаємодіють з рРНК, хоча не всі мають високу спорідне-

ність до вільної рРНК. Збирання субодиниць рибосоми з рРНК та білків (може бути здійснено *in vitro* без участі будь-яких факторів) відбувається поступово, через чотири-пять стадій: зв'язування частини білків на кожній стадії індукує конформаційну зміну рРНК у складі комплексу, що викликає спорідненість до нової порції білків.



**Рис. 8.14.** Приклади рибосомних білків у структурі рибосоми (2B9O, 2B9P).  
Унизу: взаємодія білка L15 з рРНК 23S

Головна роль рибосомних білків – підтримувати функціонально активну структуру рРНК. Хоча рРНК відповідає за майже всі активності рибосоми, активної конформації вона набуває лише в комплексі з рибосомними білками. Крім того, деякі білки беруть участь у взаємодіях з елементами системи трансляції. Наприклад, білок S1, що міститься поблизу від сайту зв'язування мРНК, сприяє розплітання дволанцюгових шпильок у складі матриці; комплекс S1–S18–S21 взаємодіє з мРНК, а також ініціаторною тРНК при ініціації трансляції; білок L10 сумісно з певною ділянкою 23S рРНК і білком L11 організує сайт зв'язування G-білків у основі стебла L7/12.

### Елонгаційний цикл

Робота рибосоми під час елонгації трансляції полягає в послідовному (потриплетно) зчитуванні інформації з мРНК і відповідному приєднанні амінокислот до поліпептидного ланцюга. Кожен такий крок складається з трьох операцій, що циклічно повторюються (*елонгаційний цикл*). Цикл розпочинається з такої конфігурації системи, коли в Р-сайті знаходиться пептидил-тРНК, А-сайт є вільним від тРНК і в його межах на маленькій субодиноці розташований черговий кодон, який має бути впізнаним (рис. 8.15). **Перша операція** циклу – **зв'язування aa-тРНК** з А-сайтом. Зв'язування має відбутися з високою специфічністю щодо взаємодій між кодоном і антикодоном – тільки споріднена до даного кодона тРНК має бути відібрана системою. Процес розміщення aa-тРНК в А-сайті часто супроводжується дисоціацією з Е-сайта деаміноацильованої тРНК, яка залишилася там з попереднього циклу. Наслідком зв'язування є належне розташування акцепторних частин aa-тРНК і пептидил-тРНК відносно одна одної та каталітичного активного центру. У результаті рибосома здійснює **другу операцію – транспептидацію** – перенесення пептидилу з пептидил-тРНК на амінокислоту у складі aa-тРНК. Наслідком є перебудова системи: в А-сайті опиняється пептидил-тРНК із подовженим на одну амінокислоту пептидилом, у Р-сайті – деаміноацильована тРНК. **Третя операція – транслокація** – полягає в переміщенні рибосоми на один кодон уздовж мРНК (молекули тРНК залишаються зв'язаними зі своїми кодонами), після чого розпочинається наступний елонгаційний цикл.

Відбір серед різних молекул aa-тРНК на першому етапі, коли неспоріднені молекули мають швидко звільнитися, а також рух рибосоми при транслокації, передбачають реалізацію певної “відкритої”, не жорстко зафіксованої, рухливої структурної форми рибосоми. Каталіз транспептидації, навпаки, вимагає жорсткої фіксації субстратів у “закритій”, жорсткій і нерухливій формі рибосоми. Ефективне вирішення цих суперечливих завдань залежить від полегшення реалізації відкритих форм рибосоми на першому та третьому етапах елонгаційного циклу завдяки двом факторам елонгації (EF – Elongation Factors): **EF1** (еукаріотичний аналог позначається як eEF1 або eEF1A, eEF – ekariotic Elongation Factor) і **EF2** (еукаріотичний аналог – eEF2). Обидва фактори належать до родини G-білків, або GTP-зв'язу-

вальних білків. Інші білки цієї родини залучені до багатьох різноманітних процесів, зокрема, є елементами асоційованих із мембранами клітинних сигнальних систем.

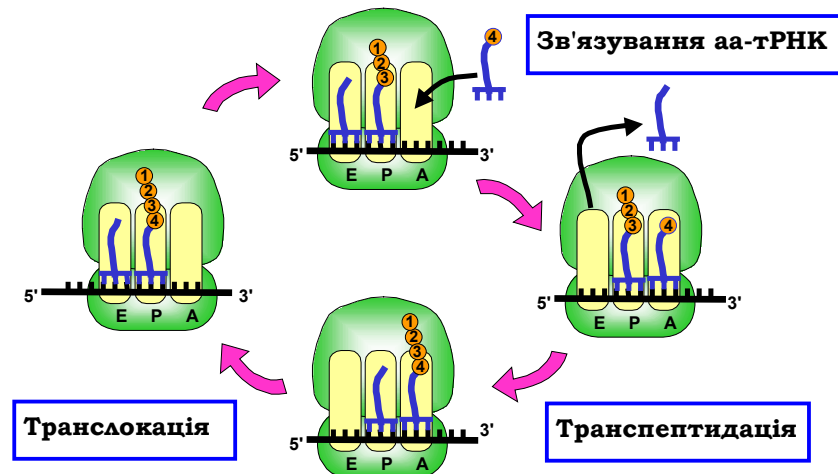


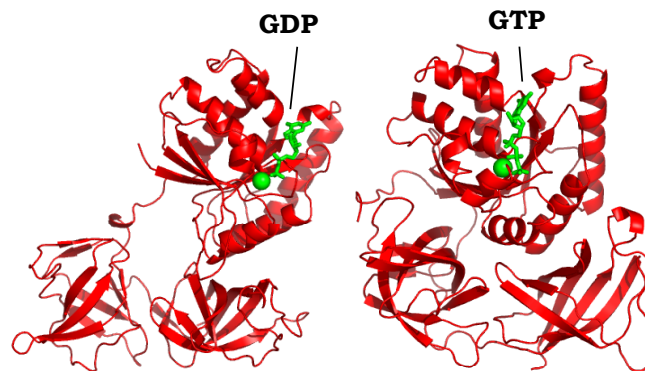
Рис. 8.15. Схема елонгаційного циклу

### Елонгаційний фактор EF1

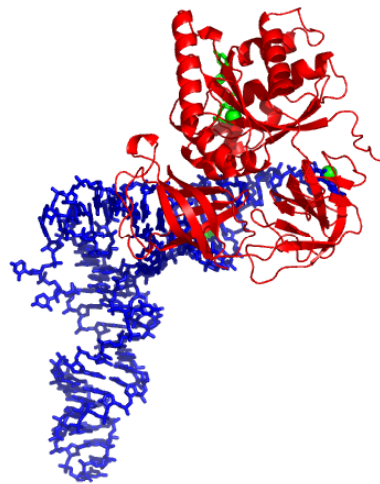
Фактор EF1 (часто також позначається як **EF-Tu**) – мономерний білок, що має три структурні домени, – може існувати у двох структурних станах (рис. 8.16). Співвідношення між вільними енергіями цих станів, а відповідно й імовірність переважної реалізації одного з них, залежить від типу ліганду (GTP чи GDP), що зв'язаний з білком (у повній відповідності до схеми на рис. 2.21). У комплексі з GDP (гуанозиндифосфат) реалізується відкрита конформація EF1 з порушеними взаємодіями між структурними домонами. Заміна GDP на GTP (гуанозинтрифосфат) приводить до локальної конформаційної перебудови в межах GTP-зв'язувального домену, унаслідок якої певні амінокислотні групи виводяться до інтерфейсу взаємодії з іншими домонами – така взаємодія відбувається і структура “замикається”. Гідроліз GTP (який здійснюється за певних умов самим білком, див. нижче) проводить до заміни ліганду і, відповідно, до зворотного перемикавання конформації.

Структурна перебудова EF1 є важливою не сама по собі. GDP- і GTP-асоційовані структурні форми мають відповідно *низьку* та *високу* спорідненість до аа-тРНК і рибосом: саме у вигляді потрібного

комплексу EF1·GTP–aa-тРНК (рис. 8.17) і відбувається зв'язування aa-тРНК з рибосомою на першому етапі елонгаційного циклу. При цьому вже потрійний комплекс відіграє роль ліганду, що перемикає структурні стани рибосоми.



**Рис. 8.16.** Структура фактора EF1 у комплексі з GDP (1TUI) і GTP (1TFT). Зелена кулька – іон  $Mg^{2+}$



**Рис. 8.17.** Комплекс EF1·GTP з aa-тРНК (1B23)

Зв'язування EF1·GTP із акцепторною частиною aa-тРНК (рис. 8.17) відбувається відразу після аміноацилювання тРНК. При завершенні процесу взаємодії з рибосомою (див. нижче) EF1 здійснює гідроліз

GTP, що веде до втрати спорідненості та дисоціації EF1·GDP. Після дисоціації молекула GDP витісняється білковим кофактором EF-Ts (еукаріотичний аналог – eEF1B), який, у свою чергу, замінюється на молекулу GTP, і знову відбувається зв'язування EF1·GTP з новою молекулою aa-тРНК.

### *Зв'язування aa-тРНК з А-сайтом рибосоми*

**Роль фактора EF1 і гідролізу GTP.** Щоб aa-тРНК могла потрапити до рибосоми, рибосома має опинитися в певному “відкритому” структурному стані з розширеним каналом між субодинамиціями з боку стебла L7/12 (див. рис. 8.9). Крім того, така відкрита форма рибосоми не допускає дуже міцної взаємодії з aa-тРНК, тобто сприяє її легкій дисоціації в разі невідповідності між кодоном і антикодоном. Однак зрозуміло, що відкрита форма рибосоми із частково порушеними контактами (енергетично вигідними взаємодіями) між субодинамиціями має характеризуватися підвищеною вільною енергією, тобто є малоймовірною. Інший важливий аспект полягає в тому, що на етапі первинного перебору aa-тРНК з різними антикодонами необхідно повністю виключити для акцепторної частини тРНК можливість випадково потрапити до пептидилтрансферазного центру. Узагалі, для такого великого ліганду, як тРНК, існує кінетичний бар'єр асоціації / дисоціації: велика кількість контактів з лігандом вимагає утворювати / руйнувати їх одночасно – проміжні стани з лімітованим набором контактів мають забезпечити прискорення обох процесів. Роль EF1 саме й полягає у вирішенні зазначених проблем:

- Взаємодія EF1 з акцепторною частиною aa-тРНК залишає тільки антикодонову частину вільною для взаємодій з рибосомою (і зв'язаною з нею мРНК).
- При зв'язуванні потрійного комплексу, EF1 (який при цьому зв'язаний з акцепторною частиною aa-тРНК) взаємодіє із сайтом у основі стебла L7/12, що виключає контакт акцепторного стебла тРНК з пептидилтрансферазним центром.
- EF1·GTP має підвищену спорідненість до відкритої форми рибосоми, тобто подібно до того, як GTP перемикає конформацію самого EF1, потрійний комплекс EF1·GTP–aa-тРНК фіксує відкриту форму рибосоми, яка стає енергетично вигідною за рахунок енергії взаємодій з EF1 (рис. 8.18).

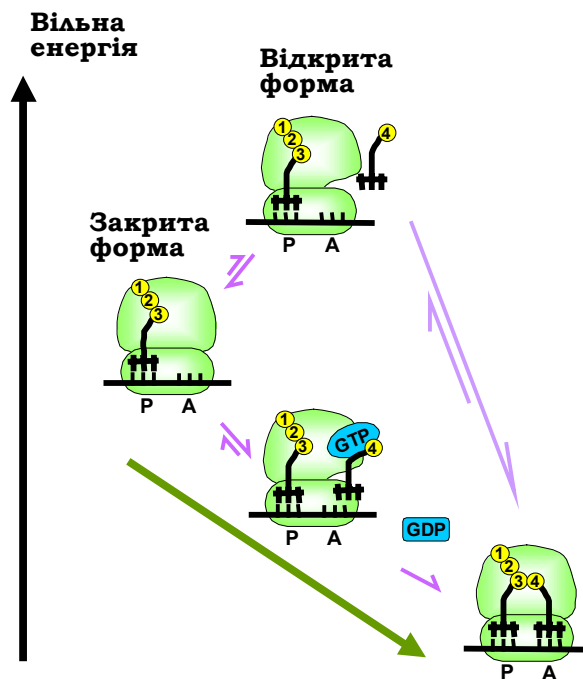
Отже, за умови присутності EF1·GTP здійснюється швидка й не дуже міцна взаємодія aa-тРНК з рибосомою (головним чином завдяки взаємодії EF1 з основою стебла L7/12), яка також швидко змінюється дисоціацією: випробування, під час яких антикодонова частина тРНК намагається впізнати кодон. Якщо таке впізнання відбувається (див. нижче), спрацьовує GTPазна активність EF1 – після гідролізу GTP фактор дисоціює, що дозволяє акцепторній частині aa-тРНК остаточно розміститися в А-сайті.

Зв'язування aa-тРНК є внутрішньою властивістю рибосоми: *in vitro* зв'язування може відбуватися без участі будь-яких факторів, оскільки супроводжується зниженням вільної енергії. Проте позафакторне зв'язування є досить повільним через наявність на його шляху енергетичного бар'єра – високоенергетичного відкритого стану рибосоми (рис. 8.18). Таким чином, EF1 *не зумовлює* зв'язування aa-тРНК, а тільки суттєво *прискорює* цей процес. Можна сказати, що EF1 здійснює каталіз конформаційних перетворень рибосоми за рахунок спорідненості до проміжного інтермедіатного стану. Щодо гідролізу GTP, то він потрібен лише для заміни ліганду з метою позбавитися цієї спорідненості й забезпечити дисоціацію EF1 і повернення рибосоми до закритого стану, необхідного для наступної стадії елонгаційного циклу. Але це має відбутися лише за умови впізнання кодона антикодоном aa-тРНК.

**Упізнання кодона.** Коли під час первинного зв'язування після кількох випробувань у межах А-сайта на маленькій субодиниці опиняється нарешті комплементарний антикодон, між ним і кодоном має утворитися коротка подвійна спіраль. Оскільки вона коротка (три пари основ), то досить нестабільна: кодон-антикодонова взаємодія за межами рибосоми є неефективною. При цьому рибосома (передусім маленька субодиниця) не тільки створює особливе середовище, яке стабілізує кодон-антикодонову спіраль, але й сприяє підвищенню специфічності – ефективно дискримінує споріднені та неспоріднені до даного кодону антикодони.

При утворенні комплементарної подвійної спіралі, з її маленьким жолобком взаємодіють два консервативні аденозини рРНК 16S (з порядковими номерами 1492 і 1493 – розташовані у верхній частині спіралі 44, див рис. 8.13), утворюючи водневі зв'язки з 2'-ОН-групами рибоз нуклеотидів кодона й антикодона (рис. 8.19). Така взаємодія, що реалізується лише за умови комплементарності, стабілізує перші дві (відносно позицій кодона) нуклеотидні пари. Друга нуклеотидна пара додатково стабілізується, також через маленький жолобок,

взаємодією з консервативним G530 рРНК 16S. Третя пара кодон-антикодонової спіралі теж взаємодіє з певним цитидином рРНК 16S, але ця взаємодія менш специфічна й допускає неоднозначність спарювання (див. підрозділ, де йдеться про тРНК).

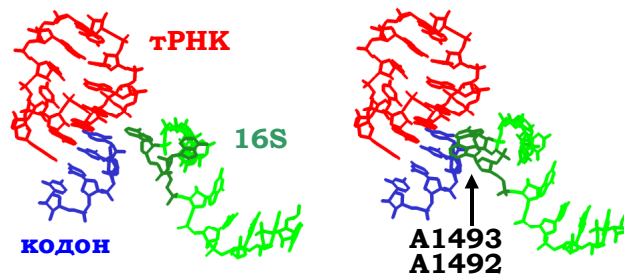


**Рис. 8.18.** Зв'язування з А-сайтом рибосоми аа-тРНК у вільному стані та в комплексі з EF1·GTP. Різні стани системи розміщені на різних рівнях відповідно до їхньої вільної енергії. Зелена стрілка вказує оптимальний шлях процесу. Показано тільки два сайти тРНК

Взаємодії нуклеотидів рРНК 16S із кодон-антикодоновою спіраллю не тільки підвищують специфічність упізнання кодона, але й мають інший важливий наслідок. Як показано на рис. 8.19, для реалізації цих взаємодій A1492 і A1493 мають бути переорієнтованими відносно сусідніх нуклеотидів. Така локальна конформаційна перебудова запускає каскад конформаційних змін в інших частинах рРНК 16S за “принципом доміно”, що спричиняє глобальну конформаційну перебудову всієї маленької субодиниці. Зокрема, головка субодиниці переміщується в бік великої субодиниці (див. рис. 8.9), міцно замикаючи



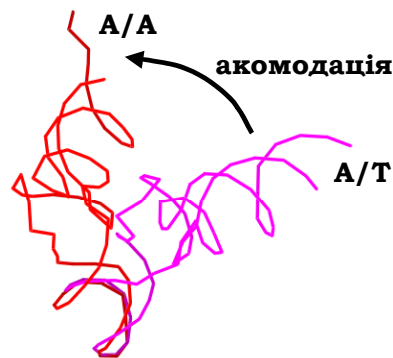
антикодонову частину aa-тРНК на кодоні мРНК. Оскільки субодиниці рибосоми взаємодіють між собою, структурна перебудова маленької субодиниці не залишається непомітною для великої: конформаційні зміни відбуваються і тут. Одним із результатів перебудови є зміна взаємодії між EF1 і відповідним сайтом його зв'язування в основі стебла L7/12, яка приводить до активації GTPази, – відбувається гідроліз GTP, конформаційна перебудова EF1 із втратою спорідненості до рибосоми та тРНК і дисоціація фактора.



**Рис. 8.19.** Взаємодія в межах А-сайта на маленькій субодиниці між кодоном, антикодоною петлею тРНК і частиною рРНК 16S у випадку неспорідненого (ліворуч, 1N36 – друга кодон-антикодонова пара основ є некомплементарною) і спорідненого (праворуч, 1IBM) антикодонів

**Акомодація aa-тРНК** завершує першу стадію елонгаційного циклу. Перед гідролізом GTP розміщення aa-тРНК на рибосомі суттєво відрізняється від “кінцевої точки” процесу зв'язування – локалізації в А-сайті. Оскільки сайти зв'язування тРНК формуються двома субодиницями (мають дві “половини”), “остаточний” А-сайт позначають як А/А. У складі комплексу з EF1 антикодонова частина розміщена в А-сайті на маленькій субодиниці, але акцепторна частина утримується в основі стебла L7/12 на великій – так зване А/Т положення (рис. 8.20). При цьому антикодонове стебло суттєво вигинається – фіксація антикодона та акцепторного стебла підтримує напружену конформацію тРНК. Після гідролізу GTP, коли EF1 дисоціює, один із фіксаторів зникає – тРНК, ніби пружина, розпрямляється, її акцепторне стебло автоматично опиняється в А-сайті на великій субодиниці – у зоні пептидилтрансферазного центру. Ефективність цього процесу *акомодації* (рис. 8.20) залежить від жорсткості фіксації антикодонової петлі тРНК на маленькій субодиниці. Якщо взаємодія між кодоном

і антикодоном не є повністю комплементарною, aa-тРНК просто дисоціює від рибосоми – це друга (після первинного зв'язування) можливість виправити помилку впізнання кодона.

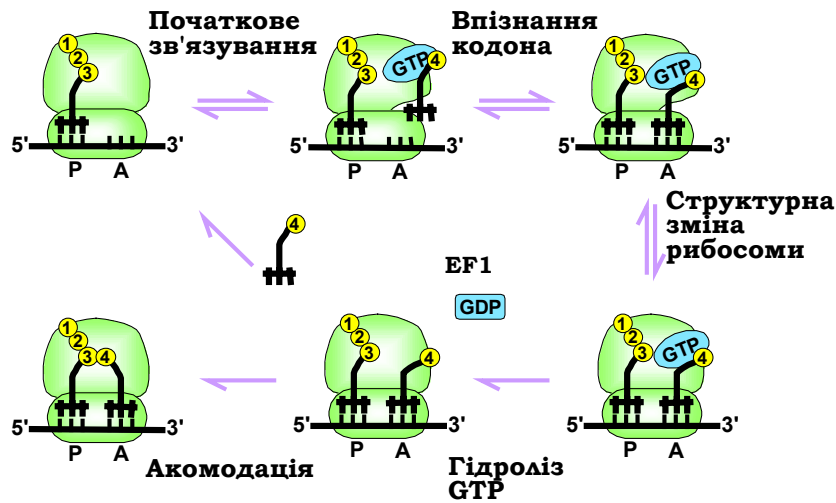


**Рис. 8.20.** А/Т (1QZA) і А/А (1QZB) положення aa-тРНК на рибосомі до й після акомодації відповідно

**Загальний сценарій зв'язування aa-тРНК з рибосомою** зображено на рис. 8.21.

- На першому етапі початкового відбору aa-тРНК відбувається швидка асоціація / дисоціація потрійних комплексів EF1-GTP-aa-тРНК з відкритою формою рибосоми, яка фіксується за рахунок взаємодій з EF1. Відкрита форма сприяє також дисоціації деаміноацильованої тРНК з Е-сайта рибосоми (не показано на рис. 8.21).
- При утворенні комплементарної кодон-антикодонової подвійної спіралі, вона стабілізується за рахунок взаємодій з нуклеотидами рРНК 16S. Локальна конформаційна зміна 16S, яка є наслідком цих взаємодій, спрацьовує як тригер структурної перебудови спочатку маленької, потім великої субодиниць. Унаслідок перебудови антикодонова петля тРНК жорстко замикається на маленькій субодиниці, акцепторне стебло – на великій у основі пальця L7/12, і EF1 набуває GTPазної активності.
- Гідроліз GTP розмикає ланцюг подій на дві частини, роблячи його необоротним. Після гідролізу EF1-GDP дисоціює, залишаючи aa-тРНК у напруженій конформації.

- Далі процес розгалужується: або, при недостатньо міцних кодон-антикодонових взаємодіях, відбувається дисоціація aa-тРНК, яка повертає систему до вихідного стану, або здійснюється акомодация – релаксация напруженої aa-тРНК, яка виводить її акцепторний кінець до пептидилтрансферазного центру.



**Рис. 8.21.** Схема основних подій при зв'язуванні aa-тРНК з А-сайтом рибосоми. Е-сайт тРНК не показано для спрощення

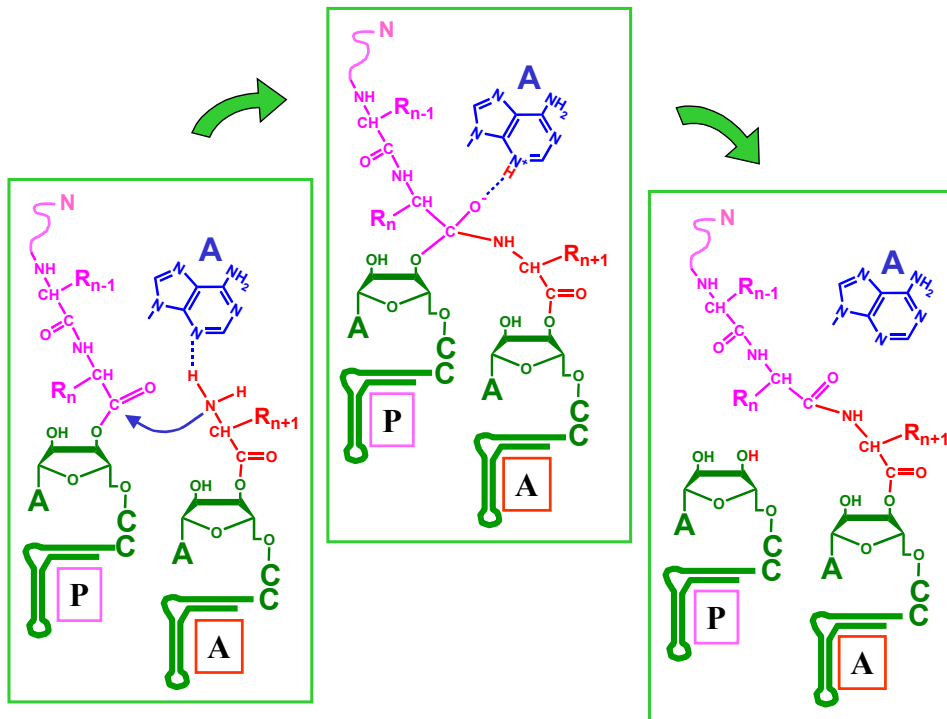
Загальним наслідком процесу зв'язування, таким чином, є повна підготовка системи трансляції до наступного етапу елонгаційного циклу.

### Транспептидація

Реакція перенесення пептидилу – транспептидація – полягає в руйнуванні ковалентного зв'язку між карбоксильною групою С-кінцевої амінокислоти пептидилу та 3'-кінцевою ОН-групою пептидил-тРНК і утворенні натомість пептидного зв'язку між звільненою карбоксильною групою та аміногрупою амінокислоти у складі aa-тРНК. Таким чином, при транспептидації пептидил переноситься з Р- до А-сайта рибосоми: в А-сайті опиняється пептидил-тРНК із подовженим на одну амінокислоту пептидилом, у Р-сайті залишається деаміноацильована тРНК (див. рис. 8.15).

Стандартна вільна енергія гідролізу зв'язку між пептидилом і тРНК оцінюється приблизно в  $-7,5$  ккал/моль, енергія гідролізу пептидного зв'язку  $-0,5$  ккал/моль. Отже, загальний енергетичний ба-

ланс транспептидації  $\sim -7,0$  ккал/моль: це спонтанний процес, який не потребує джерел енергії – енергію було заощаджено при аміноацилюванні тРНК (див. підрозділ, присвячений тРНК).



**Рис. 8.22.** Процес перенесення пептидилу з Р-сайта на аа-тРНК в А-сайті через інтермедіатну сполуку, стабілізовану взаємодією з аденозином (А синього кольору) рРНК 23S. Амінокислотні бокові залишки R пронумеровані в порядку приєднання їх до поліпептидного ланцюга

Процес транспептидації, що відбувається безпосередньо після акомодатії аа-тРНК, є найшвидчим етапом елонгаційного циклу. Каталіз реакції здійснюється пептидилтрансферазним центром великої субодиниці рибосоми, розташованим в основі центрального протуберанця (див. рис. 8.8) і формується пептидилтрансферазною петлею рРНК 23S. У радіусі  $\sim 18$  Å від субстратів реакції немає жодної білкової хімічної групи, тобто **рибосома – це рибозим**. Слід, проте, ще раз зауважити, що вільна рРНК 23S не має каталітичної активності – активна конформація рРНК реалізується лише в комплексі з певним мінімальним набором білків.

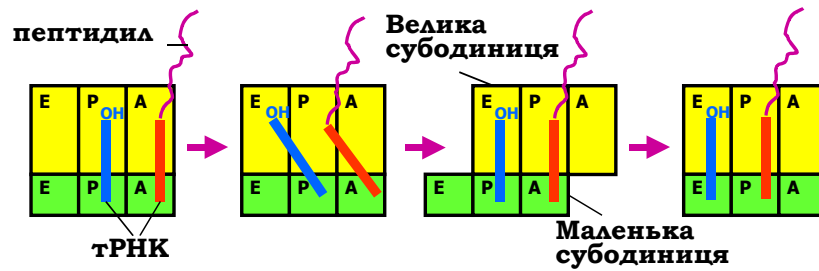
Реакція транспептидації проходить через утворення інтермедіатної сполуки з формуванням зв'язку між N аміногрупи aa-tРНК і C, яким пептидил приєднаний до tРНК, атом кисню набуває при цьому негативного заряду (рис. 8.22). Утворенню інтермедіату передують захоплення протона аміногрупи атомом N3 аденіну, що знаходиться в активному центрі. Легка іонізація N3 зумовлена його особливим мікрооточенням у складі рРНК 23S, протонований азот стабілізує негативно заряджений O. Далі протон переноситься на 3'-ОН групу пептидил-tРНК, яка, відповідно, стає деаміноацильованою – втрачає зв'язок із C, замість якого остаточно формується пептидний зв'язок.

Механізм каталізу транспептидації є цілком аналогічним до такого білкових ферментів (див. розділ 2): рибосома здійснює жорстке взаємне орієнтування субстратів у активному центрі (і ця обставина – ентропійний каталіз – є головним фактором прискорення реакції транспептидації), а також завдяки створенню специфічного мікрооточення певної хімічної групи (аденіну) забезпечує стабілізацію проміжного високоенергетичного стану.

### Транслокація

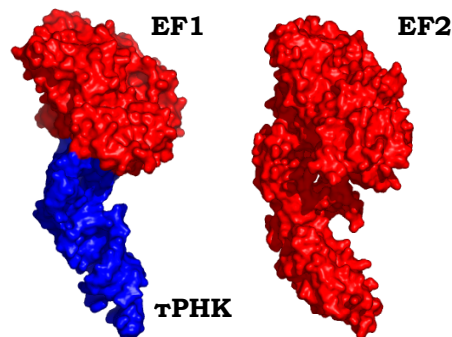
Результатом транспептидації є значне перегрупування лігандів, зв'язаних з рибосомою: пептидил-tРНК опиняється в А-сайті, деаміноацильована tРНК – у Р-сайті. Ураховуючи підвищену спорідненість акцепторної частини пептидил-tРНК до Р-сайта, рибосома має спонтанно пересунути її вздовж мРНК у напрямку до її 3'-кінця, і таке пересування – транслокація – має розпочатися з руху великої субодиниці (рис. 8.23). На проміжній стадії виникають *гібридні сайти* А/Р і Р/Е – стани, коли акцепторні частини tРНК *уже* знаходяться в Р (або Е) сайті на великій субодиниці, але антикодони частини – *ще* в А (або Р) сайті.

Зрозуміло, що рух субодиниць передбачає послаблення взаємодій між ними – певний розблокований після транспептидації, відкритий структурний стан рибосоми. Зрозуміло також, що послаблення взаємодій між субодиницями є енергетично не вигідним. Відповідно, хоча транслокація є внутрішньою властивістю рибосоми і відбувається спонтанно (вільна енергія знижується в кінцевому стані за рахунок високої спорідненості пептидил-tРНК до Р-сайта), швидкість руху є дуже повільною в позафакторній системі *in vitro*, оскільки процес проходить через проміжний високоенергетичний стан. *In vivo* фактором прискорення процесу є білок EF2 (інше позначення – **EF-G**).



**Рис. 8.23.** Схема транслокації рибосоми через проміжну стадію гібридних сайтів

Частина молекули EF2 є гомологічною білку EF1 – також має сайт зв'язування GTP/GDP і також взаємодіє з великою субодиницею рибосоми в основі пальця L7/12. Але у складі білка є ще один додатковий домен, який мімікрує під структуру антикодону частини тРНК, – структура білка в цілому нагадує комплекс EF1–тРНК (рис. 8.24).



**Рис. 8.24.** Поверхня комплексу EF1–тРНК (структура з рис. 8.17) і білка EF2 (2EFG)

Фактор EF2 існує у принаймні чотирьох структурних станах залежно від типу зв'язаного ліганду:

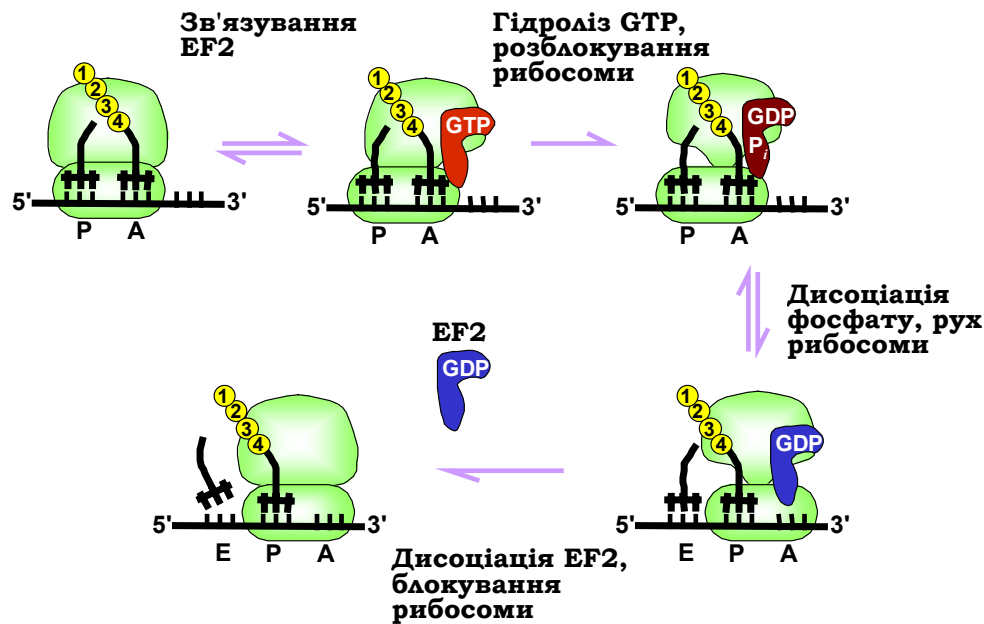
- 1) EF2·GTP поза рибосомою – має високу спорідненість до рибосоми;
- 2) EF2·GTP на рибосомі – активна GTPаза;
- 3) EF2·GDP·P<sub>i</sub> – після гідролізу GTP (P<sub>i</sub> – неорганічний фосфат);
- 4) EF2·GDP після дисоціації фосфату – має низьку спорідненість до рибосоми.

Вільний від будь-якого ліганду білок не має впливу на транслокацію, однак у комплексі з будь-яким гуаніновим нуклеотидом (включаючи також аналоги GTP, що не можуть бути гідролізовані) фактор прискорює транслокацію за рахунок своєї спорідненості до проміжного відкритого стану рибосоми. Проте у випадку гідролізу GTP (який передуює остаточній транслокації) швидкість переміщення рибосоми підвищується ще в ~50 разів: найефективнішим “стабілізатором” відкритого стану рибосоми виявляється форма EF2·GDP·P<sub>i</sub>.

Отже, сценарій транслокації можна уявити наступним чином (рис. 8.25).

- Після транспептидації рибосома перебуває в рівновазі між закритою та відкритою формою. Тільки друга форма здатна до транслокації, але вона у той самий час є більш високоенергетичною – рівновага зсунута в бік закритої форми.
- Взаємодія з EF2·GTP стабілізує відкриту форму рибосоми. Швидко після зв'язування, за рахунок взаємодії з сайтом G-білків у основі стебла L7/12, EF2 набуває GTPазної активності – відбувається гідроліз GTP і перехід EF2 у структурну форму, що найефективніше сприяє розблокуванню рибосоми.
- У розблокованій формі рибосома рухається вздовж мРНК (зв'язок мРНК з молекулами тРНК зберігається, тобто вона переміщується також відносно тРНК). Цей рух є броунівським (одномірною дифузійною під дією теплових флуктуацій), але спорідненість пептидил-тРНК до Р-сайта робить напрямок до 3'-кінця мРНК більш імовірним. Крім того, “тРНК-подібний” структурний домен EF2 взаємодіє з маленькою субодиницею в межах А-сайта, сприяючи витісненню звідти антикодонової частини пептидил-тРНК.
- Одночасно з переміщенням, і незалежно від нього, відбувається дисоціація неорганічного фосфату. EF2, який залишається в комплексі з GDP, втрачаючи спорідненість, дисоціює від рибосоми, яка, у свою чергу, повертається до закритого структурного стану.

У результаті система є готовою до наступного елонгаційного циклу: новий кодон опинився в А-сайті, сам А-сайт є вільним від тРНК. Деаміноацильована тРНК, яка в результаті транслокації опинилася в Е-сайті рибосоми, може дисоціювати звідти відразу, або залишитися до етапу зв'язування наступної aa-тРНК.



**Рис. 8.25.** Послідовність основних подій при транслокації рибосоми

Описані механізми функціонування рибосоми під час елонгаційного циклу дозволяють стверджувати, що рибосома працює за принципами молекулярної машини (див. розділ 2) – пристрою, який з високою швидкістю (у середньому 10–15 елонгаційних циклів за секунду) і точністю (частота помилок  $\sim 10^{-4}$ ) здійснює білковий синтез. Під час елонгаційного циклу рибосома осцилює між закритим структурним станом, який забезпечує каталіз реакції нарощування амінокислотного ланцюга, і відкритими станами, в яких здійснюється специфічний відбір амінокислот згідно з генетичним кодом та переміщення рибосоми вздовж матриці. Перемикання конформацій залежить від взаємодій з факторами елонгації, які фіксують проміжні відкриті стани рибосоми, забезпечуючи прискорення перебудов. Гідроліз двох молекул GTP під час елонгаційного циклу використовується для звільнення факторів і повернення рибосоми до закритої жорсткої конформації. Рухомою силою для переміщення рибосоми та її частин є хаотичний тепловий рух, але молекулярна конструкція рибосоми, її взаємодії з тРНК, матрицею та факторами елонгації, і зміни цих взаємодій уна-



слідок хімічних реакцій (транспептидація, гідроліз GTP) задають певну траєкторію рухів і каналізують їх у певних напрямках. Аналогічні механізми використовуються й при ініціації та термінації трансляції.

### Ініціація трансляції

Розпочинатися процес елонгації трансляції має зі стартового кодо- на мРНК, і в першому елонгаційному циклі, коли немає пептидил- тРНК у Р-сайті, якась стартова – ініціаторна – аа-тРНК має виконувати цю роль. Отже, сутність ініціації полягає в:

- 1) упізнанні стартового кодону, який задає початок і рамку зчитування інформації;
- 2) збиранні рибосоми з двох субодиниць на мРНК у зоні стартового кодону;
- 3) завантаженні ініціаторної аа-тРНК на стартовий кодон і водночас у Р-сайт рибосоми.

Стартовим кодоном виступає здебільшого метіоніновий кодон AUG, відповідно, ініціаторною є завжди Met-тРНК<sub>i</sub><sup>Met</sup> (індекс “i” вказує на те, що це – саме *ініціаторна* метіонінова тРНК, тобто вона відрізняється за своєю структурою від звичайної тРНК<sup>Met</sup>, яка використовується при елонгації). У прокаріотів, на відміну від еукаріотів, аміногрупа метіоніну ініціаторної тРНК є формільованою (N-формілметіонін, f

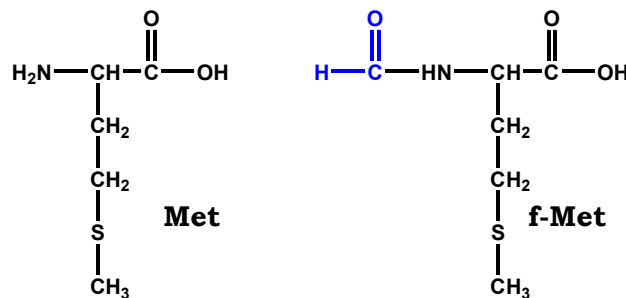


Рис. 8.26. Метіонін і N-формілметіонін

Механізми ініціації трансляції суттєво відрізняються у про- та еукаріотів. Еукаріотична мРНК після синтезу транспортується з ядра до цитоплазми, де й відбувається ініціація. Кожна молекула мРНК містить тіль-

ки одну рамку зчитування (один стартовий кодон), елементи системи ініціації взаємодіють спочатку з кепом на 5'-кінці молекули, після чого відбувається сканування матриці з метою пошуку стартового кодону – ініціація відбувається за так званим *скануючим механізмом*.

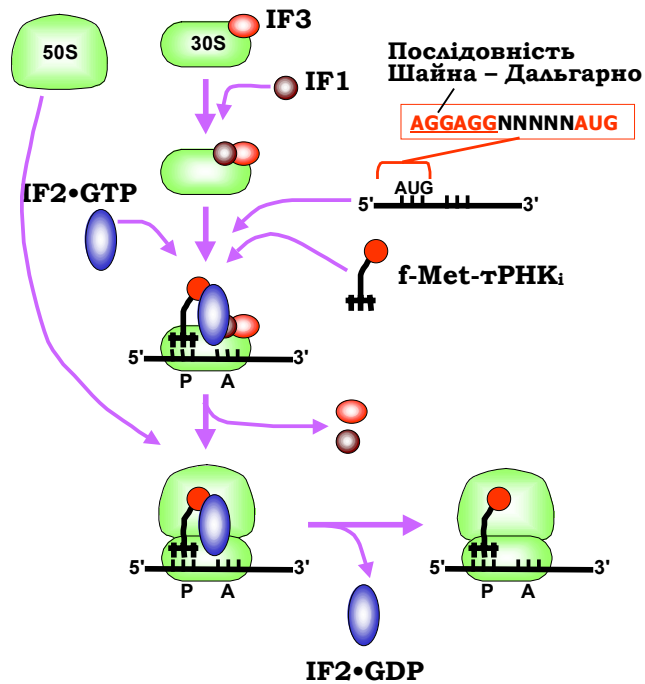
Прокаріотична трансляція тісно узгоджена з транскрипцією і відбувається під час синтезу мРНК (див. розділ 5). При цьому мРНК, що синтезується на оперонах, часто містить кілька послідовних рамок зчитування і кілька стартових кодонів. Ініціація трансляції відбувається окремо на кожному з них усередині мРНК, і впізнання цих стартових кодонів не залежить від їхнього розташування відносно 5'-кінця матриці – *внутрішня ініціація*.

### ***Ініціація трансляції у прокаріотів***

Послідовність подій, що відбуваються при ініціації трансляції у прокаріотів, показано на рис. 8.27. Рибосома вступає в процес у вигляді двох окремих субодиниць, дисоційований стан яких підтримується завдяки білкового фактору IF3 (IF – *I*nitiation *F*actor), зв'язаного з маленькою субодиницею. У присутності IF3 з маленькою субодиницею зв'язується також білок IF1, який блокує зону майбутнього А-сайта тРНК.

Потрійний комплекс маленька субодиниця–IF3–IF1 є субстратом взаємодії з трьома наступними елементами:

- Фактор ініціації IF2 – G-білок, який зв'язується з маленькою субодиницею рибосоми в комплексі з GTP і дещо нагадує за своєю структурою елонгаційний фактор EF2.
- Ініціаторна молекула тРНК – f-Met-тРНК<sub>i</sub><sup>Met</sup>, антикодонова частина якої взаємодіє з Р-сайтом маленької субодиниці, оскільки А-сайт є заблокованим фактором IF1.
- мРНК, яка взаємодіє з маленькою субодиницею таким чином, що стартовий кодон розміщується в зоні Р-сайта. Загалом з маленькою субодиницею взаємодіє ділянка мРНК довжиною ~30 нуклеотидів. Специфічне розміщення стартового кодона в Р-сайті забезпечується комплементарним упізнанням між ділянкою рРНК 16S (у 3'-кінцевій зоні) і консервативною *послідовністю Шайна – Дальгарно* (John Shine, Lynn Dalgarno), яка розміщена в мРНК за 5–9 нуклеотидів від стартового кодона в напрямку до 5'-кінця. Саме наявність послідовності Шайна – Дальгарно й робить даний кодон AUG стартовим, відрізняючи його від звичайного метіонінового кодона.



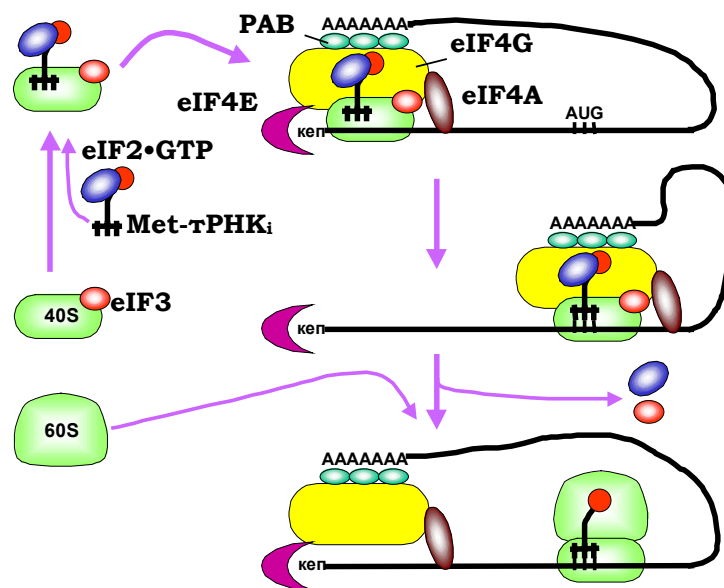
**Рис. 8.27.** Послідовність основних подій при ініціації трансляції у прокаріотів

Зв'язування названих трьох елементів, яке приводить до утворення преініціаторного комплексу, відбувається незалежно один від одного і в будь-якому порядку, зокрема ініціаторна тРНК взаємодіє з преініціаторним комплексом незалежно від стартового кодона. Наступним кроком є впізнання кодона антикодоном у Р-сайті, полегшуване факторами IF2 і IF3. Утворення кодон-антикодонової подвійної спіралі індукує конформаційну зміну маленької субодиниці, унаслідок якої відбувається звільнення факторів IF1, IF3 і зв'язування великої рибосомної субодиниці. Взаємодія між великою субодиницею та IF2 індукує GTPазну активність останнього: відбувається гідроліз GTP та дисоціація IF2.

У результаті система трансляції є готовою для першого елонгаційного циклу: А-сайт наготові прийняти aa-тРНК з другою амінокислотою поліпептидного ланцюга, на яку при транспептидації буде перенесено f-Met з ініціаторної тРНК.

### Ініціація трансляції в еукаріотів

Еукаріотична рибосома поза процесом трансляції також дисоційована на дві субодиниці, завдяки взаємодії маленької з фактором eIF3 (складається з 9–11 субодиниць). Крім того, дисоційований стан субодиниць підтримується кількома іншими факторами, що взаємодіють з маленькою (eIF3C) і великою (eIF6) субодиницями. З комплексом маленька субодиниця – eIF3 зв'язується потрібний комплекс ініціаторна метіонінова тРНК – eIF2•GTP (eIF2 є аналогом фактора елонгації eEF1).



**Рис. 8.28.** Послідовність основних подій при ініціації трансляції в еукаріотів.

Паралельно (рис. 8.28) відбувається впізнання кепа на 5'-кінці мРНК факторами ініціації групи eIF4: субодиниця eIF4E зв'язується з кепом, eIF4G є структурним модулем, що забезпечує цілісність комплексу та його взаємодію з іншими елементами ініціації, eIF4A є АТР-залежною РНК-геліказою, фактор eIF4B (не показаний на рис. 8.28)

взаємодіє з мРНК, стимулює АТФазну активність гелікази і при скануванні мРНК бере участь у впізнанні стартового кодона.

Наступним кроком є об'єднання кеп-асоційованого комплексу з таким, що асоційований з маленькою субодиницею (ініціаторна тРНК має бути присутньою на маленькій субодиниці, щоб ця остання могла приєднатися до кепу). Ефективність утворення цього преініціаторного комплексу на кепі підсилюється за рахунок взаємодії polyA-зв'язаних білків (РАВР) з eIF4G: polyA-послідовність на 3'-кінці мРНК також приєднується до преініціаторного комплексу, замикаючи мРНК у кільце.

Далі за рахунок геліказної активності eIF4A преініціаторний комплекс (залишаючи на кепі лише eIF4E) починає транслокацію: eIF4A здійснює гідроліз АТФ і пересувається вздовж мРНК у напрямку до 3'-кінця, одночасно руйнуючи дволанцюгові шпильки, якщо вони зустрічаються. Під час транслокації преініціаторний комплекс сканує послідовність мРНК, перевіряючи її на наявність стартового кодона. Зовсім не обов'язково при цьому перший кодон AUG, що зустрічається, сприймається як стартовий. Упізнання стартового кодона залежить від контексту послідовності, в якій він розташований. Найкращим контекстом, який максимально сприяє ініціації трансляції, є так звана *послідовність Козак* (Marilyn Kozak): GCC(A/G)CCA**AUGG**.

Упізнання стартового кодона викликає відповідні конформаційні зміни комплексу, звільнення факторів eIF3 і eIF4 (останній повертається на кеп разом polyA-ділянкою мРНК), приєднання великої рибосомної субодиниці та дисоціацію eIF2·GDP після гідролізу GTP. Дисоціація eIF3 і eIF2 полегшується фактором eIF5. Рибосома після цього вступає в етап елонгації трансляції, а на кепі починається збирання наступної рибосоми.

Таким чином, на мРНК одночасно працює кілька рибосом, які утворюють так звану *полісому* (рис. 8.29). Заціклення полісоми: 1) сприяє залученню зруйнованої при термінації трансляції у 3'-кінцевій зоні мРНК рибосоми до нового раунду ініціації; 2) забезпечує додатковий захист 3'-кінця мРНК (унаслідок його взаємодії з кеп-асоційованим комплексом) від нуклеазної деградації; 3) підвищує надійність системи трансляції – оскільки деградація мРНК може розпочинатися тільки з 3'-кінця (див. розділ 7), лише повноцінні матриці, що містять polyA-послідовність (і автоматично – кодуючу частину) залучаються до білкового синтезу.

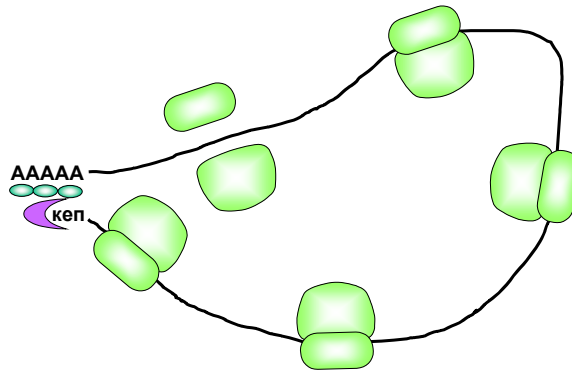


Рис. 8.29. Циркулярна полісома

### Термінація трансляції

Термінація білкового синтезу здійснюється за подібною для про- та еукаріотів схемою (рис. 8.30). Коли після чергового елонгаційного циклу (який стане останнім) в А-сайті опиняється один із трьох стоп-кодонів, він упізнається фактором термінації трансляції RF1 або RF2 (RF – Release Factor) – жодна тРНК не містить відповідних антикодонів. Цей фактор рекрутує до рибосоми комплекс RF3·GTP (аналог EF1). Під дією останнього спрацьовує пептидилтрансферазний центр рибосоми, намагаючись здійснити перенесення пептидилу. Оскільки в А-сайті немає звичайного субстрату транспептидації, відбувається перенесення на молекулу води – гідроліз зв'язку С-кінцевої амінокислоти з рибозою у складі пептидил-тРНК. На цьому закінчується власне термінація трансляції – з рибосоми звільняється поліпептид. Наступні події стосуються вже підготовки рибосоми до нового раунду трансляції.

Після гідролізу GTP здійснюється дисоціація факторів термінації та зв'язування замість них фактора відновлення рибосоми RRF (Ribosome Recycling Factor) і фактора елонгації EF2 в комплексі з GTP. Сумісна дія цих факторів у координації з гідролізом GTP спричиняє дисоціацію рибосомних субодиниць. Із маленькою субодиницею замість дисоційованих RRF і EF2 зв'язується фактор ініціації трансляції IF3 (який, таким чином, можна вважати також і фактором термінації). IF3 сприяє визволенню тРНК і мРНК з маленької субодиниці й запобігає взаємодії між субодиницями. У результаті субодиниці рибосоми знов можуть залучитися до ініціації трансляції.

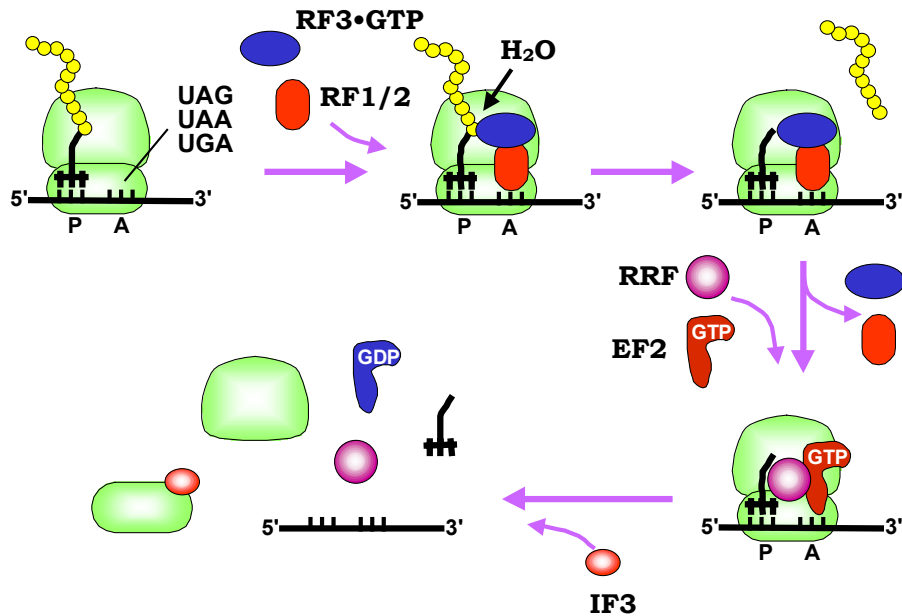


Рис. 8.30. Послідовність основних подій при термінації трансляції

### Регуляція трансляції

Прокаріотична мРНК має досить невеликий час життя – практично вона існує під час транскрипції та відразу після неї. Регуляція експресії гена у прокаріотів здійснюється переважно на рівні транскрипції (хоча білковий синтез може бути залученим до такої регуляції – див. розділ 5). Для еукаріотів регуляція трансляції є, навпаки, дуже важливим окремим елементом загальної регуляції експресії.

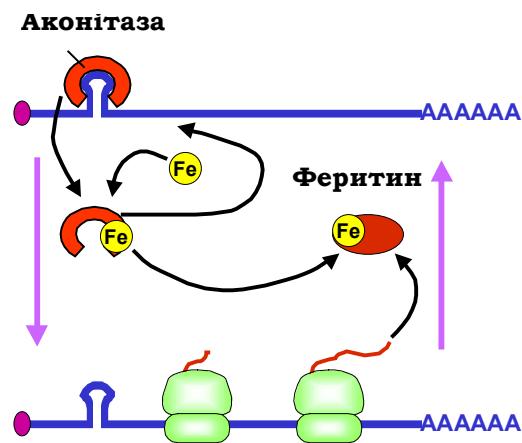
**Регуляція на рівні ініціації трансляції** – це один із найважливіших механізмів. Ефективність ініціації трансляції залежить насамперед від:

- Контексту послідовності, в якій знаходиться стартовий кодон – відхилення цього контексту від послідовності Козак (див. вище) утруднює впізнання стартової точки і, відповідно, зумовлює необхідність позитивної регуляції.

- Відстані стартового кодона від 5'-кінця мРНК – довге сканування матриці під час ініціації підвищує ймовірність руйнування преініціаторного комплексу.
- Наявності / відсутності дволанцюгових шпильок у 5'-кінцевій зоні мРНК, що не транслюється (див. рис. 7.10), які гальмують процес сканування.
- Присутності / відсутності регуляторних білків, які зв'язуються у 5'-кінцевій зоні.

Крім того, у деяких випадках (РНК пікорнавірусів, мРНК факторів росту фібробластів) процес ініціації відбувається за ефективним внутрішнім механізмом незалежно від сканування: завдяки наявності особливих послідовностей (IRES – Internal Ribosome Entry Site), що безпосередньо впізнаються елементами системи ініціації.

Один із прикладів участі білків у негативній регуляції ініціації трансляції наведено на рис. 8.31. Білок, який утримує залізо в клітині – феритин – має з'являтися негайно після проникнення в цитоплазму цього необхідного, але отруйного у вільному вигляді мікроелемента. Тому мРНК феритину завжди міститься в цитоплазмі, у відсутності заліза – у неактивному стані. Інактивація забезпечується за рахунок зв'язування репресора (який блокує процес сканування) зі шпилькою у 5'-кінцевій зоні.



**Рис. 8.31.** Регуляція синтезу феритину

Як репресор виступає аконітаза (один із ферментів циклу Кребса (Hans Adolf Krebs)). При появі заліза воно зв'язується аконітазою, що викликає звільнення аконітази від мРНК і, як наслідок, активацію



трансляції феритину. Синтезований феритин зв'язує залізо, забираючи його в аконітази, яка повертається на мРНК і зупиняє трансляцію.

Мішенню глобальної регуляції на рівні ініціації є також фактори ініціації. Наприклад, фосфорилування / дефосфорилування eIF-2 приводить, відповідно, до його деактивації / активації.

Крім регуляторних білків, до регуляції трансляції залучаються також маленькі РНК. Серед них є такі, що мають стимулюючий вплив на трансляцію: РНК деяких вірусів, РНК-продукти транскрипції Alu-повторів гальмують фосфорилування eIF-2, сприяючи активації трансляції. Молекули мікроРНК (див. розділ 6), навпаки, гальмують трансляцію (частіше на етапі сканування матриці) за рахунок комплементарної взаємодії з мРНК.

**Елонгація трансляції** також є мішенню регуляторних впливів. Швидкість елонгації варіює в досить широких межах відносно середнього значення, іноді рибосома може зупинятися на певний час. Паузи відбуваються на кодонах, які відповідають мінорним тРНК, зміна концентрації яких є одним зі шляхів регуляції швидкості трансляції. Підвищення загальної швидкості трансляції пов'язане з адаптацією наборів aa-тРНК – узгодженням між частотою найуживаніших кодонів і відповідних ізоакцепторних тРНК.

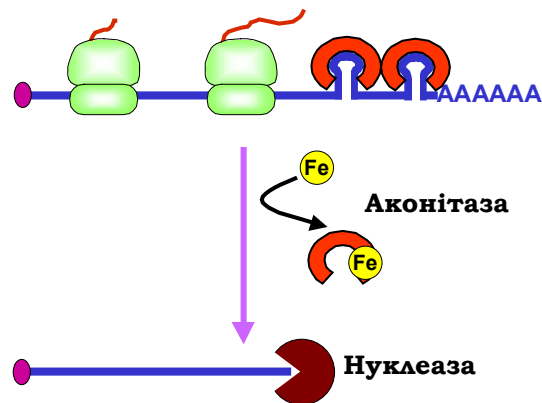
Наявність шпильок усередині кодуєчої частини мРНК також гальмує елонгацію, як і репресори, які впізнають елементи послідовності мРНК.

Однією з мішеней регуляції є також фактори елонгації, котрі (як і фактори ініціації) піддаються певним модифікаціям, що зумовлюють активацію / інактивацію факторів. Екстремальний приклад – дифтерійний токсин, який інактивує фактор eEF-2 і повністю вимикає білковий синтез.

Особливим випадком є включення в кілька важливих білків (присутні в усіх організмів ферменти, що належать до класу оксидоредуктаз) модифікованої амінокислоти селеноцистеїну (яка є елементом активного центру цих ферментів). Селеноцистеїнова тРНК (продукт особливого тРНК-гена) здатна впізнавати внутрішній стоп-кодон UGA, флаанкований (у напрямку до 3'-кінця) специфічною шпилькою. Ця тРНК спочатку акцептує Ser, який далі перетворюється на селеноцистеїн відповідними ферментами (відбувається заміна OH-групи на групу SeH, див. рис. 2.2). Селеноцистеїн-тРНК зв'язується зі специфічним аналогом елонгаційного фактора EF1 (позначається як SelB у прокариотів і як SECIS в еукаріотів). Шпилька в мРНК упізнається або самим цим специфічним фактором елонгації (прокариоти), або ще одним

білком, з яким цей фактор елонгації взаємодіє (еукаріоти), що й забезпечує зв'язування селеноцистеїн-тРНК з А-сайтом рибосоми.

**Зміна часу життя мРНК** – ще один шлях регуляції трансляції. Час життя мРНК у цитоплазмі залежить головним чином від ступеня захищеності 3'-кінця від екзонуклеаз (оскільки кеп є нечутливим до нуклеаз). Досить типову ситуацію ілюструє приклад регуляції трансляції іншого білка, який має відношення до обміну заліза, – рецептора трансферину (рис. 8.32).



**Рис. 8.32.** Регуляція синтезу рецептора трансферину

Трансферин є позаклітинним переносником заліза, транспорт заліза в цитоплазму залежить від кількості рецепторів трансферину в мембрані. У відсутності заліза всередині клітини на мРНК рецептора здійснюється синтез білка, мРНК при цьому стабілізована: аконітаза зв'язана зі шпильками в 3'-кінцевій зоні, що блокує нуклеазну деградацію. Зростання концентрації заліза викликає дисоціацію аконітази, і починається швидка деградація мРНК.

### Формування просторової структури білка

Термінація білкового синтезу на рибосомі не означає утворення функціонально активної молекули білка. Серед різних важливих операцій, яким піддається синтезований поліпептидний ланцюг, слід згадати такі, що їх можна об'єднати під назвою *білкового процесингу* – дозрівання поліпептидного ланцюга.

- Часткова протеолітична деградація – відщеплення кінцевих ділянок ланцюга, частіше на N-кінці, або іноді розрізання ланцюга на окремі фрагменти. Часткова деградація є особливо характерною для секреторних білків і ферментів, які мають гідролітичні активності, – така деградація переводить білок у активну форму.
- Утворення дисульфідних містків – ковалентних S-S-зв'язків між двома наближеними у просторі залишками Cys. Формування містків прискорюється спеціальним ферментом – дисульфідізомеразою.
- Ковалентні посттрансляційні модифікації амінокислотних залишків (фосфорилування, ацетилювання, глікозилування тощо), приєднання полісахаридів, ліпідів, простетичних груп і кофакторів небілкової природи.
- Білковий сплайсинг – досить екзотична операція, описана для кількох десятків білків різних таксономічних груп, від бактерій до хребетних. У процесі такого сплайсингу центральна частина ланцюга (*інтейн*) вирізається, а два кінцеві фрагменти (*екстейни*) зшиваються між собою в автокаталітичній реакції. Реакція не потребує не тільки зовнішнього ферменту, а й джерела енергії: подібно до сплайсингу мРНК (розділ 7), білковий сплайсинг здійснюється шляхом заміни одних зв'язків іншими.

Кожна з названих операцій не є універсальною. Проте будь-який поліпептидний ланцюг, незалежно від того, потребує він процесингу чи ні, має сформувати просторову структуру для набуття певної активності (зрозуміло, що не йдеться про неструктуровані білки, згадані в розділі 2).

### ***Закономірності укладання білкової глобули***

Перше й головне твердження щодо процесу укладання поліпептидного ланцюга в нативну глобулу полягає в тому, що ***просторова структура нативного білка і шлях її формування повністю визначаються амінокислотною послідовністю, і тільки амінокислотною послідовністю.***

Процес укладання білка в нативну структуру – це пошук конформації, яка відповідає мінімуму вільної енергії для даної послідовності амінокислот, і, відповідно, є найстабільнішою. Певне уявлення про своєрідний “енергетичний ландшафт”, що відповідає загальному набору різноманітних конформацій амінокислотного ланцюга, дає рис. 8.33 (зрозуміло, що насправді кількість конформаційних коорди-

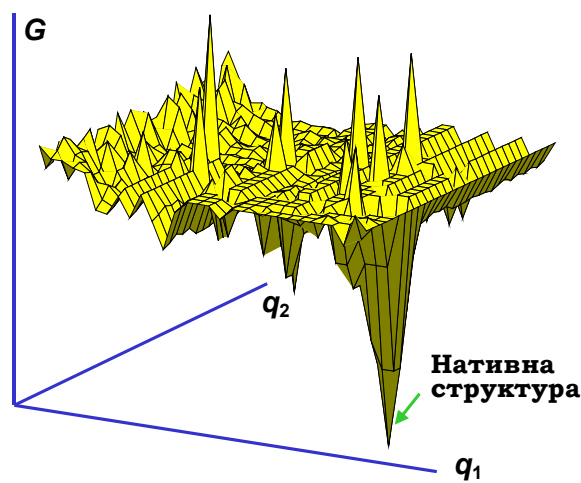
нат значно більше двох – конформаційний простір є багатовимірним). Нативна структура глобулярного білка завжди відокремлена від інших конформацій особливо глибоким мінімумом – знаходиться в енергетичній ямі (рис. 8.33). Наявність такої ями, яка відповідає стабільному щільно упакованому гідрофобному ядру (див. розділ 2), забезпечує надійність функціонування білка: теплові флуктуації не здатні вивести структуру з такої ями.

Величезна кількість конформацій, у принципі доступних для поліпептидного ланцюга, поставила проблему, відому як *парадокс Левінтала* (Cyrus Levinthal). Якщо кожен амінокислотний залишок має ~10 конформацій, ланцюг зі 100 залишків –  $10^{100}$  конформацій. За найнижчою оцінкою перемикавання однієї конформації відбувається за  $\sim 10^{-13}$  с, тобто, щоб перебрати всі конформації потрібно щонайменше  $10^{80}$  років, що значно перевищує вік Всесвіту. Аналогія конформаційного простору з ландшафтом (рис. 8.33) відразу дає рішення цього парадоксу: річка не перебирає увесь ландшафт, вона просто тече одним або кількома альтернативними низькими руслами, впадаючи врешті-решт у найнижчу яму.

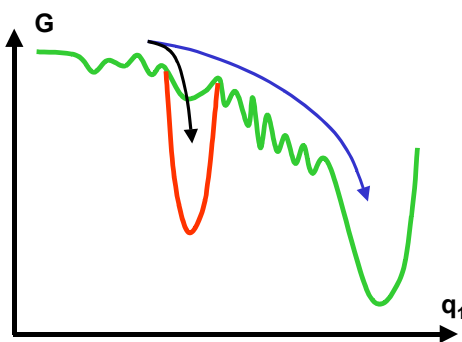
Так само й для поліпептидного ланцюга існують певні “виділені” шляхи укладки, які відповідають ієрархії структури білка (див. розділ 2). Спочатку за рахунок локальних взаємодій утворюються елементи вторинної структури з гідрофобними поверхнями. Водночас завдяки гідрофобним взаємодіям вони “злипаються” між собою, утворюючи розплавлену глобулу. Цей процес відбувається досить швидко – за кілька мікросекунд. На другому етапі, який потребує значно більшого часу (кілька секунд чи навіть хвилин), реалізується щільна упаковка глобули за рахунок вандерваальсових взаємодій з утворенням твердої нативної молекули, що відповідає глобальному мінімуму вільної енергії. Конформація ланцюга змінюється, прямуючи через локальні мінімуми енергії, які відокремлені один від одного невисокими бар’єрами. Бар’єри легко долаються за рахунок теплових флуктуацій, і ланцюг нарешті опиняється у глибокому глобальному мінімумі. Але на такому шляху можливі енергетичні пастки (рис. 8.34) – досить глибокі мінімуми, з яких неможливо вибратися за розумний час, оскільки глибина такої локальної ями перевищує енергію теплових флуктуацій.

Однією з головних причин виникнення енергетичних пасток є агрегація поліпептидних ланцюгів. У стані розплавленої глобули, яка відрізняється від нативної великомасштабними рухами своїх частин, час від часу відбувається вихід на поверхню значної частини гідрофобних залишків. Відповідно, недоструктуровані поліпептидні ланцюги будуть

прагнути утворити великі неспецифічні агрегати з метою заекранувати неполярні групи від молекул води. Агрегат і є пасткою – опиняючись у мінімумі вільної енергії, він практично не допускає дисоціації своїх елементів і не дозволяє їм продовжити процес пошуку нативної структури. Зрозуміло, що ефективність агрегації зростає при підвищенні концентрації поліпептидів, а також при зростанні температури (яке підсилює і рухливість глобули, і гідрофобні взаємодії (див. розділ 1)).



**Рис. 8.33.** Схематична тривимірна проекція конформаційного простору поліпептидного ланцюга – залежність вільної енергії ( $G$ ) від двох параметрів ( $q_1$ ,  $q_2$ ), які характеризують конформацію ланцюга



**Рис. 8.34.** Енергетична пастка на шляху укладання поліпептидного ланцюга вздовж однієї конформаційної координати (див. рис. 8.33)

Повертаючись до сформульованої вище догми про визначення шляху укладки білка його первинною структурою, слід зауважити, що процес укладки може просто не дійти до свого кінця. Численні експерименти *in vitro* свідчать про те, що догма повністю підтверджується: практично будь-який білок у денатурованому стані можна без жодних специфічних факторів ренатурувати – відновити його нативну структуру й функцію. Але успішна ренатурація є можливою, якщо штучно створити певні умови та режим її здійснення:

- Перш за все, бажано проводити ренатурацію поступово, повільно повертаючи зовнішні умови до фізіологічних. У результаті на проміжних етапах ренатурації будь-який локальний мінімум енергії не буде надто глибоким і не зможе спрацювати як енергетична пастка: ланцюг зможе швидко перебирати різні не дуже стабільні конформації, реалізуючи конформаційну рівновагу, зсунуту до певного найнижчого мінімуму енергії.
- Другою важливою умовою є здійснення ренатурації при низькій концентрації білка – з метою запобігти неспецифічній агрегації.

Крім цих загальних правил, у кожному випадку часто доводиться підбирати умови ренатурації окремо: визначати оптимальний режим зміни концентрації солі, температуру, концентрації субодиниць білка з четвертинною структурою тощо.

Зрозуміло, що все це неможливо здійснити *in vivo* – достатньо сказати, що загальна середня концентрація білків у клітині становить ~100 мг/мл або більше. Тобто клітинні умови є дуже агресивними для поліпептидного ланцюга, що перебуває у процесі пошуку своєї нативної конформації, і без спеціальних механізмів більшість білків просто не встигли б після свого синтезу таку конформацію знайти, швидко опинившись у складі агрегату. Такі механізми – створення рівноважних умов поступового пошуку нативної структури та захист від агрегації – забезпечуються спеціальними білками – *шаперонами* (від фр. *chaperon* – капюшон, або особа, що супроводжує незайману дівчину).

**Шаперони не визначають ані нативну структуру білка, ані шлях її укладання – і те, й інше детермінується амінокислотною послідовністю. Головна функція шаперонів – забезпечити умови для швидкого пошуку нативної конформації**, створюючи своєрідний “інкубатор” для неструктурованого поліпептидного ланцюга.

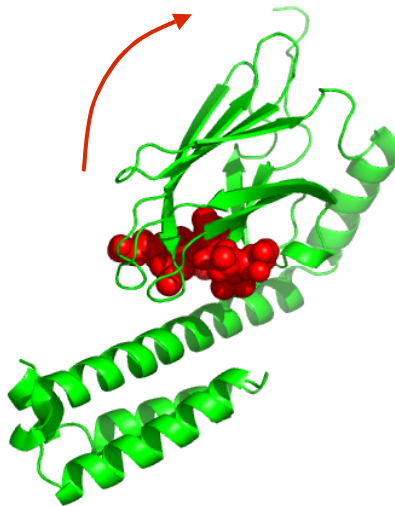
Серед досить великої кількості шаперонів є специфічні – зокрема такі, що обслуговують процеси збирання мультибілкових комплексів і субодиничних білків. Більш універсальні шаперони відносять до

родини **hsp** – білків теплового шоку (heat shock proteins). Білки hsp є загальним компонентом клітин для всіх таксономічних груп, у пойкилотермних організмів транскрипція на генах hsp активується у відповідь на зростання температури до 30–40 °C при гальмуванні транскрипції більшості інших генів (тепловий шок).

### Шаперони hsp70

Процес укладання поліпептидного ланцюга у глобулу має відбуватися саме *пост*трансляційно – поза рибосомою. Тунель, через який синтезований ланцюг виходить за межі рибосоми (рис. 8.9) має довжину ~100 Å, середній діаметр ~15 Å – у тунелі може розміститися тільки витягнутий ланцюг довжиною ~30 амінокислот. Отже, укладання не може починатися в тунелі. Білки hsp70 (тут і далі цифрами позначено молекулярну вагу в кілодальтонах) зв'язують розгорнутий поліпептидний ланцюг відразу після того, як його частина виходить з рибосоми при білковому синтезі.

Молекула має два структурні домени – пептидзв'язувальний, який взаємодіє з елементом розгорнутого поліпептидного ланцюга (рис. 8.35), та АТФазний.



**Рис. 8.35.** Комплекс пептидзв'язувального домену бактеріального hsp70 (білок DnaK) із фрагментом поліпептиду (червоний) довжиною 7 амінокислот (1DKZ). Стрілка вказує напрямок конформаційної зміни білка, яка індукується АТФ

Зв'язування з розгорнутим ланцюгом здійснюється неспецифічно за рахунок гідрофобних взаємодій з неполярними амінокислотними залишками: ланцюг вкривається “шубою” молекул шаперона. Власне, *підтримання розгорнутого стану ланцюга при запобіганні агрегації* – одна з основних функцій hsp70.

Молекула шаперону може існувати у двох структурних станах залежно від типу зв'язаного ліганду – АТФ чи АДФ. У комплексі з АТФ реалізується структурна форма, яка допускає швидку рівновагу між зв'язаним / дисоційованим поліпептидом. Гідроліз АТФ (здійснюється АТФазним доменом hsp70 за сприяння ко-шаперонів hsp40) викликає структурну зміну з міцною фіксацією поліпептиду пептидзв'язувальним доменом. За допомогою інших ко-шаперонів АДФ знову замінюється на АТФ (рис. 8.36).

У процесі такої циклічної АТФ-залежної зміни спорідненості шаперону до поліпептиду створюються умови для рівноважного пошуку нативної конформації під час короткострокової дисоціації. Цей механізм є цілком аналогічним до такого, що розглядався в розділі 4 у зв'язку зі збиранням хроматину за допомогою проміжних переносників гістонів, які також називають хроматиновими шаперонами (див. рис. 4.11): звільнений поліпептид з нестабільною структурою (з великою кількістю гідрофобних залишків на поверхні) знов розгортається hsp70 і здійснює нову спробу при наступній дисоціації. Для багатьох білків такий процес виявляється достатнім, щоб знайти нативну конформацію – структуру з гідрофобним ядром усередині, яка не може бути субстратом для hsp70. Іноді процес укладки поліпептидного ланцюга відбувається котрансляційно: укладається N-кінцева частина ланцюга, яка вже вийшла з рибосоми, тоді як решта поліпептиду ще синтезується. Зокрема, такий процес іноді спостерігається для мультидоменних білків.

Інші поліпептиди переносяться на інші шаперони (наприклад, на шапероніни, див. нижче), частина нестабільних білків, звільнюючись від hsp70, піддається протеолітичній деградації.

У складі міцного комплексу з hsp70·АДФ поліпептид транспортується крізь мембрану до інших компартментів клітини, усередину мітохондрій, до місця збірки мультибілкових комплексів тощо: близько третини всіх синтезованих білків мають бути транспортованими в інше місце клітини або секретованими в позаклітинний простір. Транспортування комплексу поліпептиду з hsp70 відбувається за рахунок наявності у складі ланцюга специфічних сигнальних послідовностей (які відщеплюються в місці призначення), що впізнаються білковими факторами та шаперонами типу hsp90.



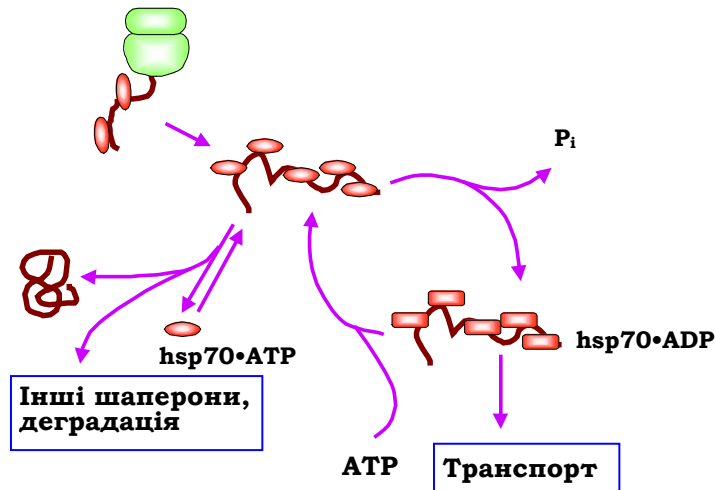


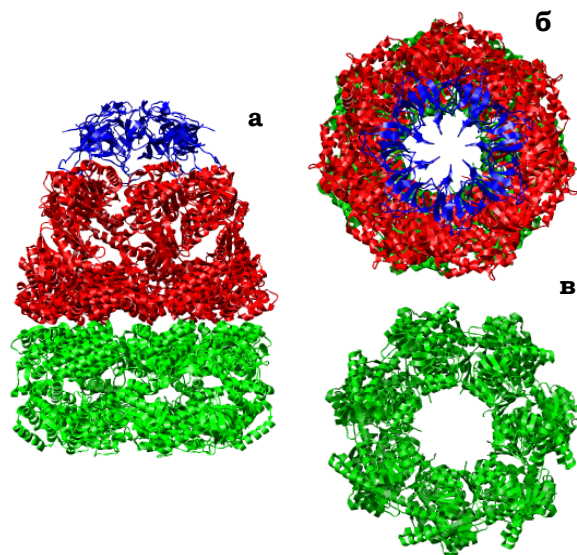
Рис. 8.36. АТФ-залежний робочий цикл hsp70

У деяких випадках hsp70 та аналогічні білки теплового шоку беруть участь у контролі активності регуляторних білків. Першим прикладом є система регуляції транскрипції самих білків теплового шоку пойкилотермних організмів. За низьких температур фактор транскрипції HSF (Heat Shock Factor) перебуває в неактивній формі – у недоструктурованому вигляді в комплексі з білками hsp40, hsp70, hsp90, які синтезуються за рахунок фонові активності своїх генів. Підвищення температури, яке зумовлює дестабілізацію багатьох білків і експонування гідрофобних груп у їхньому складі, викликає зв'язування hsp з цими білками – HSF звільняється, набуває остаточної структури і зв'язується з промоторами генів hsp, активуючи їхню транскрипцію. Підвищення концентрації hsp знов інактивує HSF. Інший приклад залучення білків теплового шоку до регуляції транскрипції вже розглядався у зв'язку з гормональними рецепторами (розділ 6): після синтезу рецептора hsp70 (що зв'язується з його ланцюгом після виходу з рибосоми) замінюється на hsp90 та кошаперон p23, і в складі цього комплексу гормонзв'язувальний домен підтримується в недоструктурованому стані до моменту зв'язування з гормоном (див. рис. 6.13).

### Шапероніни

Особливий клас шаперонів, побудованих із білків hsp60, об'єднують у групу *шаперонінів*. Бактеріальний шаперонін (шаперонін I) містить два комплекси, кожен з яких сформований сімома молекулами hsp60 (інша назва GroEL). Комплекс має форму кільця з каналом  $\sim 45$  Å у діаметрі, утворюючи своєрідну мікропробірку (рис. 8.37), яка може закриватися кришечкою, утвореною із 7 молекул hsp10 (інша назва GroES).

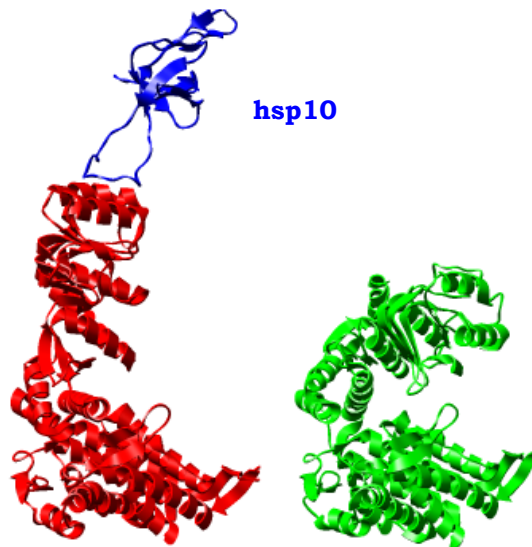
Еукаріотичні аналоги (шапероніни II) мають подібну будову, але кожне кільце формується з восьми або дев'яти субодиниць, додаткові структурні домени яких утворюють кришечку, що закривається / відкривається в результаті структурних перебудов.



**Рис. 8.37.** Структура шапероніну GroEL (1WE3) у різних проекціях.  
Два семичленних кільця hsp60 зафарбовані зеленим і червоним,  
7 субодиниць hsp10 – синім

Кожна субодиниця кільця, і на цьому базується принцип роботи шапероніну, існує принаймні у двох структурних формах (рис. 8.38). Перша (у двох варіантах, що трохи розрізняються) реалізується в комплексі з АТФ або АДФ і характеризується: 1) високою спорідненістю до hsp10 (що сприяє утворенню кришечки); 2) низькою гідро-

фобністю своєї поверхні. Дисоціація ліганду (ADP після гідролізу АТР) індукує структурну перебудову зі взаємним переміщенням трьох структурних доменів. При цьому втрачається спорідненість до hsp10 (відкриття кришечки) і на внутрішню поверхню білка (усередині каналу мікробірки) експонуються гідрофобні групи.



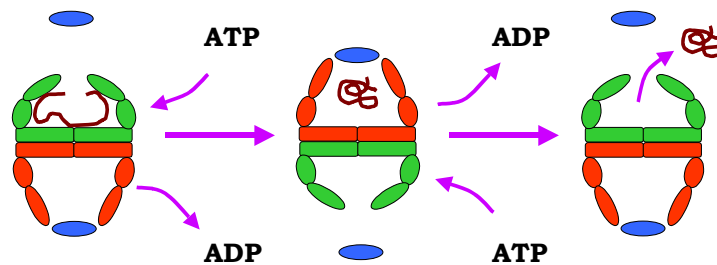
**Рис. 8.38.** Два структурні стани hsp60 (1WE3): зв'язаний з ADP у комплексі з молекулою hsp10 (червоний) і вільний від ліганду (зелений)

Зв'язування та гідроліз АТР, дисоціація ADP і відповідні перебудови відбуваються синхронно для семи субодиниць одного кільця (*позитивна* кооперативність у межах кільця), але два кільця працюють за принципом *негативної* кооперативності: перший стан одного кільця сприяє реалізації другого стану в іншому кільці (як на рис. 8.37).

Робочий цикл шапероніну показано на рис. 8.39.

Через відкритий отвір мікропробірки всередину потрапляє поліпептидний ланцюг. Кільце hsp60 (верхнє на рис. 8.39, розглянемо саме його) перебуває при цьому в стані, що не зв'язує жодного ліганду: кришечка відкрита, поліпептид фіксується всередині каналу в розгорнутому вигляді за рахунок гідрофобних взаємодій. Зв'язування семи молекул АТР викликає структурну перебудову: кришечка закривається, гідрофобність внутрішньої поверхні каналу зникає – поліпептид (повністю захищений від зовнішнього середовища) має змогу здійснювати власну укладку. На це він має ~20 с – час, поки існує зв'яза-

ний АТФ, після чого відбувається його гідроліз (за рахунок активності hsp60) і потім залишається зв'язаний АДФ. Зв'язування АТФ з нижнім кільцем індукує дисоціацію АДФ від субодиниць верхнього: кришечка відкривається, і білок зі сформованою нативною структурою – гідрофобних груп практично немає на його поверхні й він не прилипає до стінок каналу – виходить через отвір мікробірки. Для більшості білків проміжок у 20 с є цілком достатнім, щоб знайти нативну конформацію. Якщо цього не відбулося, процес може повторитися.



**Рис. 8.39.** Робочий цикл шапероніну. Структуру схематично зображено в розрізі: для кожного кільця показано по дві субодиниці, кожна складається з трьох доменів. Червоний і зелений колір відповідають різним структурним станам кільця (як на рис. 8.37, 8.38), синій – hsp10

### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. Назвіть основні структурні елементи будови тРНК? Як саме вони розташовані у просторі? Які взаємодії стабілізують просторову структуру тРНК?

2. Якою хімічною групою тРНК акцептується амінокислота? Назвіть групу амінокислоти, що задіяна в цьому зв'язку.

3. Які хімічні реакції каталізує АРСаза? Що таке активування амінокислоти?

4. Яким чином реалізується специфічність АРСаз щодо тРНК та амінокислот?

5. Що таке рибосома? З чого вона складається? Опишіть основні риси просторової будови рибосоми.

6. Які функціональні сайти має рибосома? Їхнє розташування відносно один одного та щодо субодиниць рибосоми?

7. Як розподілені функції рибосоми між двома субодиницями?

8. Порівняйте структурні та функціональні особливості рибосомних РНК і білків.
9. З яких етапів складається елонгаційний цикл рибосоми?
10. Опишіть структуру та функціональне значення факторів елонгації EF1 і EF2.
11. У чому полягає роль гідролізу GTP при елонгації трансляції?
12. Як рибосома здійснює свою декодуючу функцію?
13. Що таке акомодация aa-тРНК на рибосомі та як саме вона відбувається?
14. У чому полягає реакція подовження поліпептидного ланцюга на одну амінокислоту? Як відбувається каталіз цієї реакції?
15. Яким чином відбувається транслокація рибосоми вздовж мРНК?
16. Яка стадія елонгаційного циклу є найшвидкішою?
17. Порівняйте системи ініціації трансляції у про- та еукаріотів.
18. Опишіть основні стадії термінації трансляції.
19. За якими механізмами здійснюється регуляція білкового синтезу?
20. Назвіть основні стадії процесу укладання білкової глобули. Якою вільною енергією має характеризуватися нативна конформація білка порівняно з іншими конформаціями? Що таке енергетична пастка; як вона може виникати?
21. У чому полягає функціональна роль шаперонів?
22. Опишіть робочий цикл шаперонів hsp70.
23. Визначте структурну організацію та робочий цикл шаперонів.

#### **РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

- Киселев, Л.Л., Фаворова, О.О., Лаврик, О.И.* Биосинтез белков от аминокислот до аминоацил-тРНК. – М. : Наука, 1984.
- Финкельштейн, А.В., Птицын, О.Б.* Физика белка. Курс лекций. – М. : КДУ, 2005
- Andersen, G.R., Nissen, P., Nyborg, J.* Elongation factors in protein biosynthesis // Trends Biochem. Sci. – 2003. – Vol. 28. – P. 434–441.
- Bukau, B., Horwich, A.L.* The hsp70 and hsp60 chaperone machines // Cell. – 1998. – Vol. 92. – P. 351–366.
- Burbaum, J.J., Schimmel, P.* Structural relationships and the classification of aminoacyl-tRNA-synthetases // J. Biol. Chem. – 1991. – Vol. 266. –P. 16965–16968.

*Deuerling, E., Bukau, B.* Chaperone-assisted folding of newly synthesized proteins in the cytosol // *Crit. Rev. Biochem. Mol. Biol.* – 2004. – Vol. 39. – P. 261–277.

*Frank, J., Sengupta, J., Gao, H. et al.* The role of tRNA as a molecular spring in decoding, accommodation, and peptidyl transfer // *FEBS Letters.* – 2005. – Vol. 579. – P. 959–962.

*Laursen, B.S., Sørensen, H.P., Mortensen, K.K., Sperling-Petersen, H.U.* Initiation of protein synthesis in bacteria // *Microbiol. Mol. Biol. Rev.* – 2005. – Vol. 69. – P. 101–123.

*Mitra, K., Frank, J.* Ribosome dynamics: insights from atomic structure modeling into cryo-electron microscopy maps // *Annu. Rev. Biochem. Biomol. Struct.* – 2006. – Vol. 35. – P. 299–317.

*Moore, P.B., Steitz, T.A.* The structural basis of large ribosomal subunit function // *Annu. Rev. Biochem.* – 2003. – Vol. 72. – P. 813–850.

*Noller, H.F., Yusupov, M.M., Yusupova G.Z. et al.* Translocation of tRNA during protein synthesis // *FEBS Letters.* – 2002. – Vol. 514. – P. 11–16.

*Ogle, J.M., Ramakrishnan, V.* Structural insights into translational fidelity // *Annu. Rev. Biochem.* – 2005. – Vol. 74, 129–177.

*Preiss, T., Hentze, M.W.* Starting the protein synthesis machine: eukaryotic translation initiation // *BioEssays.* – 2003. – Vol. 25. – P. 1201–1211.

*Ramakrishnan, V.* Ribosome structure and the mechanism of translation // *Cell.* – 2002. – Vol. 108. – P. 557–572.

*Rodnina, M.V., Wintermeyer, W.* Peptide bond formation on the ribosome: structure and mechanism // *Curr. Op. Struct. Biol.* – 2003. – Vol. 13. – P. 334–340.

*Spirin, A.A.* Ribosome as a molecular machine // *FEBS Letters.* – 2002. – Vol. 514. – P. 2–10.

*Wilden, B., Savelsbergh, A., Rodnina, M.V., Wintermeyer, W.* Role and timing of GTP binding and hydrolysis during EF2-dependent tRNA translocation on the ribosome // *Proc. Natl. Acad. Sci. USA.* – 2006. – Vol. 103. – P. 13670–13675.

*Wintermeyer, W., Peske, F., Beringer, M., et al.* Mechanisms of elongation on the ribosome: dynamics of a macromolecular machine // *Biochem. Soc. Trans.* – 2004. – Vol. 32. – P. 733–737.

## Розділ 9

### РЕПЛІКАЦІЯ ДНК

It has not escaped our notice that the specific base pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

*J. Watson, F. Crick*

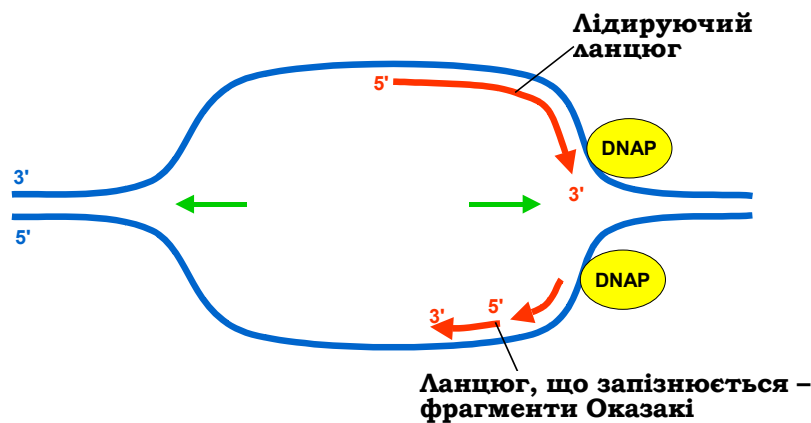
Процес подвоєння ДНК – реплікація (replication) – забезпечує відтворення спадкової інформації та передачу її до дочірніх клітин при мітозі й мейозі. Синтез ДНК відбувається при реплікації з використанням обох полінуклеотидних ланцюгів як матриць – за так званим *напівконсервативним механізмом*: дві дочірні молекули-копії містять один материнський ланцюг (що слугував матрицею) і один ланцюг, синтезований *de novo*. Включення нуклеотидів до ланцюга, що синтезується, детермінується матрицею за принципом комплементарності. Такий механізм реплікації став зрозумілим відразу після того, як Уотсоном і Кріком була сформульована модель подвійної спіралі ДНК.

Зростання ланцюга ДНК відбувається в напрямку від 5' до 3'-кінця. Субстратами реакції є 3'-кінцева ОН-група дезоксирибози зростаючого ланцюга та дезоксирибонуклеозидтрифосфати (dNTP, див. розділ 3). Фермент, що каталізує цю реакцію – *ДНК-залежна ДНК-полімераза* (DNA dependent DNA Polymerase, DNAP). Синтез ДНК за подібними механізмами здійснюється також при репарації пошкоджень і деяких інших процесах.

Молекулярні механізми реплікації є спільними для всіх живих організмів. На початку розділу буде розглянуто в основному бактеріальну систему реплікації. Далі описано особливості реплікації в еукаріотів, а також процеси синтезу ДНК при репарації.

## Реплікон

Реплікація ДНК починається з невеликої ділянки – *ориджина* (origin), де здійснюється ініціація процесу, головним моментом якої є розходження ланцюгів ДНК. Далі по ходу реплікації такий *реплікативний міхур* (рис. 9.1) розростається у двох протилежних напрямках. На кожному боці міхура існує так звана *реплікативна вилка*, у основі якої й відбувається синтез ДНК. Ділянку ДНК, де здійснюється реплікація, що розпочинається з однієї точки, називають *репліконом*. Бактеріальна хромосома, і зокрема в *E. coli*, часто містить тільки один ориджин, являє собою єдиний реплікон. У деяких бактерій може бути два реплікони на хромосому. Еукаріотична хромосома є полірепліконом – вона містить велику кількість точок ініціації.



**Рис. 9.1.** Реплікативний міхур – дві реплікативні вилки, що переміщуються у протилежних напрямках (ланцюги, які синтезуються, показано тільки для однієї з них)

У кожній реплікативній вилці працюють дві молекули ДНК-поліме-рази, що здійснюють синтез двох полінуклеотидних ланцюгів. Оскільки два ланцюги є антипаралельними (див. розділ 3), і синтез відбувається тільки в напрямку від 5'- до 3'-кінця, синтез тільки одного з ланцюгів може відбуватися (і відбувається) безперервно, починаючись від ориджина (рис. 9.1). Цей ланцюг називають *лідуючим* (leading strand), його 3'-кінець розташований поблизу від основи реплікативної вилки.



Синтез іншого ланцюга *розпочинається* від реплікативної вилки: синтезуються окремі фрагменти – так звані *фрагменти Оказакі* (Reiji Okazaki), які пізніше з'єднуються між собою (див. нижче). Довжина фрагмента Оказакі при реплікації у бактерій дорівнює 1–2 тис. нуклеотидів. Для синтезу кожного із фрагментів треба спочатку звільнити відповідний простір на матричному ланцюзі – пересунути реплікативну вилку вперед (рис. 9.1), відповідно, фрагментарний ланцюг називають *ланцюгом, що запізнюється* (lagging strand). Середня швидкість реплікації на одну реплікативну вилку становить ~750 нуклеотидів за секунду в бактерій, 60–90 нуклеотидів за секунду в еукаріотів. Синтез бактеріальної хромосоми відбувається за ~50 хвилин, повна реплікація ДНК еукаріотичної клітини – за кілька годин.

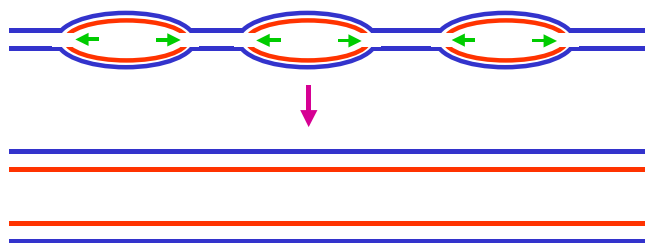
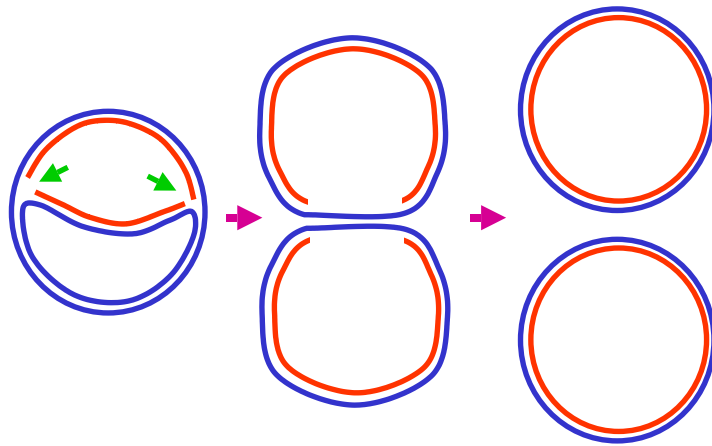


Рис. 9.2. Реплікація еукаріотичної хромосоми

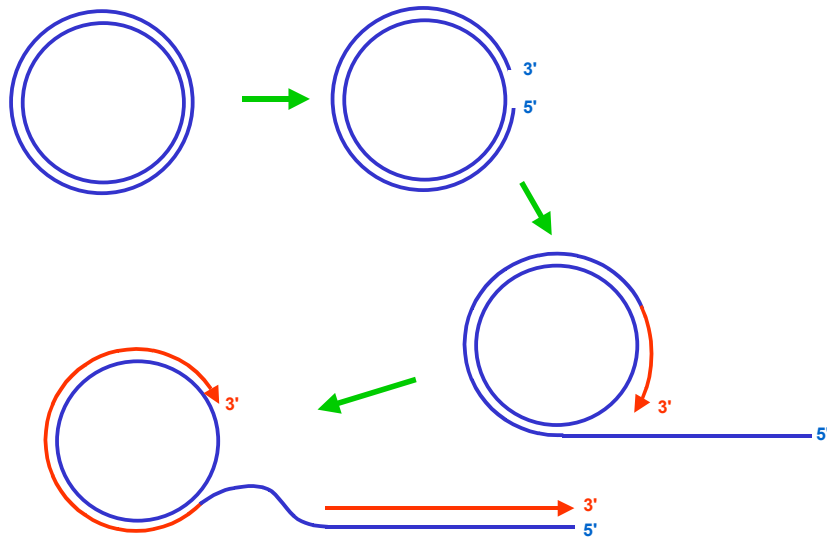
У більшості репліконів реплікація здійснюється таким чином, як зображено на рис. 9.1 – в обох напрямках. Сусідні реплікони еукаріотичної хромосоми врешті-решт “зустрічаються”, і в результаті утворюються дві копії ДНК хромосоми (рис. 9.2). Циркулярна бактеріальна хромосома також реплікується у двох напрямках з утворенням ідентичних циркулярних молекул ДНК (рис. 9.3). Аналогічно реплікуються окремі автономні реплікони бактерій – циркулярні плазмідди. Для деяких із них реплікація відбувається тільки в одному напрямку: в ориджині формується лише одна реплікативна вилка, яка рухається навколо кільця до вихідної точки.

Особливий випадок – реплікація циркулярних ДНК за механізмом *кільця, що котиться* (rolling circle), який реалізується для ДНК бактеріофагів та деяких плазмід. Один із варіантів такого механізму (для бактеріофага  $\lambda$ ) показано на рис. 9.4. В одному з ланцюгів циркулярної ДНК індукується одноланцюговий розріз, 3'-кінець, що утворився, добувається з використанням інтактного циркулярного ланцюга як матриці. Продовження цього процесу створює довгий одноланцюго-

вий 5'-кінцевий хвіст, що складається з тандемних повторів – копій кільцевої молекули. Кожен із цих повторів використовується як матриця для синтезу іншого ланцюга. Після цього лінійні дволанцюгові фрагменти вирізаються та упаковуються до частинок бактеріофага.



**Рис. 9.3.** Реплікація циркулярної бактеріальної хромосоми



**Рис. 9.4.** Реплікація циркулярної ДНК бактеріофага  $\lambda$

## ДНК-полімераза

У клітині *E. coli* працюють ДНК-полімерази трьох основних типів. Усі вони мають дві ферментативні активності: власне *полімеразну*, за рахунок якої до 3'-кінця ланцюга, що синтезується, приєднуються нуклеотиди, і *3'-екзонуклеазну*, яка використовується для редагування помилок – відщеплення помилкових нуклеотидів, щойно приєднаних до 3'-кінця.

- ДНК-полімераза I (або полімераза Корнберга (Arthur Kornberg)) – мономерний білок із мультидоменною структурою. На відміну від інших ДНК-полімераз вона має також додаткову 5'-екзонуклеазну активність. ДНК-полімераза I використовується як *допоміжна* полімераза при реплікації (див. нижче) та інших процесах синтезу ДНК.
- ДНК-полімераза II залучена до певних репараційних процесів.
- ДНК-полімераза III – основна реплікативна полімераза, дві копії якої працюють у реплікативній вилці. Складається з трьох субодиниць: субодиниця  $\alpha$  відповідає за полімеразну активність,  $\epsilon$  – за 3'-екзонуклеазну,  $\theta$  – виконує структурну роль.

## Структура ДНК-полімерази й полімеразна реакція

Структуру основної частини ДНК-полімерази I – так званій *фрагмент Кленова* (Hans Klenow) – показано на рис. 9.5 (додатковий домен полімерази відповідає за 5'-екзонуклеазну активність). Подібну структуру мають більшість інших про- та еукаріотичних ДНК-полімераз. Полімеразний активний центр розташований у межах *долоні* (palm) – щілини, яка оточена двома характерними структурними доменами: рухливими *пальцями* (fingers) і *великим пальцем* (thumb). Великий палець взаємодіє з маленьким жолобком подвійної спіралі, утвореної матричним ланцюгом і таким, що синтезується. Пальці взаємодіють із матричним ланцюгом, який при цьому трохи вигинається навкруг пальців, експонуючи чергову азотисту основу для взаємодії з НТР. Сайт зв'язування НТР утворюється пальцями, залишками активного центру, нуклеотидом матричного ланцюга та 3'-кінцем ланцюга, що синтезується. Рухливість пальців дозволяє їм існувати у двох конформаціях:

- *Відкрита* передбачає рух полімерази вздовж ДНК і швидкий перебір NTP.
- *Закрита*, коли пальці переміщуються в напрямку до великого пальця, жорстко фіксуючи NTP у активному центрі та забезпечуючи каталіз полімеразної реакції.

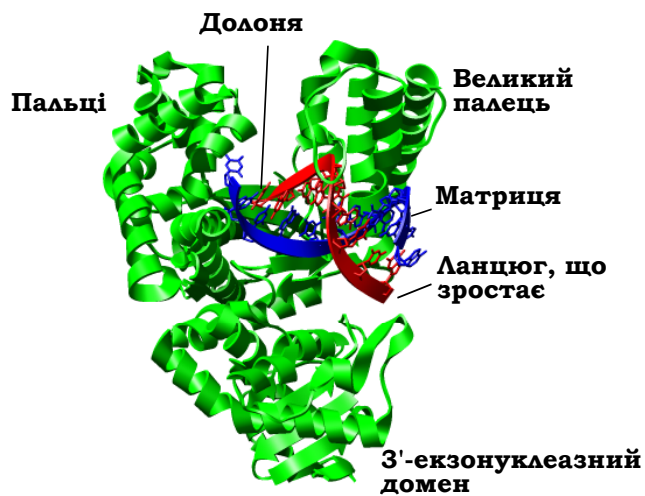


Рис. 9.5. Комплекс фрагмента Кленова ДНК-полімерази I із ДНК у полімеразному активному центрі (1L3T)

Полімеразна реакція (рис. 9.6) передбачає, що ДНК-полімераза має розташувати 3'-кінцеву ОН-групу зростаючого ланцюга в активному центрі, зв'язати NTP, за умови його комплементарності до нуклеотиду матричного ланцюга здійснити каталіз синтезу фосфодієфірного зв'язку з одночасним відщепленням пірофосфату.

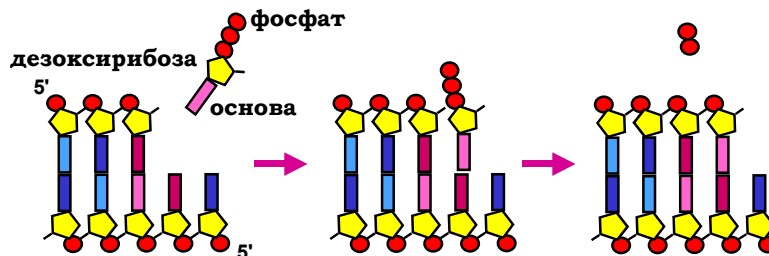


Рис. 9.6. Схема ДНК-полімеразної реакції

На першій стадії елонгаційного циклу відбувається зв'язування NTP (рис. 9.7). ДНК-полімераза при цьому реалізує свою відкриту конформацію, яка дозволяє швидку асоціацію / дисоціацію різних нуклеотидів. Якщо нуклеотид виявляється комплементарним матриці, з парою основ реалізуються численні взаємодії активного центру та його оточення. Це індукує конформаційну зміну в полімеразі (найповільніша лімітуюча стадія циклу) із переміщення пальців, наслідком чого є фіксація NTP у активному центрі (рис. 9.8) і нерухомість полімерази відносно ДНК. На третій стадії здійснюється каталіз реакції. Після приєднання нуклеотиду відбувається дисоціація пірофосфату й розмикання полімерази. Оскільки утворився новий продовжений 3'-кінець, який має спорідненість до активного центру, на останній стадії здійснюється транслокація – переміщення полімерази в її відкритій формі на один нуклеотид уздовж матриці.

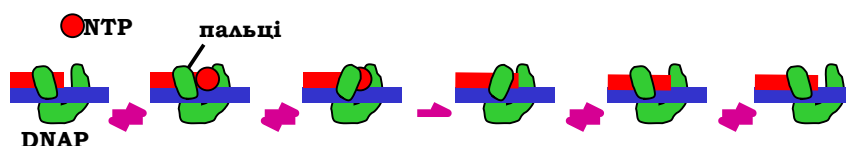


Рис. 9.7. Елонгаційний цикл ДНК-полімерази

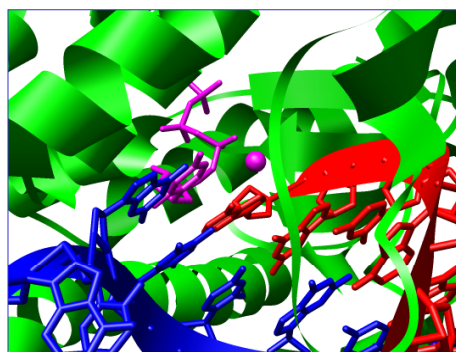


Рис. 9.8. Нуклеозидтрифосфат (СТР) та іон  $Mg^{2+}$  (фіолетовий) в активному центрі ДНК-полімерази (1LV5). Матричний ланцюг синій, зростаючий – червоний

Принципова схема елонгаційного циклу ДНК-полімерази практично не відрізняється від такої РНК-полімерази (див. розділ 5). Механізм каталізу ДНК-полімеразної реакції також є цілком аналогічним: ключову роль у каталізі виконують два іони  $Mg^{2+}$ , що утримуються в активному центрі трьома залишками Asp.

### Редагування помилок

Упізнання комплементарної пари основ активним центром ДНК-полімерази перед здійсненням каталізу полімеразної реакції забезпечує частоту помилкового включення нуклеотидів на рівні  $\sim 10^{-5}$  (основне джерело помилок – таутомерія азотистих основ, див. рис. 5.11). Оскільки ДНК синтезується “раз і назавжди” перед її передачею нащадкам, такий рівень помилок не може вважатися задовільним. Відповідно, ДНК-полімерази реалізують ефективний механізм редагування, використовуючи свій 3'-екзонуклеазний активний центр (рис. 9.9).

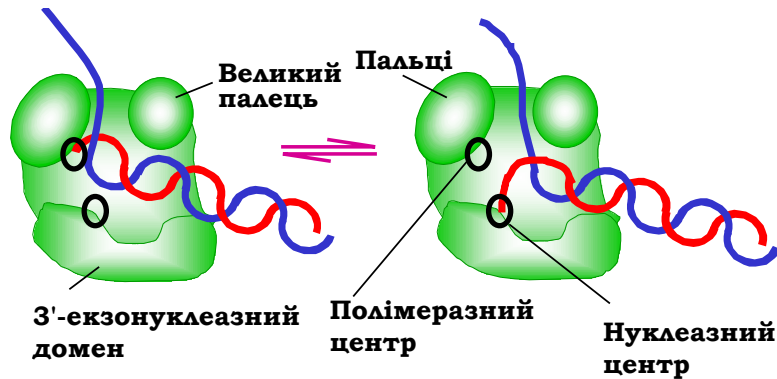


Рис. 9.9. Редагування помилок: схематичний вид “зверху” на ДНК-полімеразу (див. рис. 9.5)

Два каталітичні центри, розташовані на відстані до 30 Å, конкурують за 3'-кінець ланцюга, що синтезується. Якщо внаслідок приєднання помилкового нуклеотиду утворилася некомплементарна (нестабільна) пара основ, її спорідненість до полімеразного центру знижується, і 3'-кінець переміщується до екзонуклеазного центру. Подвійна спіраль ДНК при цьому прокручується навколо своєї осі, зберігаючи взаємодії маленького жолобка з великим пальцем, а ДНК-полімераза залишається у відкритій конформації. Помилковий нуклеотид відщеплюється, унаслідок чого відновлюється стабільність кінцевої пари основ. За рахунок більш високої спорідненості стабільної пари до полімеразного центру, 3'-кінець повертається туди, і ДНК-полімераза здійснює нову спробу його подовження.

У результаті такої осциляції полімерази з перемиканням активності між двома центрами рівень помилок знижується до  $\sim 10^{-8}$ . Остаточний рівень помилок знижується до  $\sim 10^{-10}$  за рахунок активності систем репарації.

### **Особливості ДНК-полімерази в порівнянні з РНК-полімеразою**

Незважаючи на високу подібність базових механізмів роботи двох типів полімераз, що здійснюють синтез нуклеїнових кислот, існують принципові відмінності між ними. Головна особливість полягає в тому, що для ДНК-полімерази ДНК є одночасно й матрицею, і продуктом реакції, і це створює суттєві проблеми.

Оскільки при синтезі РНК в активному центрі РНК-полімерази тимчасово існує гібридна подвійна спіраль ДНК-РНК (див. розділи 5, 6), РНК-полімераза може легко дискримінувати гібрид від звичайної подвійної спіралі ДНК. Висока спорідненість оточення активного центру РНК-полімерази до гібрида та каналу виходу транскрипту до РНК забезпечує високу процесивність ферменту – здатність працювати без дисоціації після однократного акту ініціації транскрипції. ДНК-полімераза має подвійну спіраль ДНК як у оточенні свого активного центру, так і скрізь поза полімеразним комплексом. Відповідно, існує висока ймовірність її дисоціації: процесивність ДНК-полімерази є дуже низькою – вона може синтезувати до дисоціації лише ділянку довжиною 10–20 нуклеотидів. Отже, має існувати певний додатковий механізм підвищення процесивності.

Висока спорідненість РНК-полімерази до гібрида ДНК-РНК дозволяє легко руйнувати подвійну спіраль ДНК по ходу руху полімерази при елонгації транскрипції – транскрипт просто витісняє нематричний ланцюг ДНК із дуплекса. Для ДНК-полімерази такий механізм є неможливим: дуплекси ДНК у комплексі з полімеразою та попереду від неї нічим не відрізняються один від одного, тобто ДНК-полімераза потребує наявності одноланцюгової матричної ДНК, яка має бути вилучена з подвійної спіралі.

Третя проблема полягає в тому, що ДНК-полімераза здатна робити тільки одну операцію – продовжувати (редагуючи) 3'-кінець ланцюга ДНК, вона не може ініціювати синтез, створити перший фосфодієфірний зв'язок. Це означає, що певна коротка ділянка має бути створена якимось інакше, щоб далі ДНК-полімераза могла продовжувати її синтез. Таку ділянку, без якої неможлива робота ДНК-полімерази, називають *праймером* (primer). Насправді, кор-фермент РНК-полімерази

також не здатен ініціювати синтез РНК: первинний короткий транскрипт синтезується лише за участі факторів ініціації, які допомагають додатково зафіксувати перші нуклеотиди в активному центрі (див. розділи 5, 6). При синтезі ДНК було знайдено інше еволюційне рішення: роль праймера виконує коротка ділянка РНК, що синтезується специфічною ДНК-залежною РНК-полімеразою – *праймазою*. Використання саме РНК у ролі праймера на початку синтезу ДНК має важливе значення для підвищення точності синтезу. Короткі первинні ділянки неможливо синтезувати з дуже високою точністю. На подальших етапах реплікації РНК-праймери можна легко відрізнити від ДНК, вилучити та заповнити прогалини між сусідніми ділянками ДНК за допомогою ДНК-полімерази.

Отже, ДНК-полімераза є хоч і головним, але не достатнім елементом системи реплікації. У реплікативній вилиці працює складний мультибілковий комплекс – *реплісома*, до якого крім двох молекул ДНК-полімерази входять компоненти, що забезпечують розплітання ДНК, підвищення процесивності та виконують інші важливі допоміжні операції.

## **Реплісома та інші елементи системи реплікації**

### ***Геліказа й білки SSB***

Операція розплітання подвійної спіралі (тобто утворення реплікативної вилки) здійснюється ДНК-геліказами – АТР-залежні ферменти цього типу вже згадувались у зв'язку з транскрипцією та сплайсингом (розділи 6, 7).

Основною реплікативною геліказою бактерій є білок DnaB – кільцевий комплекс шести однакових субодиниць (рис. 9.10). У каналі всередині кільця розміщується один полінуклеотидний ланцюг, уздовж якого відбувається транслокація гексамеру в напрямку 5'-3' (тобто по ланцюгу, що запізнюється – рис. 9.11), що й призводить до руйнування подвійної спіралі.

Шість субодиниць гелікази мають сайти зв'язування АТР, де відбувається його гідроліз. Структурні перебудови у відповідь на зв'язування та гідроліз АТР змінюють інтерфейс взаємодії з одноланцюговою ДНК, що й приводить до транслокації – на три-п'ять нуклеотидів за один акт гідролізу АТР.



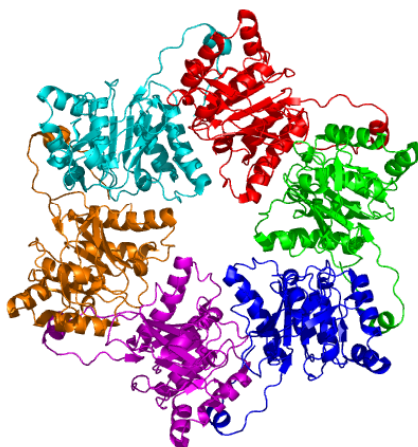
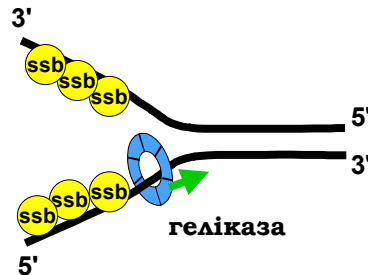


Рис. 9.10. Структура реплікативної гелікази (1G8Y)

Інші реплікативні гелікази побудовані за аналогічними принципами. У розплітанні ДНК під час реплікації беруть також участь гелікази протилежного напрямку (3'-5' – транслокація вздовж лідируючого ланцюга), але, на відміну від DnaB, вони не є життєво необхідними. Геліказа DnaB – це невід'ємний елемент реплісоми, через інші структурні модулі вона зв'язана з ДНК-полімеразами та працює в тісній координації з ними. Швидкість розплітання ДНК окремою геліказою дорівнює ~35 нуклеотидів за секунду, швидкість руху окремої ДНК-полімерази вздовж одноланцюгової ДНК – ~1 тис. нуклеотидів за секунду, при їхньому об'єднанні сумісна швидкість реплікації становить ~750 нуклеотидів за секунду: геліказа трохи гальмує полімеразу, остання – штовхає геліказу.

Енергія гідролізу АТФ використовується геліказою для здійснення енергетично не вигідного процесу розплітання дуплекса. Тобто одноланцюгова ДНК, що виникає в реплікативній вилиці внаслідок активності гелікази, має бути тимчасово стабілізована в одноланцюговому вигляді. Цю функцію виконують білки SSB (Single Strand Binding), які мають високу спорідненість до одноланцюгової ДНК (рис. 9.11). SSB білки зв'язуються з полінуклеотидним ланцюгом кооперативно, взаємодіючи не тільки з ДНК, а й між собою. У результаті ланцюг ефективно вкривається білковою “шубою”. SSB білки взаємодіють також з іншими компонентами реплісоми, підсилюючи їхню активність – у тому числі активність ДНК-полімерази.



**Рис. 9.11.** Реплікативна вилка, що виникає як результат дії гелікази та підтримується SSB білками

### Синтез ланцюга, що запізнюється

Лідуючий ланцюг при ініціації реплікації, а кожен фрагмент Оказаки під час реплікації, мають розпочинатися з РНК-праймера. Його синтез здійснює **праймаза** (білок DnaG *E. coli*). Праймаза рекрутується до основи реплікативної вилки геліказою (їхній комплекс називають *праймосомою*), робота праймази стимулюється геліказою та SSB білками. Розпочинаючи від основи реплікативної вилки, праймаза, яка має низьку процесивність, синтезує праймер – фрагмент РНК довжиною ~12 нуклеотидів (рис. 9.12). Після цього ДНК-полімераза III, подовжуючи 3'-кінець праймера, починає синтезувати решту фрагмента Оказаки, а праймаза (та сама чи інша) знову зв'язується з геліказою і згодом розпочинає синтез наступного праймера.

На відміну від ДНК-полімерази, праймаза може з однаковим успіхом включати до праймера як дезоксирибо- так і рибонуклеотиди. Праймер є фрагментом саме РНК просто тому, що внутрішньоклітинна концентрація рибонуклеозидтрифосфатів є вищою.

Зрозуміло, що РНК-праймер на початку кожного фрагмента Оказаки має бути замінений на відповідну послідовність ДНК. Цю роботу виконує **ДНК-полімераза I** (рис. 9.12): за рахунок своєї 5'-екзонуклеазної активності полімераза видаляє праймер і одночасно, використовуючи 3'-кінець попереднього фрагмента Оказаки, подовжує його, заповнюючи прогалину.

Після здійснення останньої операції між двома фрагментами Оказаки залишається одноланцюговий розрив (рис. 9.12, 9.13) – ДНК-полімераза не може його зшити, оскільки здатна тільки переносити нуклеотиди з NTP на 3'-кінцеву ОН-групу. Утворення фосфодієфірного зв'язку між сусідніми фрагментами Оказаки, яке потребує джерела енергії, здійснюється в АТР-залежній реакції, яка каталізується **лігазою** (ligase).

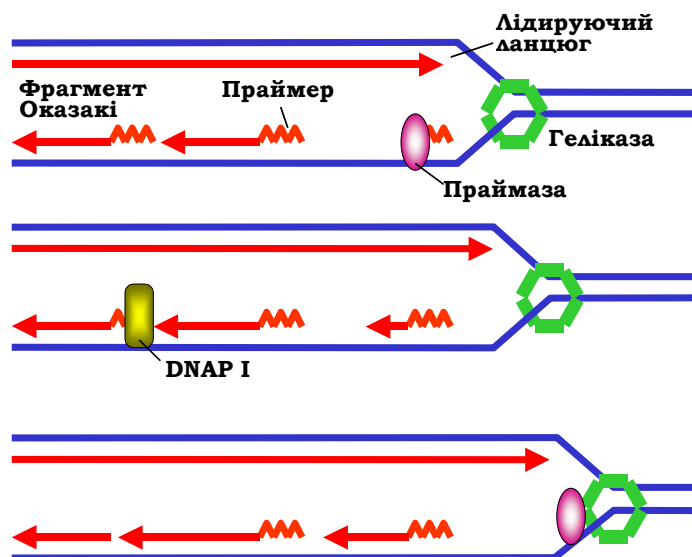


Рис. 9.12. Робота праймази та ДНК-полімерази I під час руху реплікативної вилки

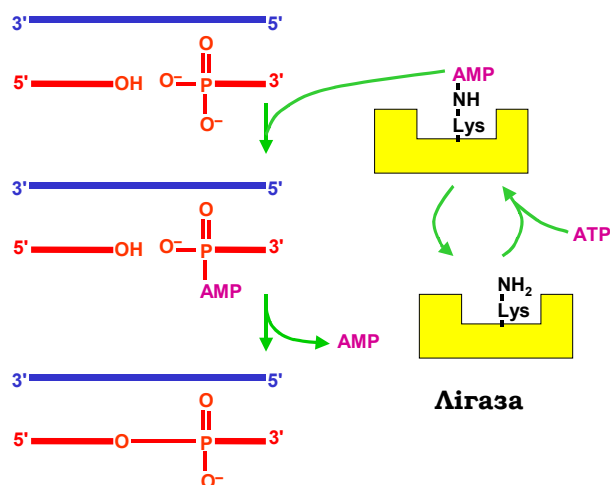


Рис. 9.13. Схема роботи лігази

Схему роботи лігази показано на рис. 9.13. До активного центру ферменту, ковалентно приєднуючись до аміногрупи Lys, переноситься AMP з молекули АТР. Далі AMP приєднується до 5'-кінцевого фос-

фату в місці розриву (своєрідна активація, що нагадує активацію амінокислоти АРСазою, див. розділ 8). Видалення АМР супроводжується утворенням фосфодієфірного зв'язку.

### *Голофермент ДНК-полімерази*

Голофермент ДНК-полімерази, який є основою реплісоми, містить дві молекули ДНК-полімерази III (по три субодиниці в кожній – кор-фермент), дві копії  $\tau$ -білка,  $\gamma$ -комплекс (п'ять субодиниць) та два  $\beta$ -білка, зв'язані з кожною полімеразою. Призначення компонентів голоферменту – утримання разом і забезпечення координації роботи компонентів реплісоми, підвищення процесивності ДНК-полімерази.

Проблема процесивності вирішується  **$\beta$ -білком**, який утворює так званий обруч (зажим), що ковзає (sliding clamp). Дві однакові субодиниці обруча, міцно взаємодіючи за принципом “голова до хвоста”, утворюють циркулярну тороподібну структуру з отвором приблизно 35 Å у діаметрі (рис. 9.14). Усередині такого кільця легко розміщується подвійна спіраль ДНК, яка практично не взаємодіє з  $\beta$ -білком. Кільце може вільно рухатись уздовж ДНК, але не може дисоціювати. С-кінцевий гідрофобний хвіст ДНК-полімерази III довжиною 11 амінокислот взаємодіє з гідрофобним карманом обруча. У результаті формується своєрідний поводек, що допускає певну свободу рухів полімерази, але міцно утримує її на ДНК.



**Рис. 9.14.** Обруч, що ковзає, – бактеріальний  $\beta$ -білок (2POL)

Оскільки обруч – білкове кільце, що оточує ДНК, його завантаження на подвійну спіраль виконується АТР-залежним шляхом, ця функція належить  **$\gamma$ -комплексу**. Відбувається завантаження на гібридній

подвійній спіралі РНК-ДНК відразу після закінчення синтезу чергового праймера.  $\gamma$ -Комплекс, що складається з п'яти субодиниць та має U-подібну структуру (рис. 9.15), зв'язує АТР і змінює свою конформацію таким чином, що взаємодіє з обручем, порушуючи його цілісність – розмикаючи кільце. Далі  $\gamma$ -комплекс впізнає дуплекс, що утворений між праймером та матрицею, забезпечуючи посадку відкритого кільця на цю подвійну спіраль. Взаємодія з дуплексом індукує АТРазну активність  $\gamma$ -комплексу. У результаті гідролізу АТР  $\gamma$ -комплекс втрачає спорідненість до  $\beta$ -білка, що призводить до замикання обруча. Далі  $\gamma$ -комплекс знову зв'язує АТР, проте обруч має вищу спорідненість до ДНК-полімерази, ніж до  $\gamma$ -комплексу: обруч залишається на ДНК і взаємодіє з полімеразою та забезпечує її процесивність.

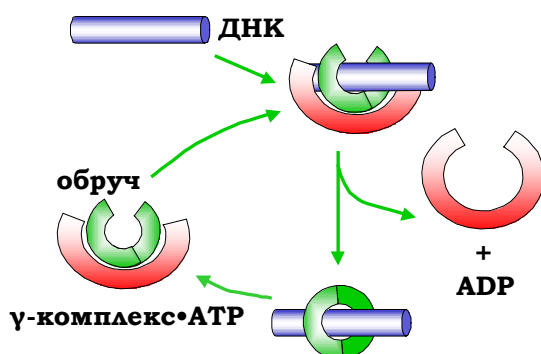
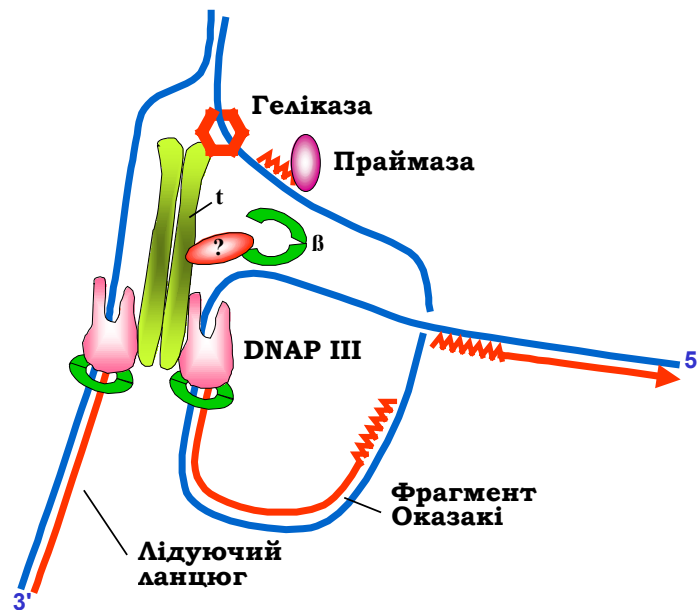


Рис. 9.15. Схема АТР-залежного завантаження обруча, що ковзає,  $\gamma$ -комплексом

Головну структурну роль в організації реплісоми виконує  $\tau$ -білок, котрий взаємодіє з ДНК-полімеразою, геліказою, праймазою та утримує на собі  $\gamma$ -комплекс. Численні взаємодії  $\tau$ -білка забезпечують функціональну координацію між елементами реплісоми.

Дві молекули  $\tau$ -білка взаємодіють між собою і кожна утримує корфермент ДНК-полімерази III, кожна з яких, у свою чергу, зв'язана з обручем (рис. 9.16). Одна з полімераз здійснює синтез лідируючого ланцюга, інша – фрагмента Оказакі. Обидві полімерази орієнтовані в реплісомі однаково. Оскільки вони при цьому ведуть синтез на антипаралельних ланцюгах, той, що запізнюється, утворює реверсну петлю: конфігурація реплікативної вилки нагадує тромбон.

На рис. 9.16 зображено момент, коли праймаза почала синтез чергового праймера, а  $\gamma$ -комплекс – наготові завантажити черговий обрuch. Коли полімераза, яка працює на ланцюзі, що запізнюється, завершить синтез фрагмента Оказакі (зустріне попередній фрагмент), вона втратить свій зв'язок з обрuchом, але збереже взаємодію з  $\tau$ -білком. Втрата спорідненості полімерази до обрucha визначається одноланцюговим розривом між фрагментами Оказакі. Новий обрuch буде при цьому вже завантажений на черговий праймер, полімераза зв'яжеться з ним і розпочне синтез нового фрагмента Оказакі. Обрuch, який щойно втратив зв'язок із полімеразою, провзаємодіє з  $\gamma$ -комплексом – останній працює також і як “розвантажувач” обрucha (див. рис. 9.15), знімаючи його з ДНК і забезпечуючи пізніше його посадку на праймер.



**Рис. 9.16.** Схема організації реплісоми

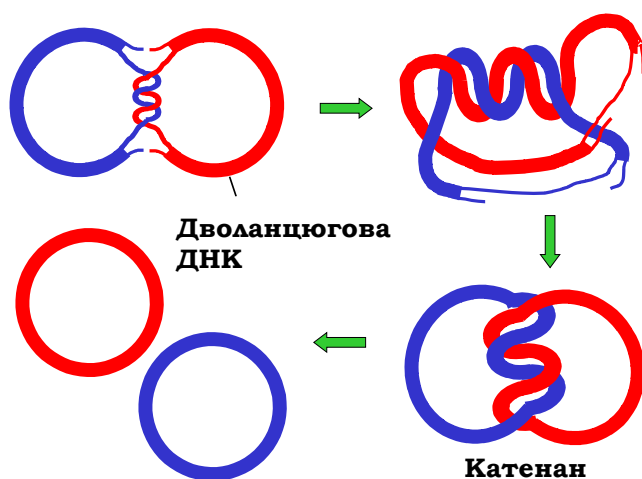
### *Топологічні проблеми, пов'язані з реплікацією*

Ще одним невід'ємним елементом системи реплікації є ДНК-топоізомерази (див. розділ 3). Рух реплікативної вилки з одночасним руйнуванням подвійної спіралі супроводжується прокруткою спіралі, що створює позитивні надспіральні витки – еластичні напруження попереду від вилки (див. рис. 3.24). Це стосується як еукаріотичних систем, де петлі ДНК зчеплені своїми кінцями з ядерним матриксом, так і прокариотич-

них циркулярних хромосом: зустрічний рух двох реплікативних вилок (рис. 9.3) створює позитивні напруження на ділянці, яка ще не є реплікованою. Еластична напруга має час від часу зніматися, інакше врешті-решт вона повністю зупинить рух реплісоми. Відповідно, процес реплікації потребує допоміжної дії топоізомераз, основна функція яких як раз і полягає в релаксації надспіралізованої ДНК: штучне вимикання активності топоізомераз повністю блокує реплікацію.

В еукаріотичній клітині завдання релаксації можуть виконувати як топоізомерази I, так і топоізомерази II. Прокаріотична топоізомераза I (топоізомераза Ia) не підходить для виконання цієї ролі, оскільки здатна знімати тільки негативну надспіралізацію. Релаксація позитивної надспіралізації при реплікації у бактерій забезпечується гіразою – ферментом, що відноситься до класу топоізомераз II. Гіраза вносить негативну надспіралізацію в циркулярну ДНК, що й компенсує позитивні напруження.

Ще одна топологічна проблема виникає при закінченні процесу реплікації в циркулярному репліконі. Коли дві реплікативні вилки зустрічаються, вони зупиняються за кілька витків подвійної спіралі одна від одної: еластичну напругу на маленькій ділянці зняти вже неможливо (рис. 9.17).



**Рис. 9.17.** Термінація реплікації у циркулярному репліконі

Відбувається руйнування короткої подвійної спіралі, яка не піддалася реплікації, але це неможливо зробити без відповідного перекручення двох дочірніх кільцевих молекул ДНК десь у іншому місці. Одноланцю-

гові прогалини далі заповнюються ДНК-полімеразами. Але після цього дві остаточно синтезовані циркулярні молекули так і залишаються топологічно зчепленими одна з одною – утворюється так званий *катенан* (рис. 9.17). Таке саме перекручення двох дочірніх молекул ДНК реалізується при закінченні реплікації на двох сусідніх еукаріотичних репліконах (див. рис. 9.2).

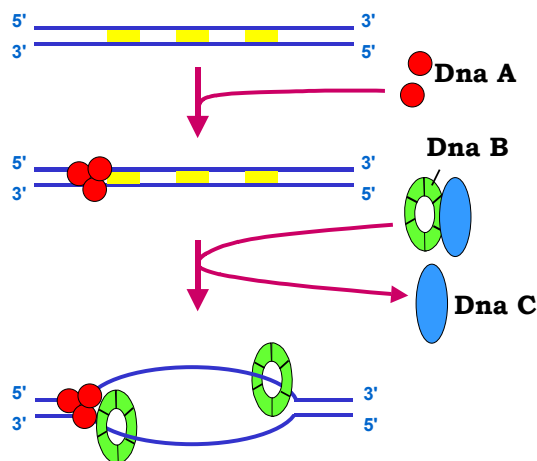
Зрозуміло, що декатенація – розділення двох молекул – є неможливим без дволанцюгового розриву в одній із них. За цю операцію також відповідають топоізомерази класу II (еукаріотична топоізомераза II, яка зв'язана з ядерним матриксом під час реплікації, і гомологічна до неї топоізомераза IV у прокариотів), що роблять тимчасовий дволанцюговий розріз.

### **Ініціація реплікації у бактерій**

Сайт ініціації реплікації *E. coli* – ориджин *oriC* – є ділянкою довжиною 245 пар основ, яка містить чотири сайти по дев'ять пар основ кожний, що мають спорідненість до білка DnaA, а також кілька АТ-збагачених ділянок по 13 пар основ зі зниженою стабільністю подвійної спіралі. Передумовою ініціації реплікації є створення за рахунок активності гірази певного рівня негативної надспіралізації в циркулярній хромосомі. По-перше, негативна надспіралізація сприяє локальному руйнуванню подвійної спіралі в межах ориджина. По-друге, оскільки введення надспіралізації є можливим тільки за умови цілісності полінуклеотидних ланцюгів (хоча б один одноланцюговий розрив знімає будь-яку еластичну напругу), це є ефективною перевіркою готовності матриці (по всій довжині) до реплікації.

На першій стадії ініціації кілька копій DnaA (у комплексі з АТР) зв'язуються з ориджином (рис. 9.18) та індують АТР-залежне локальне плавлення подвійної спіралі. Білок DnaC завантажує на одноланцюгові ділянки геліказу DnaB, яка, у свою чергу, рекрутує праймазу та інші елементи реплісоми. Праймаза здійснює синтез праймерів на лідируючих ланцюгах і замінюється на ДНК-полімеразу III: починається рух двох реплікативних вилок у протилежних напрямках. Напрямок руху задається напрямком транслокації гелікази по ланцюгу, що запізнюється. Присутність гелікази саме на цьому ланцюзі забезпечує рекрутування до нього праймази під час елонгації реплікації на початку синтезу кожного фрагмента Оказакі.





**Рис. 9.18.** Ініціація реплікації в *E. coli*. Жовтим позначено 13-мерні ділянки ДНК зі зниженою стабільністю

На діаметрально протилежному щодо *oriC* боці циркулярної хромосоми розташовано 6 ділянок по 22 пари основ – сайтів термінації реплікації (Ter A, B, C, D, E, F), які утворюють комплекс із білком Tus (Terminus utilization substance). На сайтах термінації, зв'язаних із цим білком, реплікативні вилки зупиняються. Зона між ними піддається далі локальному плавленню і після цього відбувається добудова прогалин і декатенація (див. рис. 9.17).

### Особливості еукаріотичної системи реплікації

Загальні принципи й механізми реплікації є спільними для про- та еукаріотів, розрізняючись у деталях:

- Еукаріотична реплікативна вилка містить дві ДНК-полімерази, процесивність яких забезпечується ковзним циркулярним обручем. Різниця лише в тому, що обруч (PCNA – Proliferating Cell Nuclear Antigene) складається з трьох субодиниць. Крім забезпечення процесивності ДНК-полімерази, PCNA взаємодіє ще з багатьма білками, залученими до виконання різноманітних функцій при реплікації.
- Одна з полімераз веде синтез лідируючого ланцюга, інша синтезує фрагменти Оказаки довжиною 200–400 пар основ. РНК-праймер на початку кожного фрагмента Оказаки видаляється спе-

цифічною нуклеазою. Прогалина потім заповнюється ДНК-полімеразами, одноланцюговий розрив зшивається лігазою.

- Основним структурним модулем реплісоми є мультибілковий комплекс RF-C (Replication Factor C), що виконує функції, аналогічні до бактеріального  $\tau$ -білка та  $\gamma$ -комплексу.
- Вважається, що основною реплікативною геліказою еукаріотів є моногексамерний білок МСМ (MiniChromosome Maintenance). Одноланцюговий стан ДНК у реплікативній вилиці підтримується білками RPA (Replication Protein A – аналог білків SSB).

### *Еукаріотичні ДНК-полімерази*

П'ять основних типів еукаріотичних ДНК-полімераз прийнято позначати грецькими літерами.

**Полімераза  $\alpha$**  складається з чотирьох субодиниць: одна виконує структурну роль, друга має *ДНК-полімеразну* активність, ще дві субодиниці відповідають за активність *РНК-полімеразну*. Відповідно, полімераза  $\alpha$  виконує роль *праймази*: саме вона синтезує короткий РНК-праймер (6–10 нуклеотидів), який далі трохи подовжує як ДНК. Полімераза  $\alpha$  характеризується низькою процесивністю і, на відміну від інших ДНК-полімераз, не має екзонуклеазної активності.

**Полімераза  $\beta$**  – мономерний білок, основна роль якого полягає в заповненні однонуклеотидних прогалин при ексцизійній репарації основ (див. нижче).

**Полімераза  $\gamma$**  – реплікативна ДНК-полімераза мітохондрій. Складається з двох субодиниць, крім ДНК-полімеразної має обидві (3'- та 5'-) екзонуклеазні активності.

**Полімерази  $\delta$  та  $\epsilon$**  – основні реплікативні ДНК-полімерази еукаріотів. Складаються з чотирьох ( $\epsilon$ ) або трьох–п'яти в різних видів ( $\delta$ ) субодиниць. Обидві полімерази можуть працювати в реплікативній вилиці (у комплексі з PCNA та RF-C) на обох ланцюгах (лідуючому та тому, який запізнюється). Функціональну різницю між двома полімеразами остаточно не з'ясовано. Вважається, що полімераза  $\delta$  є суто ДНК-синтезуючою машиною, тоді як полімераза  $\epsilon$  використовується також як “сенсор” пошкоджень ДНК, сприяючи активації репараційних процесів. Крім реплікативного синтезу ДНК, полімерази  $\delta$  і  $\epsilon$  здійснюють заповнення довгих багатонуклеотидних прогалин як при з'єднанні сусідніх фрагментів Оказакі, так і при репараційних процесах, які розглядатимуться далі.

Крім названих ДНК-полімераз, які відносять до групи полімераз високої точності синтезу, в еукаріотичних клітинах працює ще досить велика кількість ДНК-полімераз низької точності ( $\zeta$ ,  $\eta$ ,  $\iota$ ,  $\kappa$ ). Їхня основна функція полягає в тому, щоб забезпечити синтез ДНК у разі пошкодження матриці: зустрічаючи пошкодження, реплікативна машинерія високої точності (яка включає полімерази  $\delta$  чи  $\epsilon$ ) зупиняється. Полімерази низької точності дозволяють долати пошкодження, які пізніше можуть бути виправлені шляхом репарації, і знов замінюються на полімерази  $\delta/\epsilon$ , що продовжують високоточний синтез ДНК.

Крім того, описано ще кілька ДНК-полімераз в еукаріотичних клітинах: полімераза  $\theta$  імовірно відіграє роль у репарації міжланцюгових зшивок ДНК,  $\lambda$  є гомологічною до полімерази  $\beta$  і, ймовірно, залучена до репарації під час мейозу,  $\mu$  можливо бере участь у рекомбінації імуноглобулінових генів (розділ 10).

### *Ініціація реплікації в еукаріотів*

Суттєвою відмінністю еукаріотичної системи реплікації є те, що кожна хромосома є полірепліконом: загалом геном, наприклад ссавців, містить близько 40 тис. точок ініціації – ориджинів. Ориджини важко ідентифікувати, оскільки спільних елементів послідовності у них немає, вони лише характеризуються підвищеним вмістом АТ-пар. Розмір еукаріотичного реплікона варіює від 50 до 200 тис. пар основ, що збігається з розмірами петлевих доменів хроматину (див. розділ 4). Отже, хроматинова петля – це один реплікон, а ориджин збігається з ділянкою, асоційованою з ядерним матриксом.

Ініціація реплікації залежить від мультибілкового комплексу ORC (Origin Recognition Complex), який практично постійно є присутнім на ориджині (рис. 9.19). При підготовці до реплікації у фазі G1 клітинного циклу відбувається фосфорилування циклінозалежними кіназами факторів регуляції клітинного циклу *cdc6* та *cdc45* (*cdc* – cell division cycle). У результаті ці фактори разом із двома моногексамерами MCM (яким відводиться роль геліказ) взаємодіють з ORC та індуюють локальне плавлення подвійної спіралі. MCM разом із *cdc45* залишаються і далі в основах двох реплікативних вилок. З розплетеними полінукліотидними ланцюгами взаємодіє RPA, *cdc45* рекрутує в кожную реплікативну вилку ДНК-полімеразу  $\alpha$ . Остання починає синтез праймера, з основою вилки зв'язується RF-C, який завантажує обруч PCNA, полімераза  $\alpha$  замінюється на полімерази  $\delta/\epsilon$  (так зване перемикання полімераз). Надалі MCM продовжує розділення ланцюгів ДНК, а полімераза  $\alpha$  здійснює синтез праймерів на ланцюзі, що запізнюється, і кожного разу замінюється на другу копію полімерази  $\delta/\epsilon$ .

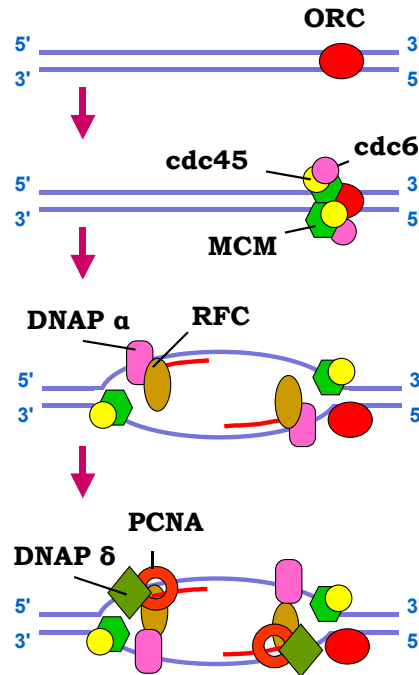


Рис. 9.19. Ініціація реплікації в еукаріотів

Реплікація ініціюється синхронно на сусідніх 25–100 репліконах, які складають так звану реплікативну фабрику. При цьому різні ділянки хроматину реплікуються не одночасно: в останню чергу відбувається реплікація гетерохроматинових зон.

### *Структурні зміни хроматину під час реплікації*

Зрозуміло, що нуклеосоми (див. розділ 4) є суттєвим бар'єром на шляху реплісоми. Попереду реплікативної вилки відбувається декомпактизація хроматинової фібрили й тимчасове видалення відносно легко взаємодіючого з ДНК гістону H1. Нуклеосоми руйнуються у два етапи (рис. 9.20): спочатку видаляється димер гістонів H2A-H2B, потім найміцніше зв'язаний із ДНК тетрамер  $(H3-H4)_2$ . Процес видалення гістонових комплексів забезпечується активністю комплексів ремоделювання хроматину (розділ 6) і присутністю проміжних акцепторів гістонів – гістонових шаперонів. Такі шаперони зв'язують гісто-

нові комплекси, що витісняються реплісомою, а потім виконують роль факторів збирання нуклеосом позаду реплікативної вилки. Необхідність таких факторів визначається надзвичайно високою спорідненістю гістонів до ДНК (див. рис. 4.11).

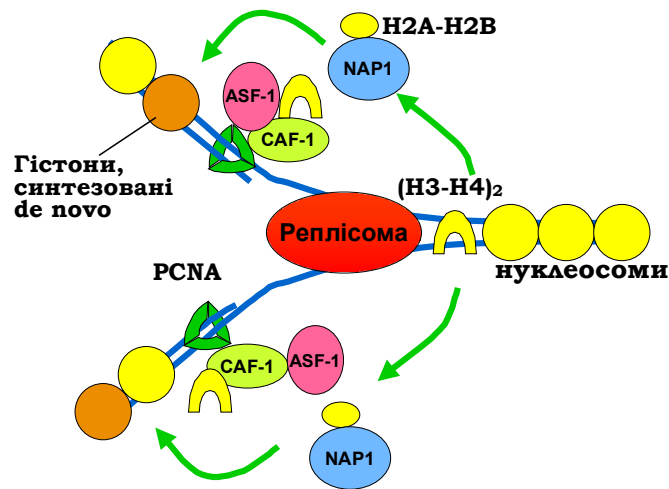


Рис. 9.20. Руйнування та відновлення нуклеосом у реплікативній вилці

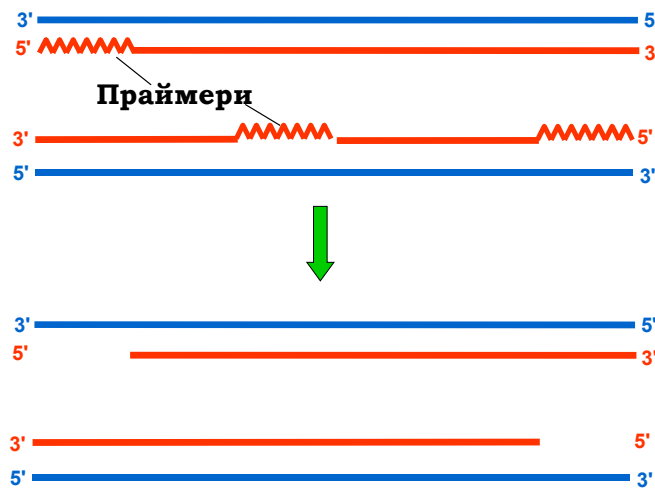
До найбільш вивчених факторів збирання нуклеосом, що працюють під час реплікації, відносять, зокрема: NAP1 (Nucleosome Assembly Protein), що має підвищену спорідненість до димерів H2A-H2B; CAF1 (Chromatin Assembly Factor) та ASF1 (Anti-Silencing Function) – споріднені до гістонів H3-H4 і PCNA, за рахунок чого тетрамери (H3-H4)<sub>2</sub> спрямовуються до основи реплікативної вилки (рис. 9.20); RSF (Remodeling and Spacing Factor), який є проміжним акцептором гістонів, а також сприяє регулярному розподілу відновлених нуклеосом на ДНК.

Відновлення нуклеосом позаду реплікативної вилки відбувається також у дві стадії: першим на ДНК повертається тетрамер (H3-H4)<sub>2</sub>, який зв'яже два димери H2A-H2B. По двох дочірніх ланцюгах ДНК гістонові комплекси розподіляються випадково, до них додаються гістони, синтезовані в цитоплазмі *de novo* під час реплікації. Таким чином, “старі” гістони, що несуть на собі певні характерні модифікації (див. розділи 4 і 6), повертаються на ту саму ділянку ДНК обох ланцюгів, на якій вони були присутні на материнській молекулі. До ділянок хроматину рекрутуються відповідні ферменти, котрі здійснюють ана-

логічні модифікації щойно синтезованих гістонів – патерн модифікацій відновлюється, що сприяє збереженню певного функціонального стану ділянки хроматину в дочірніх клітинах.

### *Подовження кінців еукаріотичної хромосоми*

Ще одна характерна відмінність еукаріотичної хромосоми полягає в тому, що вона, на відміну від прокаріотичної, є лінійною – має два кінці. Унаслідок цієї простої обставини на 3'-кінцях матричних ланцюгів ДНК залишаються одноланцюгові хвости (рис. 9.21): два РНК-праймери на 5'-кінцях синтезованих ланцюгів видаляються, а прогалина не може бути заповненою, оскільки немає 3'-кінця, який міг би бути використаним як праймер. Одноланцюгові хвости піддаються швидкій нуклеазній деградації і після кожної реплікації хромосома має вкоротитися.



**Рис. 9.21.** Дві дочірні лінійні молекули ДНК після реплікації

Кінцеві ділянки ДНК еукаріотичної хромосоми – теломери – складаються з невеликих елементів послідовності, що тандемно повторюються – теломерних повторів. Подовження теломер після реплікації здійснюється за допомогою спеціального ферменту – *теломерази*, яка є РНК-залежною ДНК-полімеразою. РНК-матриця входить до складу самого ферменту і містить ділянку, комплементарну теломерному повтору (рис. 9.22). Використовуючи цю ділянку як матрицю і 3'-кінець

як праймер, теломераза покроково добудовує до 3'-кінця кілька копій теломеразного повтору. Далі подовжений одноланцюговий хвіст використовується як матриця для синтезу іншого ланцюга за звичайним реплікативним механізмом. Видалення РНК-праймера після цього не є проблемою, оскільки хромосома вже є подовженою.

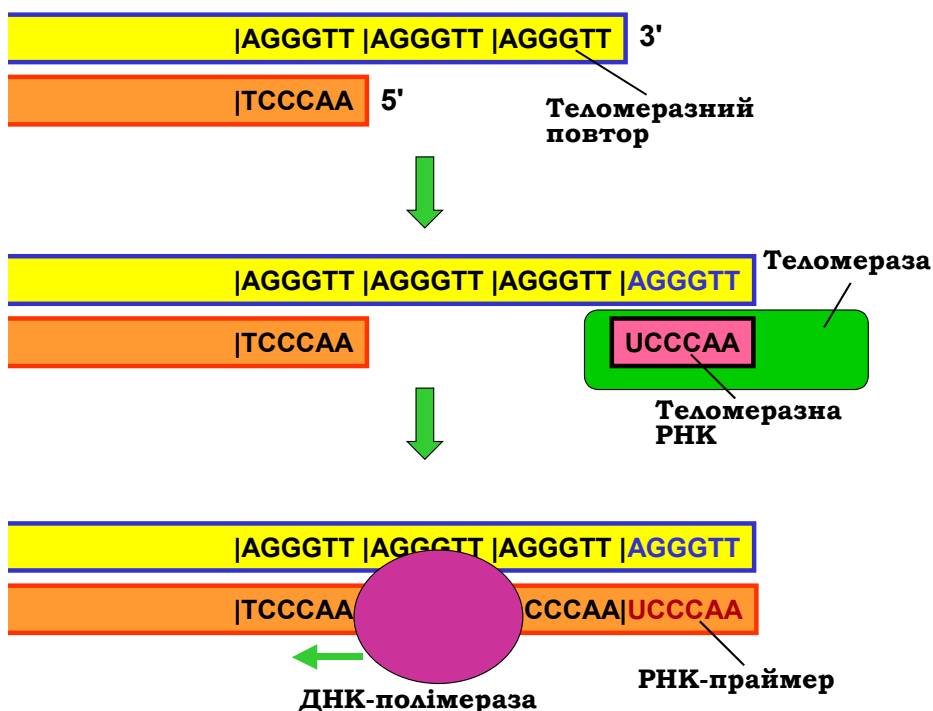


Рис. 9.22. Подовження кінців хромосоми за допомогою теломери

Теломераза є активною в клітинах, що розвиваються, і злоякісно трансформованих клітинах і неактивною – у диференційованих соматичних клітинах вищих еукаріотів. Відповідно, певне критичне скорочення теломерів, яке відбувається в таких клітинах після кількох десятків клітинних поділів, є одним із механізмів активації програми їхньої загибелі.

## Репарація ДНК

Репарація (repair) ДНК – один із загальних біологічних процесів, який спрямований на виправлення помилок синтезу ДНК при реплікації, а також численних пошкоджень, котрі виникають у ДНК унаслідок дії хімічних і фізичних факторів. До таких пошкоджень відносять різноманітні хімічні модифікації азотистих основ, ковалентні зшивки сусідніх піримідинів (утворення піримідинових, найчастіше тимінових, димерів) під дією ультрафіолетового випромінювання, одно- і дволанцюгові розриви під дією іонізуючої радіації та вільних радикалів тощо. Часто системи репарації працюють під час або відразу після реплікації. Більшість репараційних процесів передбачає видалення пошкодженої одноланцюгової ділянки з наступним синтезом ДНК за допомогою ДНК-полімераз. Але існують і процеси, пов'язані з безпосереднім “виправленням” пошкодженого елемента за рахунок прямої дії певних ферментів.

### *Пряма репарація*

Найочевиднішим випадком прямої репарації є зшивання одноланцюгового розриву ДНК лігазою (див. рис. 9.13). Іншим спільним для всіх живих організмів (за винятком ссавців) шляхом прямої репарації є *фотореактивація* – руйнування піримідинових димерів (рис. 9.23), які були індуковані ультрафіолетовим світлом, ферментом *фотоліаза*. Фотоліаза (або її власні амінокислотні залишки, або зв'язані з білком простетичні групи) здатна поглинати світло, що призводить до активації ферменту. Тобто світло, яке викликає утворення піримідинових димерів, одночасно активує фотоліазу, котра каталізує розрив ковалентних зв'язків між сусідніми піримідинами (рис. 9.23), а отже, відновлення структури ДНК.

Одним із загальних пошкоджуючих впливів на ДНК є *алкілювання* азотистих основ – ковалентне приєднання метильних чи етильних груп до атомів О або N. Пряма репарація таких пошкоджень є можливою за рахунок активності специфічних метилтрансфераз, що відщеплюють метильні групи (таким шляхом репаруються O<sup>6</sup>-метилгуанін та O<sup>4</sup>-метилтимін). Ці метилтрансферази не є ферментами: вони відщеплюють метильну групу й необернено ковалентно приєднують її до залишку Суs у своєму активному центрі – для нового акту деметилування необхідна нова молекула білка.



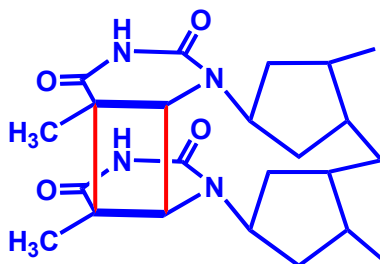


Рис. 9.23. Тиміновий димер

### Ексцизійна репарація

Більш радикальним і ефективним шляхом виправлення порушень нуклеотидів є ексцизійна репарація (excision repair), коли пошкоджена одноланцюгова ділянка вирізається з ДНК, а інший ланцюг використовується далі як матриця для нового синтезу. Існує два варіанти такої репарації.

При **ексцизійній репарації азотистих основ** (Base Excision Repair – BER), що відбувається в усіх організмів, модифікована азотиста основа видаляється ферментом *глікозилазою*. Існує певна кількість специфічних глікозилаз, що розпізнають різноманітні модифіковані основи. Глікозилаза руйнує глікозидний зв'язок між основою та C1'-атомом дезоксирибози (рис. 9.24). У ДНК залишається так званий АП-сайт (апуриний / апіримідиновий), який упізнається ендонуклеазою, що гідролізує фосфодіестерний зв'язок між 5'-фосфатом звільненої від основи дезоксирибози та попереднім нуклеотидом. Нарешті фосфодіестераза відщеплює цю фосфодезоксирибозу, і в ДНК залишається прогалина довжиною в один нуклеотид.

Ця прогалина заповнюється ДНК-полімеразою  $\beta$  (в еукаріотів), яка приєднує нуклеотид до 3'-ОН групи попереднього нуклеотиду ланцюга. Фосфодіестерний зв'язок приєданого нуклеотиду з наступним нуклеотидом ланцюга відновлюється лігазою. У прокаріотів заповнення прогалини здійснюється ДНК-полімеразою I. При цьому, за рахунок своєї 5'-екзонуклеазної активності, полімераза може руйнувати певну ділянку з 5'-кінця прогалини, одночасно продовжуючи 3'-кінець.

**Ексцизійна репарація нуклеотидів** (Nucleotide Excision Repair – NER) – процес, пов'язаний із вирізанням ділянки ДНК, яка містить пошкодження (модифіковану основу, тиміновий димер тощо).

У клітинах *E. coli* за цей шлях відповідає система *uvrABC* (*uvr* – *ultra violet repair*). Комплекс білка *uvrB* і двох білків *uvrA* впізнає пошкодження та зв'язується з ДНК у цьому місці (рис. 9.25). На наступному кроці відбувається АТР-залежна зміна конформації *uvrB*, вигин ДНК і дисоціація *uvrA*. До комплексу рекрутується білок *uvrC*. Обидва білки у складі комплексу набувають ендонуклеазної активності: *uvrC* робить одноланцюговий розріз у пошкодженому ланцюзі за кілька нуклеотидів у напрямку до 5'-кінця від пошкодження (ліворуч на рис. 9.25); *uvrB* – розріз з іншого боку від пошкодження. Довжина ділянки між розрізами дорівнює 12 (або 13 у випадку для тимінового димеру) нуклеотидам. Далі геліказа *uvrD* руйнує подвійну спіраль між двома розрізами, тобто видаляє пошкоджену ділянку. Прогалина, що залишилася, заповнюється ДНК-полімеразою I, лігаза остаточно відновлює цілісність ланцюга.

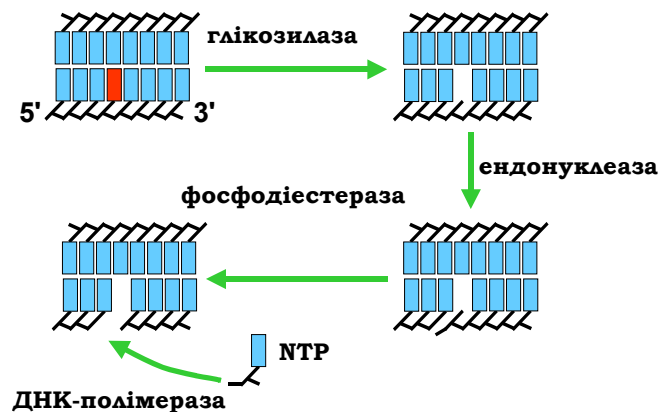


Рис. 9.24. Екцизійна репарація азотистих основ.  
Пошкоджена основа забарвлена червоним

Аналогічна система екцизійної репарації працює в еукаріотичних клітинах. До неї залучено близько 17 білків, причому за руйнування подвійної спіралі відповідає геліказна частина загального фактора транскрипції TFIIH (див. розділ 6). Це забезпечує тісну координацію репарації з транскрипцією: і підвищена ймовірність пошкоджень, і першочергова необхідність виправляти їх виникає насамперед у транскрипційно активних ділянках. Пошкодження розпізнаються або особливими білковими факторами, або РНК-полімеразою, яка робить зупинку на пошкодженому нуклеотиді. Після цього геліказа

руйнує ділянку подвійної спіралі довжиною 24–32 пари основ, пошкоджена ділянка вирізається ендонуклеазами, і прогалина заповнюється ДНК-полімеразами  $\delta/\epsilon$ .

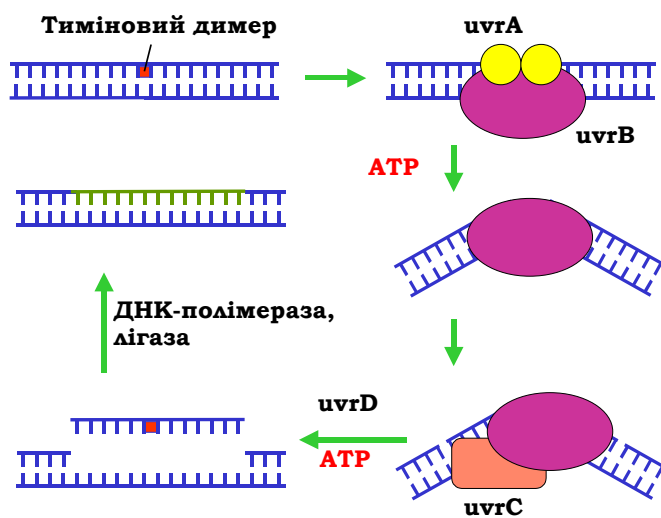


Рис. 9.25. Система *uvrABC* ексцизійної репарації нуклеотидів у *E. coli*

### Репарація некомплементарних пар основ – місметчів

Незважаючи на редагування помилок під час реплікації, певна кількість невірних спарених основ залишається в синтезованих ланцюгах ДНК. Зрозуміло, що при репарації таких місметчів (mismatch) із двох некомплементарних нуклеотидів замінити слід саме той, що входить до синтезованого, а не до матричного ланцюга.

У бактеріальній клітині за репарацію місметчів відповідає так звана система *mutHLSU*. По бактеріальному геному розподілені (на середній відстані 256 пар основ) короткі паліндромні послідовності *СТАG*, у складі яких аденін піддається постреплікативному метилюванню. Але певний час після реплікації метильованим є лише матричний (материнський) ланцюг. Саме за цей час і спрацьовує система репарації (рис. 9.26): білок, що позначається як *mutS*, упізнає місметч і рекрутує білок *mutL*, останній взаємодіє з двома білками *mutH*, які зв'язуються з тетра nukлеотидними паліндромними сайтами по обидва боки від місметча. У складі утвореного комплексу *mutH* набуває ендонуклеазної активності й робить одноланцюговий розріз у *немети-*

льованому ланцюзі в межах одного із сайтів (один із двох сайтів обирається випадково). Далі геліказа mutU (той самий білок, що й *uvrD*) розплітає подвійну спіраль, а екзонуклеаза руйнує ланцюг від розрізу до місметча й трохи далі. Нарешті прогалина заповнюється ДНК-полімеразою III, і одноланцюговий розрив зшивається лігазою.

Система mutHLSU є консервативною, гомологічні білки присутні також і в еукаріотів. Метильовання аденіну не використовується для дискримінації ланцюгів: білки еукаріотичної системи репарації місметчів пов'язані з реплісосою та ланцюгами ДНК, що синтезуються, тобто спрацьовують безпосередньо під час реплікації.

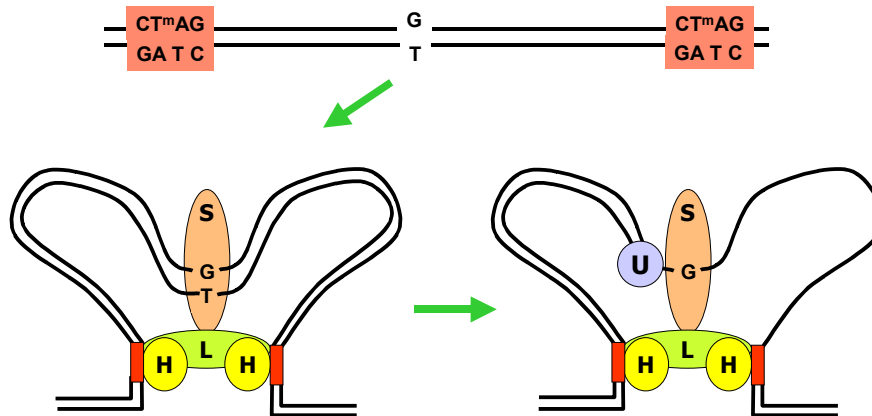


Рис. 9.26. Система репарації місметчів mutHLSU

### Репарація без репарації

Іноді в клітині активуються процеси, які прийнято також називати репарацією, хоча насправді вони є засобом здійснити реплікативний синтез ДНК, “не звертаючи уваги” на пошкодження її структури. Реплікативна машинерія зазвичай зупиняється, зустрічаючи пошкодження у складі матриці. Якщо таких пошкоджень надто багато, й істинні репараційні системи не встигають їх виправити, перемикання на неточний синтез ДНК дає клітині шанс на виживання. Пошкодження при цьому залишаються і, як наслідок, дають велику кількість мутацій. Усі процеси такого типу зазвичай об'єднують під назвою **SOS-репарації** або (що точніше) – механізмів синтезу ДНК, толерантних до пошкоджень (damage tolerance mechanisms).

У прокаріотів перемикавання на неаккуратний синтез, який дозволяє долати перешкоди, відбувається завдяки заміні ДНК-полімерази III на полімераза низької точності синтезу – ДНК-полімераза V (полімераза активується у відповідь на несприятливі умови, наприклад, ультрафіолетове опромінювання). Ця полімераза вставляє напроти тимінового димеру дві довільні нуклеотиди (виникає мутація), після чого замінюється на ДНК-полімераза III, яка продовжує точний синтез. Крім того, долати невеликі одноланцюгові прогалини при SOS-репарації допомагає ДНК-полімераза II – полімераза низької точності синтезу. Аналогічно перемикавання на полімерази низької точності (див. вище) відбувається і в еукаріотів за наявності в матриці пошкоджених основ, піримідинових димерів тощо.

Інший шлях здійснення реплікації через тиміновий димер у складі матриці отримав назву **постреплікативної (рекомбінаційної) репарації**. У разі наявності тимінового димеру реплісома може його обійти, залишивши прогалину в ланцюзі, що синтезується (рис. 9.27). У цьому випадку на місце прогалини шляхом гомологічної рекомбінації (див. наступний розділ) вставляється ділянка сестринської молекули ДНК. Прогалина, що залишається при цьому в сестринській молекулі, легко заповнюється ДНК-полімеразою. Таким чином, як і при SOS-репарації, пошкодження залишається і може бути виправлено пізніше завдяки ексцизійній репарації.

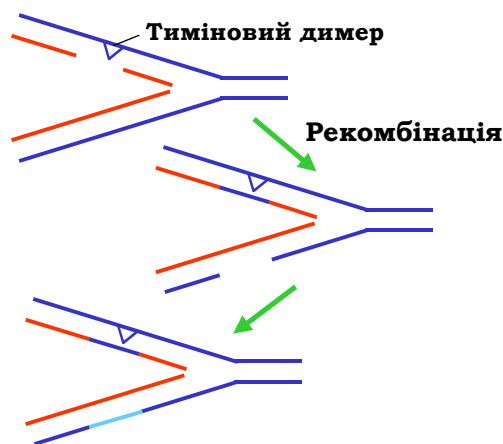


Рис. 9.27. Постреплікативна (рекомбінаційна) репарація

### Репарація дволанцюгових розривів

Точна репарація дволанцюгового розриву ДНК можлива лише під час реплікації, коли розірвана ділянка може бути відновлена завдяки використанню сестринської молекули як матриці за механізмом гомологічної рекомбінації (яка розглядатиметься в наступному розділі).

В інших випадках цілісність розірваних молекул ДНК відновлюється в консервативному для всіх організмів процесі *негомологічного з'єднання кінців* (NHEJ – Non-Homologous End Joining). З'єднуються при цьому будь-які кінці ДНК, що, за наявності великої кількості розривів, призводить до об'єднання ділянок різних хромосом, транслокацій тощо.

Ключову роль у процесі NHEJ відіграють білки Ku, які (у гетеродимерній формі) упізнають кінці ДНК, об'єднують їх у нековалентний комплекс і рекрутують до цього комплексу низку інших білків. У складі комплексу відбувається розплітання кінцевих ділянок ДНК і пошук мікрогомології – коротких комплементарних (чи майже комплементарних) одноланцюгових ділянок, які належать різним молекулам (рис. 9.28). У результаті утворюється короткий дуплекс, зайві одноланцюгові ділянки відщеплюються нуклеазами, лігаза зшиває одноланцюгові розриви. Таким чином, у процесі з'єднання відбувається втрата кількох пар основ. Ця особливість NHEJ-системи використовується при рекомбінації імуноглобулінових генів (див. наступний розділ) з метою підвищення розмаїття їхніх послідовностей.

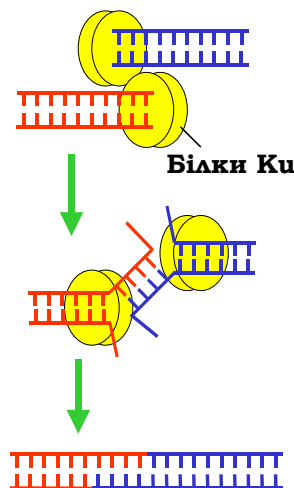


Рис. 9.28. Негомологічне з'єднання кінців ДНК

**КОНТРОЛЬНІ ЗАПИТАННЯ**

1. Що таке реплікон? Охарактеризуйте основні типи репліконів.
2. Дайте визначення реплікативної вилки. Яка різниця між двома ланцюгами ДНК, що синтезуються під час реплікації? Що таке фрагменти Оказаки?
3. Які ферментативні активності мають бактеріальні ДНК-полімерази? Порівняйте особливості й функціональне значення ДНК-полімераз I і III.
4. Опишіть основні риси просторової структури ДНК-полімераз.
5. Які спільні риси мають РНК- і ДНК-полімерази? У чому полягають принципові відмінності між ними?
6. Як здійснюється редагування помилок при синтезі ДНК?
7. Що таке реплісома? Назвіть її основні компоненти та їхнє функціональне значення.
8. Дайте визначення гелікази. Яку роль вона відіграє в реплікації? Що таке білки SSB?
9. Що таке праймер? Навіщо він потрібен, яка його хімічна природа, який елемент реплісоми його синтезує?
10. Яку роль виконує лігаза при реплікації ДНК?
11. Як забезпечується висока процесивність ДНК-полімерази?
12. Які топологічні проблеми виникають під час реплікації та за допомогою чого вони розв'язуються?
13. Дайте визначення ориджина. Як саме здійснюється ініціація реплікації в бактерій?
14. Укажіть особливості еукаріотичної системи реплікації порівняно з прокариотичною.
15. Які ДНК-полімерази працюють в еукаріотичній клітині? Яка полімераза здійснює ініціацію реплікації?
16. Яку функціональну роль виконує теломераза? Опишіть механізм роботи цього ферменту.
17. Порівняйте ексцизійну репарацію основ і нуклеотидів.
18. Як здійснюється репарація місметчів у ДНК?
19. Яким чином відбувається репарація дволанцюгових розривів ДНК?

**РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

*Agarwal, S., Tafel, A.A., Kanaar, R.* DNA double-strand break repair and chromosome translocations // DNA Repair (Amst). – 2006. – Vol. 5. – P. 1075–1081.

*Baker, T.A., Bell, S.P.* Polymerases and the replisome: machines within machines // Cell. – 1998. – Vol. 92. – P. 295–305.

*Demeret, C., Vassetzky, Y., Méchali, M.* Chromatin remodelling and DNA replication: from nucleosomes to loop domains // Oncogene. – 2001. – Vol. 20. – P. 3086–3093.

*Donmez, I., Patel, S.S.* Mechanisms of a ring shaped helicase // Nucleic Acids Res. – 2006. – Vol. 34. – P. 4216–4224.

*Dutta, A., Bell, S.P.* Initiation of DNA replication in eukaryotic cells // Annu. Rev. Cell Dev. Biol. – 1997. – Vol. 13. – P. 293–332.

*Friedberg, E., Walker, G., Siede, W.* DNA repair and mutagenesis. – Washington, DC : ASM Press, 1995.

*Hübscher, U., Maga, G., Spadari, S.* Eukaryotic DNA polymerases // Annu. Rev. Biochem. – 2002. – Vol. 71. – P. 133–163.

*Joyce, C.M., Steitz, T.A.* Function and structure relationships in DNA polymerases // Annu. Rev. Biochem. – 1994. – Vol. 63. – P. 777–822.

*Kornberg A., Baker T.A.* DNA replication. – New York : W.H. Freeman and Company, 1992.

*Kuriyan, J., O'Donnell, M.* Sliding clamps of DNA polymerases // J. Mol. Biol. – 1993. – Vol. 234. – P. 915–925.

*Latt, W.L., de, Jaspers, N.G.J., Hoeijmakers, J.H.J.* Molecular mechanism of nucleotide excision repair // Genes Dev. – 1999. – Vol. 13. – P. 768–785.

*Lindahl, T., Wood, R.D.* (1999) Quality control by DNA repair // Science. – Vol. 286. – P. 1897–1905.

*Lohman, T.M., Bjornson, K.P.* Mechanisms of helicase-catalyzed DNA unwinding // Annu. Rev. Biochem. – 1996. – Vol. 65. – P. 169–214.

*O'Donnell, M.* Replisome architecture and dynamics in *Escherichia coli* // J. Biol. Chem. – 2006. – Vol. 281. – P. 10653–10656.

*Stillman, B.* Cell cycle control of DNA replication // Science. – 1996. – Vol. 274. – P. 1659–1664.

*Zakian, V.A.* Telomeres: beginning to understand the end // Science. – 1995. – Vol. 270. – P. 1601–1606.



## Розділ 10

### РЕКОМБІНАЦІЯ ДНК

Чекалинский снова стасовал карты: игра пошла своим чередом.

*А. Пушкин. Пиковая дама*

Суттєвим моментом існування ДНК у живих системах є не тільки розглянуті в попередньому розділі процеси відновлення та збереження інформації, що міститься в послідовності нуклеотидів, а й різноманітні операції, спрямовані на перетасування цієї інформації з метою створення нових комбінацій генів. Серед процесів рекомбінації (recombination), які приводять до генетичного розмаїття, розрізняють такі:

- Гомологічна рекомбінація – обмін ділянками між досить довгими молекулами ДНК із гомологічними послідовностями пар основ. Відбувається в усіх організмів, які розмножуються статевим шляхом, між гомологічними хромосомами при мейозі (кросинговер), а також при мітозі та в прокаріотів – наприклад, після кон'югації двох бактеріальних клітин і проникнення ДНК з однієї в іншу.
- Сайт-специфічна рекомбінація – вирізання / вбудовування однієї молекули ДНК із/в іншу, або зміна орієнтації (інверсія) фрагмента ДНК у межах однієї молекули. Рекомбінація цього типу здійснюється за рахунок упізнання специфічними білками коротких елементів послідовності ДНК.
- Незаконна рекомбінація – об'єднання двох молекул ДНК, які не мають ані гомології, ані специфічних елементів послідовності.
- Переміщення в межах геному мобільних елементів послідовності ДНК.

## Гомологічна рекомбінація

Необхідною умовою для здійснення гомологічної рекомбінації є загальна гомологія між двома молекулами ДНК по всій довжині.

Загальну модель початкового етапу гомологічної рекомбінації зображено на рис. 10.1. Ініціюючою подією є *дволанцюговий розріз* в одній з гомологічних молекул. Цей розріз “розширюється” шляхом 5'-екзонуклеазної деградації ДНК – у результаті в місці розрізу залишаються два одноланцюгові 3'-хвости. Один із них здійснює *інвазію* – утворює подвійну спіраль з антипаралельним ланцюгом інтактної гомологічної молекули ДНК. Інший ланцюг цієї останньої, відповідно, виштовхується з дуплекса у вигляді одноланцюгової D-петлі. D-петля (від displacement), разом із 3'-хвостом, здатна переміщуватись у пошуку гомології – максимальної комплементарності в межах подвійної спіралі, що утворилася між ланцюгами двох гомологічних молекул ДНК. На наступному кроці відбувається репараційний синтез ДНК: два 3'-кінці розірваної молекули ДНК подовжуються ДНК-полімеразами з використанням у ролі матриць двох ланцюгів інтактної молекули. До цього моменту схема на рис. 10.1 є одночасно *схемою точної репарації дволанцюгового розриву* в одній із двох сестринських молекул ДНК під час реплікації.

Під час рекомбінації відновлення цілісності ДНК є тільки завершенням початкового етапу. У результаті інвазії та репараційного синтезу дві молекули ДНК об'єднуються в чотириланцюгову структуру з двома перехрестями ланцюгів – **структурами Холідея** (Robin Holliday). Кожна структура Холідея може переміщуватись (так звана *міграція гілки*), результатом чого є подовження **гетеродуплекса** – подвійної спіралі між двома *майже* комплементарними ланцюгами двох гомологічних молекул ДНК.

Вважається, що за схемою на рис. 10.1 відбуваються початкові стадії гомологічної рекомбінації в більшості про- та еукаріотів. Але механізми цих процесів найкраще вивчені для бактерій. В *E. coli* два кінці розірваної молекули ДНК упізнаються комплексом трьох білків – **recBCD** (рис. 10.2), який має геліказну та екзонуклеазні активності. Комплекс починає рухатися, руйнуючи подвійну спіраль і одночасно деградує ДНК за рахунок своїх 5'- та 3'-екзонуклеазних активностей. У так званому  $\chi$ -сайті (сайт підвищеної частоти рекомбінації в *E. coli* – послідовність довжиною вісім пар основ, що зустрічається із середньою періодичністю у 5 тис. пар основ) 3'-екзонуклеазна актив-

ність гальмується, результатом чого є утворення одноланцюгового 3'-кінцевого хвоста. Цей хвіст відразу вкривається білком **гесА**. Останній має два сайти зв'язування з ДНК, за рахунок першого з них і відбувається (у присутності АТР як кофактора) кооперативна взаємодія з одноланцюговою ДНК: молекули білка оточують полінуклеотидний ланцюг, формуючи на його поверхні праву спіраль. Саме цей комплекс і здійснює інвазію в подвійну спіраль інтактної гомологічної молекули ДНК: другий сайт взаємодії **гесА** з ДНК використовується на цій стадії для утворення комплексу з двома ланцюгами гомологічних молекул ДНК. За умови комплементарності двох ланцюгів здійснюється гідроліз АТР, після чого **гесА** втрачає спорідненість до ДНК, залишаючи сформований гетеродуплекс.

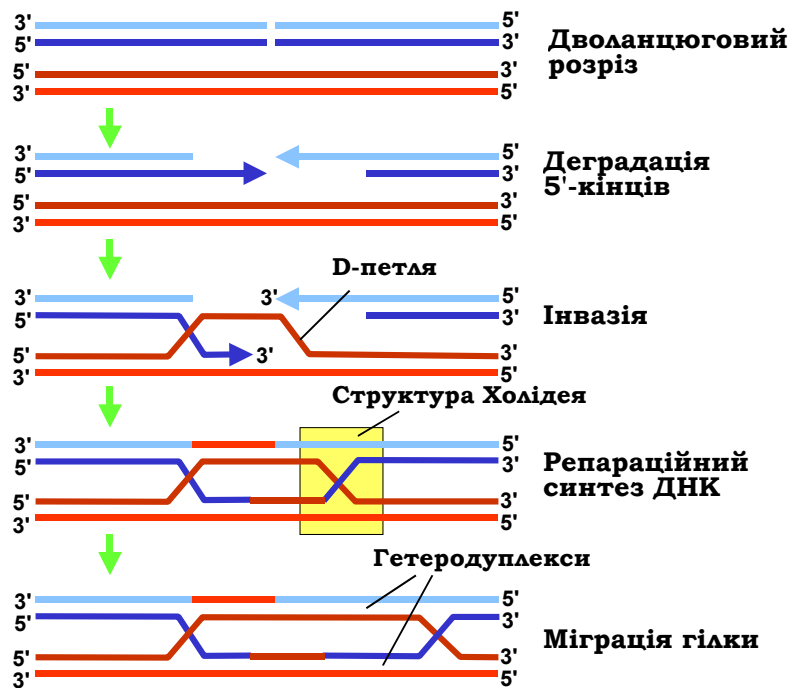


Рис. 10.1. Початкові стадії гомологічної рекомбінації

Повертаючись до останньої панелі на рис. 10.1, розглянемо окрему структуру Холідея. Як показано на рис. 10.3, її можна піддати ізомеризації, повернувши два дволанцюгові кінці на 180°. Саме у формі без перехрестя ланцюгів (праворуч на рис. 10.3) структура Холідея

фіксується завдяки її взаємодії з гомотетрамерним білком **ruvA**: білок зв'язується з центром хреста, утримуючи чотири одноланцюгові ділянки у приблизно планарній квадратній конфігурації (рис. 10.4, а). З **ruvA** та двома дуплексами, що виходять з хреста у протилежних напрямках, взаємодіє білок **ruvB** (шість субодиниць, які оточують подвійну спіраль кільцем). RuvB працює як АТР-залежна геліказа (або, швидше, як комплекс ремоделювання хроматину, див. рис. 6.16): два гексамери **ruvB** генерують обертальні рухи подвійної спіралі у протилежних напрямках, що приводить до протягування ланцюгів через комплекс **ruvA/ruvB** (рис. 10.4, б). Саме активність **ruvB** і забезпечує переміщення структури Холідея – міграцію гілки з одночасним подовженням гетеродуплекса.

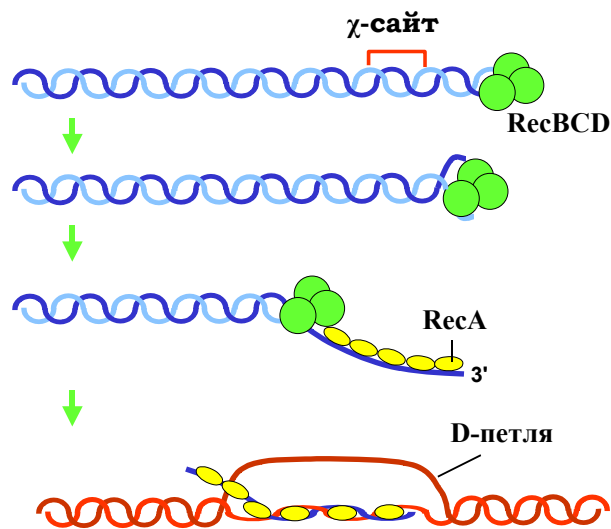


Рис. 10.2. Механізм початкових стадій гомологічної рекомбінації в *E. coli*

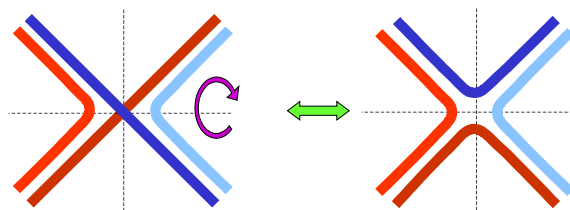
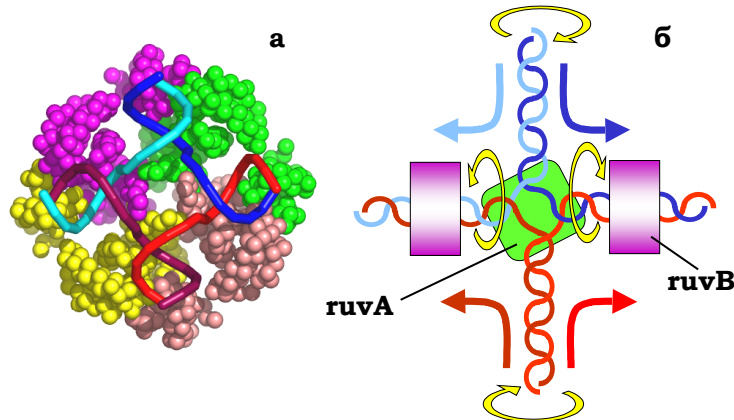


Рис. 10.3. Ізомеризація структури Холідея



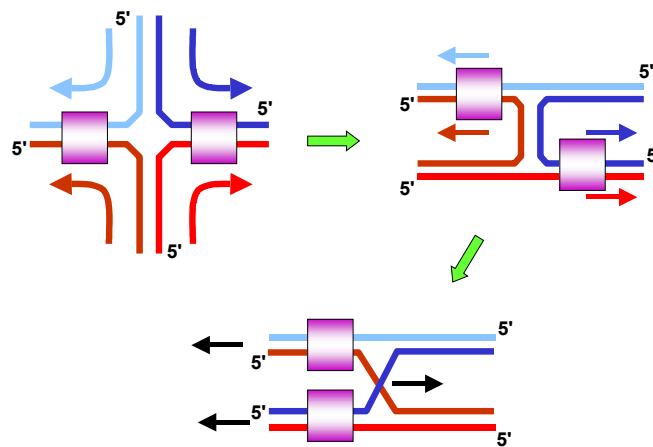
**Рис. 10.4.** (а): Комплекс *ruvA* зі структурою Холідея (1BDX, показано лише  $C^{\alpha}$ -атоми білка та Р-атоми ДНК). (б): Схема переміщення структури Холідея за рахунок активності *ruvB*: стрілки вказують напрямок обертання дуплексів та напрямок руху ланцюгів через комплекс *ruvA/ruvB* (колір стрілок збігається з таким ланцюгів)

Зі схеми на рис. 10.5 видно, як можна ототожнити протягування чотирьох полінуклеотидних ланцюгів із міграцією гілки. На схемі у два етапи проведено формальну ізомеризацію планарної конфігурації структури Холідея в конфігурацію з перехрестям: на першому етапі два ланцюги у складі хреста випрямляються, на другому – нижня частина отриманої конфігурації обертається на  $180^{\circ}$ . Очевидно, що зображене на рис. 10.4, 10.5 протягування ланцюгів через комплекс *ruvA/ruvB* є еквівалентним зсуву перехрестя у правий бік.

Останньою подією рекомбінації є розділення (resolution) структури Холідея **резольвазою** – білком *ruvC*. Резольваза є ендонуклеазою, дві молекули якої взаємодіють з комплексом *ruvA/ruvB* і двома ланцюгами хреста, розташованими один навпроти одного: є два рівноймовірні варіанти такого зв'язування. Отже, резольваза робить дволанцюговий розріз через хрест двома можливими шляхами (рис. 10.6). Після наступного зшивання розривів лігазою залишаються дві дволанцюгові молекули ДНК.

На рис. 10.6 показано ізомеризовану чотириланцюгову структуру з рис. 10.1 – дві структури Холідея з перехрестями перетворені на планарні (ділянки, синтезовані шляхом репарації, не позначено) – і результат її розділення. Із чотирьох можливих комбінацій розділення

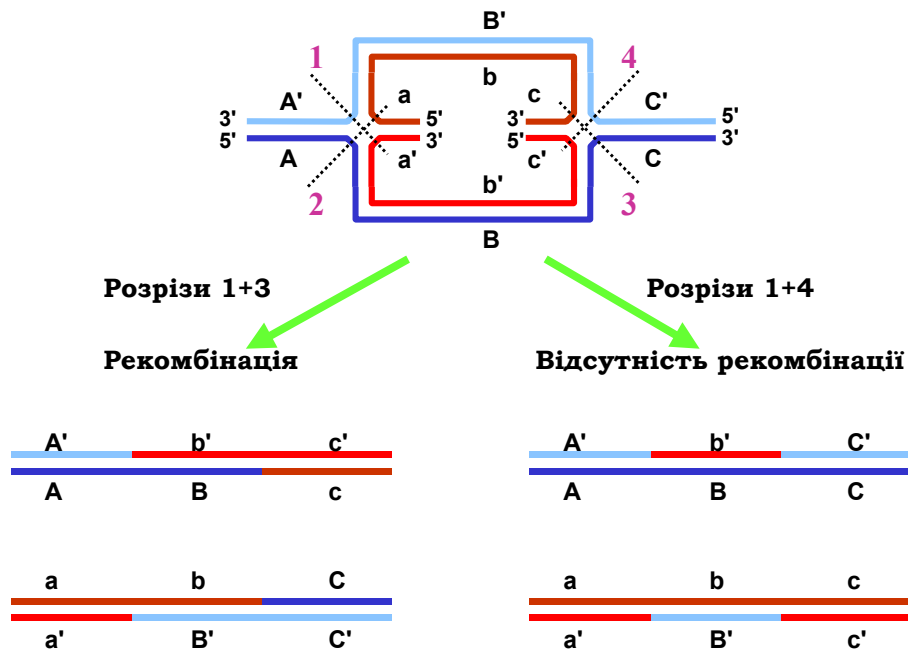
двох структур Холідея показано дві. Одна з цих комбінацій викликає рекомбінацію (аналогічно, рекомбінантною є і пара розрізів 2 + 4): дві гомологічні молекули ДНК обмінялися ділянками й умовна “абетка”, якою позначено фрагменти ДНК, змінює регістр – великі літери замінюються на маленькі й навпаки. Пари розрізів 1 + 4 і 2 + 3 не приводять до рекомбінації. Таким чином, рекомбінація при розділенні структур Холідея відбувається з імовірністю 50 %.



**Рис. 10.5.** Спрощена схема комплексу *ruvB* зі структурою Холідея з рис. 10.4 та її формальні перетворення у дві топологічно еквівалентні конфігурації. Остання конфігурація збігається з правою структурою Холідея з рис. 10.1. Кольорові стрілки вказують напрямок руху ланцюгів через *ruvB*, чорні – напрямок руху дуплексів і перехрестя

Незалежно від того, чи відбулася рекомбінація, усі продукти містять гетеродуплекси (середня частина всіх кінцевих молекул на рис. 10.6). Оскільки гетеродуплекси складаються з майже комплементарних ланцюгів, вони містять місметчі. Відповідно, кінцевою операцією, яка завершує процес гомологічної рекомбінації, є репарація цих місметчів описаною в попередньому розділі системою *mutHLSU*. На відміну від того, як ця система спрацьовує після реплікації, після рекомбінації ланцюг, де відбувається заміна нуклеотидів, обирається випадково. Наприклад, центральний фрагмент першої молекули на рис. 10.6 перетворюється або на *B/B'*, або на *b/b'*. Якщо на цій ділянці розташований ген, шляхом репарації обирається один із його алелів – *B* або *b*. Отже, побічним ефектом рекомбінації є відоме в генетиці явище *конверсії гена*. Білки системи *mutHLSU* якимось чином

блокують рекомбінацію на ранніх етапах, коли місметчів виявляється надто багато (так звана *гомеологічна рекомбінація*): міграція гілки зупиняється і чужорідна ДНК виштовхується з гетеродуплекса.



**Рис. 10.6.** Схема розділення двох структур Холідея.

Угорі: конфігурація чотирьох ланцюгів еквівалентна конфігурації на нижній панелі з рис. 10.1. Літерами позначено ділянки ланцюгів (великі й маленькі літери відповідають гомологічним ділянкам двох молекул, літери зі штрихом і без – комплементарним ділянкам вихідних дуплексів). Цифрами 1–4 позначено можливі розрізи резольвазою. Унизу: дві пари дволанцюгових молекул після розділення структур Холідея, отримані в результаті відповідних розрізів

Еукаріотичні клітини містять білки, гомологічні бактеріальним білкам RecA, RecBCD, RuvA, B і C. Наприклад, білок RAD51 дріжджів і людини є гомологічним білку RecA й виконує подібні функції. Ініціюючий рекомбінацію дволанцюговий розріз індукується у дріжджів білком Spo11, що належить до родини ДНК-топоізомераз II. Гомологічні білки знайдено в інших еукаріотів і бактерій. Отже, механізми гомологічної рекомбінації є, імовірно, спільними для всіх біологічних систем.

### Сайт-специфічна рекомбінація

Як приклад сайт-специфічної рекомбінації розглянемо інсерцію (інтеграцію) ДНК бактеріофага  $\lambda$  у бактеріальну хромосому (див. також розділ 5). У складі фагової ДНК є сайт attP – ділянка довжиною 270 пар основ, у складі бактеріальної ДНК – сайт attB (23 пари основ (рис. 10.7). Обидва сайти мають ділянку спільної (або гомологічної) послідовності довжиною 15 пар основ. Два білки – бактеріальний IHF (Integration Host Factor) і продукт одного з генів бактеріофага *інтеграза* – зв'язуються із цими сайтами.

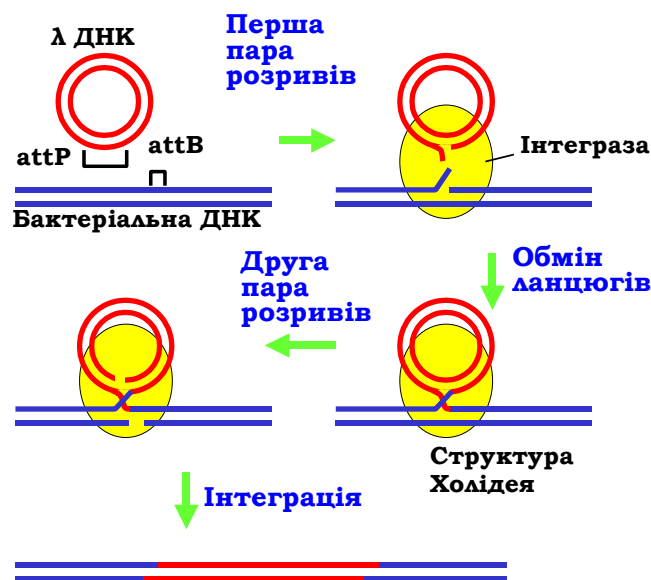


Рис. 10.7. Інтеграція ДНК бактеріофага  $\lambda$  у бактеріальний геном

За механізмом своєї дії інтеграза є сайт-специфічною ДНК-топоізомеразою I (див. розділ 3). Чотири субодиниці білка містять чотири активні центри, у кожному з яких є залишок Туг. На першому етапі інтеграції два активних центри роблять два одноланцюгові розрізи в бактеріальній і фаговій ДНК, одночасно ковалентно приєднуючи кінцеві фосфати до залишків Туг. Далі відновлюється фосфодієфірний



зв'язок, але з відповідним кінцем іншого дуплекса – відбувається обмін ланцюгами. На цьому етапі утворюється конфігурація чотирьох ланцюгів, еквівалентна структурі Холідея. Після цього друга пара субодиниць робить другу пару розривів (на відстані семи пар основ від першого розриву в кожному ланцюзі), і здійснюється друга пара обмінів ланцюгами, що й приводить до інтеграції.

Подібно до інтегрази працює *інвертаза* – сайт-специфічна топоізомераза I, яка здійснює інверсію ділянки ДНК бактеріофага  $\mu$ . Ділянка довжиною 3 тис. пар основ, що піддається інверсії, містить на кінцях короткі інвертовані повтори (рис. 10.8). Чотири субодиниці інвертази здійснюють чотири тимчасові розриви, відбувається обмін кінців і відновлення зв'язків, результатом чого є зміна напрямку. Інвертована ділянка містить два гени, що перекриваються, на різних ланцюгах ДНК (тобто такі, що піддаються транскрипції в різних напрямках). У результаті інверсії той чи інший ген (залежно від виду бактерії, в яку потрапив фаг) підставляється під промотор і транскрибується.

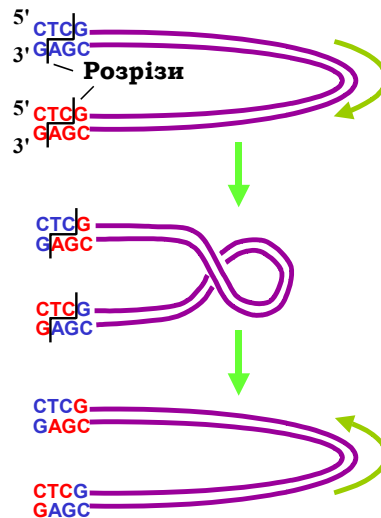


Рис. 10.8. Інверсія в бактеріофага  $\mu$

Розглянуті приклади вказують на основну різницю між двома типами рекомбінації: якщо при гомологічній рекомбінації дві молекули ДНК упізнають одна одну шляхом порівняння своїх послідовностей по великій довжині, сайт-специфічна рекомбінація передбачає наяв-

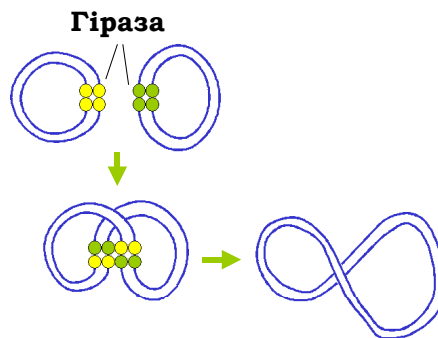
ність також гомологічних, але коротких елементів послідовності, що впізнаються специфічними білками.

Процеси сайт-специфічної рекомбінації за механізмами, цілком аналогічними до розглянутих, дуже поширені в бактеріофагів, при інсерції плазмід в основний геном у бактерій і дріжджів, вирізанні плазмід. Що стосується вищих еукаріотів, одним із головних процесів, у ході якого використовуються механізми сайт-специфічної рекомбінації, є дозрівання імуноглобулінових генів (див. нижче).

### Незаконна рекомбінація

Під незаконною рекомбінацією зазвичай розуміють об'єднання двох молекул ДНК без гомології послідовностей і без участі механізмів сайт-специфічної рекомбінації, коли з'єднання молекул відбувається в будь-якому місці послідовності.

Наприклад, повністю негомологічні молекули ДНК фага  $\lambda$  та плазмиди pBR322 можуть об'єднатися за допомогою гірази (див. розділ 3). Дві молекули гірази (по чотири субодиниці в кожній) роблять тимчасовий дволанцюговий розрив у кожній з молекул ДНК (розірвані кінці утримуються гіразою). Після цього дві гірази обмінюються парами субодиниць і відновлюють фосфодієфірні зв'язки (рис. 10.9).



**Рис. 10.9.** Незаконна рекомбінація двох циркулярних молекул ДНК

Загальним механізмом незаконної рекомбінації є репарація дволанцюгових розривів системою негомологічного з'єднання кінців (див. рис. 9.28), яка приводить до з'єднання чужорідних молекул ДНК.

Механізми незаконної рекомбінації (поряд із такими сайт-специфічною) використовуються при дозріванні імуноглобулінових генів з метою підвищення їх розмаїття.

### V(D)J-рекомбінація імуноглобулінових генів

Основою імунної відповіді при появі чужорідної молекули чи частинки – антигену – є синтез імуноглобуліну певного типу, який має до цього антигену високу специфічну спорідненість. Кількість типів імуноглобулінів, як і антигенів, є практично необмеженою. Зрозуміло, що закодувати таке розмаїття у вигляді окремих імуноглобулінових генів також неможливо. Еволюційним розв'язанням цієї проблеми став особливий рекомбінаційний процес дозрівання імуноглобулінових генів.

Молекула імуноглобуліну складається з двох так званих важких і двох легких поліпептидних ланцюгів (рис. 10.10). Кожен із ланцюгів має константну С-кінцеву (спільну для усіх імуноглобулінів) і варіабельну N-кінцеву частини. Варіабельні частини кожної пари важкого та легкого ланцюгів формують два антигензв'язувальні сайти на поверхні молекули.

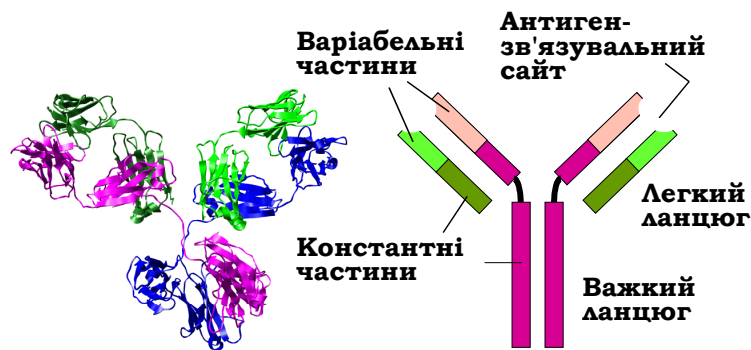
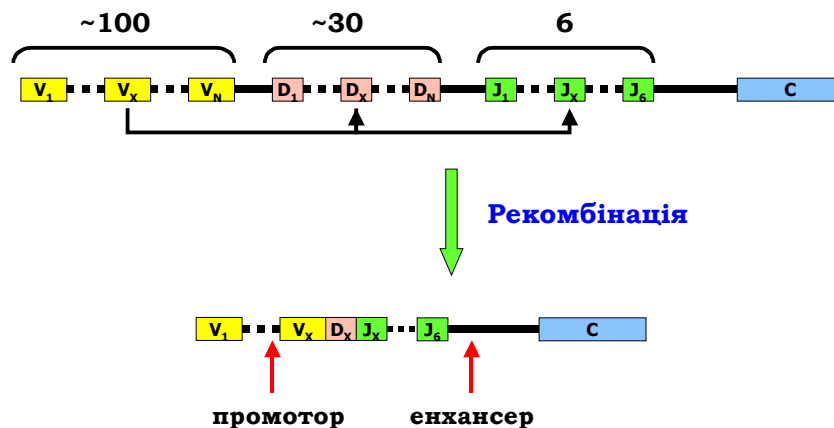


Рис. 10.10. Структура імуноглобуліну IgG1 (1IGY) і схема будови молекули

Незрілі імуноглобулінові гени мають вигляд кластерів окремих елементів нуклеотидної послідовності – блоків, з яких шляхом рекомбінації в імунокомпетентній клітині збирається активний імуноглобуліновий ген. Кластер важкого ланцюга містить ~100 V-сегментів (від variable), що тандемно повторюються (послідовності всіх сегментів роз-

різняються між собою), ~30 D-сегментів (від diversity), шість J-сегментів (від joining) і C-ділянку, яка кодує константну частину ланцюга (рис. 10.11). Аналогічно побудований кластер легкого ланцюга, який містить тільки два типи варіабельних сегментів (~100 V- та чотири J-сегменти). Активний ген “збирається” із сегментів трьох (або двох) типів, як із кубиків, шляхом V(D)J-рекомбінації: один випадковий D-сегмент з'єднується з випадковим J-сегментом (ділянка між ними вирізається), до них приєднується один з V-сегментів (рис. 10.11). Аналогічно, для легкого ланцюга об'єднуються один V- з одним із J-сегментів. Перед кожним V-сегментом розташований промотор, у спейсері перед C-ділянкою – енхансер. Їхнє зближення після рекомбінації активує транскрипцію, зайві спейсери та інтрони константної частини видаляються з мРНК шляхом сплайсингу.



**Рис. 10.11.** Будова кластера генів важкого ланцюга імуноглобуліну та схема збирання активного гена. Константна частина (C) містить інтрони, які не показано

Для важкого ланцюга може реалізуватися ~ 18 тис. комбінацій між сегментами трьох типів, для легкого – ~ 400 комбінацій, і загалом за рахунок рекомбінації для обох ланцюгів утворюється ~7 млн варіантів послідовності. Додаткова варіабельність забезпечується за рахунок індукції мутацій у межах V-сегментів, а також завдяки застосуванню механізмів незаконної рекомбінації при з'єднанні сегментів (див. нижче).

Рекомбінація сегментів залежить від сигнальних послідовностей RSS (Recombination Signal Sequences), що фланкують із двох боків кожен із сегментів (рис. 10.12) і є сайтами рекомбінації. Сигнальні по-

слідовності містять дві консервативні ділянки по сім та дев'ять пар основ, розділені неконсервативним спейсером. Послідовність спейсера не має значення – важливий лише його розмір: два типи спейсерів довжиною 23 і 12 пар основ визначають два типи RSS. Об'єднання сегментів при рекомбінації можливе лише за умови виконання правила “12/23”: об'єднуються тільки сегменти, фланковані RSS різних типів. Завдяки цьому виключається об'єднання однотипних сегментів (наприклад, V-V) чи пропущення D-сегмента.

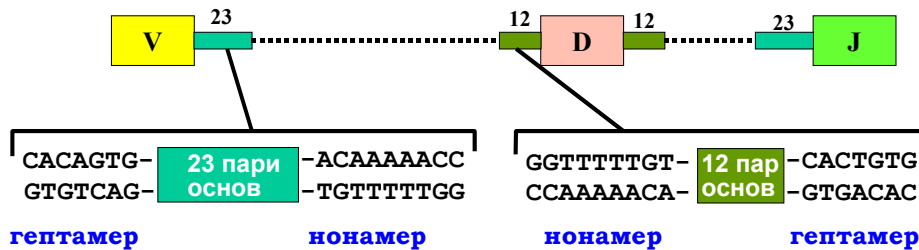
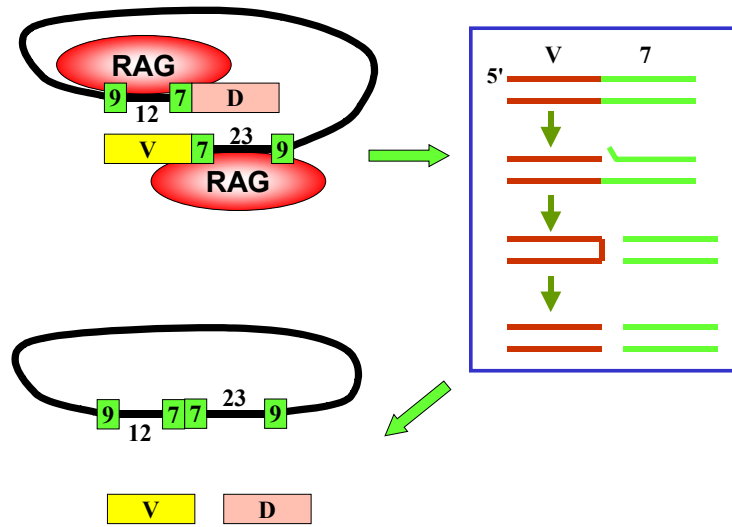


Рис. 10.12. Організація сигнальних послідовностей, які контролюють рекомбінацію в імуноглобулінових генах

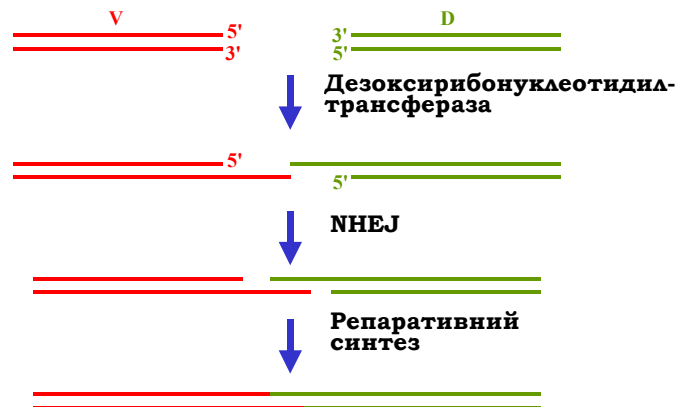
Консервативні елементи сигнальних послідовностей упізнаються білками RAG 1 та RAG 2 (продукти генів *RAG* – *Recombination Activating Gene*, активних лише у клітинах-попередниках В- і Т-лімфоцитів). Білки об'єднують у єдиний комплекс дві сигнальні послідовності при виконанні умови “12/23” і каталізують низку реакцій, яку схематично показано на рис. 10.13. Перша реакція – одноланцюговий розріз між сегментом і сигнальним гептамером. Далі відбувається трансестерифікація – розрив зв'язку в іншому ланцюзі та ковалентне замикання кінців сегмента у шпильку. Унаслідок таких самих реакцій у другій RSS вирізається, замикається в кільце та видаляється ділянка між двома сегментами.

Ковалентні шпильки на кінцях сегментів є нестабільними – вони швидко розмикаються (рис. 10.13). При цьому необов'язково утворюється тупий (без одноланцюгових виростів) кінець, оскільки шпилька може розімкнутися асиметрично. На останніх стадіях 3'-кінці сегментів добудовуються дезоксирибонуклеотидилтрансферазою (рис. 10.14) – ферментом, який без участі матриці приєднує випадкові нуклеотиди. Зрозуміло, що цей процес ще значно підвищує розмаїття послідовностей зрілих генів імуноглобулінів. Частина цих генів буде неак-

тивною завдяки стоп-кодонам, що з певною імовірністю виникають при такому випадковому синтезі, проте це є прийнятною платою за загальне збільшення варіантів.



**Рис. 10.13.** Залежне від білків RAG вирізання ділянки між двома сегментами імуноглобулінового гена



**Рис. 10.14.** Завершальні стадії процесу з'єднання сегментів імуноглобулінового гена

Нарешті кінці сегментів з'єднуються за допомогою системи репарації дволанцюгових розривів NHEJ (див. рис. 9.28): відбувається пошук мікрогомології між двома 3'-кінцевими виростами та репаративний синтез ДНК, що остаточно заповнює прогалини. Таким чином, на останніх стадіях рекомбінації процес з'єднання сегментів імуноглобулінових генів навмисно “загрубується” з метою підвищення варіабельності.

### **Транспозиції мобільних елементів**

Переміщення (транспозиції) мобільних (інтерсперсних) елементів (див також розділ 4) є ще одним особливим типом рекомбінації ДНК, яка відбувається в усіх організмів. Мобільні елементи можна поділити на два основні класи: *ДНК-транспозони*, що переміщуються за принципом "cut and paste", і елементи, переміщення яких відбувається шляхом синтезу РНК із наступним синтезом ДНК на РНК-матриці. Основними типами мобільних елементів, які належать до другого класу, є: *LTR-ретропозони* з довгими кінцевими повторами (Long Terminal Repeats); елементи *LINE* (Long Interspersed Elements); елементи *SINE* (Short Interspersed Elements).

Переміщення мобільних елементів є досить рідкою подією – у бактерій відбувається одна транспозиція на  $10^5$ – $10^6$  клітин. Проте активність мобільних елементів має важливе біологічне значення. Вбудовування мобільних елементів у кодуючі частини генів очевидно приводить до мутацій, у регуляторні – до зміни транскрипційної активності. У процесі свого переміщення мобільний елемент може захопити той чи інший регуляторний елемент послідовності та вбудувати його в інший промотор. Переміщення генетичного матеріалу разом із мобільними елементами розглядається сьогодні як один із найважливіших еволюційних факторів.

#### *ДНК-транспозони*

Мобільні елементи цього типу часто містять один або два гени, що кодують один або два білки під спільною назвою *транспозаза*. Транспозаза здійснює каталіз хімічних реакцій, що забезпечують транспозицію елемента – його вирізання з донорного сайту та вбудовування в сайт-мішень. Крім генів транспозази мобільний елемент може містити інші гени, або навпаки – гени транспозази можуть бути пошкодженими. В останньому випадку транспозиція даного елемента відбувається з використанням транспозази, закодованої іншим ДНК-транспозоном.

Кодуюча частина транспозона фланкована невеликими повторами, котрі розпізнає транспозаза, вирізаючи фрагмент ДНК з тупими кінцями (розрізи у двох полінуклеотидних ланцюгах розташовані точно один напроти одного). Сайт-мішень – невелика специфічна послідовність ДНК, яка теж упізнається транспозазою і теж розрізається – але так, що в місці розрізу залишаються одноланцюгові 5'-кінцеві вирости (рис. 10.15). Далі транспозаза каталізує ковалентне приєднання 3'-кінців транспозона до цих виростів, прогалини заповнюються шляхом репаративного синтезу.

Процес транспозиції (який, по суті, є одним із варіантів сайт-специфічної рекомбінації) залишає дволанцюговий розріз у місці, де знаходився транспозон. У випадку незалежної від реплікації транспозиції (*нереплікативна транспозиція*), цей розріз піддається репарації за механізмом негомологічного з'єднання кінців, тобто транспозон просто “стрибає” в інше місце. Але досить часто транспозиція відбувається під час реплікації (*реплікативна транспозиція*) – тоді розріз репарується за рекомбінаційним механізмом (див рис. 10.1): сестринська молекула ДНК використовується як матриця, і ділянка, що містила транспозон, відновлюється – транспозон “розмножується”.

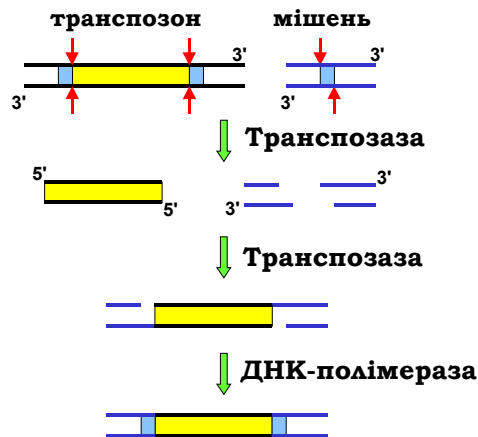


Рис. 10.15. Механізм переміщення ДНК-транспозона

### LTR-ретропозони

Кодуюча ділянка LTR-ретропозона (5–8 тис. пар основ) містить кілька генів: *зворотної транскриптази* (РНК-залежна ДНК-полімераза), інтегрази (аналог транспозази), РНКазы Н. На кінцях розташовані



довгі кінцеві повтори (200–600 пар основ), які й надали назву цим мобільним елементам.

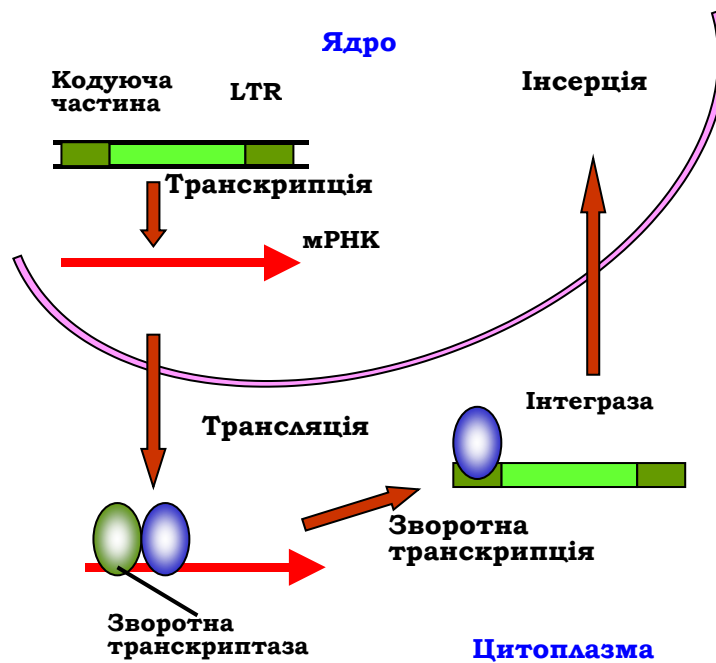


Рис. 10.16. Механізм переміщення LTR-ретропозона

Процес переміщення копії LTR-ретропозона нагадує життєвий цикл ретровірусів. Першим етапом є транскрипція ретропозона: синтез мРНК РНК-полімеразою II (рис. 10.16). Ця мРНК транспортується до цитоплазми, де піддається трансляції. Зворотна транскриптаза, яка є продуктом цієї трансляції, здійснює синтез ДНК: мРНК використовується як матриця, 3'-кінець молекули тРНК – як праймер. Комплекс синтезованої ДНК-копії ретропозона з інтегразою повертається до ядра, де здійснюється інсерція цієї ДНК у геном.

### Мобільні елементи LINE та SINE

Мобільні елементи LINE довжиною приблизно 6 тис. пар основ містять принаймні два гени – зворотної транскриптази та ендонуклеази – і невеликі специфічні послідовності на кінцях. Процес переміщення копії елемента починається так само, як і у випадку ретропозонів –

із транскрипції РНК-полімеразою II (рис. 10.17). Після трансляції в цитоплазмі синтезовані білки зв'язуються з мРНК, цей комплекс повертається до ядра, де й відбувається зворотна транскрипція: ендонуклеаза здійснює одноланцюговий розріз у геномній ДНК, 3'-кінець у місці розрізу використовується як праймер. Після синтезу другого ланцюга ДНК нова копія елемента остаточно вбудовується в геном.

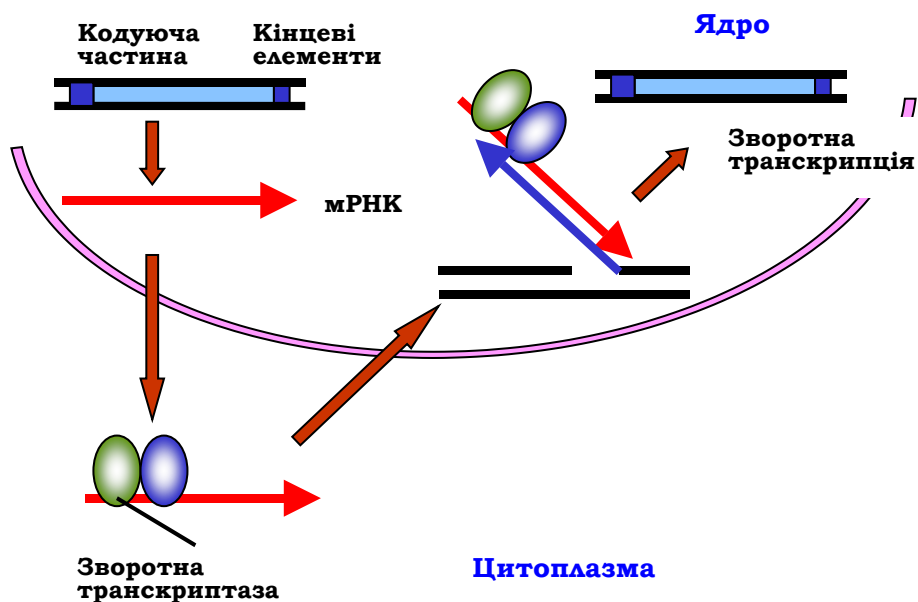


Рис. 10.17. Механізм переміщення мобільного елемента LINE

Синтез мРНК при транскрипції елемента LINE, як і для більшості інших еукаріотичних мРНК, закінчується на polyA-сигналі (див. розділ 7). Цей polyA-сигнал є слабким, що дозволяє елементу вбудовуватись в інтрони звичайних генів без особливих перешкод для експресії цих генів: система процесингу часто не помічає слабкий polyA-сигнал. Елементи LINE є, відповідно, найпоширенішими мобільними елементами в геномі вищих еукаріотів (~20 % геному людини представлено послідовностями LINE) і основним джерелом зворотної транскриптази у клітинах. З часом накопичуються мутації й окремі елементи LINE стають неактивними (тобто вже не кодують зворотну транскриптазу),

але переміщуються, використовуючи білки, що кодуються активними елементами. Найчастіше в геномі людини зустрічається елемент типу LINE, котрий позначається як L1 (кілька сотень тисяч копій).

Іноді мобільні елементи є не просто ділянками “егоїстичної ДНК”, що автономно розмножуються в геномі, а виконують певні конкретні біологічні функції. Наприклад, у дрозофіли відсутня теломеразна система й елемент LINE певного типу виконує функцію подовження кінців хромосом після реплікації: зворотна транскриптаза виступає як теломераза, мРНК мобільного елемента – як теломеразна матрична РНК (див. рис. 9.22). Цікаво, що послідовності ДНК гена теломерази та елементів LINE характеризуються високою гомологією: цілком можливо, що теломеразна система походить від мобільних елементів LINE.

Мобільні елементи SINE – короткі (100–300 пар основ) елементи послідовності, які не кодують білків і містять лише промотор для РНК-полімерази III. Процес їхнього переміщення є таким самим, як зображено на рис. 10.17: елементи SINE “паразитують” на системі LINE, використовуючи її ферменти. На відміну від усіх інших мобільних елементів, елементи SINE сконцентровані в ділянках геному, збагачених на гени. Імовірною причиною цієї закономірності є те, що РНК-продукт SINE має спорідненість до певної протеїнкінази, яка гальмує процес трансляції. Відповідно, зв'язування РНК із протеїнкіназою блокує останню, що приводить до загального підвищення рівня білкового синтезу. Таким чином, підвищення рівня транскрипції SINE разом з активацією інших генів стимулює трансляцію.

### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. Як саме починається гомологічна рекомбінація? Як здійснюється точна репарація дволанцюгових розривів за рекомбінаційним механізмом?
2. Що таке гетеродуплекс? Яку роль відіграє білок гесА при гомологічній рекомбінації?
3. Дайте визначення структури Холідея. В якій формі вона реалізується під час гомологічної рекомбінації? За рахунок чого стабілізується така форма?
4. За яким механізмом відбувається міграція гілки під час гомологічної рекомбінації?

5. Як здійснюється поділ структури Холідея під час гомологічної рекомбінації? З якою імовірністю відбувається обмін ділянками між двома гомологічними молекулами ДНК і від чого залежить реалізація такого обміну?

6. Що відбувається з гетеродуплексом після завершення гомологічної рекомбінації?

7. Чим відрізняється сайт-специфічна рекомбінація від гомологічної?

8. Що таке незаконна рекомбінація?

9. Як побудовані кластери імуноглобулінових генів? Які події при дозріванні імуноглобулінового гена можна віднести до сайт-специфічної рекомбінації?

10. Якими механізмами забезпечується необмежене зростання розмаїття послідовностей імуноглобулінових генів при їхньому дозріванні?

11. Яка різниця між реплікативним і нереплікативним механізмом переміщення ДНК-транспозонів?

12. Яка різниця між механізмами переміщення LTR-ретропозонів і мобільних елементів типу LINE?

13. Чим різняться мобільні елементи типів LINE і SINE?

#### **РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

*Boeke, J.D., Chapman, K.B.* Retrotransposition mechanisms // *Curr. Opin. Cell Biol.* – 1991. – Vol. 3. – P. 502–507.

*Grindley, N.G.F., Leschziner, A.E.* DNA transposition: from a black box to a color monitor // *Cell.* – 1995. – Vol. 83. – P. 1063–1066.

*Haniford, D.B., Chaconas G.* Mechanistic aspects of DNA transposition // *Curr. Opin. Genet. Dev.* – 1992. – Vol. 2. – P. 698–704.

*International Human Genome Sequencing Consortium.* Initial sequencing and analysis of the human genome // *Nature.* – 2001. – Vol. 409. – P. 860–921.

*Kazazian, H.H., Moran, J.V.* The impact of L1 retrotransposons on the human genome // *Nature Gen.* – 1998. – Vol. 19. – P. 19–24.

*Kuzimov, A.* Unraveling the late stages of recombinational repair: metabolism of DNA junctions in *E. coli* // *BioEssays.* – 1996. – Vol. 18. – P. 757–765.

*Labhart, P.* Nonhomologous DNA end joining in cell-free systems // *Eur. J. Biochem.* – 1999. – Vol. 265. – P. 849–861.

*Landy, A.* Dynamic, structural, and regulatory aspects of lambda site-specific recombinations // *Annu. Rev. Biochem.* – 1989. – Vol. 58. – P. 913–949.

*Levin, K.L.* It's prime time for reverse transcriptase // *Cell.* – 1997. – Vol. 88. – P. 5–8.

*Sadofsky, M.J.* The RAG proteins in V(D)J recombination: more than just a nuclease // *Nucl. Acids Res.* – 2001. – Vol. 29. – P. 1399–1409.

*Singer, M. F.* LINE1: a human transposable element // *Gene.* – 1993. – Vol. 135. – P. 183–188.

*Stahl, F.* Meiotic recombination in yeast: coronation of the double-strand-break repair model // *Cell.* – 1996. – Vol. 87. – P. 965–968.

*West, S.F.* Enzymes and molecular mechanisms of genetic recombinations // *Annu. Rev. Biochem.* – 1992. – Vol. 61. – P. 603–640.

*West S.* Processing of recombination intermediates by the Ruv ABC proteins // *Annu. Rev. Genet.* – 1997. – Vol. 31. – P. 213–244.

## Розділ 11

# МЕТОДИ МОЛЕКУЛЯРНОЇ БІОЛОГІЇ

... блоха действительно была на все ноги подкована на настоящие подковы, а Левша доложил, что и это еще не всё удивительное.

*Н. Лесков. Левша*

## Клонування, ампліфікація і секвенування ДНК

### Клонування ДНК

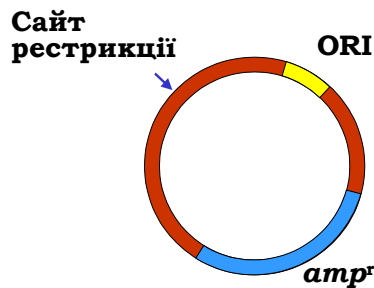
Перша проблема, котра постає перед дослідником, – це отримання в достатній кількості предмета свого дослідження. Ефективним підходом, що дозволяє розмножити будь-який конкретний фрагмент ДНК, є техніка клонування цієї ДНК у бактеріальних клітинах: фрагмент, що цікавить, вбудовується у ДНК-вектор з утворенням *рекомбінантної молекули ДНК*, яка вводиться у клітину (так звана *трансформація*). Далі залишається зачекати зростання бактеріальної культури – бактерія використовується як своєрідний біореактор.

Як вектор часто використовують бактеріальні плазміди – порівняно невеликі циркулярні молекули ДНК, що існують у клітині незалежно від бактеріальної хромосоми. Основними вимогами до плазміди як вектора є наявність у її складі ориджина реплікації, унікального (одного на плазміді) сайту, що впізнається певною рестриктазою, і гена стійкості до одного з антибіотиків (рис. 11.1).

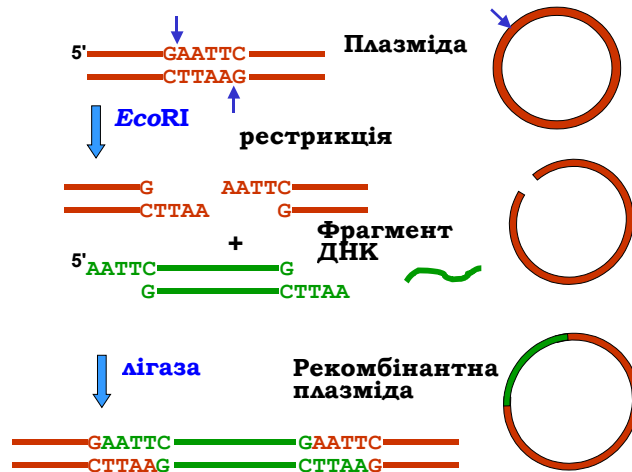
**Рестриктази** – специфічні до невеликих елементів послідовності ендонуклеази, що здійснюють розрізання обох ланцюгів ДНК. Назви рестриктаз (існує кілька сотень таких ферментів) утворюють за таким принципом: перша велика літера позначає рід мікроорганізму, дві маленькі – вид, римські цифри та іноді великі літери – порядко-

вий номер рестриктази серед інших рестриктаз даної бактерії. Наприклад, *EcoRI* – рестриктаза RI із *Escherichia coli*.

Сайтом рестрикції, що впізнається рестриктазою, є невеликі (чотири, шість, іноді трохи більше пар основ) паліндромні послідовності (див., наприклад, сайт *EcoRI* на рис. 11.2). Залежно від типу рестриктази, два розрізи, які вона здійснює, можуть бути розташованими точно один напроти одного у двох ланцюгах, що приводить до утворення так званих тупих (blunt) кінців. Частіше рестриктази залишають взаємно комплементарні 5'-кінцеві (а іноді 3'-кінцеві) одноланцюгові вирости (як на рис. 11.2) – липкі (sticky) кінці.



**Рис. 11.1.** Схема організації плазмідного вектора (Ori – ориджин реплікації, *amp<sup>r</sup>* – ген стійкості до ампіциліну)



**Рис. 11.2.** Виготовлення рекомбінантної ДНК на основі плазмідного вектора

Із метою **створення рекомбінантної ДНК** очищену плазмиду, яка містить єдиний сайт певної рестриктази, обробляють цією рестриктазою і отримують лінійний вектор з липкими кінцями (рис. 11.2). Далі додають фрагмент ДНК, який було вилучено за допомогою тієї самої рестриктази. За рахунок комплементарної взаємодії між липкими кінцями фрагмента й вектора утворюється циркулярний нековалентний комплекс двох молекул ДНК. Завдяки використанню іншого ключового ферменту рекомбінантної технології – ДНК-лігази (див. розділ 9) – полінуклеотидні ланцюги зшиваються. Найзручнішими є плазмідні вектори, які містять так званий *полілінкер* – ділянку з певним набором унікальних рестриктних сайтів, що дозволяє підібрати одну з рестриктаз, найпридатнішу для кожного випадку.

Фрагменти ДНК із тупими кінцями також можна вбудувати у вектор за допомогою лігази, хоча така реакція є на порядок менш ефективною. За допомогою *термінальної нуклеотидилтрансферази* (фермент, який без участі матриці приєднує нуклеотиди до 3'-кінця) до 3'-кінців фрагмента ДНК можна приєднати, наприклад, одноланцюговий poly(dA)-хвіст, а до 3'-кінців лінійного вектора – poly(dT). При змішуванні між комплементарними одноланцюговими кінцями відбувається спарювання, остаточне зшивання завершує ДНК-лігаза.

Можна також утворити тупі кінці на фрагментах із липкими кінцями. Для цього або здійснюють видалення одноланцюгових виростів за допомогою нуклеази S1 (нуклеаза, що гідролізує тільки одноланцюгову ДНК), або липкі кінці забудовують за допомогою фрагмента Кленова ДНК-полімерази I *E. coli* (див. розділ 9). Утворений фрагмент із тупими кінцями вбудовують у вектор за щойно описаною процедурою.

**Трансформація** бактеріальних клітин рекомбінантною плазмидою здійснюється зазвичай у розчині CaCl<sub>2</sub> або шляхом електропорації (через суспензію клітин проводиться короткий імпульс електричного струму). В обох випадках підвищується проникність клітинної стінки – і рекомбінантна плазміда потрапляє всередину. Зрозуміло, що далеко не всі бактерії отримують плазмиду при трансформації. І тут стає в нагоді ген стійкості до антибіотика: достатньо обробити бактеріальну культуру цим антибіотиком, щоб залишити тільки трансформовані клітини. Далі відбувається автономна реплікація плазмиди та розмноження самих клітин, що призводить до значного зростання загальної кількості плазмід (рис. 11.3). Клоновані плазмиди виділяють із бактеріальної культури, а обробка їх тією самою рестриктазою, що була використана при виготовленні рекомбінантної молекули, дозволяє вирізати з вектора клонований фрагмент ДНК.



Трансформація бактерій плазмідами є тим ефективнішою, чим меншою за розміром є плазміда. Відповідно, існує обмеження в розмірі фрагментів, що їх можна клонувати описаним шляхом – до 10 тис. пар основ. Альтернативною, але цілком аналогічною технікою, що дозволяє працювати з фрагментами довжиною ~20 тис. пар основ, є клонування ДНК із використання векторів на основі бактеріофага  $\lambda$  (див. розділ 5). За допомогою рестриктази фрагмент ДНК вбудовується у фагову ДНК, додаються порожні фагові оболонки, і здійснюється збирання фагових частинок *in vitro*. Рекombінантними бактеріофагами заражають бактеріальну культуру, де відбувається їхнє розмноження.

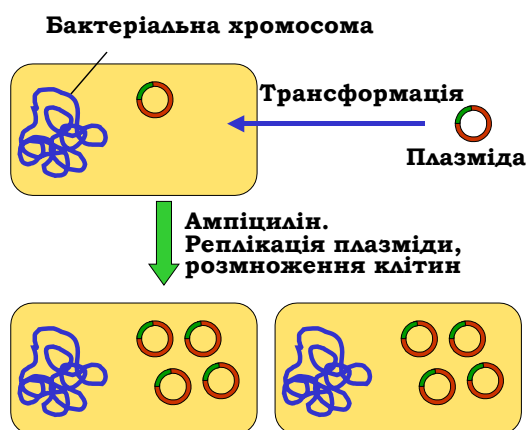


Рис. 11.3. Розмноження рекombінантної плазміди в бактеріальних клітинах

Використовуючи вектори на основі *космід*, можна клонувати фрагменти ДНК до 40 тис. пар основ. Косміда є плазмідною, яка крім ориджину, сайтів рестрикції та генів стійкості до антибіотиків містить два *cos*-сайти: липкі кінці лінійної молекули ДНК бактеріофага  $\lambda$ . Саме за рахунок *cos*-сайтів лінійна молекула фагової ДНК циркуляризується після її проникнення в бактеріальну клітину. У лінійну космиду з такими липкими кінцями вбудовується фрагмент ДНК, який бажано клонувати. Рекombінантна косміда упаковується *in vitro* у фагові частинки, якими обробляють бактеріальну культуру. У цьому випадку бактеріофаг використовується як ефективний засіб трансформації: лінійна косміда проникає у клітину, де циркуляризується за рахунок

cos-сайтів і бактеріальної лігази. Далі циркулярна косміда розмножується як звичайна плазміда (див. схему на рис. 11.3).

Для клонування фрагментів ДНК від 100 тис. пар основ використовують спеціально сконструйовані вектори ВАС (Bacterial Artificial Chromosome) і УАС (Yeast Artificial Chromosome). ВАС-вектори створено на основі F-плазмід бактерій, УАС-вектор являє собою штучну дріжджову мініхромосому, яка містить центромеру, теломери й точку початку реплікації. У такий вектор можна ввести чужорідний фрагмент ДНК розміром понад 100 тис. пар основ. Така мініхромосома, уведена в дріжджову клітину, буде реплікуватися й поводити себе аналогічно іншим дріжджовим хромосомам при мітотичному поділі. Більшість сучасних векторних систем є поліфункціональними – придатними не тільки для клонування ДНК, а й для експресії рекомбінантних білків, про що йтиметься нижче.

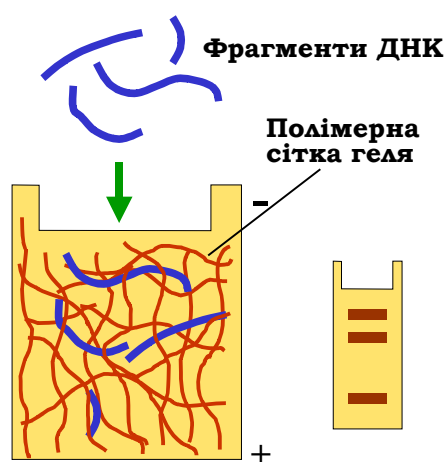
### *Гель-електрофорез*

Після виділення клонованого рекомбінантного вектора та розрізання його рестриктазою постає проблема відокремити фрагмент ДНК від вектора. З цією метою найчастіше застосовують гель-електрофорез – надзвичайно важливий метод у молекулярній біології.

Принцип електрофорезу дуже простий (рис. 11.4). Суміш фрагментів ДНК (або РНК) наносять у лунку пластинки гелю – середовища, сформованого полімерною сіткою. Зазвичай гель виготовляють на основі агарози або поліакриламід. Через нього пропускають електричний струм, завдяки чому негативно заряджені молекули нуклеїнової кислоти рухаються в електричному полі. Полімерна сітка гелю створює для цього руху суттєвий опір, який долається тим легше, чим меншим є розмір фрагмента. Через певний час фрагменти розділяються на окремі зони – смуги гелю, які містять гомогенний набір фрагментів одного розміру. Для візуалізації смуг гель забарвлюють флуоресцентним барвником (найчастіше використовують бромистий етидій), який зв'язується з ДНК. Після цього смугу можна вирізати з гелю, та екстрагувати ДНК.

Гель-електрофорез широко використовують також і як аналітичний засіб із метою з'ясування складу суміші. Для візуалізації смуг (особливо якщо розділяють невеликі кількості ДНК) часто застосовують радіоактивне мічення ДНК шляхом уведення радіоактивного ізотопу  $^{32}\text{P}$  у склад 5'-кінцевих фосфатних залишків. Після електрофорезу здійснюється *авторадіографія*: на гель накладають фотопластинку, яка засвічується напроти смуг унаслідок їхньої радіоак-

тивності. Альтернативою авторадіографії є експонування гелю на спеціальний фосфорний екран, який потім сканується сполученим із комп'ютером приладом – *фосфоріміджером*. Крім вищої, якщо порівнювати з авторадіографією, чутливості, використання фосфоріміджера має ту перевагу, що дозволяє зі значно вищою точністю оцінювати кількість матеріалу в смузі.



**Рис. 11.4.** Розділення фрагментів ДНК за розміром шляхом гель-електрофорезу. Праворуч – окремі зони, що містять розділені фрагменти

У найпростішому варіанті – для розділення фрагментів дволанцюгової ДНК за розміром – найчастіше використовують агарозні гелі. Поліакриламідний гель (ПААГ) має значно різноманітніше застосування. Зокрема, саме ПААГ використовують для розділення сумішей білків. Найпопулярнішим є варіант електрофорезу білків у присутності додецилсульфату натрію – зарядженого детергента, який розгортає поліпептидний ланцюг і вкриває його своєю «шубою». У результаті відбувається розділення ланцюгів різної довжини за принципом, що ілюструє рис. 11.4.

На рухливість макромолекул під час електрофорезу впливає також їхня форма. Відповідно, за допомогою електрофорезу можна, зокрема, розділити:

- Молекули ДНК, що містять перманентні вигини, зумовлені особливою послідовністю пар основ – вигнута ДНК більшою мірою гальмується полімерною сіткою гелю.

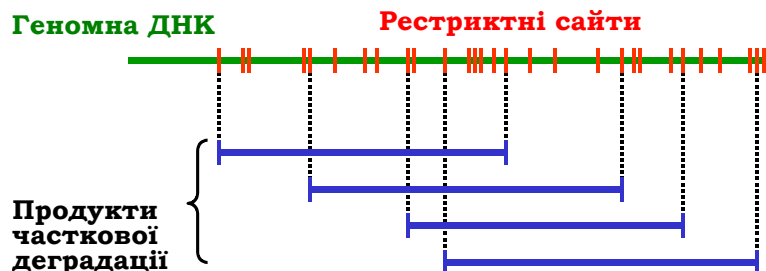
- Циркулярні та лінійні молекули ДНК, а також окремі циркулярні топоізмери (див. розділ 3), оскільки вони розрізняються за ступенем компактності.
- Молекулу вільної ДНК і комплекс цієї ДНК із білком – зрозуміло, що комплекс більше гальмується сіткою гелю.

Особливим варіантом електрофорезу нуклеїнових кислот є електрофорез одноланцюгових фрагментів ДНК у денатуруючих умовах (підвищена температура, висока концентрація сечовини). Якщо такий гель (поліакриламідний) досить тонкий (~0,4 мм – *секвенуючий гель*), він дозволяє розділити фрагменти, довжина яких різниться на один нуклеотид.

### *Створення та скринінг геномних бібліотек*

Раніше було розглянуто принципи клонування фрагмента ДНК. Але звідки береться цей фрагмент? Для того, щоб працювати з певною ділянкою ДНК (наприклад, певним геном) спочатку здійснюють клонування великої кількості різноманітних фрагментів геному – створюють бібліотеку клонів, серед яких уже шукають потрібний.

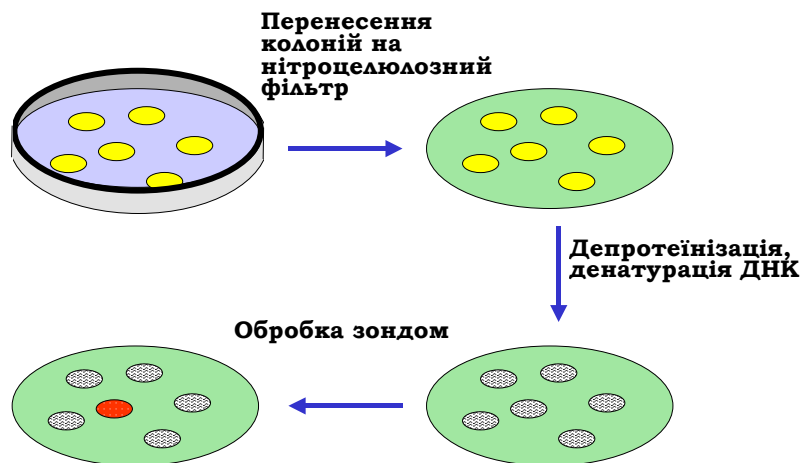
Стандартна процедура створення бібліотеки – метод дробовика (shotgun) – розпочинається з часткової (протягом обмеженого часу) обробки всієї геномної ДНК певною рестриктазою – таким чином, щоб отримати набір різноманітних фрагментів, які перекриваються, середня довжина яких становить ~20 тис. пар основ (рис. 11.5). Усі такі фрагменти вбудовують у вектори, і здійснюють трансформацію. Бактерії висівають на тверде середовище так, що кожна колонія походить від однієї клітини. Отже, набір колоній – це набір різноманітних клонів, кожен з яких містить один із фрагментів геному.



**Рис. 11.5.** Метод дробовика

Аналогічно створюють бібліотеки клонів кДНК (комплементарна ДНК-копія молекули мРНК). Для цього спочатку виділяють сумарну мРНК із клітин певного типу. Використовуючи зворотну транскриптазу та оліго(dT), які є комплементарними 3'-кінцевим polyA-хвостам мРНК, як праймери, на цих мРНК синтезують комплементарні ланцюги ДНК. Видаляють ланцюг РНК РНКазою і добудовують другий ланцюг ДНК за допомогою фрагмента Кленова ДНК-полімерази I. Далі отримані молекули кДНК піддають клонуванню. На відміну від геномної бібліотеки, бібліотека клонів кДНК містить тільки кодуючі послідовності (екзони) генів і тільки таких, котрі є активними в клітинах даного типу.

Потрібну нуклеотидну послідовність серед бібліотеки клонів можна знайти за допомогою гібридизації клонованої ДНК із радіоактивно-міченим ДНК-зондом (рис. 11.6). Колонії бактерій у чашці Петрі переносять на нітроцелюлозний фільтр – роблять своєрідну репліку. Здійснюються лізис клітин, депротейнізація та денатурація ДНК у лужному розчині. Після висушування фільтра одноланцюгова ДНК необоротно фіксується на ньому. Фільтр обробляють радіоактивно-міченою ДНК певної послідовності – зондом. Якщо зонд є комплементарним ДНК клону, відбувається гібридизація – ренатурація дволанцюгової ДНК. Детектування цього факту за допомогою авторадіографії дозволяє ідентифікувати розшукуваний клон.



**Рис. 11.6.** Гібридизація клонованої ДНК із радіоактивним зондом на нітроцелюлозному фільтрі

Зрозуміло, що описаний скринінг бібліотеки клонів залежить від наявності зонда. Одна з можливостей приготувати зонд базується на відомостях про амінокислотну послідовність білка того гена, клон якого потрібно відшукати (або білка-гомолога іншого організму). Хімічний синтез короткого олігонуклеотиду (~20 нуклеотидів) із послідовністю, що відповідає фрагменту амінокислотної послідовності білка й тому має бути комплементарною ділянці ДНК клону, дозволяє отримати зонд. Після внесення радіоактивної мітки такий олігонуклеотид можна використати для гібридизації. Проте, зважаючи на виродженість генетичного коду, бажано синтезувати набір послідовностей, які відповідають усім можливим комбінаціям кодонів для даної ділянки послідовності амінокислот. Суміш таких альтернативних олігонуклеотидів, один із яких є напевно комплементарним, і використовують як зонд.

Замість синтезу набору олігонуклеотидів можна обійтися синтезом лише одного зонда, використовуючи базу даних EST (Expressed Sequence Tag). База є загальнодоступною колекцією великої кількості коротких (200–400 пар основ) фрагментів послідовностей кДНК багатьох організмів. Щоб приготувати зонд, достатньо знайти в EST-базі коротку ділянку, яка відповідає фрагменту послідовності білка, і синтезувати її.

### **Полімеразна ланцюгова реакція**

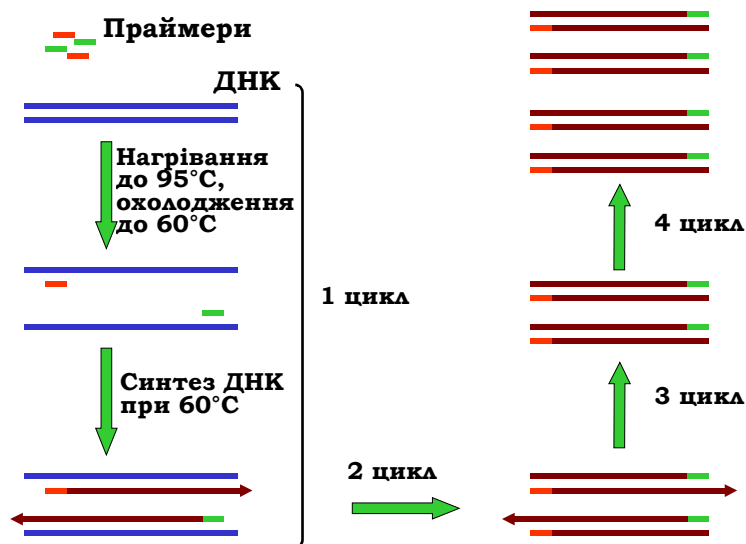
Альтернативним і додатковим до клонування методом збільшення кількості бажаного фрагмента ДНК – ампліфікації – є полімеразна ланцюгова реакція (ПАР, або PCR – Polymerase Chain Reaction). Практично ПАР – це реакція синтезу ДНК *in vitro*, яка повторюється багато разів: синтезовані ланцюги стають матрицями для синтезу в наступному циклі реакції. Для здійснення ПАР треба знати принаймні короткі послідовності ДНК на кінцях того фрагмента, котрий має бути ампліфікований.

Реакційна суміш містить слідові кількості вихідної ДНК, надлишкові кількості двох синтетичних праймерів до кінців фрагмента, нуклеозидтрифосфати чотирьох типів і ДНК-полімеразу *Taq* термофільної бактерії *Thermus aquaticus*. Використання саме полімерази *Taq* зумовлено тим, що вона не втрачає своєї активності при нагріванні до високих температур.

У першому циклі ПАР (рис. 11.7) відбувається нагрівання суміші до 95 °С (денатурація ДНК) і охолодження до 60 °С. Після охолодження до такої все ще досить високої температури ланцюги ДНК не ренатують, але праймери (оскільки вони присутні в досить високих кон-

центраціях) знаходять свої комплементарні ділянки й зв'язуються з ними. Полімераза *Taq* починає працювати, подовжуючи ці праймери.

Усі наступні цикли ПАР (реакція здійснюється в автоматичному режимі на приладі, що називається ампліфікатором) полягають у нагріванні суміші до 95 °С, яке змінюється охолодженням до 60 °С. Після другого циклу будуть знов синтезовані дві такі самі молекули, як після першого, а також дві молекули, котрі зображено на рис. 11.7. Після третього циклу – дві такі самі молекули, що й після другого, а також дві молекули, довжина яких чітко обмежена двома праймерами (рис. 11.7). Далі кількість таких коротких матриць буде зростати, і після 30-го циклу їх буде в мільйон разів більше, ніж вихідних довгих матриць.



**Рис. 11.7.** Полімеразна ланцюгова реакція.  
Деякі матриці, довжина яких виходить за межі ділянки, обмеженої двома праймерами, не показано (кількість таких матриць після багатьох циклів залишається дуже малою)

Метод ПАР має надзвичайно широке застосування. Він використовується кожного разу, коли є необхідність детектувати та дослідити невелику кількість ДНК, у тому числі – у складі неочищеного біологічного матеріалу. ПАР застосовують також у комбінації з клонуванням: ампліфікований ДНК-продукт можна клонувати й використати, наприклад,

для експресії білка (див. нижче); ампліфікація за допомогою ПАР стає в нагоді для збільшення невеликої кількості клонованої ДНК.

ПАР застосовують також для ампліфікації мРНК у вигляді її кДНК-копій – після синтезу кДНК на мРНК за допомогою зворотної транскриптази: варіант ПАР, що позначається як RT-PCR (Reverse Transcriptase Polymerase Chain Reaction). Аббревіатура RT-PCR використовується також у іншому значенні – Real Time Polymerase Chain Reaction (інше позначення: QRT-PCR – Quantitative Real Time Polymerase Chain Reaction): процедура, в якій ампліфікація ДНК супроводжується визначенням її кількості після кожного раунду за допомогою флуоресцентних барвників.

### **Секвенування ДНК**

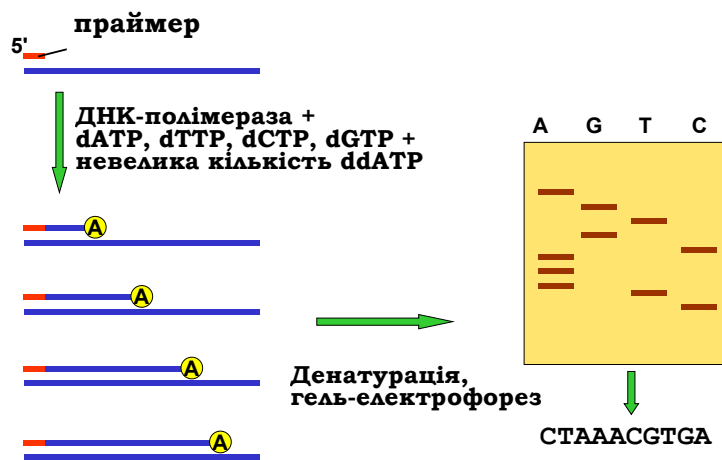
Клонований або ампліфікований фрагмент ДНК можна дослідити різними способами, однак найвичерпнішу інформацію дає встановлення нуклеотидної послідовності (sequence) фрагмента – секвенування.

На рис. 11.8 показано схему найбільш популярного сьогодні методу Сангера (Frederick Sanger). До одноланцюгової ДНК-матриці додається радіоактивно мічений праймер, повний набір дезоксинуклеозидтрифосфатів (dNTP), ДНК-полімераза та невелика кількість дидезоксинуклеозидтрифосфату одного із чотирьох типів (наприклад, ddATP). Дидезоксинуклеотид відрізняється тим, що містить атом Н замість ОН-групи не тільки при 2'-, а також і при 3'-атомі пентози (див. рис. 3.2). Включення такого нуклеотиду в ланцюг, що синтезується, приведе до зупинки подальшого зростання ланцюга внаслідок відсутності 3' ОН-групи на його кінці (див. рис. 3.4). Оскільки ddATP присутній у невеликій кількості, така подія буде відбуватися в різних точках ланцюга – в усіх, де аденін розташований напроти тиміну в складі матриці. Шляхом денатурації продуктів реакції отримаємо набір мічених одноланцюгових фрагментів від праймера до кінцевого аденіну. Довжина цих фрагментів у нуклеотидах визначить порядковий номер аденіну в складі ланцюга.

З метою визначення довжини фрагментів проводять гель-електрофорез у денатуруючих умовах, на сусідні лунки гелю наносять також продукти синтезу у присутності інших дидезоксинуклеотидів. Як показано на рис. 11.8, після електрофорезу та візуалізації смуг із такого гелю можна прочитати нуклеотидну послідовність.



В іншому методі секвенування – методі Максама – Гілберта (Allan M. Maxam, Walter Gilbert) замість ферментативного синтезу застосовують хімічне розривання ланцюга на нуклеотиді певного типу. Отримані фрагменти різної довжини так само розділяють за допомогою електрофорезу.



**Рис. 11.8.** Секвенування ДНК за Сангером: схема синтезу ДНК у присутності дидезоксиАТР (ліворуч). Шляхом аналогічної процедури для інших трьох дидезоксинуклеотидів отримуємо набір одноланцюгових фрагментів, що аналізуються за допомогою гель-електрофорезу в денатуруючих умовах (праворуч) – розподіл смуг дозволяє прочитати послідовність (праворуч унизу)

Метод Сангера використовується сьогодні автоматичними секвенаторами, в яких замість радіоактивно-мічених застосовуються флуоресцентно-мічені праймери. Для кожної з чотирьох реакцій беруть чотири різні флуоресцентні мітки, котрі випромінюють світло в різних спектральних діапазонах. Продукти всіх чотирьох реакцій наносять на гель для електрофорезу разом. Сканування гелю після електрофорезу лазерним променем, що збуджує флуоресценцію, дозволяє дискримінувати продукти різних реакцій, тобто різні кінцеві нуклеотиди, і, таким чином, негайно прочитати послідовність.

Здійснення секвенування кожного із фрагментів, які містяться в геномній бібліотеці клонів, і порівняння послідовностей фрагментів, що перекриваються (рис. 11.5), дозволяє розмістити клоновані фрагменти в порядку їхньої локалізації в геномі, тобто встановити повну послідовність геному.

Інший сучасний підхід у секвенуванні геномів (так зване *піросеквенування*), який реалізується на автоматизованих секвенаторах, дозволяє визначити послідовність значно швидше, дешевше і при цьому не потребує ані клонування ДНК, ані електрофорезу. Одноланцюгові фрагменти, отримані з невеликої кількості геномної ДНК, пришивають 5'-кінцями до мікрокульок (один фрагмент на кульку) та ампліфікують за допомогою ПАР. Кожну кульку із пришитими до неї ампліфікованими ідентичними фрагментами розміщують у мікрореакторі, де здійснюється ДНК-полімеразна реакція. Нуклеозидтрифосфати подаються в реакційну суміш імпульсно один за одним. Якщо нуклеотид певного типу виявляється комплементарним матриці та включається у зростаючий ланцюг, пірофосфат, що при цьому звільняється, залучається до низки хімічних реакцій, де остання реакція супроводжується випромінюванням світла (хемілюмінесценція). Світловий сигнал фіксується оптичною системою, і послідовність таких сигналів читається як нуклеотидна послідовність. Реакція здійснюється паралельно у 200 тис. мікрореакторів (для 200 тис. фрагментів, які перекриваються), що дозволяє встановити послідовність приблизно 200 млн пар основ за 4,5 години.

Останнім часом з'явилася нова перспектива секвенування ДНК у процесі протягування молекули через нанопору. Пóra діаметром ~1 нм утворюється у кремнієвій мембрані, і одноланцюгова ДНК досить повільно проходить крізь цю пору під дією електричного поля. Виявилось, що характеристики електричного струму через мембрану залежать від того, який нуклеотид знаходиться в порі в даний момент. Є сподівання, що з розвитком цього методу з'явиться можливість швидкого та дешевого секвенування геному однієї конкретної людини.

### **Біоінформатика**

Велика кількість послідовностей (у тому числі повних послідовностей геномів), які вже встановлені та продовжують накопичуватись, створює проблему їхнього збереження та аналізу. Зрозуміло, що таке питання, яке постає вже при порівнянні послідовностей окремих клонованих фрагментів з метою з'ясування нуклеотидної послідовності геному, може бути вирішеним лише за допомогою комп'ютерів. Біоінформатика – галузь, пов'язана з розв'язанням подібних проблем, є сьогодні невід'ємною складовою частиною молекулярної біології.

Найбільші загальнодоступні через Інтернет бази даних нуклеотидних і амінокислотних послідовностей створено в Європейській Лабораторії Молекулярної Біології (EMBL Sequence Data Base) і Національних Інститутах Здоров'я США (GenBank). На відповідних порталах розміщено

також програмні засоби порівняння послідовностей, пошуку промоторів, стартових точок транскрипції, інтронів та екзонів, визначення білкових послідовностей із послідовностей нуклеотидів тощо. Серед баз даних on-line слід згадати також бази даних просторових структур білків (Protein Data Bank) і нуклеїнових кислот (Nucleic Acid Database).

На основі біоінформатичного аналізу можна встановити, наприклад, функціональне значення невідомого щойно клонованого гена за його гомологією з відомим геном іншого організму, структуру гена (знайти промотор, екзони, точки термінації транскрипції), структурно-функціональні особливості нових білків, функціональні зв'язки між різними білками та генами, філогенетичні стосунки між різними таксонами. Завдяки розвитку біоінформатики існує можливість (у доповнення до результатів, що отримують *in vivo* та *in vitro*) розв'язувати різноманітні проблеми просто за допомогою комп'ютера – *in silico*.

### **Експресія рекомбінантних білків**

Оскільки переважна більшість білків синтезується клітиною в невеликих кількостях, просте виділення цих білків із метою їхнього дослідження є практично неможливим. Технології рекомбінантних ДНК вирішують це питання. Ген, що кодує даний білок, можна клонувати, вбудувати (наприклад, у бактеріальну плазмиду) і змусити бактеріальну культуру продукувати цей білок. Виділення та очищення білка з культури біохімічними методами вже не є принциповою проблемою, зважаючи на його велику кількість.

Одну з найпростіших схем експресії білка в *E. coli* зображено на рис. 11.9. Зрозуміло, що еукаріотичний ген не має сенсу вводити в прокаріотичну клітину – прокаріоти не мають системи сплайсингу. Тому беруть лише кодуючу частину гена, яку можна отримати з бібліотеки клонів кДНК. Використовуючи придатну рестриктазу та лігазу, потрібну кДНК вбудовують у плазмиду поряд із промотором – наприклад, лактозного оперона (див. розділ 5). Здійснюють трансформацію рекомбінантної плазмиди в бактеріальні клітини, до бактеріальної культури додають синтетичний індуктор *lac*-оперона IPTG. Промотор активується, і здійснюється транскрипція гена та трансляція білка.

Більш ефективна двостадійна система експресії використовує промотор РНК-полімерази бактеріофага T7. У бактеріальному геномі ні таких полімераз, ні відповідних промоторів немає. У клітину вводяться два плазмідні вектори: один містить *lac*-промотор і ген РНК-полі-

мерази T7, інший – сильний промотор полімерази T7 разом із геном білка, що має бути експресований. IPTG індукуює експресію полімерази, яка зв'язується тільки з промотором у складі другої плазміди (інших промоторів немає) і забезпечує синтез великої кількості білка.

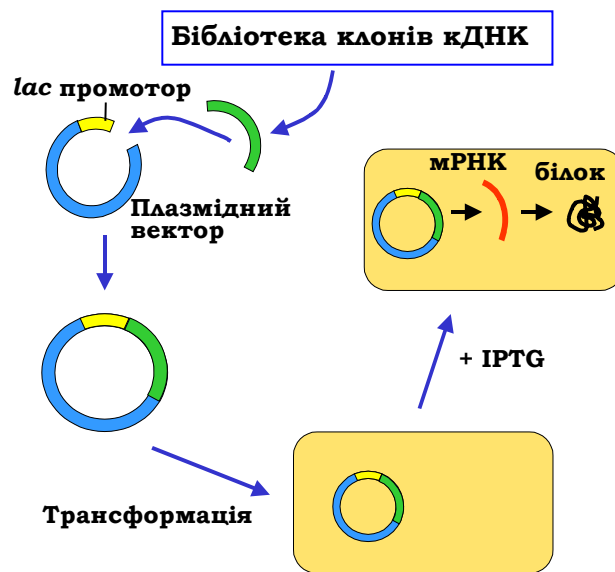
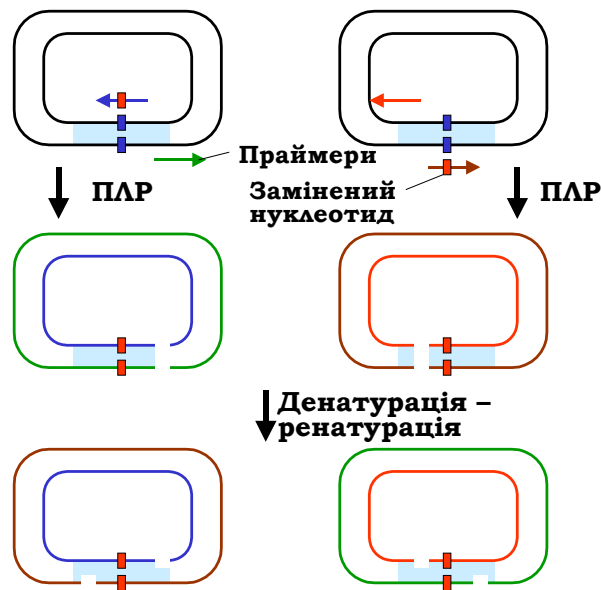


Рис. 11.9. Експресія білка в бактеріальній клітині.  
*lac*-Промотор – промотор лактозного оперона,  
IPTG – синтетичний індуктор оперона

Вивчення структури й особливостей функціонування багатьох білків було б просто неможливим без такої штучної експресії. Крім того, використання рекомбінантних ДНК для експресії білків відкриває принципово нові можливості: послідовність ДНК можна практично як завгодно змінити й отримати мутантну форму білка. До білкового гена можна пришити ген іншого білка, який буде виконувати роль мітки. Частина гена, що відповідає частині білка (скажімо, структурному домену) можна вилучити, й отримати редукований білок чи окремий домен для структурних досліджень. Будь-який нуклеотид можна замінити на інший методами спрямованого мутагенезу, завдяки чому отримати білок з амінокислотною заміною. Такі підходи мають велике значення для вивчення механізмів активності ферментів, специфічного впізнання білками інших молекул тощо.

Одну з можливостей замінити нуклеотид за допомогою ПАР ілюструє рис. 11.10. Як матриця для ампліфікації використовується плазміда, котра містить послідовність-мішень. Реакція проводиться у двох варіантах – із різними наборами праймерів. Один із двох праймерів у кожному наборі містить замінений нуклеотид, як показано на рисунку. Результатом ампліфікації є лінійні молекули ДНК із заміненою нуклеотидною парою. Об'єднання продуктів ампліфікації, денатурація і потім ренатурація приводять до того, що половина молекул обмінюється ланцюгами: оскільки дволанцюговий розрив має різне положення у продуктах ампліфікації, утворюються дві циркулярні молекули з односторонніми розривами. Після трансформації в бактеріальну клітину лінійні молекули (друга половина матеріалу) руйнуються бактеріальними нуклеазами, а односторонні розриви в циркулярних молекулах репаруються. У результаті відбувається клонування послідовності-мішені, яка містить замінену нуклеотидну пару.



**Рис. 11.10.** Спрямований мутагенез за допомогою ПАР. Послідовність-мішень зафарбована блакитним кольором

Велика кількість білків (у тому числі, ферменти, що використовуються в рекомбінантних технологіях) виробляються шляхом експресії в бактеріальних клітинах. Поряд із прокаріотичними розроб-

ляються й використовуються також системи експресії рекомбінантних білків в еукаріотичних клітинах. Особливо важливими еукаріотичні системи експресії є для еукаріотичних білків, котрі не можуть бути синтезовані бактеріальною клітиною в активній формі. Це, зокрема, стосується білків, які піддаються суттєвим посттрансляційним модифікаціям – наприклад, глікопротеїдів.

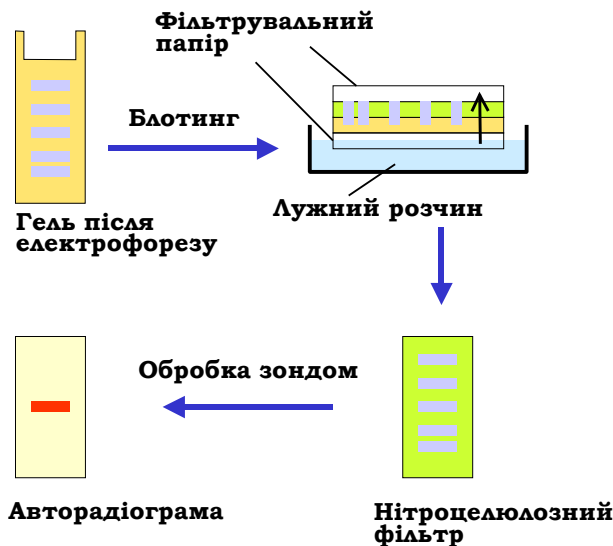
## **Методи аналізу структури й експресії генів і геномів**

### **Блот-гібридизація**

**Саузерн-блотинг.** Потужним засобом аналізу складних сумішей ДНК щодо наявності там специфічних елементів послідовності є блот-гібридизація на нітроцелюлозних фільтрах за Саузерном (Edward Southern). Назва процедури, яку схематично зображено на рис. 11.11, походить від слова blotting (промакування): фрагменти ДНК розділюються за допомогою гель-електрофорезу (залишаючись невидимими в гелі), після чого на гель накладають нітроцелюлозний фільтр, а під та над цим “сендвічем” розміщують фільтрувальний папір і занурюють нижній шар паперу в лужний розчин. Під дією капілярних сил розчин піднімається до верхнього шару паперу, “захоплюючи” при цьому ДНК і переносячи її з гелю на нітроцелюлозу. Одночасно при цьому ДНК денатурується лугом. У результаті одноланцюгова ДНК опиняється на фільтрі – середовищі, придатному для подальшої гібридизації, а сам фільтр є точною реплікою вихідного гелю. Так само можна перенести ДНК на нітроцелюлозу, застосувавши електрофорез у поперечному напрямку – із гелю на фільтр. Далі проводять обробку фільтра зондом – одноланцюговим фрагментом ДНК певної послідовності, який містить радіоактивну мітку. Зонд гібридизується з комплементарною ДНК у певних досі невидимих смугах, що можна зафіксувати за допомогою авторадіографії.

У такий спосіб можна встановити, наприклад, наявність у геномі додаткових копій послідовності, що є гомологічними вже відомій; присутність у невивченому геномі генів, гомологічних відомим генам; присутність специфічних послідовностей ДНК у препаратах, що отримані з білково-нуклеїнових комплексів. Прикладом одного з численних застосувань Саузерн-блотингу є метод **фінгерпринтингу ДНК** (DNA fingerprinting). Метод базується на факті наявності в еукаріотичних геномах мінісателітних повторів (див. розділ 4) – невеликих елементів

послідовності, що тандемно повторюються в різних місцях геному кілька разів. Розподіл локусів за кількістю повторів є індивідуальним – так само, як відбитки пальців. З метою ідентифікації особи (чи особи – у криміналістиці, судових справах тощо) геномну ДНК обробляють рестриктазою, що не має своїх сайтів усередині повтору. Фрагменти розділяються шляхом електрофорезу, здійснюється блотинг і гібридизація з радіоактивно міченим елементом послідовності мінісателіта. У результаті на авторадіограмі представлено специфічний для особи набір фрагментів різної довжини, тобто різної кількості повторів мінісателіта – своєрідний молекулярний відбиток (DNA fingerprint).



**Рис. 11.11.** Блот-гібридизація

**Нозерн-блотинг** відрізняється від описаної процедури блотингу за Саузерном лише тим, що на гель для електрофорезу наносять сумарний препарат виділеної мРНК. Гібридизація з міченим фрагментом ДНК (наприклад, κДНК із бібліотеки клонів) дозволяє встановити наявність певної мРНК, тобто активність гена, у клітинах певного типу, після дії активуючих / репресуючих факторів тощо, а також оцінити рівень цієї активності (концентрацію мРНК) за інтенсивністю забарвлення смуги на авторадіограмі. Назва методу (nothern) є просто жартиливою аналогією з буквальним значенням прізвища Саузерна.

**Вестерн-блотинг** – назва, що продовжує цю аналогію – це метод перенесення білків з гелю після електрофорезу на фільтр з метою аналізу складних білкових сумішей. Фільтр обробляють антитілами, специфічними до певного білка, потім міченими (наприклад, флуоресцентною міткою) антитілами до цих антитіл. Практично, це єдиний спосіб детектувати наявність малої кількості конкретного білка в тотальному білковому препараті.

**Визначення стартових і кінцевих точок та рівня активності транскрипції**

Кілька описаних нижче прикладів ілюструють деякі підходи, розроблені для характеристики транскриптів і дослідження рівня транскрипційної активності.

**Картування за допомогою нуклеази S1** дозволяє не тільки визначити старт і кінцеву точку транскрипції певного гена, а й оцінювати рівень транскрипційної активності цього гена. Процес починається з клонованого фрагмента гена, який містить кінцеву точку, яка має бути встановленою. Навкруг неї обирають два рестриктні сайти однієї рестриктази, і ще один – іншої, як показано на рис. 11.12. Перша рестриктаза вирізає фрагмент, залишаючи одноланцюгові вирости на 5'-кінцях. Використовуючи їх як матрицю, фрагмент Кленова ДНК-полімерази I добудовує 3'-кінці, приєднуючи радіоактивно мічені нуклеотиди – у результаті 3'-кінці обох ланцюгів містять мітку. Щоб позбавитись однієї з них, використовують сайт другої рестриктази, який все ще залишається всередині. Нарешті маємо мітку на 3'-кінці тільки матричного ланцюга (який із ланцюгів є матричним, треба знати при підборі рестриктаз або встановити за кінцевим результатом). Денатурація цього фрагмента дає радіоактивно-мічений одноланцюговий зонд, для якого є точно відомими його довжина й локалізація в гені. На останніх етапах проводять гібридизацію цього зонда із сумарним препаратом клітинної РНК у розчині – комплементарна частина відповідної мРНК утворює із зондом подвійну спіраль. Після цього й використовують нуклеазу S1, яка здатна гідролізувати тільки одноланцюгову нуклеїнову кислоту. Для локалізації кінцевої точки транскрипції залишається визначити точну довжину отриманого міченого фрагмента ДНК за допомогою електрофорезу: ця довжина дорівнює відстані від мітки до кінцевої точки. Інтенсивність смуги після



електрофорезу дає інформацію про транскрипційну активність гена: чим вищою є інтенсивність, тим більше мРНК містилося у препараті, виділеному із клітини.

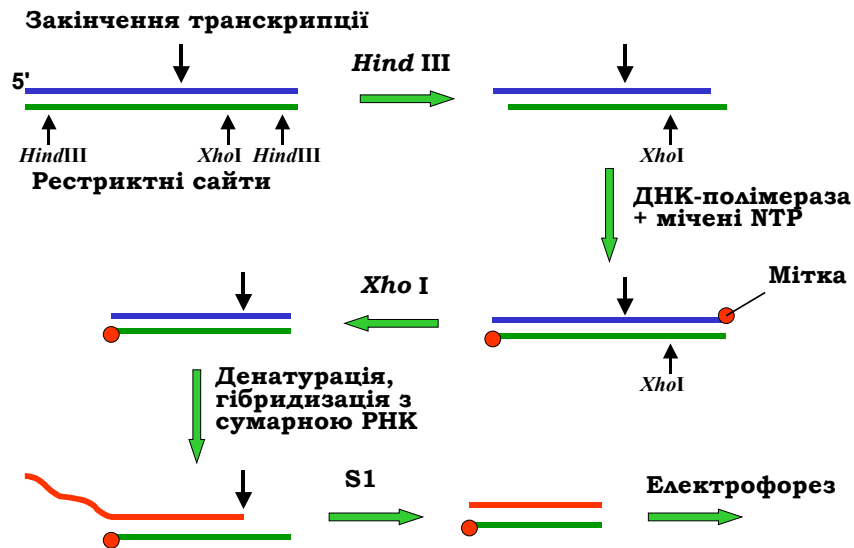


Рис. 11.12. S1-картування точки закінчення транскрипції

Так само, підібравши інші рестриктази та здійснивши на відповідному етапі мічення 5'-кінців, можна визначити стартову точку транскрипції.

**Подовження праймера** (primer extension) – альтернативний метод визначення стартової точки. Для його реалізації необхідно знати будь-яку невелику послідовність усередині кодуєчої частини гена й синтезувати відповідний праймер, помітивши його (рис. 11.13). Праймер додається до сумарної мРНК і гібридується з нею, після чого здійснюється реакція зворотної транскрипції. Синтез ДНК зворотною транскриптазою припиняється на 5'-кінці мРНК, який і відповідає стартовій точці. Електрофорез у денатуруючих умовах дозволяє точно встановити довжину фрагмента, тобто локалізувати стартову точку відносно місця гібридизації праймера. Як і у випадку S1-картування, інтенсивність електрофоретичної смуги дозволяє оцінити кількість мРНК у препараті, тобто рівень транскрипційної активності.

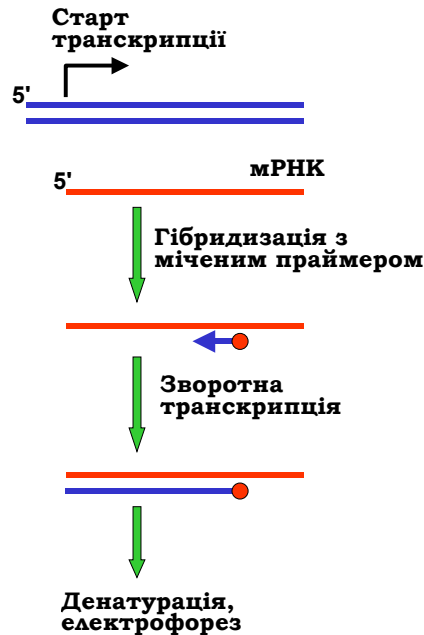


Рис. 11.13. Метод подовження праймера для визначення старту транскрипції

**Визначення транскрипційної активності *in vivo*.** Описані методи визначають рівень присутності мРНК певного типу в клітині. Цей рівень залежить не тільки від рівня транскрипційної активності, але й від швидкості деградації мРНК. Один із методів, який дозволяє встановити, чи відбувається транскрипція в даний момент, полягає в тому, що клітинні ядра виділяються і до них додаються мічені нуклеозидтрифосфати. Нові акти ініціації транскрипції вже не відбуваються в таких відокремлених ядрах, і РНК-полімераза просто продовжує добудовувати транскрипти, синтез яких розпочався до виділення. Зрозуміло, що такі подовжені транскрипти виявляються міченими. Після виділення РНК їх треба ідентифікувати, що можна зробити шляхом гібридизації з ДНК генів, які цікавлять дослідника.

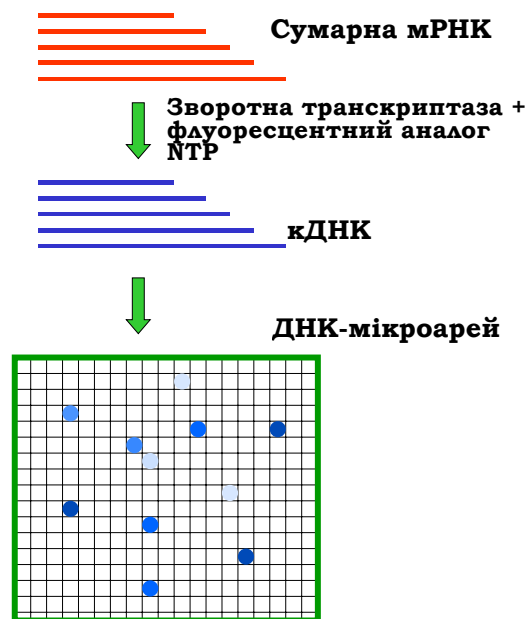
### Аналіз експресії геному

Проаналізувати повну програму транскрипції організму чи клітин певного типу за певних фізіологічних умов або у процесі розвитку, а також вирішувати інші завдання, пов'язані з вивченням функціонування цілого геному, дозволяють методи, що базуються

на використанні так званих ДНК-мікроареїв (DNA-microarrays) або ДНК-чипів (DNA-microchips).

Фрагмент ДНК довжиною до 1 тис. пар основ, для якого відомо його розташування в геномі, ампліфікується, і одноланцюгові продукти ампліфікації пришивають до невеликої зони на поверхні предметного скла мікроскопа. Скло розміром 2 × 2 см – ДНК-мікроарей – покрито сіткою ~6 тис. таких мікроплям, кожна з яких містить ДНК певної геномної ділянки. На поверхні ДНК-чипів синтезуються короткі (до 20 нуклеотидів) олігонуклеотиди: кілька олігонуклеотидів, які походять із однієї ділянки геному, синтезуються поряд. Таким чином, невелика зона містить ДНК, комплементарну індивідуальній геномній ділянці.

Одну з типових схем використання мікроарея зображено на рис. 11.14. Сумарна мРНК, отримана з клітин певного типу, використовується як матриця в реакції зворотної транскрипції. Поряд зі звичайними, до реакційної суміші додається флуоресцентний аналог одного з NTP.



**Рис. 11.14.** Аналіз сумарної мРНК за допомогою ДНК-мікроарея

У результаті отримують препарат флуоресцентно міченої кДНК. Після гібридизації з цією кДНК мікроарей аналізують за допомогою флуоресцентного мікроскопа: наявність флуоресцентної плями свідчить про активність певного гена, інтенсивність флуоресценції – про рівень цієї активності.

Експерименти такого типу дозволяють з'ясувати зміни загальної програми експресії генів при змінах зовнішніх умов, активність різних генів у різних тканинах багатоклітинного організму, зміни активності груп генів у процесі диференціювання клітин. Техніка ДНК-мікроарей широко використовується також у дослідженнях ДНК-білкових взаємодій.

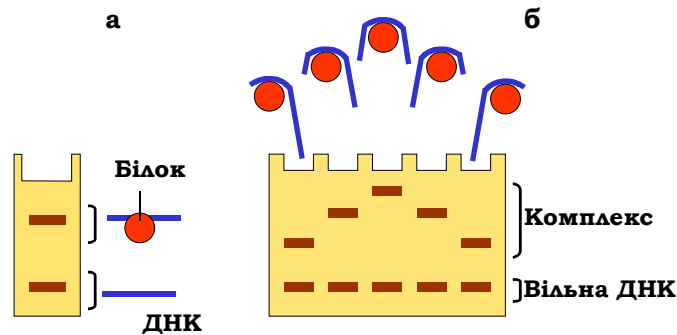
### **Методи дослідження ДНК-білкових взаємодій**

Вивчення тонких механізмів білково-нуклеїнових взаємодії є здебільшого доступним для різноманітних фізичних методів (див. нижче). У цьому підрозділі описано підходи, що дозволяють встановити сам факт взаємодії та з'ясувати локалізацію певного білка на ДНК *in vitro* та *in vivo*.

#### ***Гель-електрофорез білково-нуклеїнових комплексів***

Оскільки рухливість макромолекул при електрофорезі залежить від їхньої маси та форми, факт зв'язування певного білка з певним фрагментом ДНК легко встановити за допомогою шифт-тесту (shift assay). Принцип методу є дуже простим: ДНК-білковий комплекс (якщо він утворюється) суттєво гальмується в гелі порівняно з вільною ДНК, що приводить, як правило, до наявності двох смуг ДНК після електрофорезу (рис. 11.15, а).

При цьому, якщо білок, як це часто буває, вигинає ДНК у сайті взаємодії, рухливість комплексу буде залежати від позиції білка на фрагменті ДНК: максимальну рухливість мають комплекси, у складі яких білок розташований на одному з кінців фрагмента, мінімальну – ті, де білок займає центральну позицію (рис. 11.15, б). Причиною цієї закономірності є те, що довші (і тому рухливіші) кінцеві ділянки вільної ДНК легше проходять через пори гелю. Відповідно, це дозволяє оцінити локалізацію білка відносно кінців фрагмента ДНК або розділити комплекси, що розрізняють за такою локалізацією.



**Рис. 11.15.** Гальмування рухливості ДНК-білкового комплексу при гель-електрофорезі (а) та залежність ступеня гальмування від позиції білка на ДНК у випадку вигину ДНК на поверхні білка (б)

### Футпринтинг

Іншим підходом, який використовують для точної локалізації білка на фрагменті ДНК, є *футпринтинг* ДНК (DNA footprinting). Принцип методу схематично зображено на рис. 11.16. Кінець одного з ланцюгів мітять радіоактивною міткою. Зазвичай  $^{32}\text{P}$  у складі фосфатного залишку приєднується до попередньо дефосфорильованих 5'-кінців полінуклеотидкіназою, після чого один із кінців фрагмента відрізається рестриктазою (аналогічно відповідному етапу процедури на рис. 11.12). Якщо внести в цю ДНК обмежену кількість одноланцюгових розрізів (наприклад, один розріз на фрагмент у середньому), отримаємо випадковий набір мічених фрагментів різної довжини, який можна розділити за допомогою гель-електрофорезу в денатуруючих умовах. Як агенти, що вносять розриви, найчастіше використовують ДНКазу I або гідроксил-радикали: обидва агенти розрізають один з ланцюгів ДНК через маленький жолобок. У випадку, коли з фрагментом ДНК взаємодіє білок, ДНК виявляється захищеною в сайті взаємодії: на гелі після електрофорезу можна побачити зону, де відсутні смуги ДНК, або інтенсивність смуг є нижчою порівняно з контрольним результатом деградації вільної ДНК. Довжини верхньої та нижньої смуг, які оточують цю зону, є межами сайту взаємодії відносно мітки.

Аналогічний підхід у комбінації з блот-гібридизацією (так зване *непряме кінцеве мічення*) дозволяє встановити ступінь захищеності ДНК білками в конкретних геномних зонах. Першим етапом процедури є обмежена деградація ДНК нуклеазою у клітинних ядрах. Після цього

виділяють сумарну ДНК і обробляють її рестриктазою: у результаті препарат містить фрагменти ДНК від рестриктного сайту до місця розрізу нуклеазою. Усі ці фрагменти розділюють шляхом електрофорезу, здійснюють блотинг і обробляють нітроцелюлозний фільтр радіоактивно міченим зондом до певної ділянки генома. На авторадіограмі візуалізуються (або ні, якщо ДНК захищена білком) фрагменти ДНК з цієї зони.

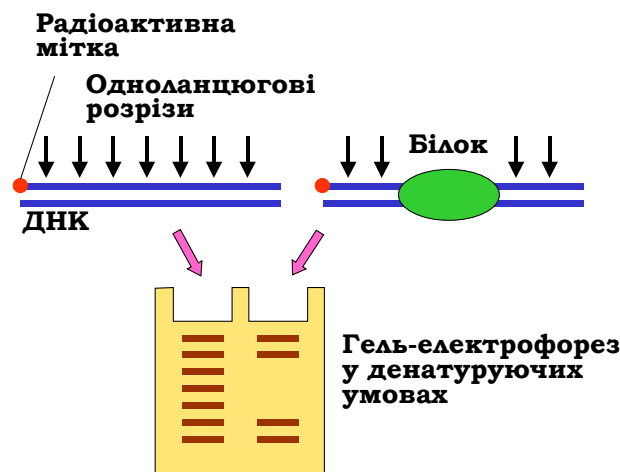


Рис. 11.16. Футпринтинг ДНК у складі білково-нуклеїнового комплексу

### Імунопреципітація хроматину

Метод імунопреципітації (ChIP – Chromatin ImmunoPrecipitation) є потужним засобом установлення наявності білка на певній ділянці ДНК *in vivo* (рис. 11.17). Для цього клітини обробляють формальдегідом із метою ковалентного пришивання всіх білків до ДНК. Далі здійснюють лізис клітин, виділяють ДНК і частково її фрагментують. Ковалентні ДНК-білкові комплекси осаджують антитілами, специфічними до конкретного білка (або пропускають через колонку, на якій іммобілізовані антитіла). Отримані комплекси руйнують, видаляючи зшивку хімічним шляхом, і аналізують ДНК, яка була зв'язаною у клітині з даним білком.

Одним із варіантів аналізу може бути ампліфікація ДНК за допомогою ПАР і наступне секвенування або гібридизація з міченим зондом. Якщо, наприклад, зробити зшивання в різні моменти після регуляторного сигналу, то можна встановити послідовність зв'язування білків з регуляторними ділянками промотора певного гена.

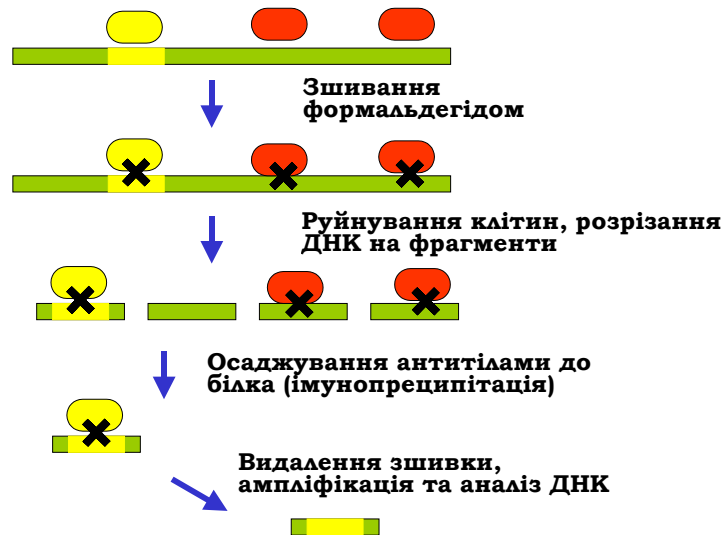


Рис. 11.17. Імунопреципітація хроматину

Якщо білок має численні сайти взаємодії в геномі, до всіх отриманих фрагментів ДНК пришивають промотор РНК-полімерази бактеріофага T7 і за допомогою полімерази здійснюють транскрипцію *in vitro* – синтез РНК-копій фрагментів. Далі використовують ці РНК як матриці для зворотної транскрипції у присутності флуоресцентного аналога одного з NTP і отримані флуоресцентно мічені кДНК гібридизують з ДНК-мікроареем, з'ясовуючи рівень присутності білка в різних ділянках геному.

### Методи дослідження протеому

Кінцевою точкою експресії спадкової інформації є набір білків. Аналіз складу цього набору (протеому), відносної кількості окремих білків і складної картини взаємодій між ними є одним із найактуальніших завдань.

Історично першим підходом (який зберігає своє значення) для аналізу складної білкової суміші став **двовимірний електрофорез** у поліакриламідному гелі. У цьому методі білкова суміш послідовно розганяється у двох взаємно перпендикулярних напрямках у різних електрофо-

ретичних буферах (з різним рН, іонним складом тощо). Часто в першому напрямку відбувається розділення білків за їхнім зарядом, у другому – за молекулярною масою (у присутності додецилсульфату натрію), що дозволяє ефективно розділити до 1 тис. білків одночасно.

Зручним і потужним методом розділення білкових сумішей є **високоєфективна рідинна хроматографія** (High Performance Liquid Chromatography – HPLC). Найчастіше розчин денатурованих білків наноситься на колонку, заповнену твердими пористими частинками, які адсорбують білки за рахунок гідрофобних взаємодій. Змивання (елюція) розчинником, який пропускається через колонку під високим тиском і полярність якого варіюється (наприклад, лінійно зростає), дозволяє розділити окремі білки, що незначною мірою розрізняються один від одного за своєю гідрофобністю. Метод використовують як для аналізу, так і з метою виділення та очистки білків.

Найпотужнішим засобом аналізу складних білкових сумішей є **мас-спектрометрія**, принцип якої полягає в розділенні пучка заряджених частинок в електричному та магнітному полі на окремі фракції з однаковим відношенням маси до заряду. Сучасні методи іонізації зразків дозволяють іонізувати й переводити в газову фазу великі органічні сполуки, у тому числі олігопептиди та білки. Наявність  $10^{-15}$  моль/л білка з молекулярною вагою до 200 тис. може бути детектована за допомогою сучасних мас-спектрометрів.

Проте визначення маси білка ще не дозволяє точно ідентифікувати його. З цією метою білок (після його виділення за допомогою двовимірного електрофорезу або рідинної хроматографії) піддають обмеженій протеолітичній деградації. Отримана комбінація олігопептидів, маса яких визначається за допомогою мас-спектрометрії, є своєрідним “відбитком” даного білка. Наступний аналіз геномних послідовностей дозволяє знайти відповідну амінокислотну послідовність білка, що практично позбавляє від необхідності хімічного визначення такої послідовності. З іншого боку, так звана тандемна мас-спектрометрія дозволяє отримати пряму інформацію щодо послідовності олігопептидів завдяки їхній вторинній фрагментації в камері спектрометра та наступного визначення розподілу отриманих фрагментів за масою. Поєднання мас-спектрометра з рідинним хроматографом – протеоліз білкової суміші, розділення олігопептидів методами рідинної хроматографії, аналіз їх за допомогою мас-спектрометрії та наступний пошук у базах даних з метою їхньої ідентифікації – це найперспективніший шлях розвитку кількісної протеоміки.



Наступним кроком у вивченні протеому є з'ясування складної картини білок-білкових взаємодій. Одним із найефективніших методів дослідження таких взаємодій є так званий **pull-down**: молекула певного білка, що використовується як пастка (bait), адсорбується на колонці (завдяки пришитому до білка специфічному “ярику” – tag), через колонку пропускається білкова суміш, утворені білкові комплекси аналізуються за допомогою гел-електрофорезу. Ідентифікацію білків, вилучених із комплексів і розділених шляхом електрофорезу, можна також здійснити за допомогою мас-спектрометрії. Цілком аналогічно до ДНК-мікроареїв, у дослідженнях білок-білкових взаємодій використовують також білкові мікроареї – скельця з пришитими до них білками (понад 5 тис. білків), які дозволяють здійснювати скринінг протеому.

### **Фізичні методи дослідження структури й активності біомакромолекул**

Широкого застосування в молекулярній біології набули різноманітні фізичні методи, важливість яких важко переоцінити. Детальний розгляд принципів, іноді досить складних, фізичних методів дослідження структури й функціонування біополімерів та їхніх комплексів виходить за рамки цієї книги. Нижче наведено перелік основних таких методів із зазначенням деяких особливостей і характеру інформації, яку можна отримати за їхньою допомогою.

#### **Методи безпосереднього спостереження**

Історично першим методом спостереження досить великих макромолекулярних комплексів із роздільною здатністю у нанометровому діапазоні стала **електронна мікроскопія**. При застосуванні цього методу зразок розміщують у вакуумній камері мікроскопа, опромінюють променем електронів, який далі проектується на флуоресцентний екран. Зразок при цьому має бути “зафарбований” певною речовиною, непрозорою для електронів.

Одна з поширених технік полягає в тому, що макромолекули наносять на карбонову підкладку, практично прозору для електронного променя (як і макромолекули), і фіксують на ній. Потім підкладку промивають розчином солі важкого металу, наприклад уранілацетатом. Після висушування така підкладка стає непрозорою для електронів скрізь, крім ділянки, де розміщена макромолекула, й уранілаце-

тат не зв'язався з підкладкою. Така техніка *негативного контрасту* дозволяє спостерігати під електронним мікроскопом молекули ДНК, комплекси ДНК із білками (такі як нуклеосоми), мультибілкові комплекси тощо. Розділення при негативному контрасті лімітоване розміром частинок важкого металу.

Важливим варіантом електронної мікроскопії, який дозволяє уникнути фіксації та контрастування зразка, є **кріоелектронна мікроскопія**. Тонкий шар (~100 нм) розчину макромолекул на підкладці швидко заморожується до  $-160\text{ }^{\circ}\text{C}$  і при цій температурі розміщується в охолодженій вакуумній камері мікроскопа. Виявляється, що такі зразки дають достатньо високий рівень контрасту без додаткової обробки. Кріоелектронна мікроскопія відіграла важливу роль у дослідженнях структури та механізмів функціонування рибосоми.

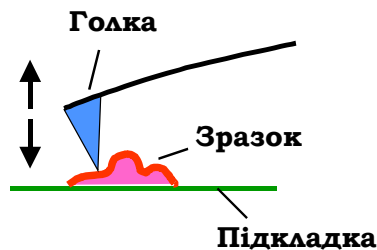


Рис. 11.18. Принцип мікроскопії атомних сил (AFM)

Сучасною альтернативою електронної мікроскопії є **мікроскопія атомних сил** (AFM – Atomic Force Microscopy). Макромолекули адсорбуються на підкладці, після чого зразок може бути дослідженим як після висушування, так і в разі, коли залишається шар рідини. Підкладка сканується голкою з надзвичайно тонким вістрям, яке за рахунок міжатомних (вандерваальсових) сил взаємодіє з поверхнею, відстежуючи її “ландшафт” (рис. 11.18). Рух голки реєструється оптичною системою (лазерний промінь, що віддзеркалюється від голки на фотодіод), і перетворюється комп'ютером у зображення поверхні. Роздільна здатність є порівнянною з такою електронного мікроскопа, однак, на відміну від електронної мікроскопії, AFM не потребує фіксації та контрастування зразка, дозволяє працювати з макромолекулами, що перебувають в умовах, близьких до фізіологічних. Крім того, існує можливість кількісної оцінки висоти зміщення голки відносно поверхні підкладки.

### *Рентгеноструктурний аналіз*

Головною технікою, яка дозволяє встановити просторову структуру макромолекули з роздільною здатністю до кількох ангстремів, є рентгеноструктурний аналіз. Об'єктом аналізу є кристал, у складі якого однакові макромолекули чи макромолекулярні комплекси утворюють впорядковану тривимірну решітку. Кристали, придатні для рентгеноструктурного аналізу (розміром щонайменше 0,3 мм), готують із перенасичених розчинів макромолекул, контролюючи швидкість зростання кристала за рахунок варіювання концентрації солі, температури, додаючи органічні розчинники тощо. У кристалах білків і нуклеїнових кислот, як правило, залишається значна кількість упорядкованих молекул води.

Вузький рентгенівський промінь буде частково розсіюватись на атомах кристала. Оскільки групи атомів періодично повторюються у кристалі, розсіяні промені будуть інтерферувати між собою, підсилюючи один одного в певних точках на детекторі, що реєструє рентгенівські промені після проходження їх крізь зразок. У результаті на детекторі формується впорядкований розподіл так званих рефлексів – точок високої інтенсивності.

Картина рентгенівського розсіювання містить інформацію щодо просторових позицій атомів у кристалі. Вилучення цієї інформації є досить складним математичним завданням, хоча сьогодні комп'ютерна техніка дозволяє аналізувати рентгенограми досить швидко: лімітуючим кроком рентгеноструктурного аналізу є практично вирощування кристалів. Безпосереднім результатом аналізу рентгенограми є складна тривимірна карта електронної щільності. Інтерпретація цієї карти є можливою за умови, якщо послідовність біополімеру в складі кристала є відомою. Шляхом численних підгонок на комп'ютері ця послідовність (чи послідовності кількох елементів досліджуваного комплексу) узгоджується з електронною картою, результатом чого є встановлення просторової координати кожного атома в макромолекулі. Саме таким шляхом отримано інформацію про структуру численних білків і міжмолекулярних комплексів, частина яких зображена на рисунках цієї книги.

Проте, як правило, великі макромолекули виявляються надто складними для того, щоб рентгенівське розсіювання на їхніх кристалах дозволило отримати тривимірну структуру. Для визначення структури більшості білків, нуклеїнових кислот і комплексів застосовують *метод ізоморфних похідних*, який потребує набору кристалів макромолекули:

вихідний кристал і принаймні два–три кристали, які були б ідентичними, але містили б атоми важких металів, що значно інтенсивніше розсіюють рентгенівське світло. Важкі атоми вводять у кристал або шляхом дифузії важких металів у нього, або використовуючи мутантні форми білка. В останньому випадку методами спрямованого мутагенезу здійснюють заміну однієї амінокислоти на Cys, до SH-групи якого можна легко приєднати атом важкого металу. Таким чином, розвиток методів рекомбінантної ДНК та експресії рекомбінантних білків став важливою передумовою структурних досліджень: завдяки рекомбінантним технологіям стало можливим не тільки отримувати достатню кількість високоочищених білків, придатних для подальшої кристалізації, а й готувати їхні ізоморфні похідні. Порівняння рентгенограм, отриманих від немодифікованого кристала та його похідних, дозволяє врешті-решт розрахувати структуру макромолекули.

Численні експериментальні результати свідчать, що структура макромолекули в кристалі, не тільки у своїх головних загальних рисах, а й у багатьох деталях, збігається зі структурою в розчині. Проте за рахунок сил кристалічної упаковки можливі деякі деформації просторової структури в кристалах. Крім того, кристал дає статичну картину молекулярної структури, не дозволяючи досліджувати конформаційну динаміку макромолекул за умов, наближених до фізіологічних. Тому важливим доповненням до рентгеноструктурного аналізу є різноманітні методи дослідження структури макромолекул у розчині.

### *Дослідження структури макромолекул у розчині*

Вивчити особливості конформаційної рухливості та взаємодії макромолекул у розчині дають змогу різноманітні спектральні методи. Загальний принцип будь-якої спектроскопії полягає в наступному: розчин опромінюється тим чи іншим типом електромагнітного випромінювання, яке *мінімально* взаємодіє зі зразком. Відповідно, жорстке іонізуюче випромінювання, яке здатне руйнувати ковалентні зв'язки, не є придатним для спектральних досліджень. Різні типи молекулярної спектроскопії використовують діапазони випромінювання, які характеризуються не дуже високою енергією – від ближнього ультрафіолету до високих радіочастот.

Як правило, лише деякі хімічні групи зразка, котрі можна розглядати як реперні, здатні до специфічної взаємодії з випромінюванням певного типу, і, що є ключовою обставиною, характер взаємодії залежить від мікрооточення цих груп. Певні параметри випромінювання,

які змінюються внаслідок його взаємодії з реперними групами, реєструються. Їхня інтерпретація дозволяє отримати інформацію про мікрооточення груп, тобто про структуру макромолекули або, частіше, її перебудови внаслідок зміни умов, взаємодій з іншими молекулами тощо.

Найчастіше для досліджень макромолекул застосовують спектрофотометрію, круговий дихроїзм, флуоресцентну спектроскопію та ядерний магнітний резонанс.

Спектрофотометр – прилад для вимірювання ефективності поглинання світла в ультрафіолетовій і видимій областях – є неодмінним атрибутом будь-якої молекулярно-біологічної лабораторії. У першу чергу, **спектрофотометрія** використовується з метою визначення концентрацій молекул: ефективність поглинання є пропорційною до концентрації. Але крім того, ефективність поглинання може залежати від мікрооточення хімічних груп, які відповідають за поглинання – хромофорів. Найяскравіше така залежність виявляється для нуклеїнових кислот: поглинання азотистих основ в ультрафіолеті значно знижується (так званий *гіпохромний ефект*) за умови реалізації між ними стекінг-взаємодій (див. розділ 3). Дослідження денатурації нуклеїнових кислот (плавлення подвійної спіралі) за допомогою протилежного, *гіперхромного*, ефекту (зростання в поглинанні при руйнуванні стекінг-взаємодій) дає інформацію про стабільність подвійних спіралей залежно від зовнішніх умов, нуклеотидного складу, іноді – особливостей послідовності тощо.

Спектроскопія **кругового дихроїзму** (КД) базується на різниці в поглинанні оптично активними молекулами (якими є, зокрема, амінокислоти та нуклеотиди) світла, що є поляризованим по колу ліворуч чи праворуч. Для амінокислот і нуклеотидів така різниця залежно від довжини хвилі (спектр КД) виявляється у ближній ультрафіолетовій зоні (220–300 нм). Форма спектра КД залежить від взаємної орієнтації однотипних оптично активних молекул (амінокислот чи нуклеотидів) у складі макромолекули. Відповідно, аналіз спектра дозволяє оцінити відносний вміст різних типів вторинної структури в молекулі білка, зафіксувати структурні перебудови білка, РНК чи ДНК.

Певні хромофори після поглинання світла (збудження хромофора з переходом електрона на вищий енергетичний рівень) здатні втрачати енергію збудження шляхом випромінювання. Випромінювання такого типу, яке називається флуоресценцією, є основою **флуоресцентної спектроскопії**. У складі біополімерів до флуоресценції в ультрафіолеті здатні дві амінокислоти – Трп і Туг. Однак широке використання флуоресцентної спектроскопії зумовлено, головним чином, великою кількістю синтетичних флуоресцентних зондів (які нековалентно взаємо-

діють із макромолекулами) і міток (які ковалентно пришиваються до макромолекул). Висока інтенсивність флуоресценції зондів і міток дозволяє, по-перше, детектувати молекули певного типу в невеликих кількостях: у процесі кількісного ПАР, у гелях після електрофорезу, при гібридизації з ДНК-мікроареями, у живих клітинах за допомогою флуоресцентного мікроскопа. По-друге, різноманітні параметри флуоресценції залежать від мікрооточення зондів, міток і природних білкових хромофорів, що дає багату інформацію про характер конформаційної рухливості макромолекул, структурні перебудови, взаємодію між макромолекулами тощо.

Вимірювання кінетики флуоресценції під час певного процесу (коли процес запускається шляхом швидкого змішування реагентів) дозволяє вивчати тонкі механізми каталітичної активності ферментів, білково-нуклеїнового впізнання, роботи молекулярних моторів. Наприклад, використання флуоресцентних похідних aa-тРНК дозволило досить детально з'ясувати послідовність стадій елонгаційного циклу рибосоми (див. рис. 8.21, 8.25).

Серед різних флуоресцентних підходів особливе місце посідає метод резонансного перенесення енергії (FRET – Fluorescence Resonance Energy Transfer). Перенесення енергії є можливим на відстані для певних пар здатних до флуоресценції хромофорів: як правило, використовують певні флуоресцентні мітки, які пришивають до конкретних хімічних груп макромолекули. Одна з міток є донором, інша – акцептором енергії. При збудженні флуоресценції донора перенесення енергії можна легко зафіксувати за зменшенням інтенсивності флуоресценції донора та / або зростанням флуоресценції акцептора. Оскільки ефективність перенесення залежить від відстані (для різних донорно-акцепторних пар відстань, на якій ефективність перенесення становить 50 %, варіює в межах ~2–5 нм), метод є ефективною молекулярною “лінійкою”, що дозволяє вимірювати відстані в нанометровому діапазоні та зміни в цих відстанях при структурних перебудовах.

Спектроскопія **ядерного магнітного резонансу** (ЯМР) використовує той факт, що деякі атомні ядра (найчастіше йдеться про ядра гідрогену) мають власний магнітний момент (*spin*), який орієнтується вздовж зовнішнього постійного магнітного поля. Електромагнітне випромінювання певної частоти в радіочастотному діапазоні здатне змінювати цю орієнтацію на протилежну: відбувається резонансне поглинання енергії, що приводить до зміни магнітних характеристик зразка, яку можна зафіксувати. Значення резонансної частоти залежить від того, до складу якої хімічної групи належить даний гідроген,

і якого типу хімічні групи містяться поруч. Тобто спектр ЯМР – залежність ефективності поглинання від частоти – несе інформацію про типи хімічних груп, їхнє взаємне розташування та ступінь їх рухливості у складі макромолекули.

Особливо цінною є техніка **двовимірного ЯМР**, коли сигнали резонансу детектують після складної комбінації кількох радіочастотних імпульсів і відображають як функцію двох координат (наприклад, у найпростішому варіанті – це частота та інтервал між двома імпульсами). На спектрі такого типу можна розрізнити сигнали гідрогенів різних амінокислот і виміряти зсуви цих сигналів, зумовлені іншими наближеними амінокислотами. Величина зсуву є мірою відстані між амінокислотами: порівняння цієї інформації з амінокислотою послідовністю дозволяє з'ясувати просторову структуру білка. Сьогодні двовимірний ЯМР – єдиний метод, за допомогою якого можна встановити структуру білка в розчині. На жаль, це можливо тільки для порівняно невеликих білків (до 200 амінокислот). Та оскільки результатом двовимірного ЯМР є набір дуже подібних одна до одної структур, що відображає конформаційну рухливість білка, і оскільки метод у принципі позбавлений можливих артефактів, пов'язаних із кристалічною упаковкою, двовимірний ЯМР є важливим доповненням до рентгеноструктурного аналізу. Нині банк даних білкових структур містить понад 36 тис. структур, отриманих методом рентгеноструктурного аналізу, і понад 6 тис. – методом двовимірного (чи останнім часом мультिवимірного) ЯМР.

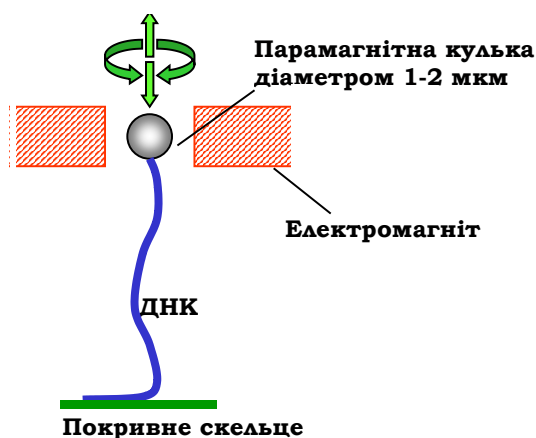
Ще одним важливим доповненням до експериментальних методів вивчення структури макромолекул є комп'ютерні розрахунки в рамках підходу **моделювання молекулярної динаміки**. Метод розрахунку базується на використанні другого закону Ньютона  $F = ma$ , де  $F$  – сила, що діє на частинку,  $m$  – її маса,  $a$  – прискорення руху частинки. У програму розрахунку вводяться дані про координати атомів макромолекули, отримані методами рентгеноструктурного аналізу чи ЯМР. Знаючи силу, яка діє на кожен атом (на основі між-атомних відстаней та енергії нековалентних взаємодій різних типів), можна обчислити прискорення кожного атома в системі: отримати систему диференціальних рівнянь руху. Наступне інтегрування цих рівнянь дозволяє обчислити залежності траєкторій і швидкостей руху атомів від часу. Таким чином, розрахунки молекулярної динаміки можуть дати детальну інформацію про флуктуації та конформаційні перетворення білків і нуклеїнових кислот. Застосування методу лімітується потужністю комп'ютерів. Та оскільки їхня швидкодія постійно зростає, метод молекулярної динаміки набуває все більшого

поширення. Напевно, найбільшим досягненням на сьогодні є детальний розрахунок на суперкомп'ютері конформаційних перебудов рибосоми під час акомодатії aa-tРНК в А-сайті.

### **Методи дослідження одиничних макромолекул**

Більшість методів дослідження (за винятком мікроскопії різних типів) передбачає роботу з досить великою кількістю – ансамблем – однакових молекул. Останнім часом розвиваються методи, що дозволяють слідкувати за поведінкою одиничних (звичайно, досить великих) макромолекул і макромолекулярних комплексів: ДНК та її комплексів з білками, хроматинової фібрили, актоміозинової фібрили тощо.

Одним із прикладів таких підходів є застосування магнітного пінцета (magnetic tweezers). У випадку ДНК (рис. 11.19), молекули (довжиною ~5 тис. пар основ чи більше) пришивають за один кінець до покривного скельця мікроскопа, за інший – до парамагнітної кульки ~2 мкм у діаметрі. Розмір кульки дозволяє слідкувати за нею (тобто за індивідуальною молекулою ДНК) за допомогою звичайного оптичного мікроскопа. Кулька знаходиться в полі електромагніта, який можна переміщувати або обертати, як показано на рис. 11.19, прикладаючи тим самим зовнішню силу до кінця молекули. Відповідь на зміну зовнішньої сили, що проявляється у відстані кульки від скельця (детектується відеокамерою), несе інформацію про еластичні властивості молекули, тобто про сили, що цю молекулу стабілізують.



**Рис. 11.19.** Принципова схема експерименту з ДНК за допомогою магнітного пінцета



За допомогою підходів такого типу отримано цінну інформацію, зокрема про жорсткість ДНК щодо розтягування та кручення, структурні особливості деяких білково-нуклеїнових комплексів, конформаційні перебудови нуклеосом, тонкі механізми роботи РНК-полімерази в реальному часі.

### **КОНТРОЛЬНІ ЗАПИТАННЯ**

1. У чому полягає процес клонування ДНК? Яким вимогам має відповідати плазмідний вектор для клонування і чому?
2. Що таке рестриктази та як їх використовують у рекомбінантних технологіях?
3. Поясніть принцип і охарактеризуйте основні типи гел-електрофорезу.
4. Що таке геномна бібліотека? Як створюються такі бібліотеки?
5. Дайте визначення кДНК. Як створюють бібліотеку клонів кДНК?
6. Що таке гібридизація? Як здійснюється гібридизація ДНК на нітроцелюлозних фільтрах?
7. Опишіть основні принципи полімеразної ланцюгової реакції.
8. Як здійснюють секвенування ДНК за Сангером?
9. Опишіть принципову схему експресії рекомбінантних білків у бактеріальних клітинах.
10. Що таке блот-гібридизація? Яка різниця між Саузерн-, нозерн- і вестерн-блотингом?
11. Як здійснюють фінгерпринтинг ДНК?
12. Яким чином можна визначити стартові та кінцеві точки транскрипції?
13. Як аналізують активність геному за допомогою ДНК-мікроареїв?
14. Що таке футпринтинг ДНК?
15. Як і з якою метою здійснюють імунопреципітацію хроматину?
16. Охарактеризуйте основні фізичні методи дослідження біологічних макромолекул.
17. Дайте визначення гіперхромного ефекту. Як його використовують?
18. Яким чином можна з'ясувати просторову структуру макромолекули з високою роздільною здатністю?

**РЕКОМЕНДОВАНА ЛІТЕРАТУРА**

- Глик, Б., Пастернак, Дж.* Молекулярная биотехнология. Принципы и применение. – М. : Мир, 2002.
- Кантор, Ч., Шиммель, П.* Биофизическая химия : в 3 т.– М. : Мир, 1984.
- Остерман, Л.А.* Методы исследования белков и нуклеиновых кислот. – М. : МЦНМО, 2002.
- Branden, C.-I., Tooze, J.* Introduction to protein structure. – New York : Garland Science, 1999.
- Cheung, V.G., Morley, M., Aguilar, F. et al.* Making and reading micro-arrays // Nature Genetics Supplement. – 1999. – Vol. 21. – P. 15–19.
- Chiu, W.* Electron microscopy of frozen, hydrated biological specimens // Annu. Rev. Biophys. Biophys. Chem. – 1986. – Vol. 15.– P. 237–257.
- Cooke, R.M., Cambell, I.D.* Protein structure determination by nuclear magnetic resonance // Bioessays. – 1988. – Vol. 8. – P. 52–56.
- Gerstein, M., Lan, N., Jansen, R.* Proteomics: integrating interactomes // Science. – 2002. – Vol. 295. – P. 284–287.
- International Human Genome Sequencing Consortium.* Initial sequencing and analysis of the human genome // Nature. – 2001. – Vol. 409. – P. 860–921.
- Leibowitz, M. J., Barone, F. P., Georgopoulos, D. E.* In vitro protein synthesis // Meth. Enzymol. – 1991. – Vol. 194. – P. 536–545.
- Lesk, A.M.* Introduction to bioinformatics. – New York : Oxford University Press, 2002.
- Lionnet, T., Dawid, A., Bigot, S. et al.* DNA mechanics as a tool to probe helicase and translocase activity // Nucl. Acids Res. – 2006. – Vol. 34. – P. 4232–4244.
- PCR technology: principle and applications for DNA amplification ; ed. H.Erlich. – New York : W.H.Freeman and Company, 1992.
- Sambrook, J., Fritsch, E.F., Maniatis, T.* Molecular cloning: a laboratory manual. – Cold Spring Harbor, New York : CSHL Press, 2001.
- Service, R.F.* Microchip arrays put DNA on the spot // Science. – 1998. – Vol. 282. – P. 396–399.
- Wolfsberg, T.G., Wetterstrand, K.A., Guyer, M.S. et al.* A user's guide to the human genome // Nature Genetics Supplement. – 2002. – Vol. 32. – P. 4–79.

## ПЕРСПЕКТИВИ МОЛЕКУЛЯРНОЇ БІОЛОГІЇ

У цій книзі розглянуто базові молекулярні механізми біологічних процесів. Подальше вивчення таких механізмів, залишаючись одним із найважливіших завдань сучасної молекулярної біології, далеко не вичерпує її змісту. Великі зусилля концентруються останнім часом на з'ясуванні загальної картини функціонування всієї сукупності біологічних макромолекул у клітині та цілісному організмі.

Першою дисципліною, що пов'язана з інтегральним підходом до молекулярних систем клітини, стала **геноміка**, яка вивчає структуру та функціонування цілих геномів. Роком її народження можна вважати 1990, коли у США було офіційно розпочато роботу за проектом “Геном людини” і створено міжнародний консорціум IHGSC (International Human Genome Sequencing Consortium). Через п'ять років був оприлюднений перший бактеріальний геном *Haemophilus influenzae*, через шість – перший еукаріотичний геном *Saccharomyces cerevisiae*, у 2001 р. було в основному завершено секвенування геному людини. Як відомо з бази даних GOLD, <http://www.genomesonline.org/gold.cgi>, за станом на травень 2008 р. повністю секвеновано геноми 660 бактерій (1769 у стані розробки) і 86 еукаріотичних організмів (943 у розробці).

Одним із найважливіших результатів робіт із секвенування геному людини стало суттєве вдосконалення методичних підходів геноміки. Нещодавно було здійснено секвенування першого індивідуального геному – геному Джеймса Уотсона, людини, яка стояла на початку молекулярної біології. Геном Уотсона було встановлено методом піросеквенування за два місяці, що коштувало менше 1 млн дол. (для порівняння: загальне фінансування проекту “Геном людини” з боку американського уряду склало приблизно 3 млрд дол. за 11 років). Швидкий розвиток сучасних технологій геноміки робить секвенування індивідуальних геномів кожної людини реальністю у близькому майбутньому.

Установлення геномних послідовностей – це лише перший крок системної молекулярної біології. Безпосереднім продуктом активно-

сті геному є сукупність транскриптів – **транскриптом**. З метою ретельного аналізу інформації, яка записана в геномі та реалізується внаслідок транскрипції, чотири роки тому Національним інститутом США з вивчення геному людини (National Human Genome Research Institute) було започатковано міжнародний проєкт ENCODE (Encyclopedia of DNA Elements), одним із основних завдань якого є тотальний аналіз транскриптому людини. Нещодавно було завершено перший етап цієї роботи – проаналізовано функціонування 1 % геному людини (~30 млн пар основ), і перші результати принесли низку несподіванок (The ENCODE Project Consortium // *Nature*. – 2007. – Vol. 447. – P. 799–816).

Частка ДНК, що піддається транскрипції, виявилася неочікувано високою – на рівні 80 %. Напевно, частина цих первинних транскриптів є просто наслідком неспецифічної хаотичної активності РНК-полімераз. Проте значна частка транскриптів містить консервативні (для ссавців і серед популяцій людини) елементи послідовності – приблизно 60 % консервативних нуклеотидів існує поза межами відомих раніше білкових генів чи регуляторних ділянок. Імовірно, у багатьох випадках ці транскрипти являють собою невідомі раніше РНК, які не піддаються трансляції, для більшості з них їхнє функціональне значення ще має бути з'ясовано. Крім того, виявилася, що часто піддаються транскрипції регуляторні ділянки (промотори, енхансери). Можливо, така транскрипція є просто одним із засобів підтримувати регуляторну ділянку в доступному стані деконденсованої хроматинової фібрили.

Для відомих білкових генів (399 у дослідженій частині геному) було продемонстровано наявність великої кількості невідомих раніше стартових точок транскрипції. Приблизно половина генів мають альтернативні стартові точки, що відстоять на 100 тис. пар основ від анованих раніше стартів транскрипції цих генів. Деякі з цих стартових точок використовують промотори інших генів: одна стартова точка може бути спільною для двох чи трьох генів, і первинний транскрипт іноді містить кілька генних локусів (груп екзонів).

Крім того, для більшості білкових генів аналіз їхніх транскриптів показує наявність невідомих раніше екзонів. Деякі з цих екзонів розташовані на відстані до кількох тисяч пар основ від усіх інших екзонів гена, іноді опиняючись у межах іншого гена, тобто більшим за

очікуване виявилася кількість різних ізоформ мРНК, котрі виникають як унаслідок альтернативного сплайсингу, так і транс-сплайсингу.

Результати проекту ENCODE вказують на диспергований характер розподілу регуляторних елементів по всьому геному: багато регуляторних елементів розташовано всередині екзонів та інтронів, при цьому вони можуть бути елементами системи регуляції зовсім іншого гена.

Загалом проект ENCODE продемонстрував наявність дуже складної “надбудови” над геномом у вигляді розгалуженої мережі транскриптів. Ретельне вивчення загальної картини регуляції цієї мережі залишається одним із важливих завдань майбутніх досліджень. Однак уже сьогодні зрозуміло, що основне для молекулярної біології поняття гена потребує певної ревізії. Висловлюється навіть думка про те, що воно взагалі є архаїчним, і одиницями спадковості є транскрипти. Проте таке судження слід визнати надто радикальним: фундаментальне уявлення про ДНК як сховище спадкової інформації зберігає свою силу. Відповідно до одного з нещодавно запропонованих визначень, “ген – це об'єднання (union) геномних послідовностей, що кодують зв'язний (coherent) набір функціональних продуктів, які можуть частково перекриватися” (Gerstein, M. *et al.* // *Genome Res.* – 2007. – Vol. 17. – P. 669–681). Під когерентністю набору продуктів мається на увазі, що у випадку білкових генів кожен екзон є спільним хоча б для двох продуктів цього набору. Згідно з цим визначенням, регуляторні елементи не є компонентами гена (пропонується називати їх “елементами, асоційованими з генами” – gene-associated): система регуляції є складнішою, ніж просте співвідношення “один до одного” між регуляторними елементами та кодуючими ділянками. Наведене визначення, звичайно, не може вважатися остаточним. Очевидно, що поняття гена є надто складним, щоб його можна було чітко сформулювати: різні визначення, які акцентують увагу на різних властивостях гена, мають право на існування і є, відповідно до принципу додатковості, одночасно справедливими.

Складна система регуляції транскрипції та сплайсингу створює специфічний для клітин кожного типу набір матриць для білкового синтезу. “Інвентаризацією” всіх білків клітини та організму займається **протеоміка**. Завдання протеоміки є значно складнішими по-

рівняно з геномікою: 20–30 тис. білкових генів створюють близько 100 тис. транскриптів (у середньому п'ять ізоформ на один ген за рахунок альтернативного сплайсингу), які у свою чергу породжують близько мільйона білків (кількість білкових ізоформ зростає за рахунок посттрансляційних модифікацій).

Першим завданням протеоміки є ідентифікація білків (визначення амінокислотної послідовності) за допомогою їхнього розділення шляхом двовимірного електрофорезу та наступного аналізу методами рідинної хроматографії, мас-спектрометрії та біоінформатики. Знання амінокислотної послідовності білка й порівняння цієї послідовності з відомими аналогами часто дозволяє зробити висновки щодо його структурних особливостей, присутності у складі білка стандартних структурних доменів, можливих взаємодій, в які вступає даний білок, функціональної ролі, яку він може виконувати.

Друге завдання протеоміки – визначення загальної картини взаємодій (**інтерактому**) між білками у складі протеому. Одним із найбільших досягнень у цьому напрямі є експериментальна ідентифікація кількох тисяч білок-білкових взаємодій у межах протеому *S. cerevisiae*. За сучасними оцінками, кожен білок протеому дріжджів вступає в середньому в п'ять таких взаємодій (не беручи до уваги гомотипічних взаємодій між однаковими білковими субодинаціями). Зважаючи на те, що білки вищих еукаріотів, зокрема людини, відрізняються в середньому більшою кількістю структурних доменів у своєму складі, мережа білок-білкових взаємодій, яка, власне, і визначає розмаїття біологічних функцій, у протеомах вище організованих організмів має бути ще складнішою.

Паралельно протеоміка спрямована на визначення просторової структури та, відповідно, функціонального значення кожного білка. Установлення структури кожного білка методами рентгеноструктурного аналізу чи двовимірного ЯМР є надзвичайно важким завданням. Тому великого значення для структурної протеоміки набувають методи теоретичного передбачення просторових структур (у тому числі такі, що використовують структурну інформацію, яка вже міститься в банках даних білкових структур). Сучасні методи передбачення просторової структури поки що не дозволяють установити структуру білка з такою самою надійністю, яку дають рентгеновська кристалографія та ЯМР. Але теоретичні методи, завдяки швидкому розвитку комп'ютерної техніки (зокрема виникненню

глобальної комп'ютерної мережі, що дозволяє об'єднати розрахункові можливості багатьох комп'ютерів по всьому світу – World Community Grid), уже починають відігравати важливу роль у створенні структурних анотацій цілих протеомів.

Нещодавно таку роботу було виконано для протеому *S. Cerevisiae* (Malmström et al. // PLoS Biol. – 2007. – Vol. 5(4). – P. e76 (www.plosbiology.org)). Дослідники розділили усі білки (понад 6 тис.) протеому на приблизно 15 тис. структурних доменів, і просторова структура доменів, котрі не мають гомологій до відомих просторових структур, була побудована теоретично за допомогою найкращих сучасних алгоритмів. Узгодження теоретичних результатів з експериментальними даними щодо функціонального значення певних білків, залучення їх до певних біологічних процесів, внутрішньоклітинної локалізації дозволило здійснити віднесення вивчених структурних доменів до родин еволюційно близьких білків. Таке передбачення просторових структур у масштабі протеому створює основу для подальшого вивчення механізмів функціонування великої кількості дріжджових білків.

Аналогічне завдання передбачення структури білків протеому людини здійснюється сьогодні в рамках відповідного міжнародного проекту (Human Proteome Folding Project). Дві основні категорії білків розглядаються як найважливіші у практичній площині. По-перше, це білки, послідовності яких закодовані в геномі людини і для яких невідомі структурні аналоги. По-друге, білки, закодовані в геномах патогенів людини – бактерій і вірусів, котрі викликають різноманітні захворювання. Саме білки, які часто відіграють ключову роль у розвитку патологій, можуть розглядатися як мішені для лікарських препаратів.

Розвиток структурної біоінформатики зумовляє появу нових можливостей комп'ютерної розробки специфічних інгібіторів активних центрів білків з метою створення нових ліків (drug design). Якщо просторова структура білка відома (як результат експериментального вивчення структури або теоретичного передбачення), можна провести моделювання взаємодії активного центру з малими лігандами: підібрати певну органічну сполуку, яка стерично відповідає активному центру й може вступати з ним у ефективні міжмолекулярні взаємодії. У більшості випадків проводять скринінг величезної кількості вже існуючих синтетичних сполук, відбирають

найкращі щодо спорідненості до білка-мішені, здійснюють конструювання хімічних похідних, які б мали підвищену спорідненість. Хімічний синтез, експериментальне вивчення взаємодій такого сконструйованого ліганду з білком і клінічні випробування завершують процес створення нового лікарського препарату.

Дизайн ліків є тільки одним із аспектів нових підходів у лікуванні захворювань, які можна об'єднати під назвою **молекулярної медицини**. Уже сьогодні практичне застосування молекулярної медицини є достатньо різноманітним.

У першу чергу це стосується молекулярної діагностики спадкових захворювань, у тому числі – до народження (пренатальна діагностика). Гени практично всіх спадкових захворювань уже відомі, методи їхнього визначення широко застосовуються в медицині з метою запобігти народженню хворої дитини. Методи молекулярної діагностики дозволяють виявити не тільки гени спадкових (моногенних) захворювань (гемофілія, муковисцидоз, міодистрофія Дюшенна, фенілкетонурія тощо), а й гени схильності до того чи іншого захворювання: хвороб, які розвиваються в більш похилому віці (хвороба Альцгеймера, рак молочної залози, нейродегенеративні хвороби), і таких, що виникають за дії певних зовнішніх факторів (діабет, атеросклероз, деякі онкологічні захворювання). Молекулярна діагностика дає можливість поставити діагноз задовго до появи симптомів і, відповідно, розпочати профілактику або лікування. Розвиток методів геноміки дозволить у майбутньому проводити тотальну “генетичну паспортизацію” з метою виявлення індивідуальних ризиків кожної людини.

Іншим напрямом молекулярної медицини є генна терапія спадкових захворювань. Основна ідея полягає в лікуванні патологій шляхом введення у клітини відповідних генетичних конструкцій. Зрозуміло, що тут виникає велика кількість проблем, однак сьогодні на стадії розробки існує кілька сотень проектів генної терапії, спрямованих на лікування онкологічних, інфекційних (СНІД, гепатит, туберкульоз) і моногенних захворювань.

Переважає більшість захворювань пов'язана не стільки з дефектами генів, скільки з порушеннями систем регуляції їхньої експресії, які є причиною змін у транскриптомі та протеомі. Установлення кореляцій між змінами в наборі білків і початком захворювання є також одним із завдань сучасної протеоміки.



Одним із прикладів своєрідного контрапункту, до якого сходяться інтереси фундаментальних молекулярно-біологічних досліджень і медицини, є **апоптоз** – програмована загибель клітин. Апоптоз – фізіологічний процес, він активується як внутрішньоклітинними, так і зовнішніми сигналами, які сприймаються рецепторними системами. Фізіологічне значення апоптозу полягає у видаленні дефектних клітин. Зовнішні ліганди (такі як FasL і фактор некрозу пухлин) взаємодіють із рецепторами клітинної мембрани (наприклад, Fas), що приводить до активації специфічних протеаз, задіяних у процесах апоптозу, – каспаз. На внутрішніх шляхах активації апоптозу як центральний елемент системи запуску програмованої загибелі використовуються мітохондрії: різноманітні стреси, накопичення пошкоджень ДНК, порушення систем внутрішньоклітинної регуляції активують проапоптичний білок Вах, який зв'язується з мітохондріальною мембраною та індукує визволення в цитозоль цитохрому с та інших мітохондріальних білків. Активність Вах контролюється кількома інгібіторами, зокрема білком Bcl2. Цитохром с утворює в цитозолі мультибілковий комплекс із кількома іншими білками – апоптосому, у складі якої відбувається активація так званої каспази 9, і, як наслідок, інших каспаз. Протеолітична активність каспаз спричинює загибель клітини. Паралельно при цьому активується фрагментація хроматину нуклеазами, апоптична клітина перетворюється на кілька апоптичних тілець, які фагоцитуються макрофагами. Важлива роль в індукції програми апоптозу належить білку р53: активація гена р53 відбувається у відповідь на накопичення пошкоджень у ДНК – якщо ефективність репараційних систем є недостатньою, то індукується апоптоз як механізм захисту організму від присутності пошкоджених, здатних до злоякісної трансформації клітин. Сам білок р53 є активатором транскрипції кількох проапоптичних генів, зокрема генів білків Вах, Fas та інших, продукти яких запускають апоптичну програму. Крім того, р53 активує транскрипцію р21 – інгібітора клітинного поділу, який сприяє зупинці клітинного циклу на стадії G1.

Наведений дуже поверхневий огляд молекулярних механізмів апоптозу демонструє складність системи його регуляції. Порушення цієї системи призводять до виникнення різноманітних захворювань, пов'язаних із підсиленням або гальмуванням апоптозу. Одним із під-

ходів у лікуванні таких захворювань є створення лікарських препаратів вибіркової дії, які б дозволяли спрямовано впливати на процеси проліферації та програмованої загибелі клітин. Найперспективнішими в цьому плані вважаються так звані антизмістовні олігонуклеотиди, які є комплементарними нуклеотидним послідовностям апоптоз-специфічних генів і діють за принципами РНК-інтерференції (на рівні транскрипції або трансляції). Наприклад, використання антизмістовних олігонуклеотидів, комплементарних певним ділянкам гена Bcl2, має привести до інгібування синтезу цього білка і, відповідно, до активації апоптозу цільових клітин-мішеней.

Успіх такої терапії залежить від способу доставки антизмістовного олігонуклеотиду (чи іншого препарату спрямованої дії) до цільових клітин. Останнім часом як вектори-переносники використовують певні вірусні системи та білки-переносники, а також наночастинки (карбонні нанотрубки, фулерени тощо). Інтенсивний розвиток нанотехнологій і прогрес у розумінні принципів функціонування біологічних молекулярних машин вже зумовили народження **біонанотехнології**.

Сьогодні біонанотехнологія робить лише перші кроки. Створення нанороботів, які б могли лікувати хвороби безпосередньо в місці, де розвиваються патологічні процеси, і репарувати пошкодження, що виникли в організмі, ще не є реальністю, хоча вже й не виглядає абсолютно фантастичним. Живі клітини мають тисячі працюючих наномашин, здатних перебудовувати матерію на атомному рівні: мотори молекулярного розміру, сенсори, транспортні й інформаційні пристрої та інші корисні механізми. Завданням біонанотехнології буде розвиток у бажаних напрямках властивостей таких машин шляхом їхньої модифікації та конструювання їхніх нових аналогів. Це дозволить створювати нові матеріали з точно контрольованою структурою на атомарному рівні, наночіпи та комп'ютерні пристрої мінімальних розмірів, прецизійні сенсорні системи. Біонанотехнології будуть насамперед застосовувати в медицині, адже біонаномашини найкраще працюють у живих клітинах. Молекулярні комплекси, здатні знаходити злоякісні клітини або здійснювати локальну лікувальну дію нині уже існують. Білкова інженерія вже сьогодні дозволяє створювати ферменти з підвищеною термостабільністю або зміненою специфічністю. Діагностичні біонаносенсори, молекулярні машини для здійснення терапевтичних завдань, інші медичні застосування біонанотехнології – справа недалекого майбутнього.

У найближче десятиліття слід очікувати остаточного розв'язання проблеми визначення просторової структури білків, виходячи з їхньої амінокислотної послідовності. Це дозволить не тільки передбачати структуру та функцію будь-якого білка, а й ефективно вдосконалювати природні білки, здійснювати дизайн нових, пристосованих під різноманітні потреби технології та медицини. Наступним етапом, напевно, стане клітинна інженерія. Уже тепер певним чином модифіковані клітини широко використовують для виробництва білків. На основі зростаючої інформації щодо організації геномів і протеомів, а також глибокого розуміння молекулярних механізмів взаємодій і функціонування біомакромолекул, з'явиться можливість модифікувати клітини під ширше коло прикладних завдань.

За трохи більше ніж 50 років свого розвитку молекулярна біологія стала однією із ключових галузей фундаментального природознавства та основою майже всіх біологічних наук. Саме на молекулярно-біологічних підходах базується сучасна біологія клітини, фізіологія, таксономія, антропологія тощо. Швидкий розвиток молекулярних біотехнологій і народження на наших очах біонанотехнології та молекулярної медицини робить молекулярну біологію однією з найважливіших прикладних дисциплін.

## **ПРЕДМЕТНИЙ ПОКАЖЧИК**

**А**зотисті основи 65-66,72  
Акомодація aa-тРНК 245-247  
Активний центр 49-50, 55-56  
Амінокислоти 33-35  
Антикодон 223-225  
Антитермінація 149-150  
АРСаза 225-229  
Атенуація 150-151  
Ацетилювання гістонів 121-123, 176-177, 189-192

**Б**актеріофаг  $\lambda$  152-155, 316, 334  
Бібліотека клонів 337-339  
Білково-нуклеїнове впізнання 88-92  
Блот-гібридизація  
– вестерн 349-350  
– нозерн 349  
– Саузерн 348-349  
Бромодомен 122-123, 177

**В**андерваальсові взаємодії 23-25, 42-43, 71  
Відкрита рамка зчитування 105  
Вільна енергія 13-18  
Водневий зв'язок 27-29, 37-40, 70-73, 90-91  
Вторинна структура 37-39

**Г**еліказа 146, 165-166, 284-286, 294, 302-304  
Ген 105-106  
Генетичний код 103-105, 225  
Гени домашнього господарства 110  
Геном 106-109  
Гетеродуплекс 310  
Гетерохроматин 127, 189-192, 196  
Гібридизація нуклеїнових кислот 338, 348-349  
Гідроліз АТФ 57-59

Гідрофобні взаємодії 29-30, 40-43, 71, 87  
Гістони 85, 115-125, 176-177, 296-298  
Гістоновий код 123  
Глобула 40-43  
Гормоновий рецептор 83-84, 174-175, 269

**Дезоксирибоза** 66-68  
ДНК-полімераза 279-284, 294-295  
Дуплекс 70

**Екзон** 109-110, 113, 204-205  
Екзонуклеаза 69, 279, 282  
Електростатичні взаємодії 25-26, 89-90  
Електрофорез 335-337, 353-355  
Елонгація  
– транскрипції 139-145, 186-188  
– трансляції 239-247  
Ендонуклеаза 69  
Енергія активації 22-23, 54-55  
Ентальпія 14-15  
Ентропія 15  
Енхансер 162  
Енхансосома 172

**Зворотна транскрипція** 114, 325-327

**Імунопреципітація хроматину** 355-356  
Ініціація  
– реплікації 292-293, 295-296  
– транскрипції 132, 136-139, 162, 167-170  
– трансляції 253-257  
Інтрон 109-110, 113, 204-205

**Каталіз** 54-56, 141, 207-208, 249  
кДНК 338  
Кеп 201-202  
Кількість зчеплень 93-96  
Кіназа 121-122, 166-167  
Клонування ДНК 331-335

Ковалентний зв'язок 19-23  
Кодон 103-105, 225, 243-245  
Комплементарність 70-72  
Константа рівноваги 14, 16-18  
Конформація 13, 20-22, 36-37  
Конформаційна рухливість 20-22, 51-53, 80, 91-92  
Кооперативність 153, 171-172  
Косміда 334-335

**Л**ейциновий зіпер 83-84  
Лігаза 152, 286-288

**М**аленька ядерна РНК 205-206  
Маленька ядерцева РНК 230  
Медіатор 166  
Мембранні білки 50-51  
Метилювання  
– ДНК 192-194  
– гістонів 121-122, 191-192  
Мінісателіти 114, 347  
Мікро-РНК 196, 261  
Мікросателіти 114  
Місметч 303  
Мобільні елементи 114-115, 323-327  
Молекулярна машина 59-62, 145, 207-208, 252-253  
мРНК 106, 212

**Н**адспіралізація 94-97  
Нуклеаза 69-70  
Нуклеозид 66  
Нуклеосома 115-121  
Нуклеотид 66-67

**О**перон 108, 147-151  
Ориджин 292-293, 295-296

**П**іримидини 65-66  
Пентоза 65-68  
Пептидил-трансферазний центр 232-233, 248-249

Пептидний зв'язок 35-37  
Плазміда 109, 331-332  
ПАР 339-341  
Поіаденілування мРНК 209-211  
Подвійна спіраль 70-75  
Послідовність  
– Козак 257  
– Шайна – Дальгарно 254-255  
Праймаза 284, 286-287, 290, 292, 294  
Преініціаторний комплекс 162, 166-167  
Промотор  
– бактеріальний 131-132, 137  
– еукаріотичний 162-164, 168-170, 173  
Процесивність 140, 283-284, 288  
Процесинг 201, 222, 262-263  
Псевдоген 111  
Пурини 65-66

#### **Райзинг 95**

Редагування  
– мРНК 217  
– помилок при  
– аміноацилюванні тРНК 229  
– реплікації 282-283  
– транскрипції 142-144  
Резольваза 313  
Рекомбінантна ДНК 332-333  
Рекомбінація  
– гомологічна 309-315  
– імуноглобулінових генів 319-323  
– незаконна 309, 318  
– сайт-специфічна 309, 316-318  
Ремоделювання хроматину 178-181  
Рентгеноструктурний аналіз 360-361  
Репарація  
– дволанцюгових розривів 306  
– ексцизійна 301-303  
– місметчів 303-304  
– пряма 300

Реплісома 284-286, 288-290  
Реплікон 276-278  
Репресор 147  
Рестриктаза 69, 331-332  
Ретропозони 114, 323, 325  
Рибоза 66-68  
Рибозим 208-209  
Рибосома 230-238  
РНК-інтерференція 194-197  
РНК-полімераза 131-136, 159-161  
Розплавлена глобула 42-43, 264  
рРНК 106, 235-237

**Сайленсер** 162  
Секвенування ДНК 341-343  
Сигнальна трансдукція 175  
 $\alpha$ -Спіраль 37-39  
Сплайсинг 109-110, 201, 204-209  
Сплайсосома 205-209  
Стекінг-взаємодії 71  
 $\beta$ -Структура 39-40  
Структура Холідея 310-315  
Структурні мотиви 44-48  
Структурні форми подвійної спіралі 76-78

**Тандемні повтори** 111-114  
Таутомерія 143-144  
Твіст (кут) 75  
Твіст (кількість витків) 93  
Теломераза 298-299  
Теломери 109, 190, 298  
Термінація  
– реплікації 291-293  
– транскрипції 145-146, 209-212  
– трансляції 258-259  
Топоізомер 94-95  
Топоізомерази 97-100, 291-292  
Топологія 93  
*Транс*-елемент 147



**Транслокація**

- ДНК-полімерази 281
- рибосоми 239-240, 249-252
- РНК-полімерази 142-143, 160-161

Транспептидація 239-240, 247-249

Транспозони 114, 323-324

Трансформація ДНК 331, 333-334

Третинна структура 40

тРНК 106, 221-225

**Фактори**

- процесингу мРНК 209-213
- транскрипції 147, 170-173
  - базальні 162-166, 168-170
  - елонгації 187
- трансляції 239-242, 249-250, 254-259

Фінгерпринтинг ДНК 347-348

Флуоресценція 362-363

Фосфодієфірний зв'язок 68

Фрагменти Оказаки 276-277, 286-287

Футпринтинг ДНК 354-355

**Хроматинова фібрила** 125-127

Хромодомен 123, 191

**Центромери** 109, 191

Цинковий палець 83-84

Циркулярна ДНК 93-97

*Цис*-елемент 147

Цукрофосфатний остов 68-69

**Четвертинна структура** 40

**Шаперон** 266-269

Шаперонін 270-272

**Ядерний матрикс** 127-128

ЯМР 363-364

**Alu-повтор** 114, 261  
ATP 57-62

**СpG-острівці** 110, 193  
CTD 161, 168, 201, 203

**DPE** 162-163

**GTP** 240

**HAT** 121  
HMG 85-86, 183-185  
HD 122  
HP1 191-192  
Hsp 174, 267  
HTH 81-82

**LINE** 114, 326-327  
LTR-ретропозони 114, 325

**MAR** 127

**SINE** 114, 327  
Swi/Snf 178

**TATA-бокс** 162  
TAF 163-164  
TBP 87, 163, 164  
TFIIB 164-165  
TFIID 163  
TFIIF 165-166

**Навчальне видання**

**СИВОЛОБ** Андрій Володимирович

# **МОЛЕКУЛЯРНА БІОЛОГІЯ**

Підручник

Редактор *В.Р. Філь*

Оригінал-макет виготовлено Видавничо-поліграфічним центром "Київський університет"



Підписано до друку 26.12.08. Формат 70x100<sup>1/16</sup>. Вид. № 17. Гарнітура Bookman Old Style. Папір офсетний.  
Друк офсетний. Наклад 200. Ум. друк. арк. 31,0. Обл.-вид. арк. 27,4. Зам. № 28-4683

Видавничо-поліграфічний центр "Київський університет"  
01601, Київ, б-р Т.Шевченка, 14, кімн. 43

☎ (38044) 239 32 22; (38044) 239 31 72; тел./факс (38044) 239 31 28  
Свідоцтво внесено до Державного реєстру ДК № 1103 від 31.10.02