Молекулярная биология для информатиков

- **Летняя школа по биоинформатике**

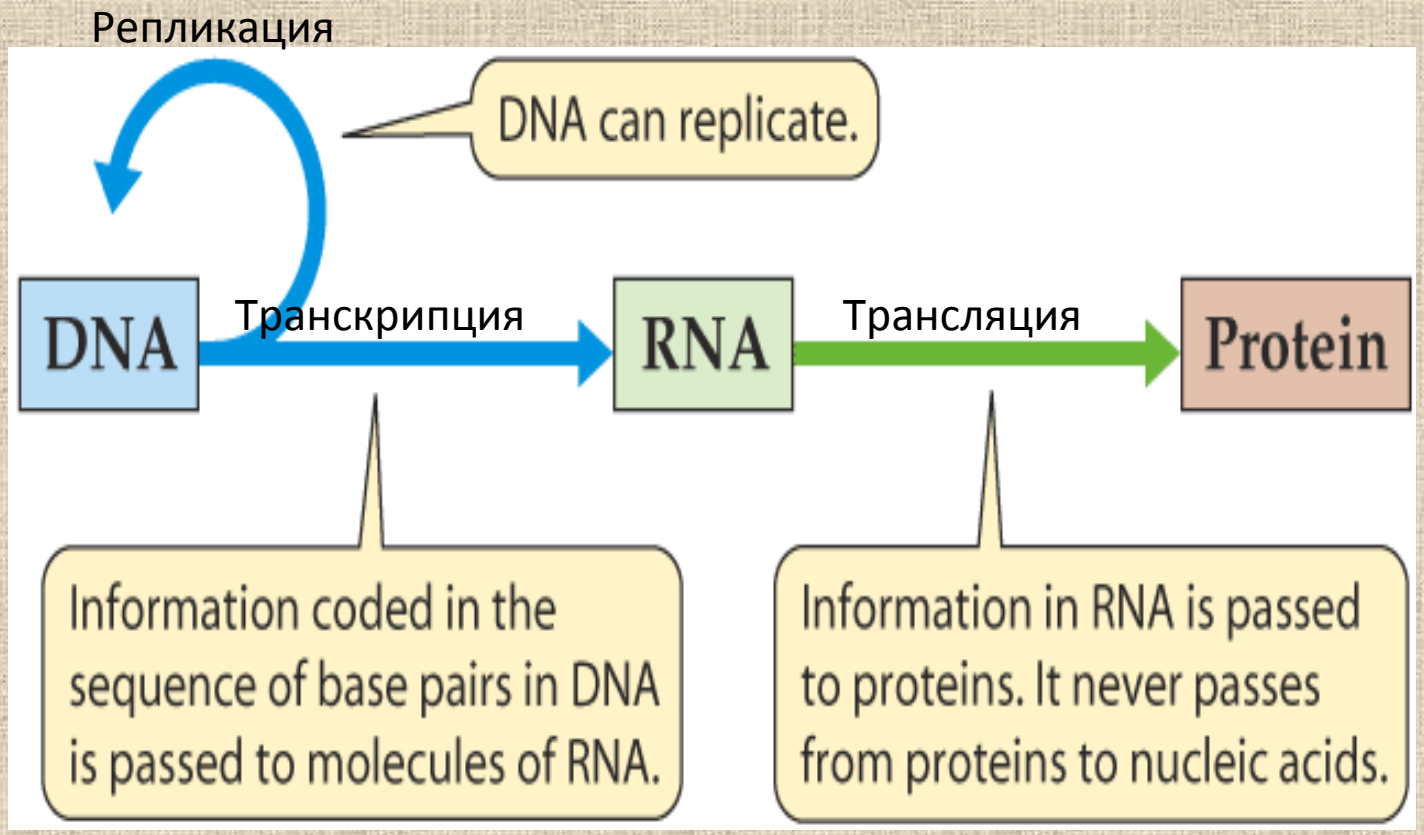
- Володина Наталья Яковлевна, Ph.D.

Академический Университет РАН, Алкор-био

Матричные процессы

- Репликация
- Транскрипция
- Трансляция

Перенос генетической информации

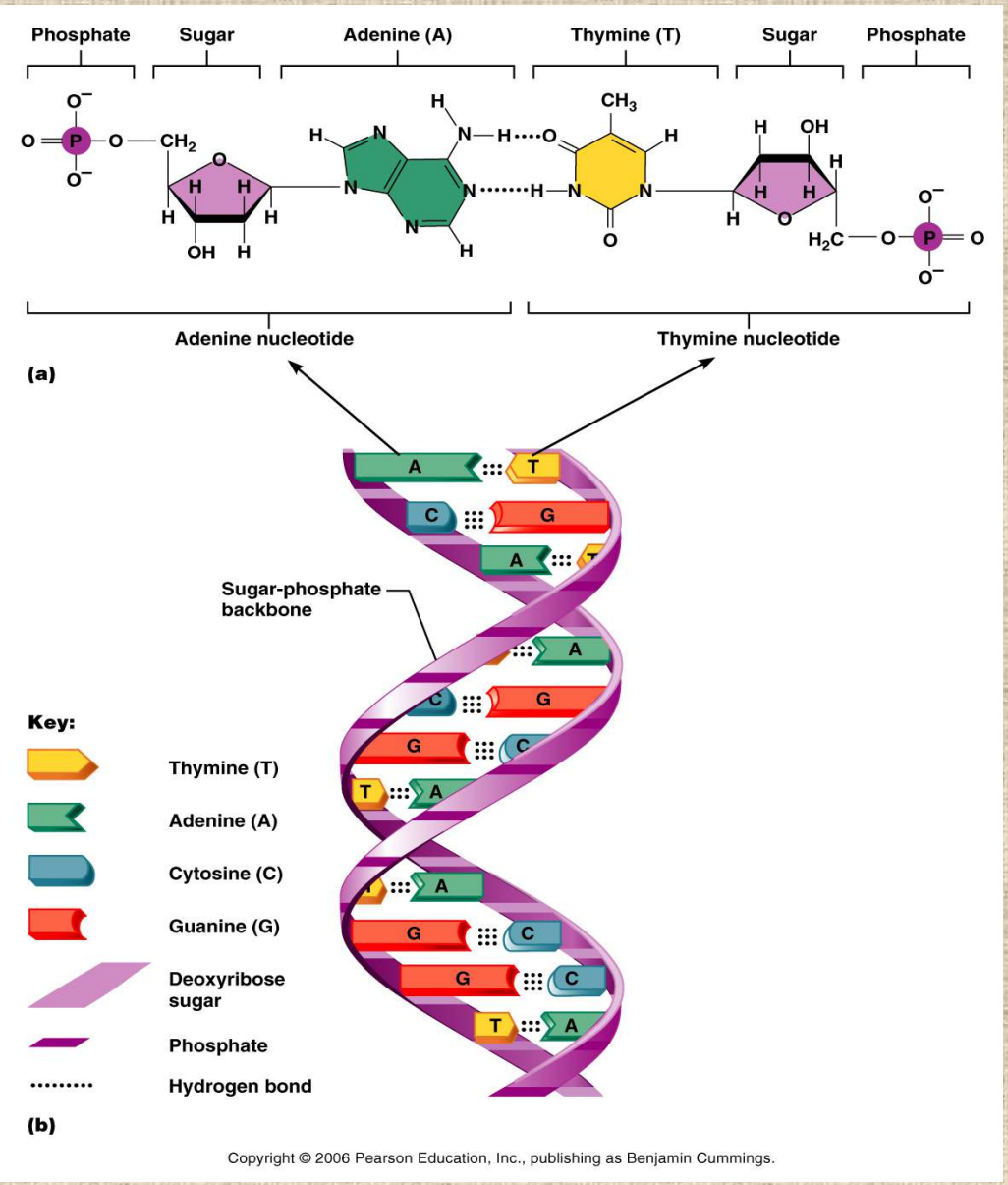


Центральная догма: ДНК-РНК-белок

Жизнь зависит от 3 молекул

- ДНК
  - Содержит информацию о работе клетки
- РНК (мРНК)
  - Передает эту информацию в различные части клетки
  - Служит матрицей для синтеза
- Белки
  - Образуют ферменты, катализирующие различные процессы в клетке
  - Образуют основные структурные компоненты живого организма

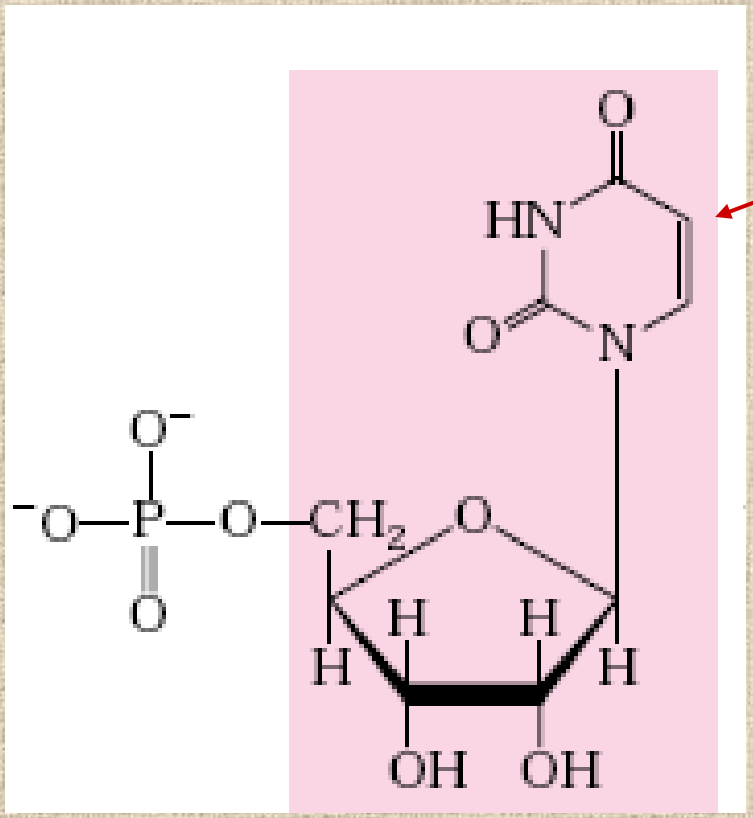
Структура ДНК



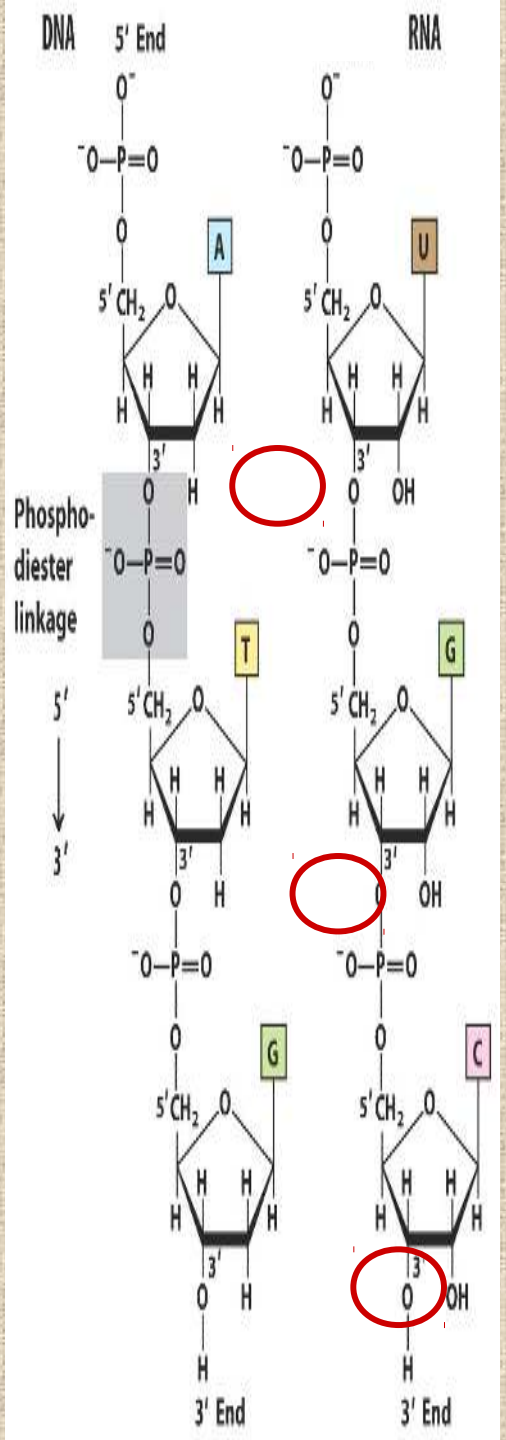
Copyright © 2006 Pearson Education, Inc., publishing as Benjamin Cummings.

Figure 2.22b

Полимерная структура РНК



Уридинмонофосфат



Виды РНК прокариот

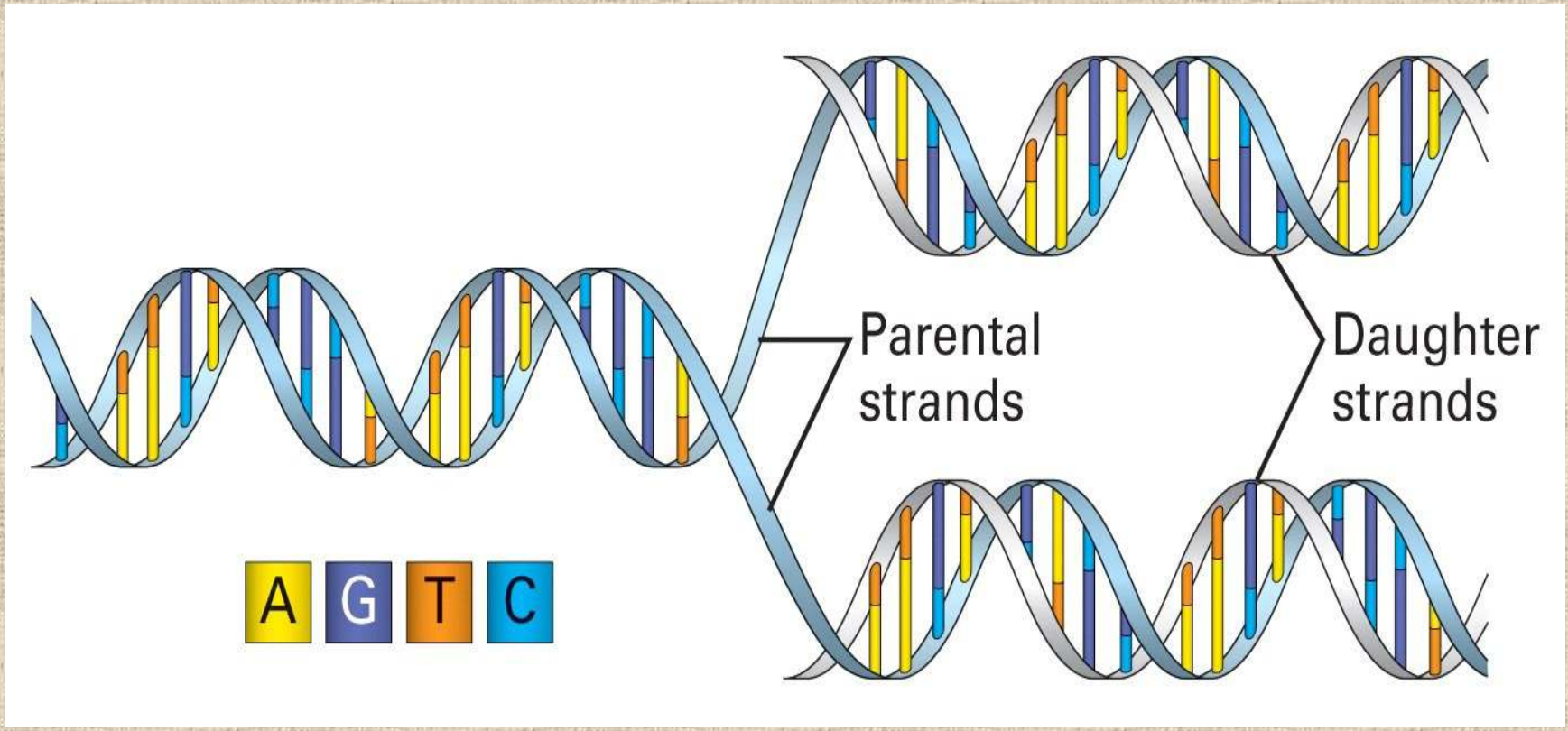
Матричные РНК (мРНК) – короткоживущие посредники в процессе переноса информации от ДНК к белку

- **Рибосомные РНК (рРНК)** – компонент рибосом и рибонуклеопротеиновых частиц
- **Транспортные РНК (тРНК)** – доставка аминокислот к рибосомам

Виды РНК ЭУКАРИОТ

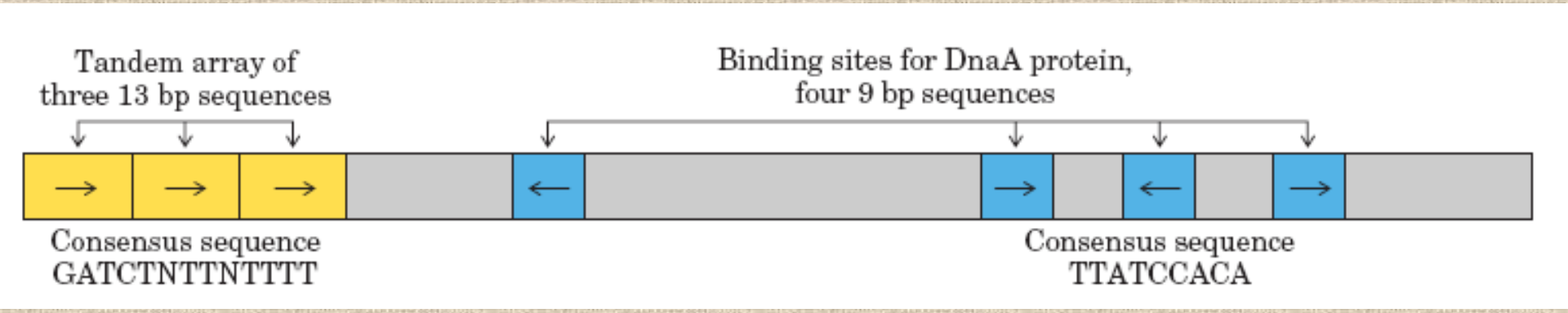
- **Гетероядерные РНК**– предшественники мРНК (mRNA)
- **Рибосомные РНК (rRNA) –**компонент рибосом
- **Транспортные (tRNA) –** переносят АК к месту синтеза пептидной цепи
- **Малые ядерные (snRNA)** – компоненты сплайсосомы и др.функции
- **Малые ядрышковые РНК (snoRNA)** – процессинг рРНК и тРНК
- **МикроРНК (miRNA)** – регуляторная роль
- **Малые интерферирующие РНК (siRNA)**,регуляторная роль

Репликация ДНК

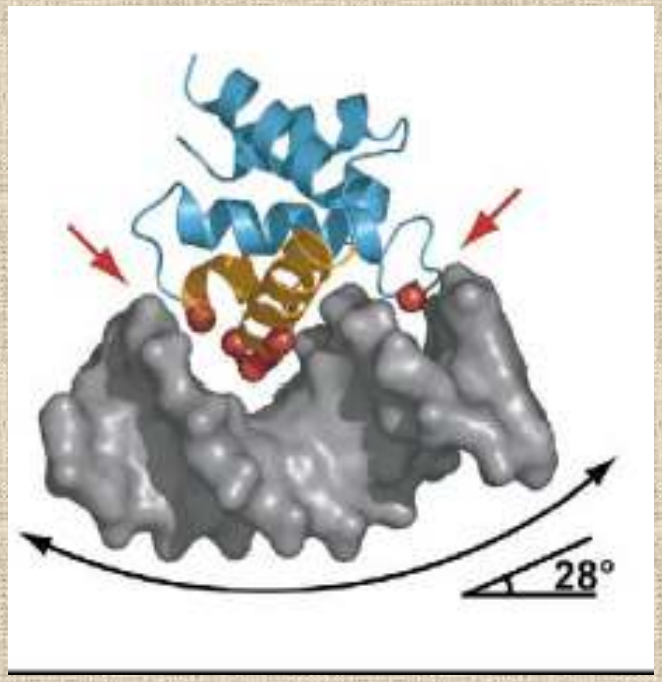


Инициация репликации (*E.coli*)

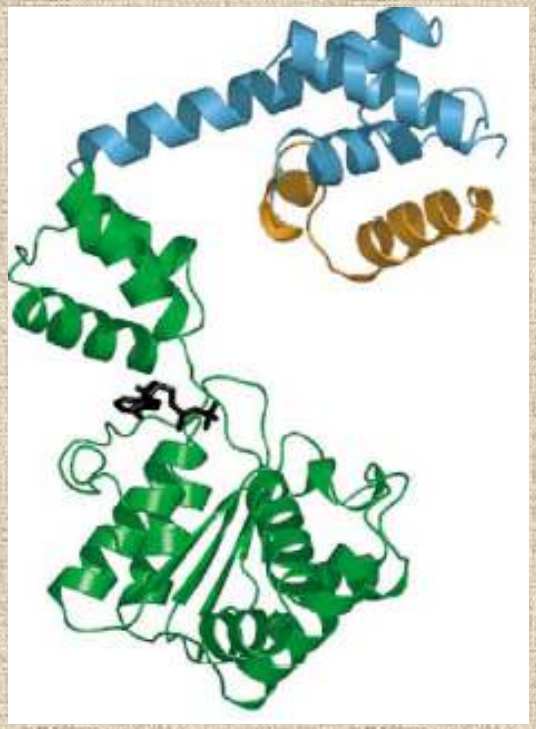
Структура ориджина репликации (ori) *E.coli*



Инициация репликации (*E.coli*)

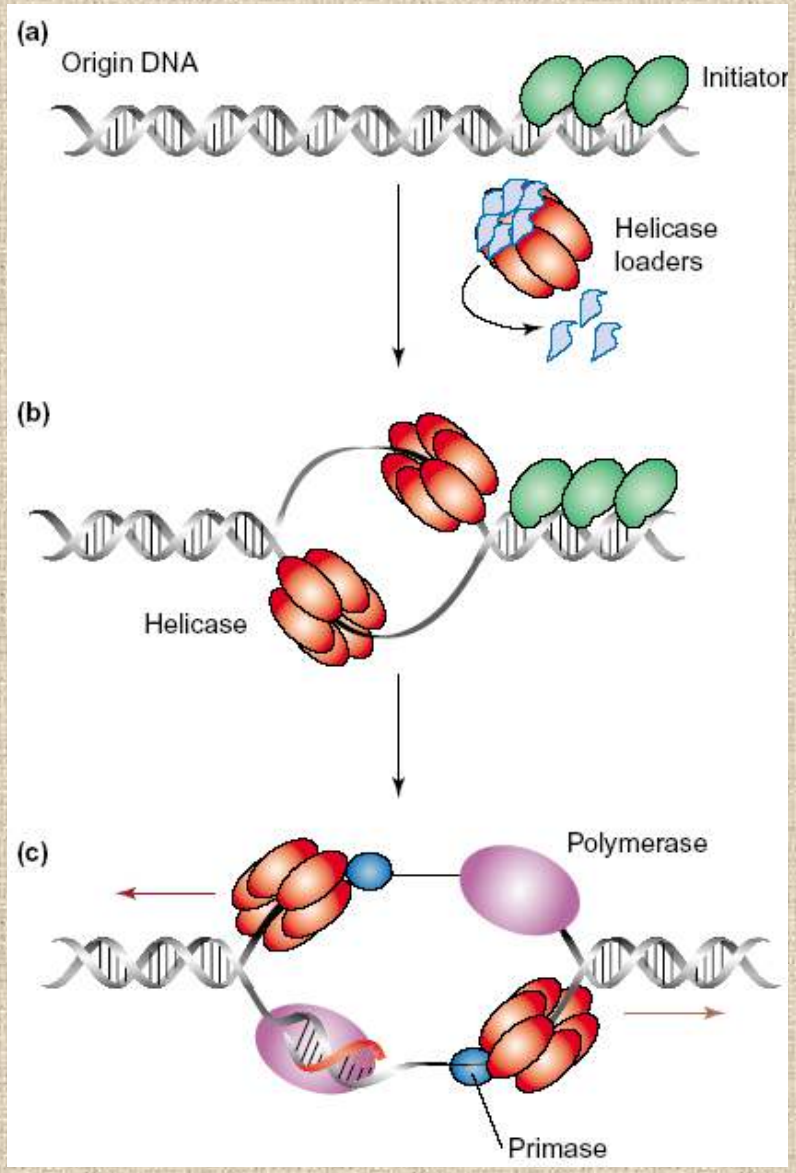


OriC:



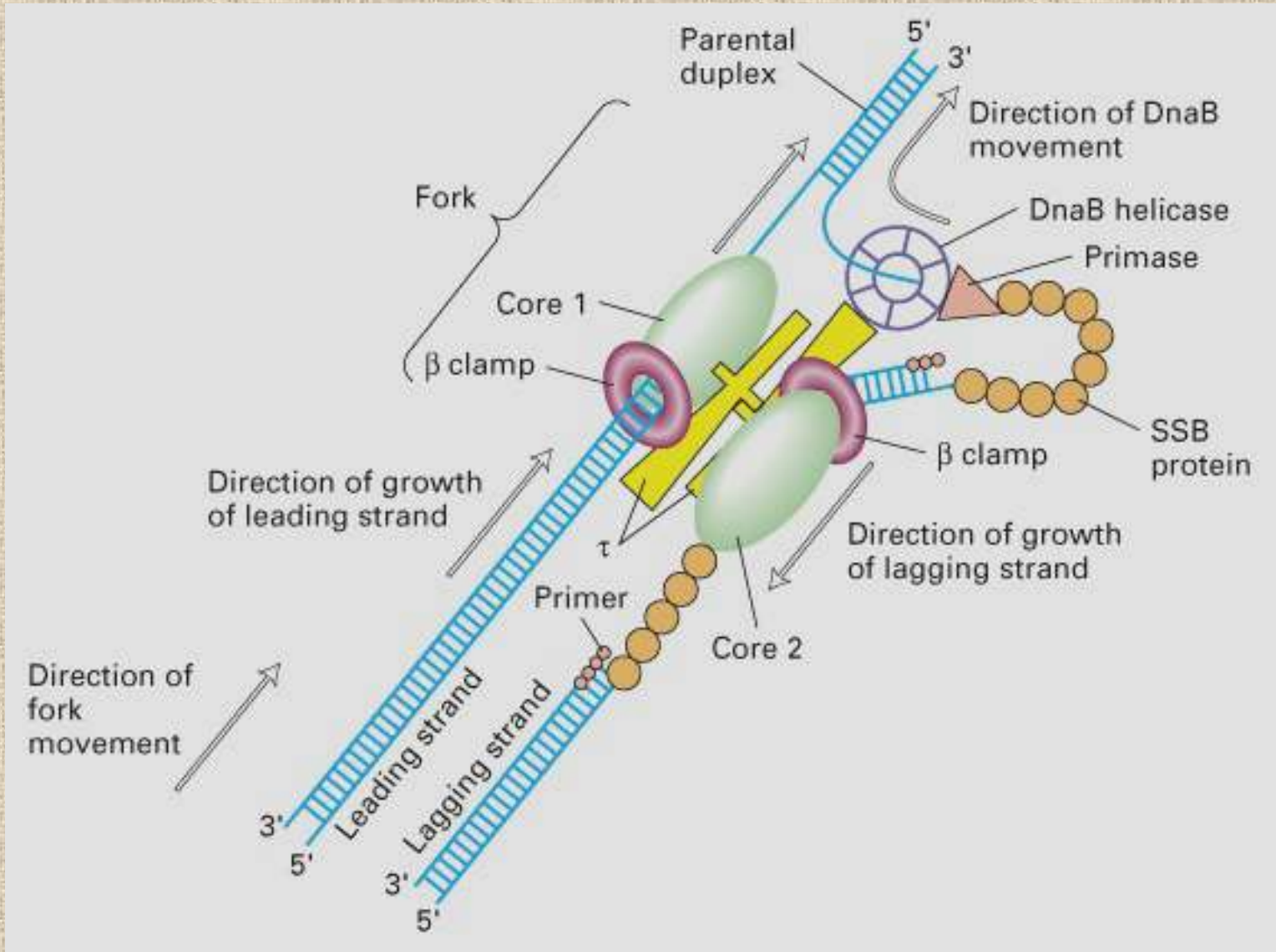
Связывание инициаторного белка dnaA с ДНК

Присоединение геликазы и праймазы

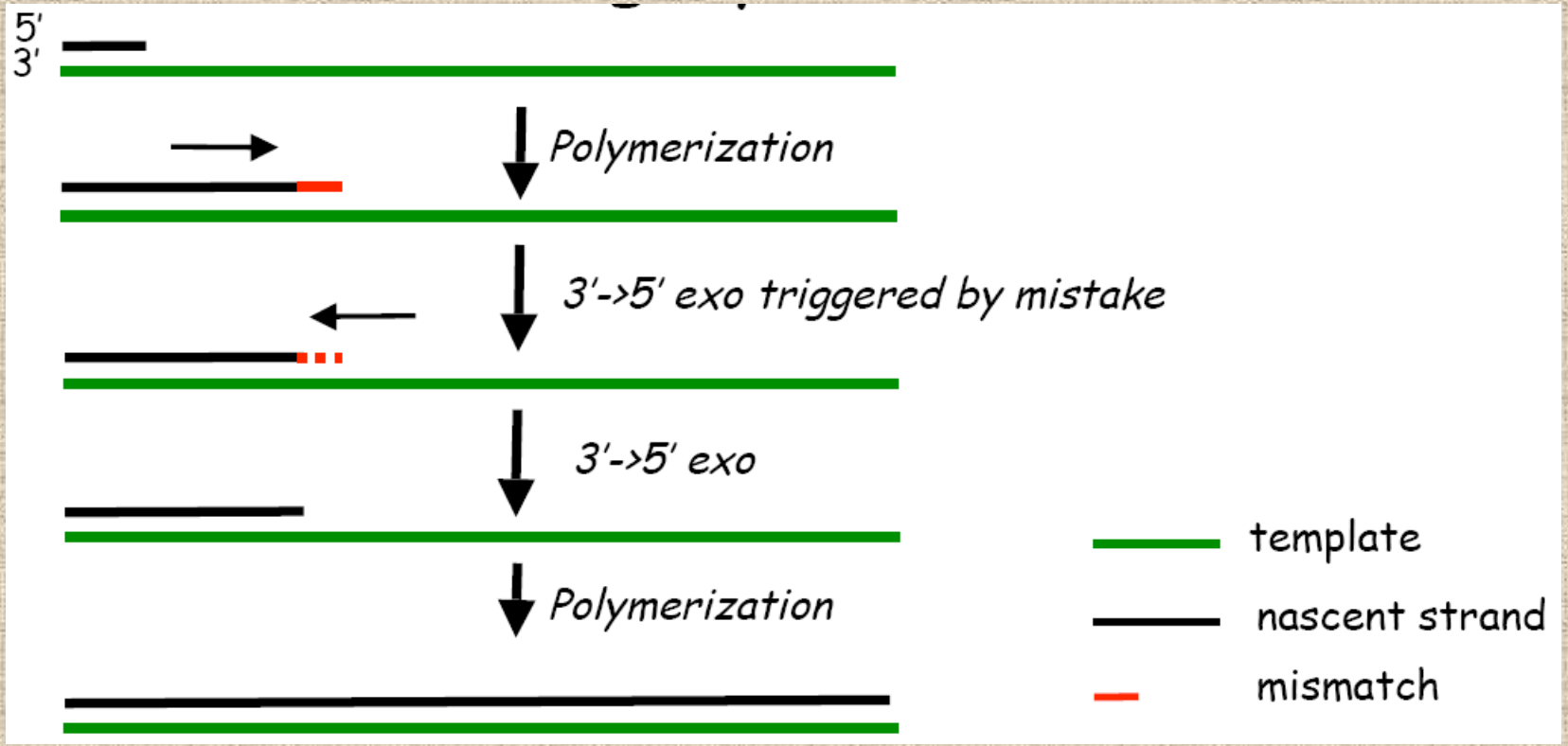


Общая схема репликации ДНК. Элонгация репликации

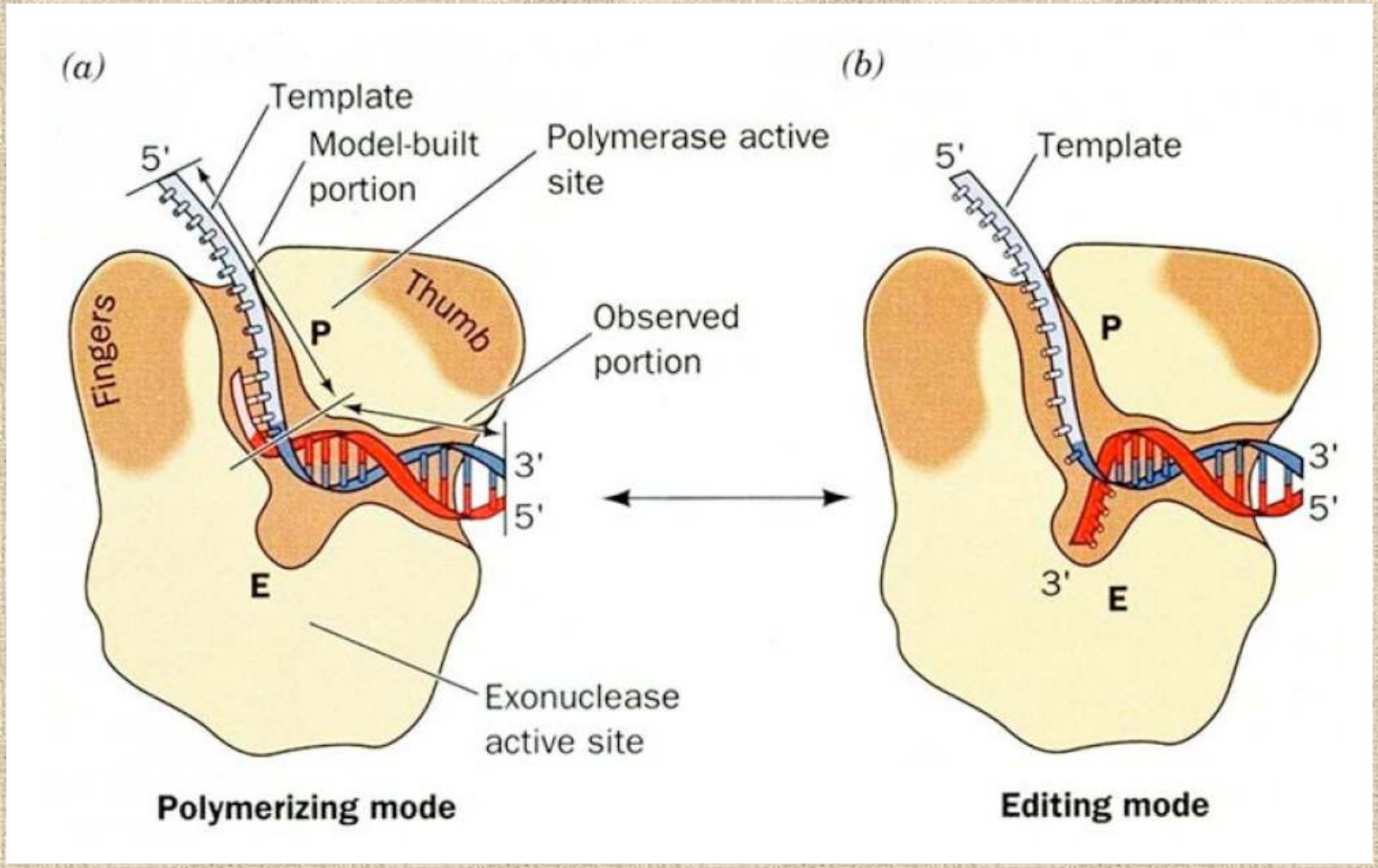
Как ДНК-полимераза перемещается к точке начала синтеза нового фрагмента Оказаки?

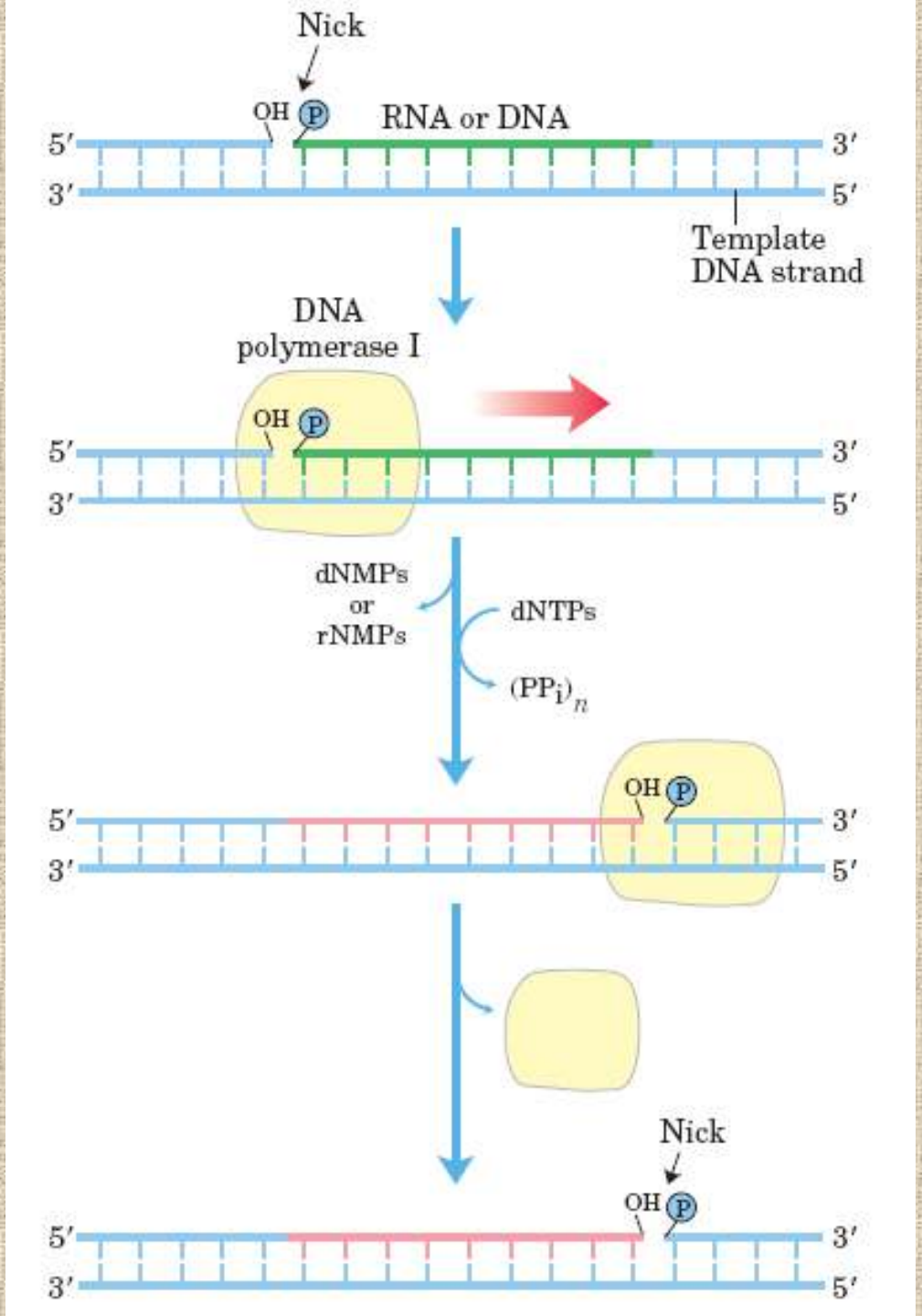


Коррекция ошибок ДНК-полимеразой III



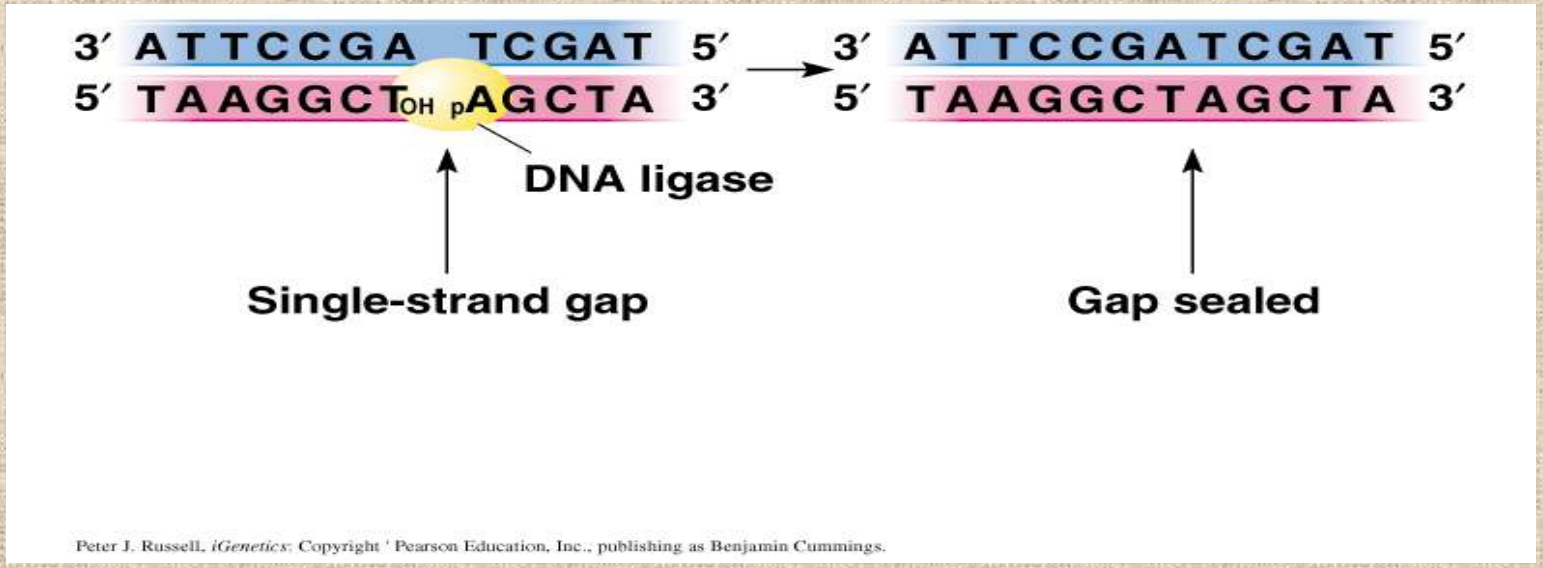
За корректирующую активность ответственна ε-субъединица ДНК-полимеразы III





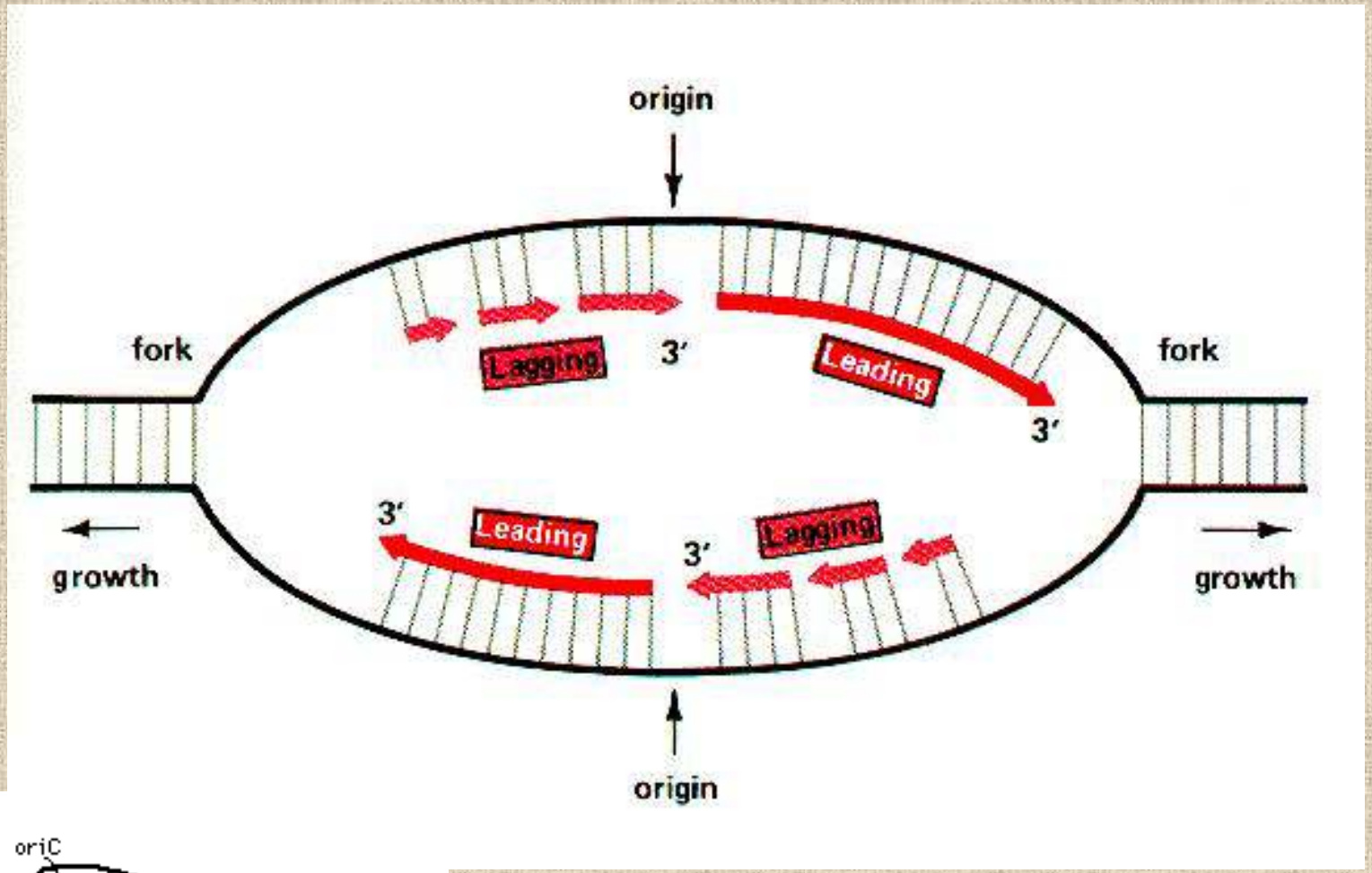
Удаление РНК-праймеров ДНК-полимеразой I

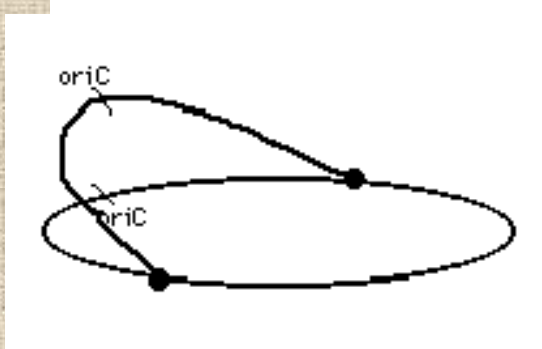
Соединение фрагментов Оказаки



Peter J. Russell, *iGenetics*: Copyright ' Pearson Education, Inc., publishing as Benjamin Cummings.

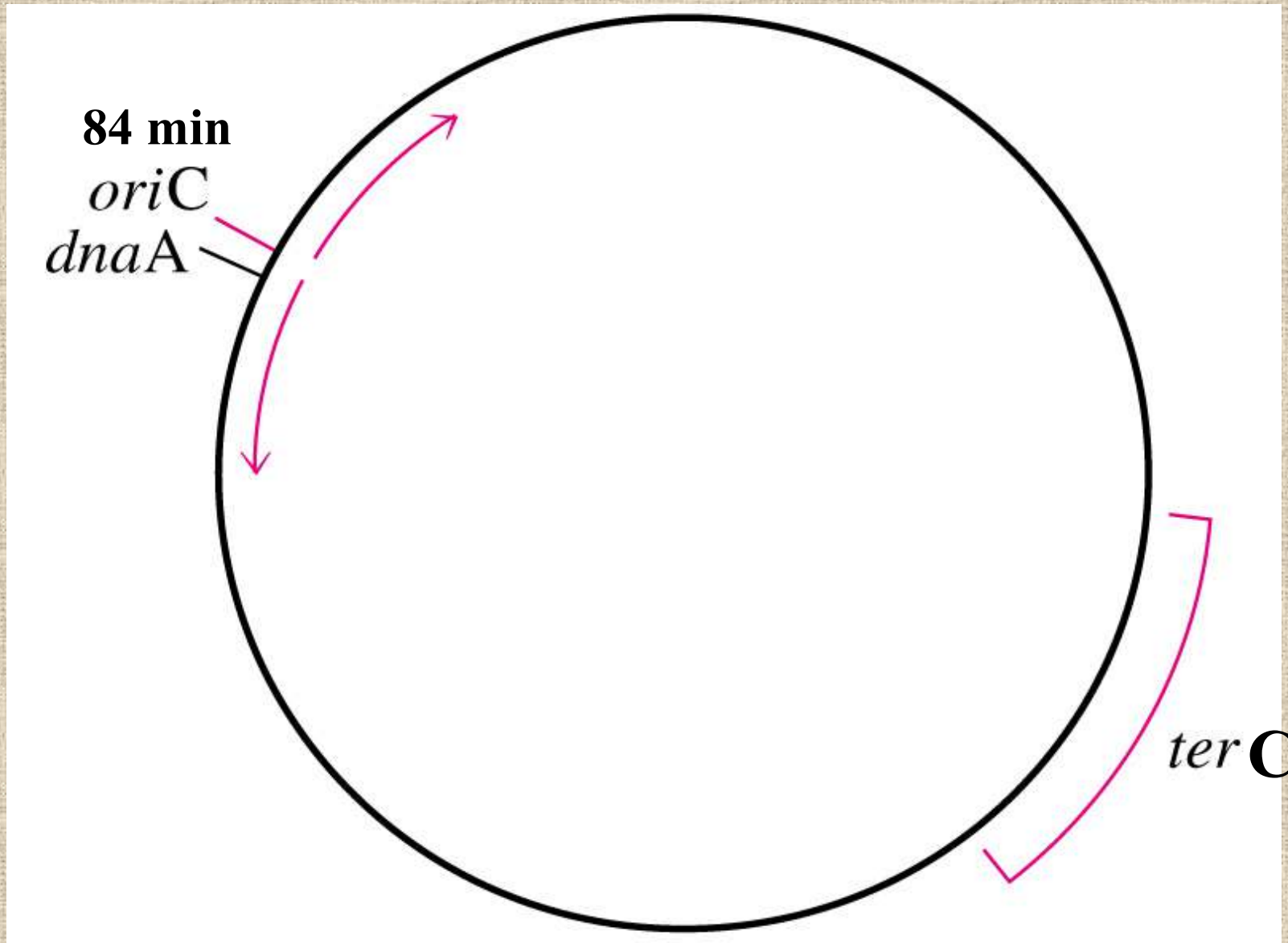
Терминация репликации



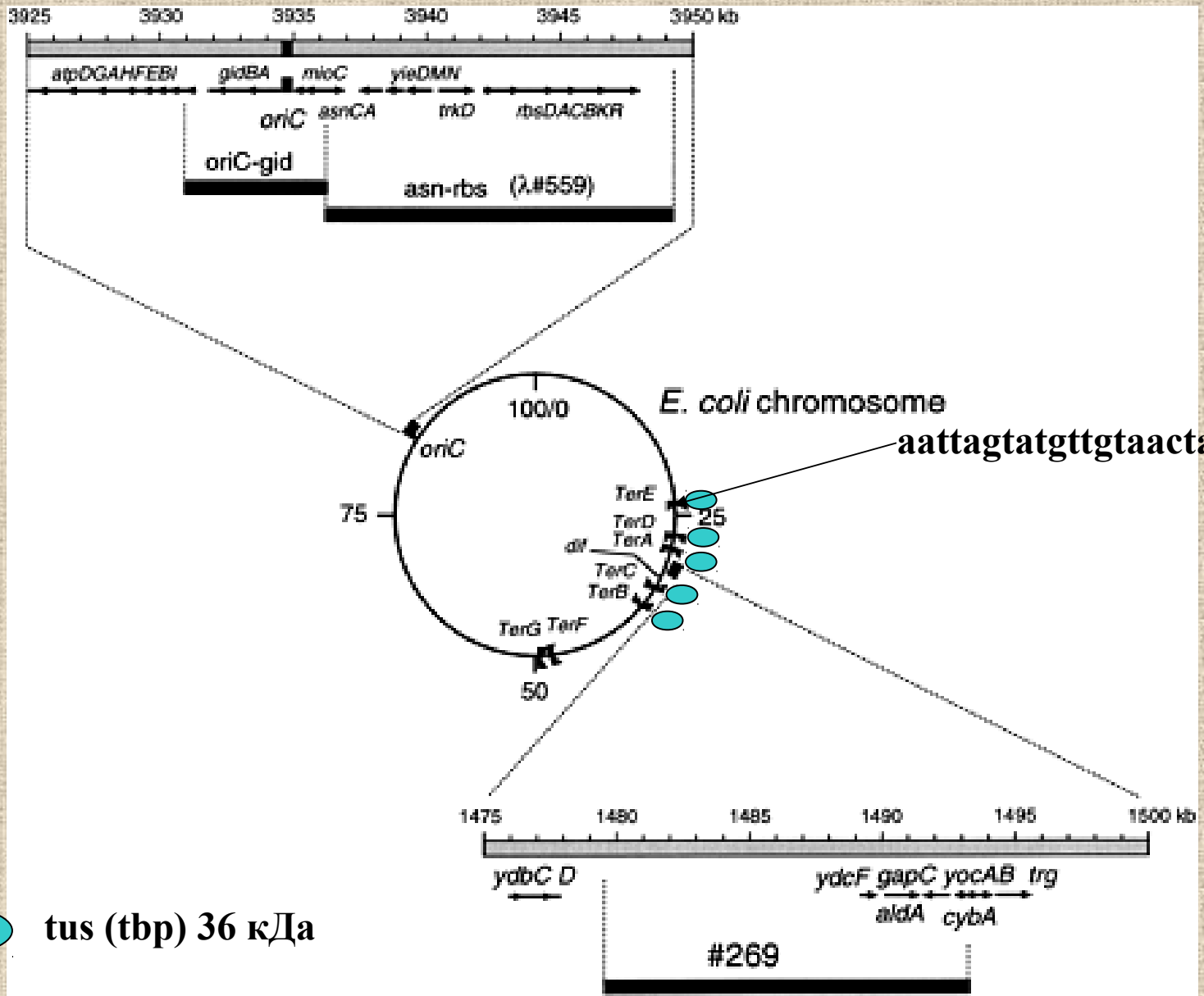


«Вилка репликации»

Хромосома *E. coli* (4,2 млн. нуклеотидов) – один репликон



Организация области терминации репликации у E. coli



tus (tbp) 36 кДа

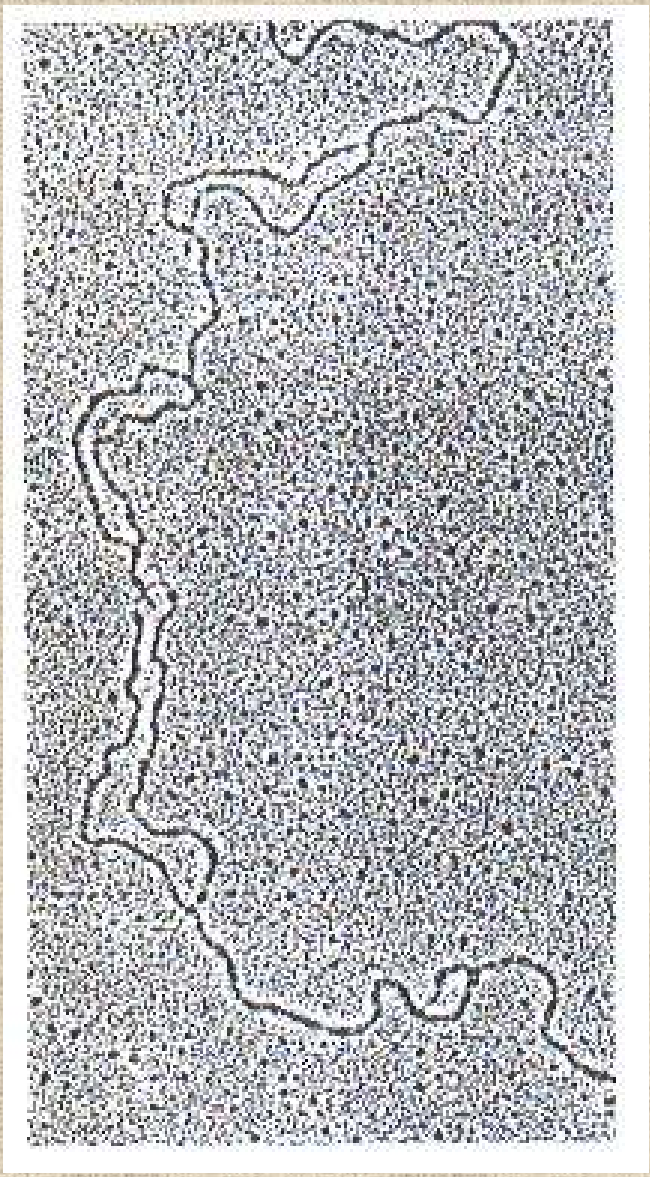
- **Репликация кольцевой ДНК у прокариот начинается в локусе OriC, заканчивается в области TerC.**
- **Инициатором репликации является белок DnaA.**
- **Репликацию ДНК осуществляет реплисома - динамичный мультисубъединичный белковый комплекс.**
- **Взаимодействие субъединиц ДНК-полимеразы III с другими белками реплисомы увеличивает скорость и процессивность фермента.**
- **Репликация происходит в направлении 5’⟶3’, одной цепи (лидирующей) неперерывно, другой (отстающей) – прерывисто.**
- **Репликация отстающей цепи включает дополнительные этапы: узнавание сайтов связывания праймазы, синтез праймеров, синтез фрагментов Оказаки, диссоциацию полимеразы, удаление праймеров, лигирование.**
- **Обе цепи реплицируются синхронно, что достигается взаимодействием ДНК-геликазы, ДНК-праймазы и ДНК-полимеразы и торможением синтеза лидирующей цепи во время синтеза праймеров.**

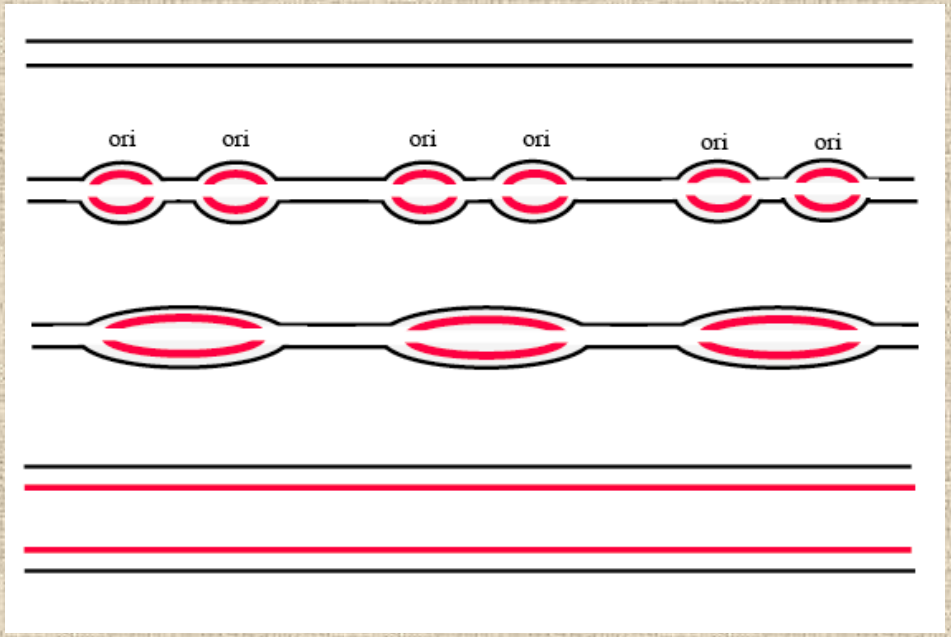
Эукариотические ДНК-полимеразы:

1. Polymerase α (alpha): nuclear, DNA replication, no proofreading
2. Polymerase β (beta): nuclear, DNA repair, no proofreading
3. Polymerase γ (gamma): mitochondria, DNA repl., proofreading
4. Polymerase δ (delta): nuclear, DNA replication, proofreading
5. Polymerase ε (epsilon): nuclear, DNA repair (?), proofreading

- Different polymerases for nucleus and mtDNA
- Some proofread; others do not.
- Some used for replication; others for repair.

Множественные ориджины эукариот





В ядрах млекопитающих при репликации обнаруживается до 150 центров репликации - «репликативных единиц», «репликативных фокусов» или «репликативных фабрик»

В центрах репликации аккумулируются все белки, участвующие в этом процессе.

Инициация репликации у эукариот

- **Возможно, ориджины высших эукариот не содержат консенсусных последовательностей.**
- **Функционирование ориджина может определяться контекстом, структурой хроматина, локализацией нуклеосом.**
- **Общая закономерность структуры инициаторных участков – обогащенность АТ-парами.**

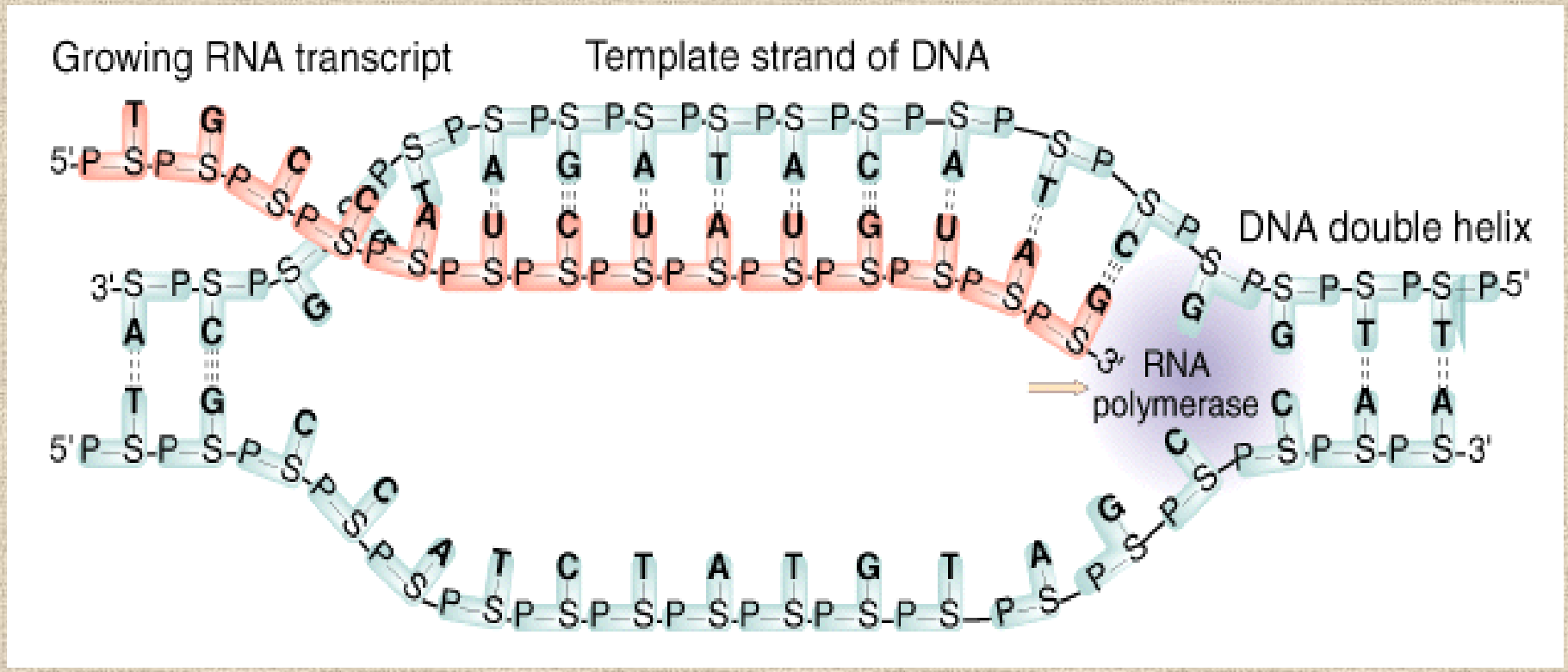
Элонгация репликации ДНК у эукариотических организмов

Replisome		
Polα primase holoenzyme	Pol1 Cdc17	Polymerase priming; replication of lagging strands
Polδ holoenzyme	Pol3 Cdc2	Bidirectional replication
Polε holoenzyme	Pol2 Cdc20	Bidirectional replication
Rfc1-5 complex		PCNA clamp loading
PCNA	Pol30	Polymerase clamp
RPA1-3 complex	SSB	SSB coats ssDNA

Транскрипция

- ДНК-зависимый синтез РНК – **транскрипция**
- Ферментативный аппарат транскрипции – **ДНК-зависимая РНК-полимераза**
- Этапы транскрипции:
- **инициация,**
- **элонгация,**
- **терминация**

Транскрипция



ДНК

A T G C

| | | | | | | |

U A C G

РНК

Матричный процесс

Образуются Уотсон-Криковские пары

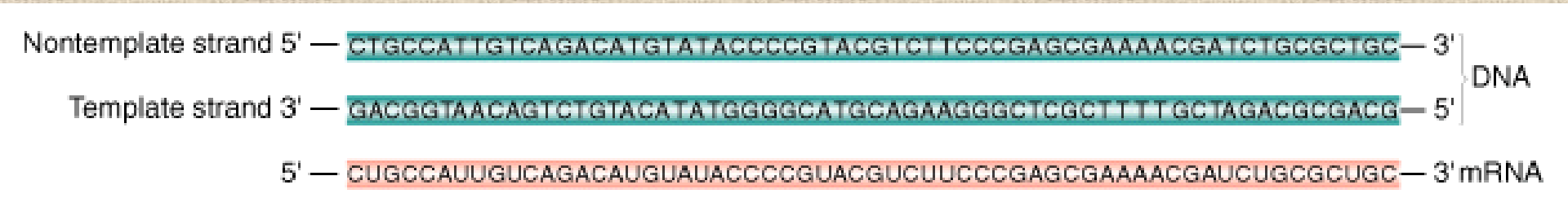
РНК-продукт не остается комплементарно связанным с матрицей;

по окончании синтеза двойная спираль восстанавливается;

одноцепочечная РНК высвобождается;

синтезированные РНК – копии ограниченных Участков ДНК

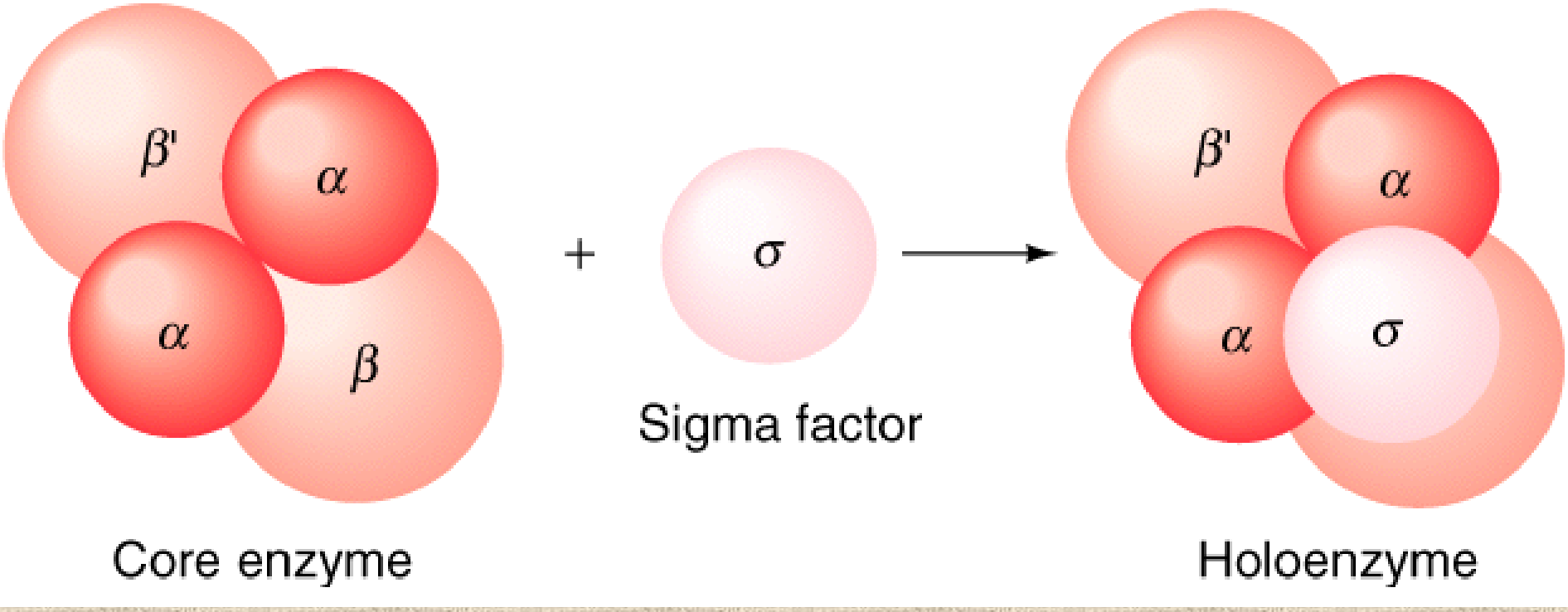
Матрица


Nontemplate strand 5' — CTGCCATTGTCAGACATGTATACCCCGTACGTCTTCCCGAGCGAAAACGATCTGCGCTGC — 3'
Template strand 3' — GACGGTAACAGTCTGTACATATGGGGCATGCAGAAGGGCTCGCTTTTGCTAGACGCGACG — 5'
DNA
5' — CUGCCAUUGUCAGACAUGUAUACCCCGUACGUCUUCCCGAGCGAAAACGAUCUGCGCUGC — 3' mRNA

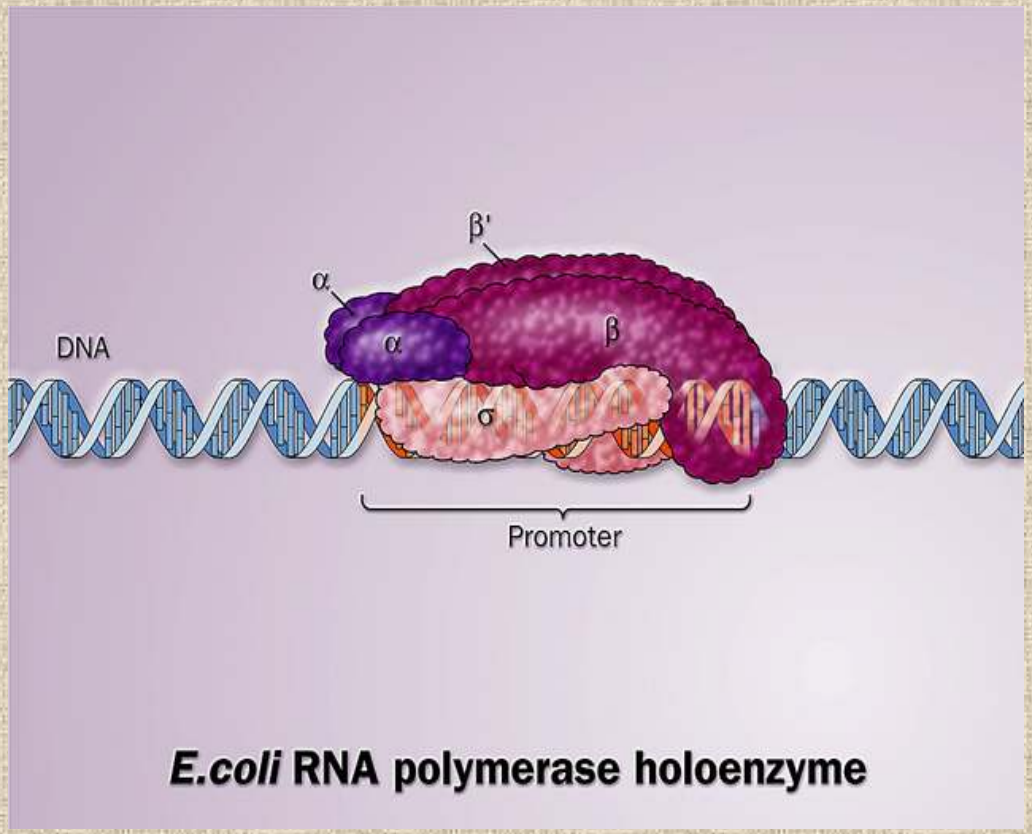
ПРОДУКТЫ ТРАНСКРИПЦИИ ПРОКАРИОТ

- **Матричные РНК (мРНК)** – короткоживущие посредники в процессе переноса информации от ДНК к белку
- **Рибосомные РНК (рРНК)** – компонент рибосом и рибонуклеопротеиновых частиц
- **Транспортные РНК (тРНК)** – доставка аминокислот к рибосомам

РНК полимераза (прокариоты)



Инициация транскрипции у прокариот



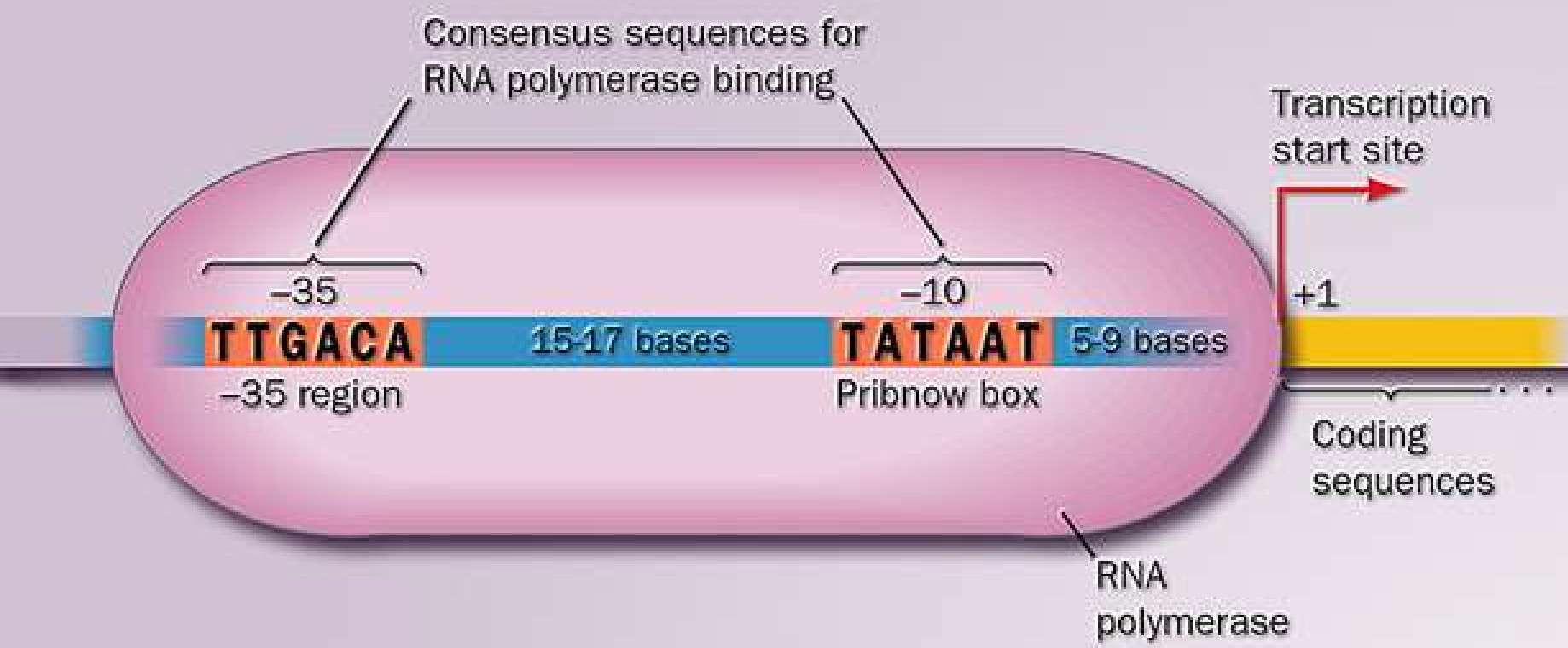
Не нужна затравка

Синтез начинается с определенного нуклеотида (+1)

Начало синтеза определяется *промотором*

За узнавание промотора ответственна σ-субъединица

Направление синтеза 5’→3’

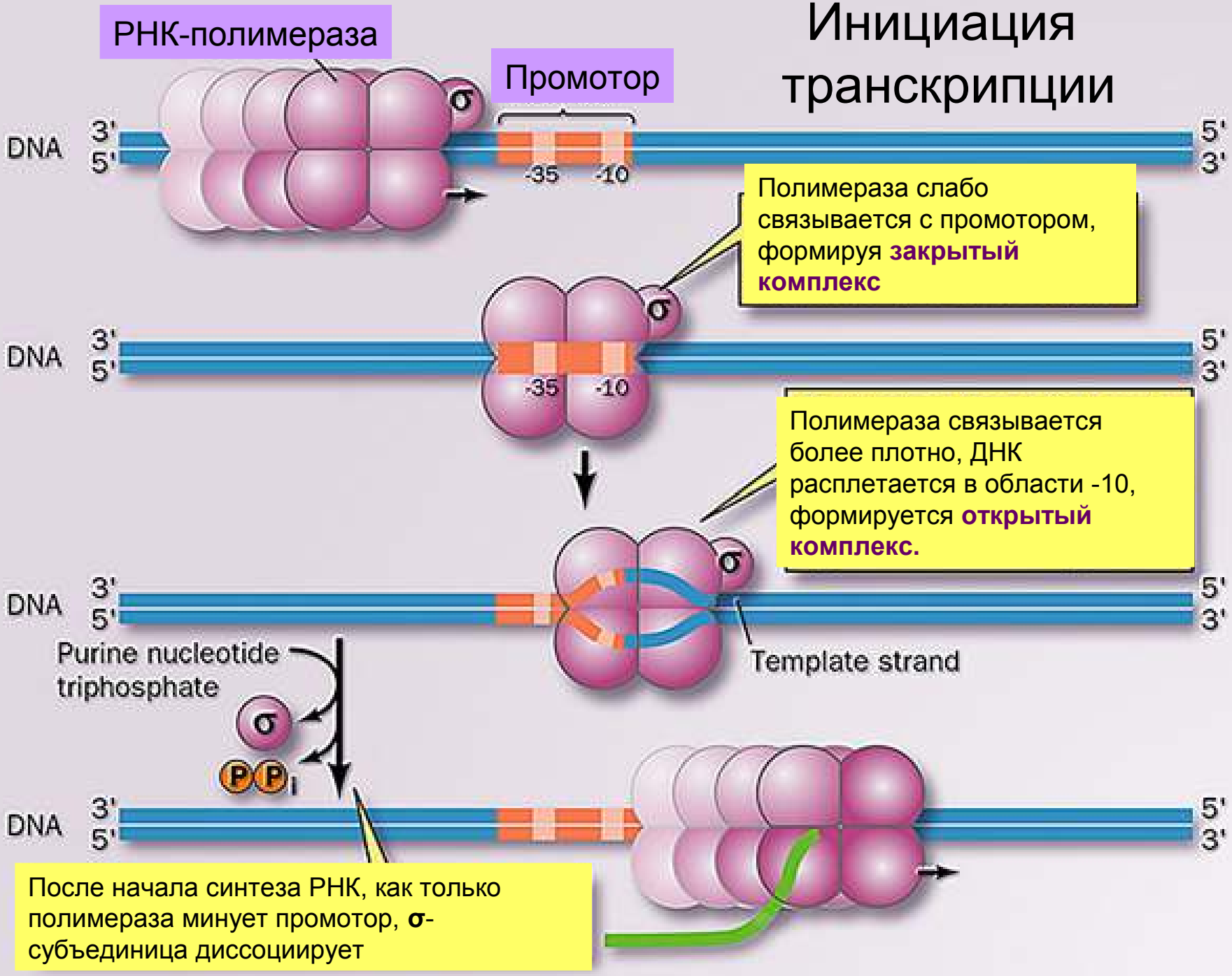


Структура прокариотического промотора

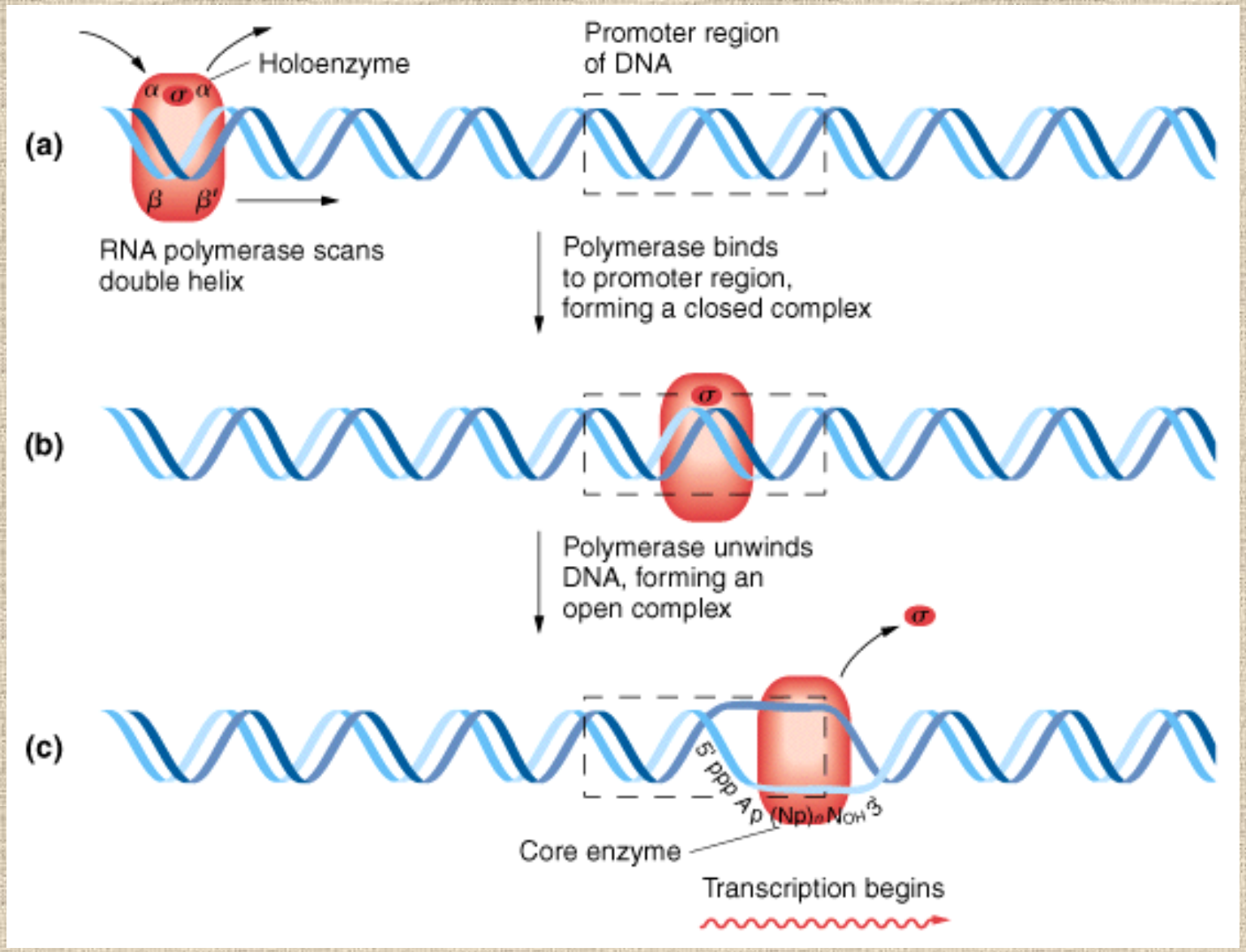
Consensus sequence

Promoters for:

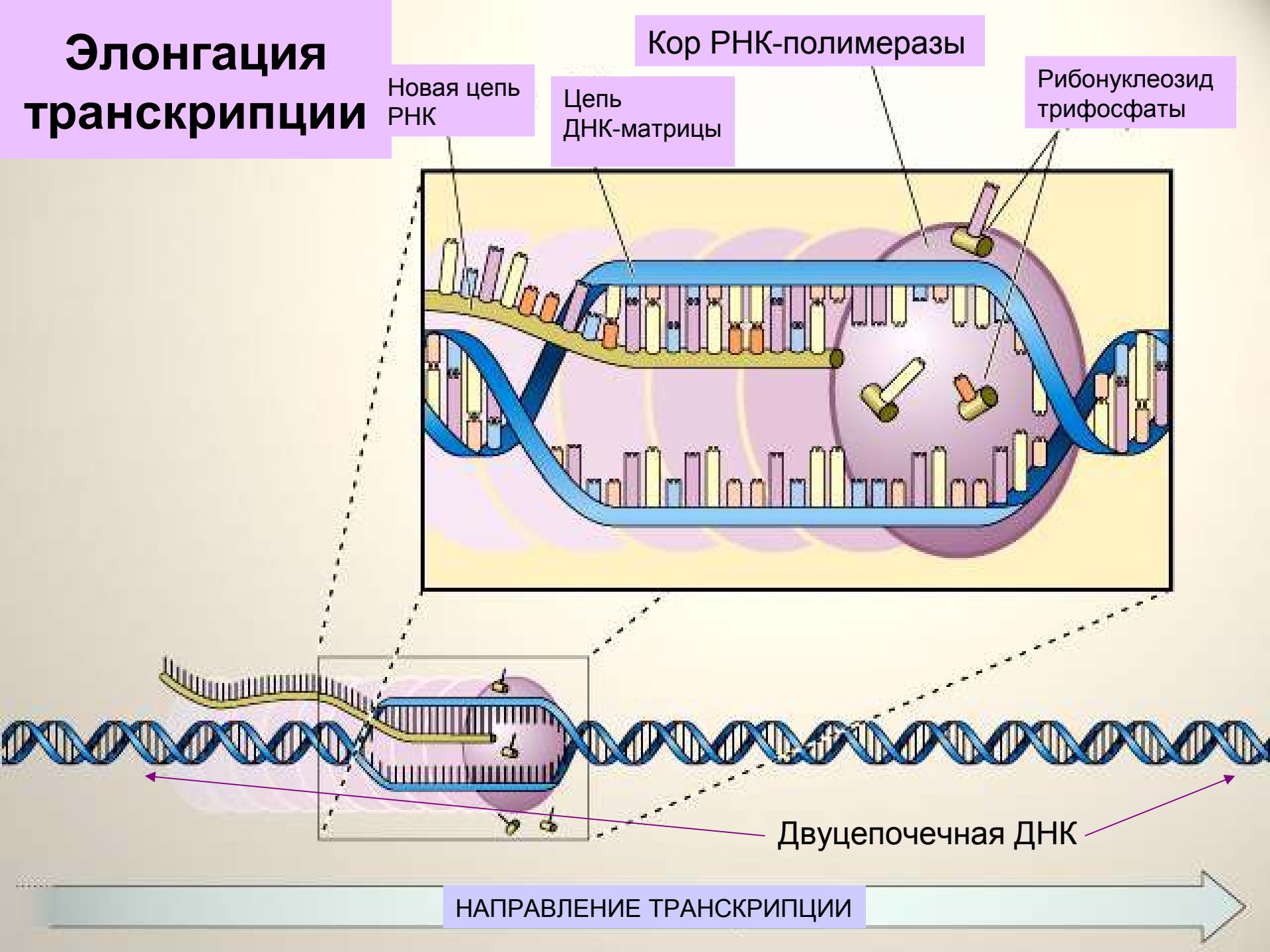
Инициация транскрипции

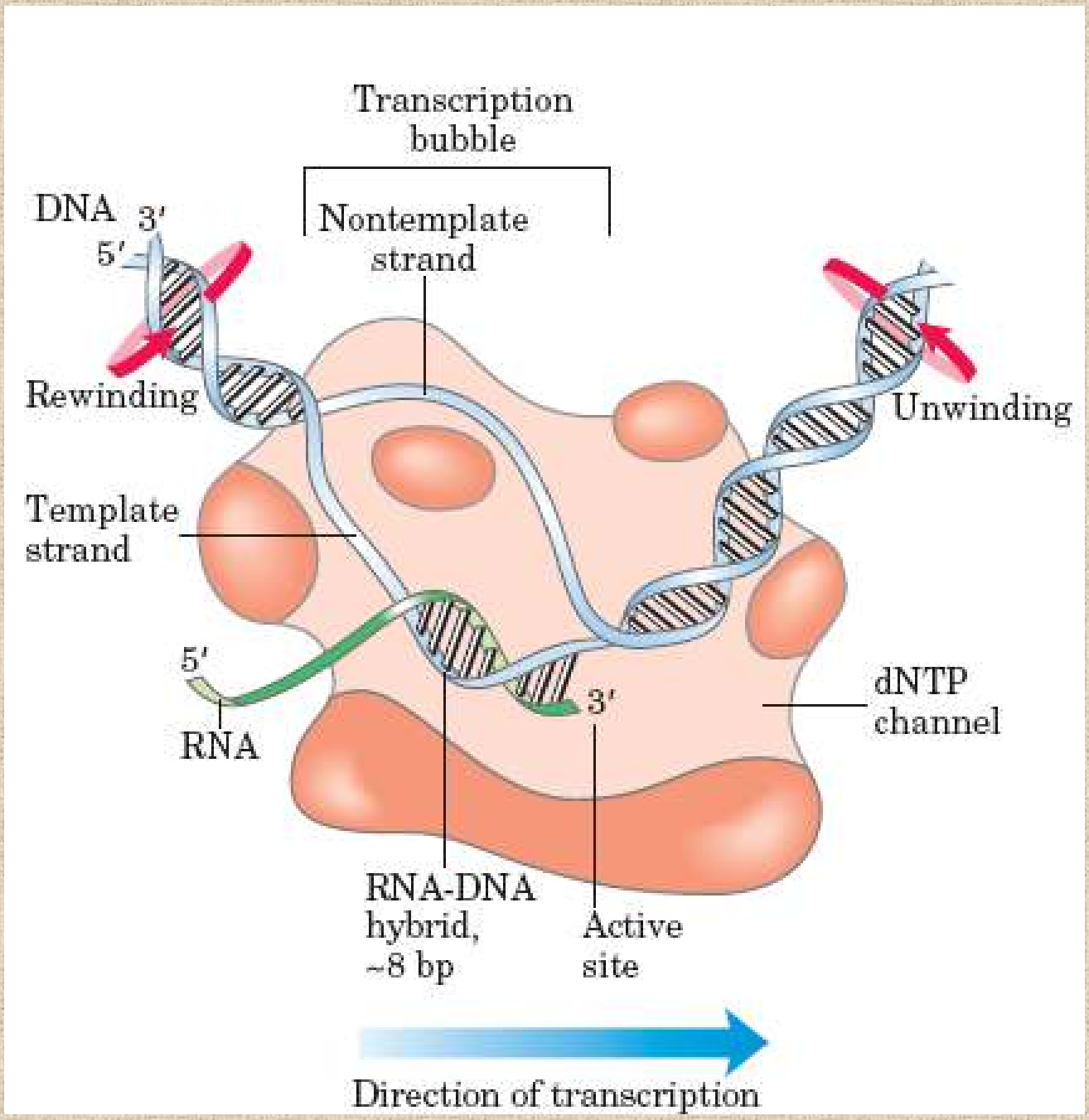


Инициация транскрипции

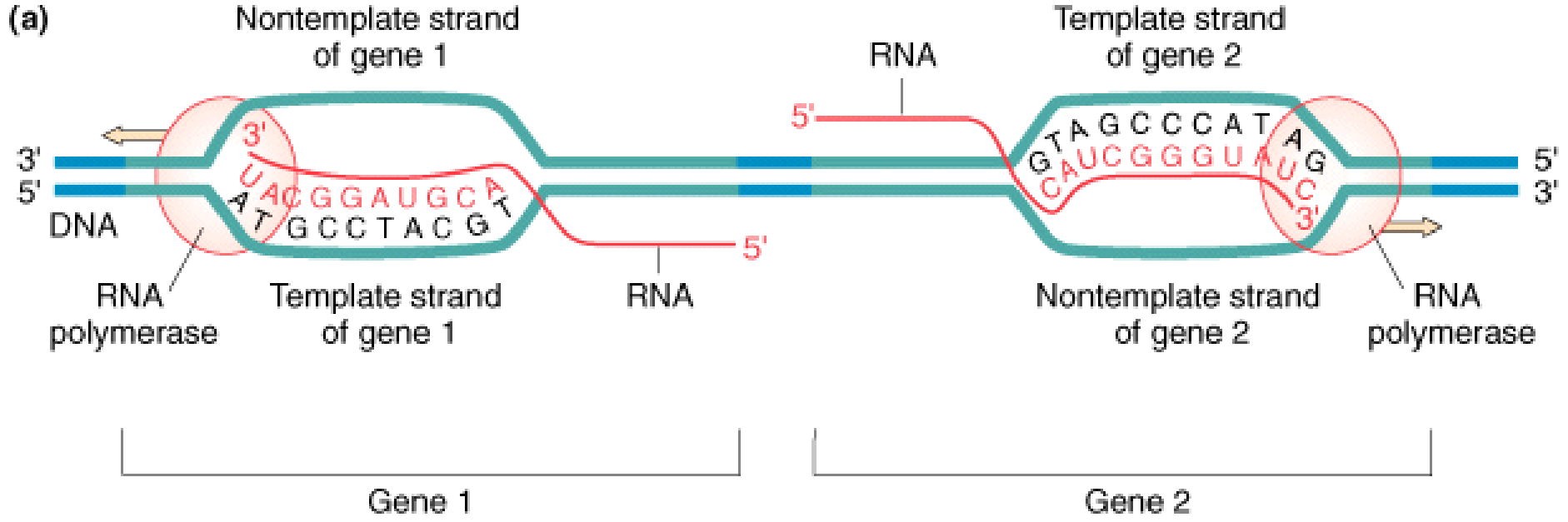


Элонгация транскрипции

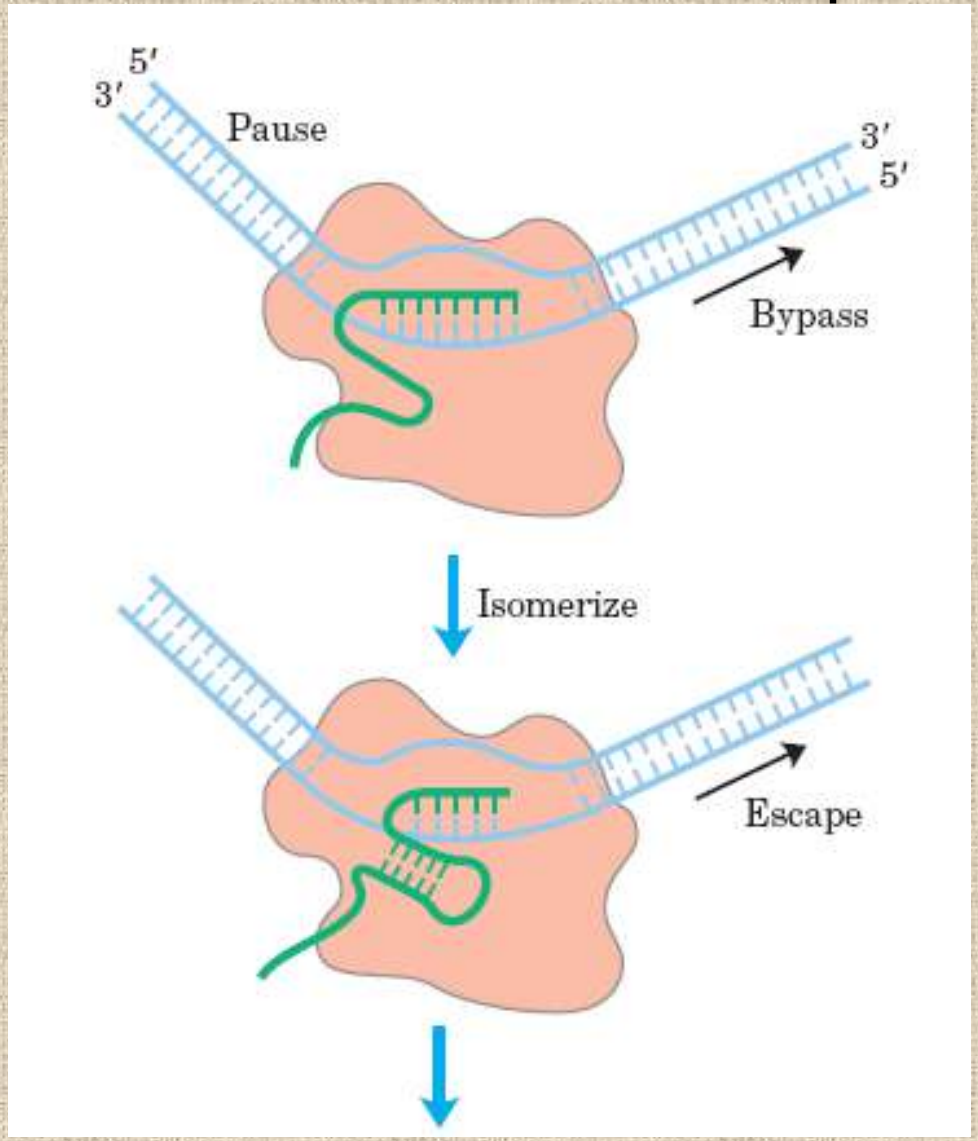


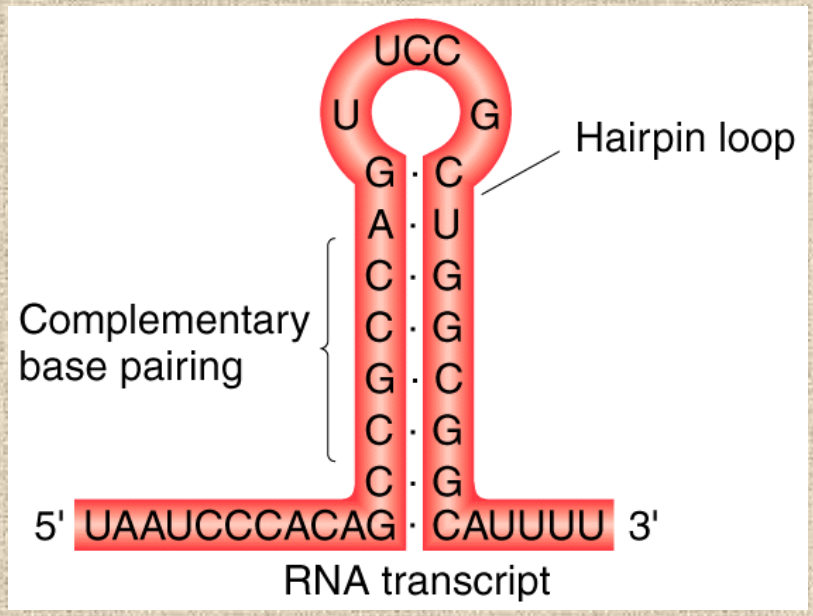

Transcription bubble
DNA
3′
5′
Nontemplate strand
Rewinding
Unwinding
Template strand
5′
RNA
3′
dNTP channel
RNA-DNA hybrid, ~8 bp
Active site
Direction of transcription

Цепи ДНК


(a)
Nontemplate strand of gene 1
Template strand of gene 2
RNA
3'
5'
DNA
5'
3'
C A C G G A U G C A
A T G C C T A C G T
G T A G C C C A T A G
C A U C G G G U A U C
3'
5'
5'
3'
RNA polymerase
Template strand of gene 1
RNA
Nontemplate strand of gene 2
RNA polymerase
Gene 1
Gene 2

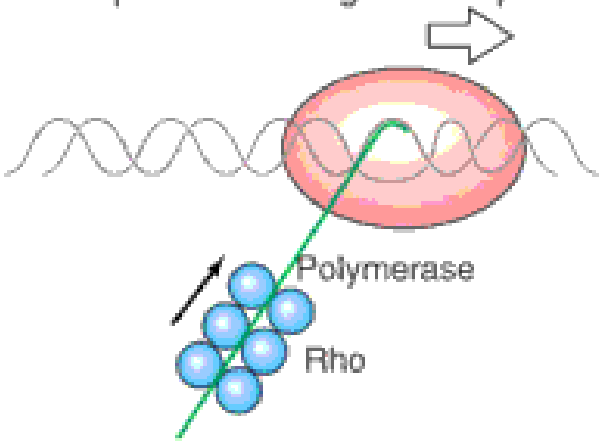
Rho-независимая терминация транскрипции



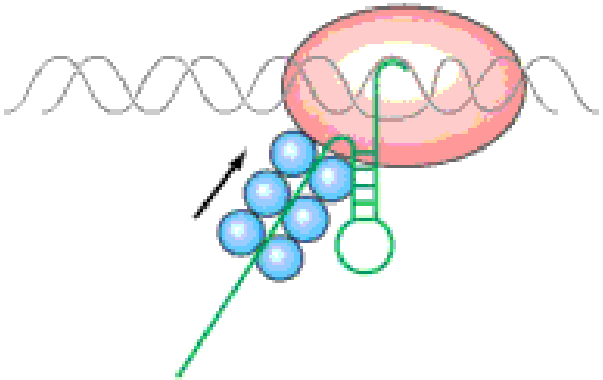


терминация

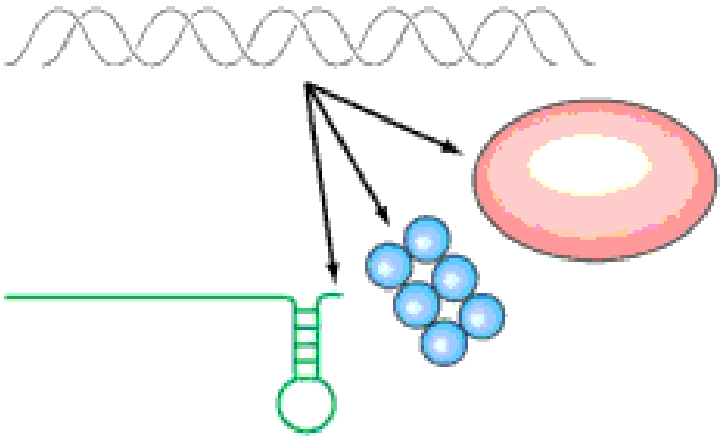
(a) Rho binds to transcript at rho loading site and pursues polymerase.



(b) Hairpin forms; polymerase pauses; rho catches up.



(c) Rho helicase releases transcript and causes termination.



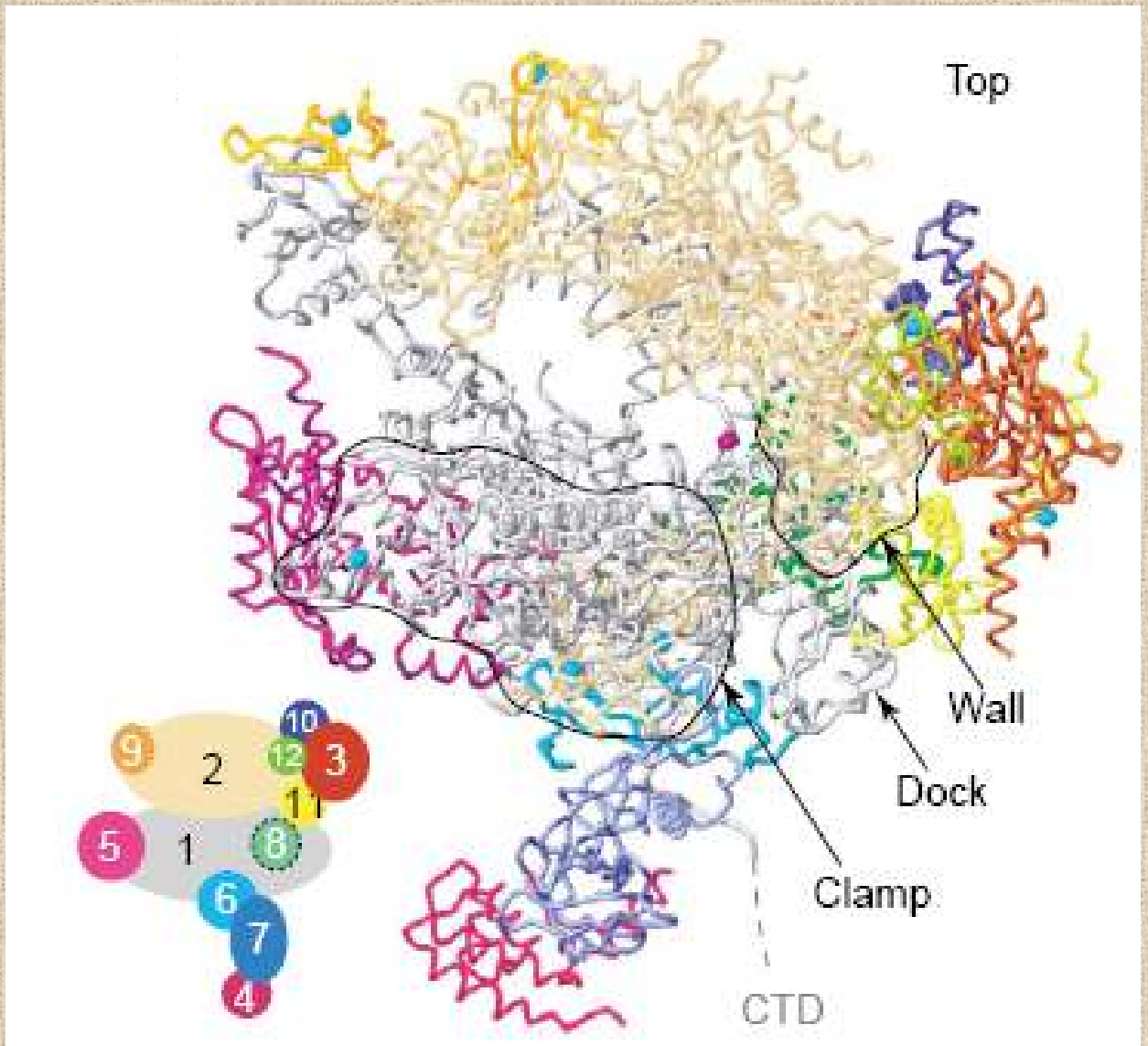
Rho-зависимая терминация транскрипции

Транскрипция у эукариот

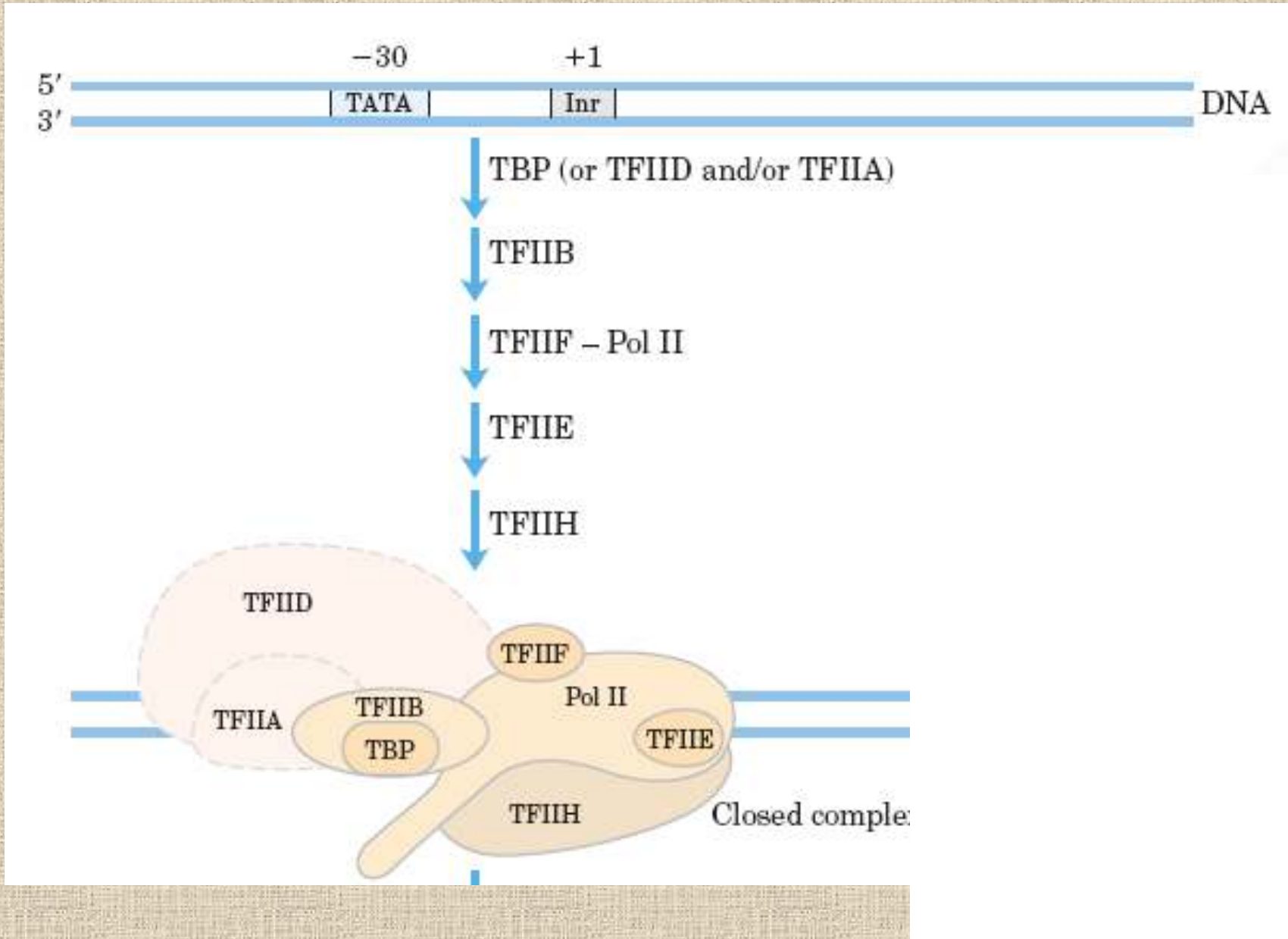
ДНК-зависимые РНК-полимеразы эукариот

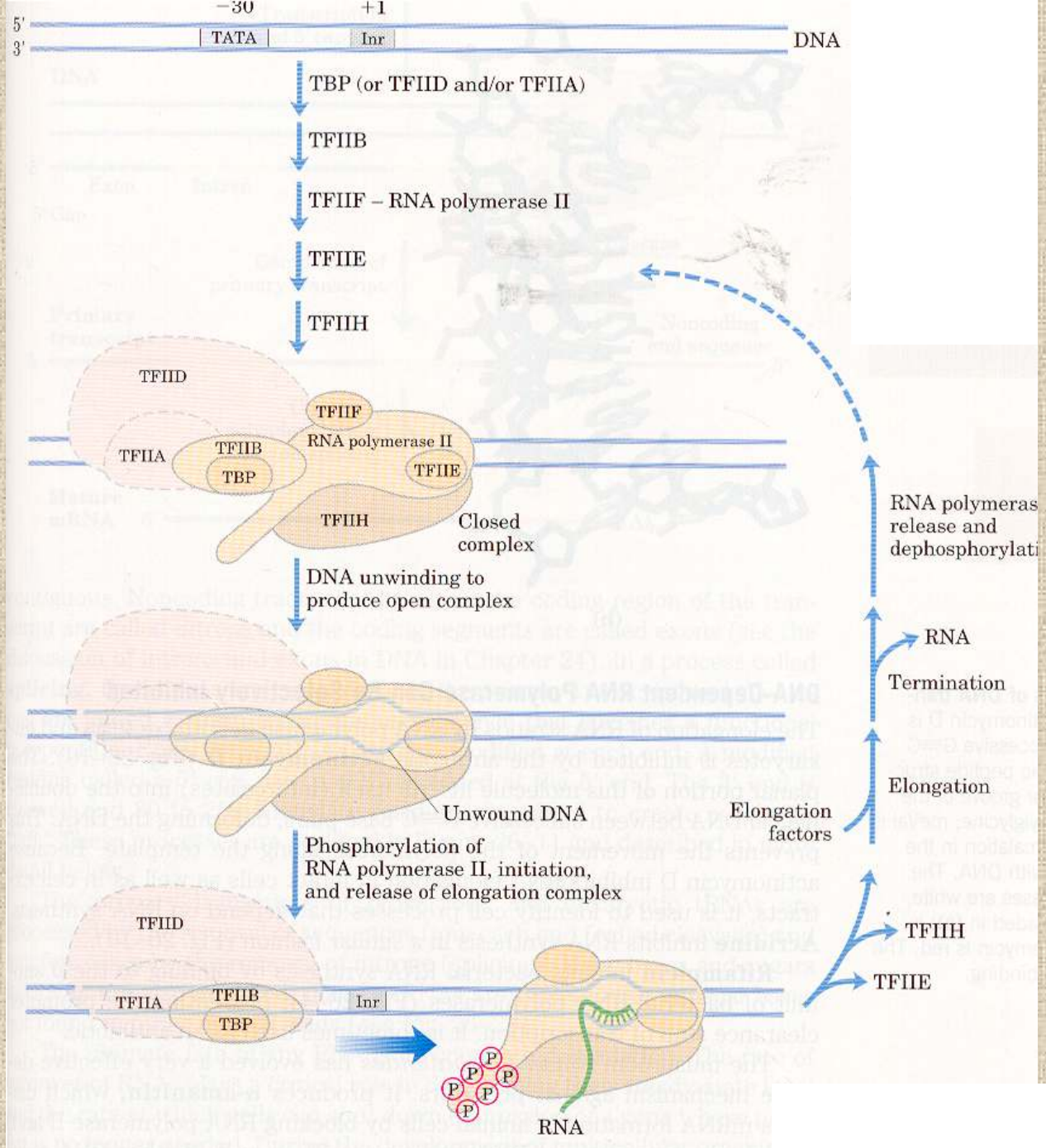
Type	Localization	Cellular transcripts
I	Nucleolus	18S, 5.8S, and 28S ribosomal RNAs (rRNA)
II	Nucleoplasm	mRNA precursors and small nuclear RNAs (snRNA)
III	Nucleoplasm	tRNAs and 5S rRNAs

Субъединичная структура эукариотической РНК-полимеразы II



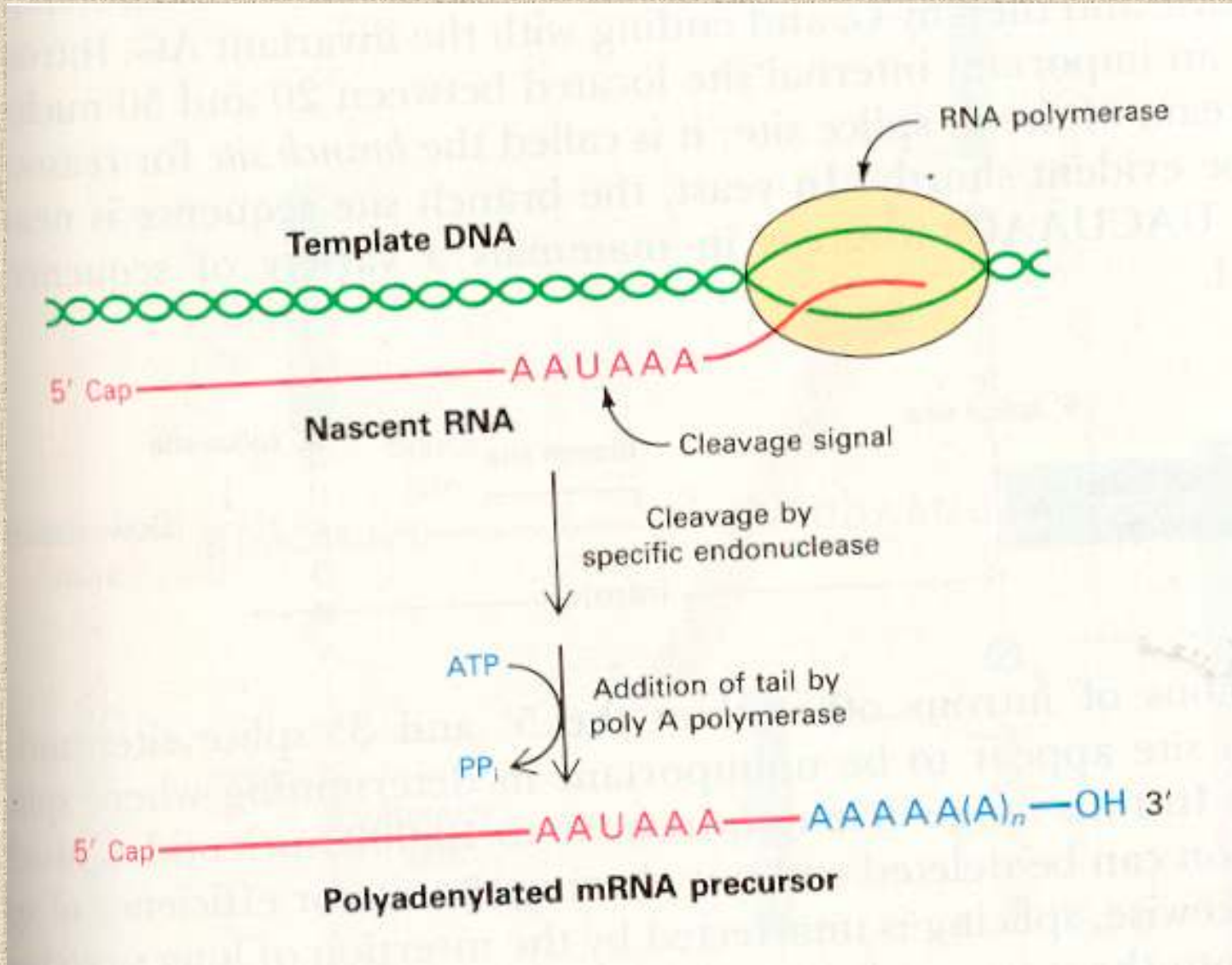
Инициация транскрипции




−30
+1
5'
3'
TATA
Inr
DNA
TBP (or TFIID and/or TFIIA)
TFIIB
TFIIF – RNA polymerase II
TFIIE
TFIIH
TFIID
TFIIF
RNA polymerase II
TFIIA
TFIIB
TBP
TFIIE
TFIIH
Closed complex
DNA unwinding to produce open complex
Unwound DNA
Phosphorylation of RNA polymerase II, initiation, and release of elongation complex
TFIID
TFIIA
TFIIB
TBP
Inr
P
RNA
RNA polymeras
release and
dephosphorylati
RNA
Termination
Elongation
Elongation factors
TFIIH
TFIIE
(a)

Элонгация цепи, терминация транскрипции

PLAY



Процессинг РНК:

1. **Процессинг рРНК и тРНК:** выщепление под действием эндонуклеаз из высокомолекулярного предшественника (прокариоты, эукариоты)

2. **Процессинг иРНК** (только эукариоты):
   кэпирование
   полиаденилирование
   сплайсинг
   редактирование

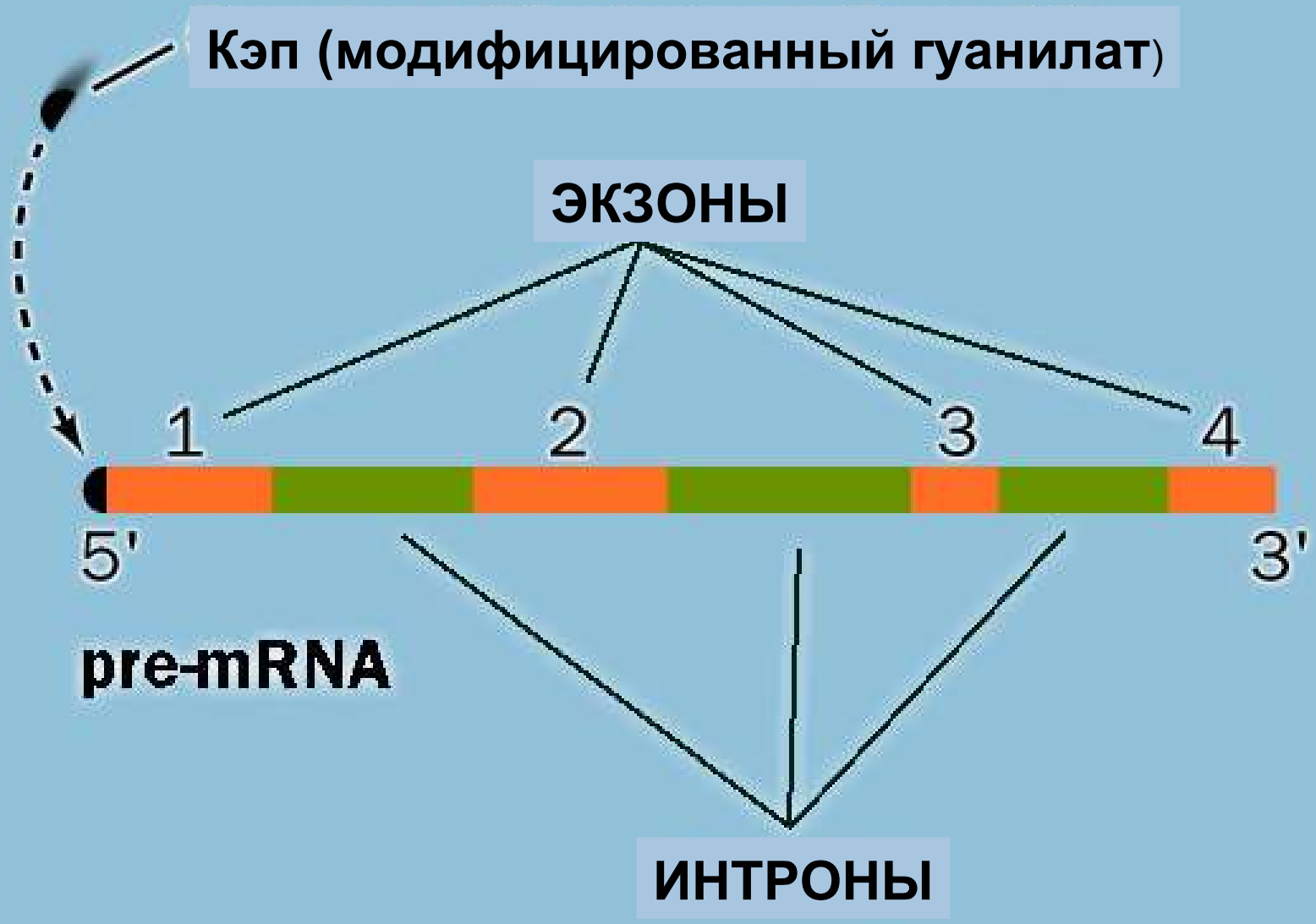
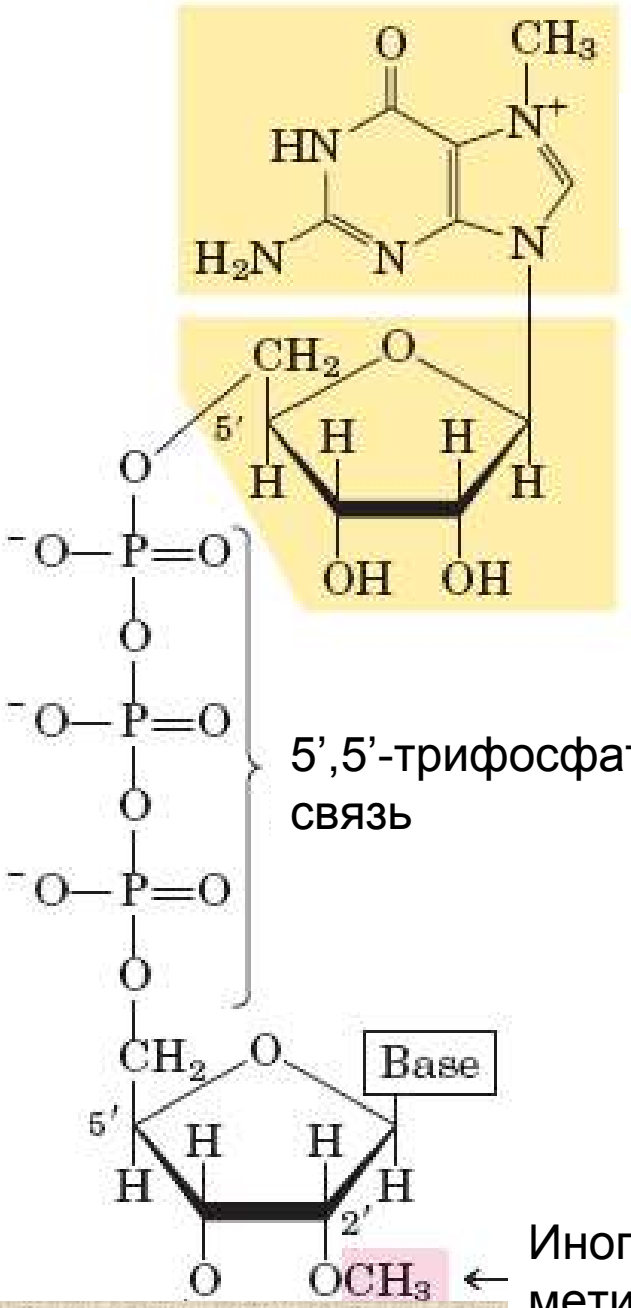
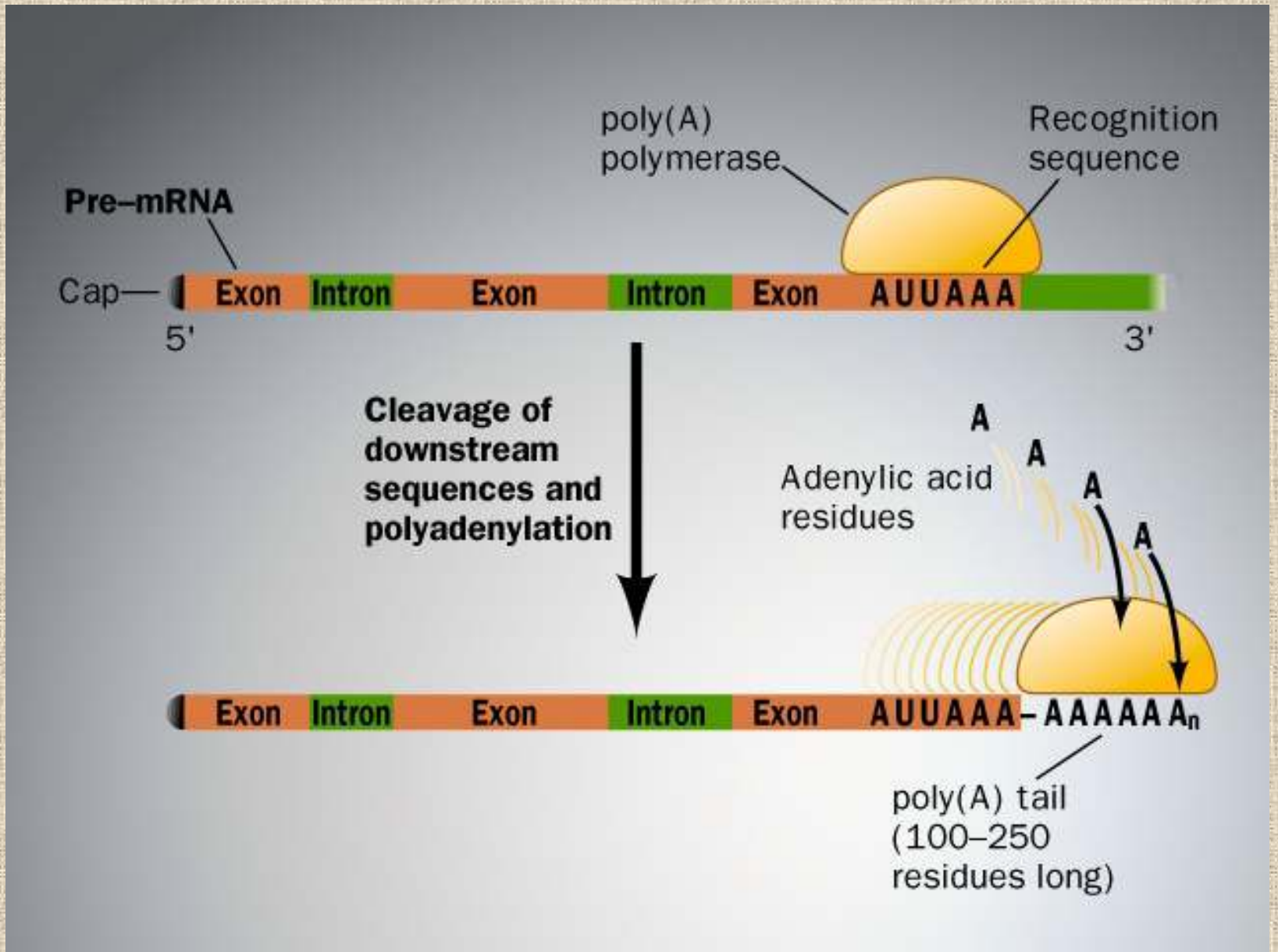


СХЕМА ЭУКАРИОТИЧЕСКОЙ ПРЕ-МРНК

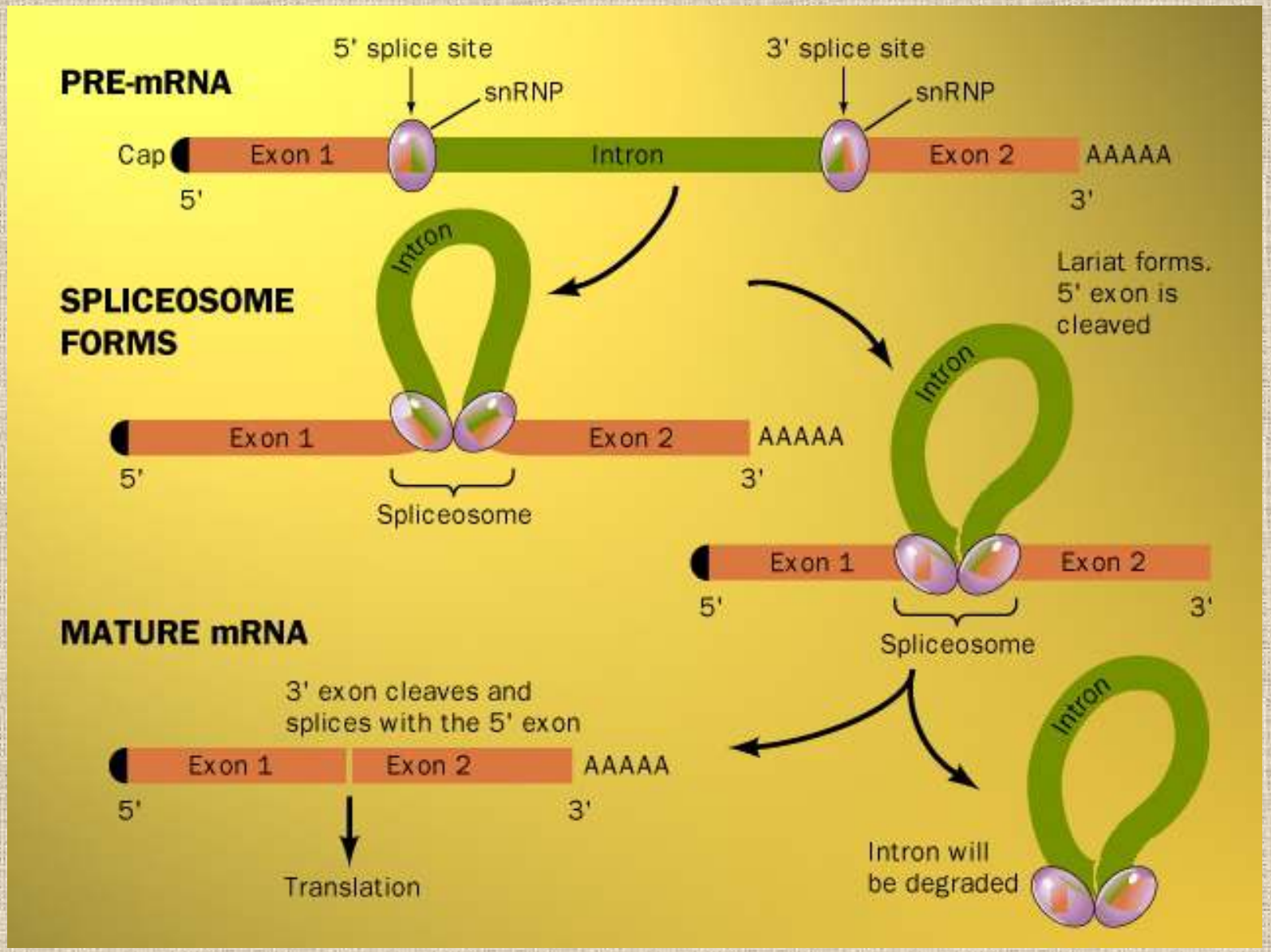
СТРУКТУРА КЭПА



Полиаденилирование
(до сигнала терминации)



Сплайсинг пре-мРНК

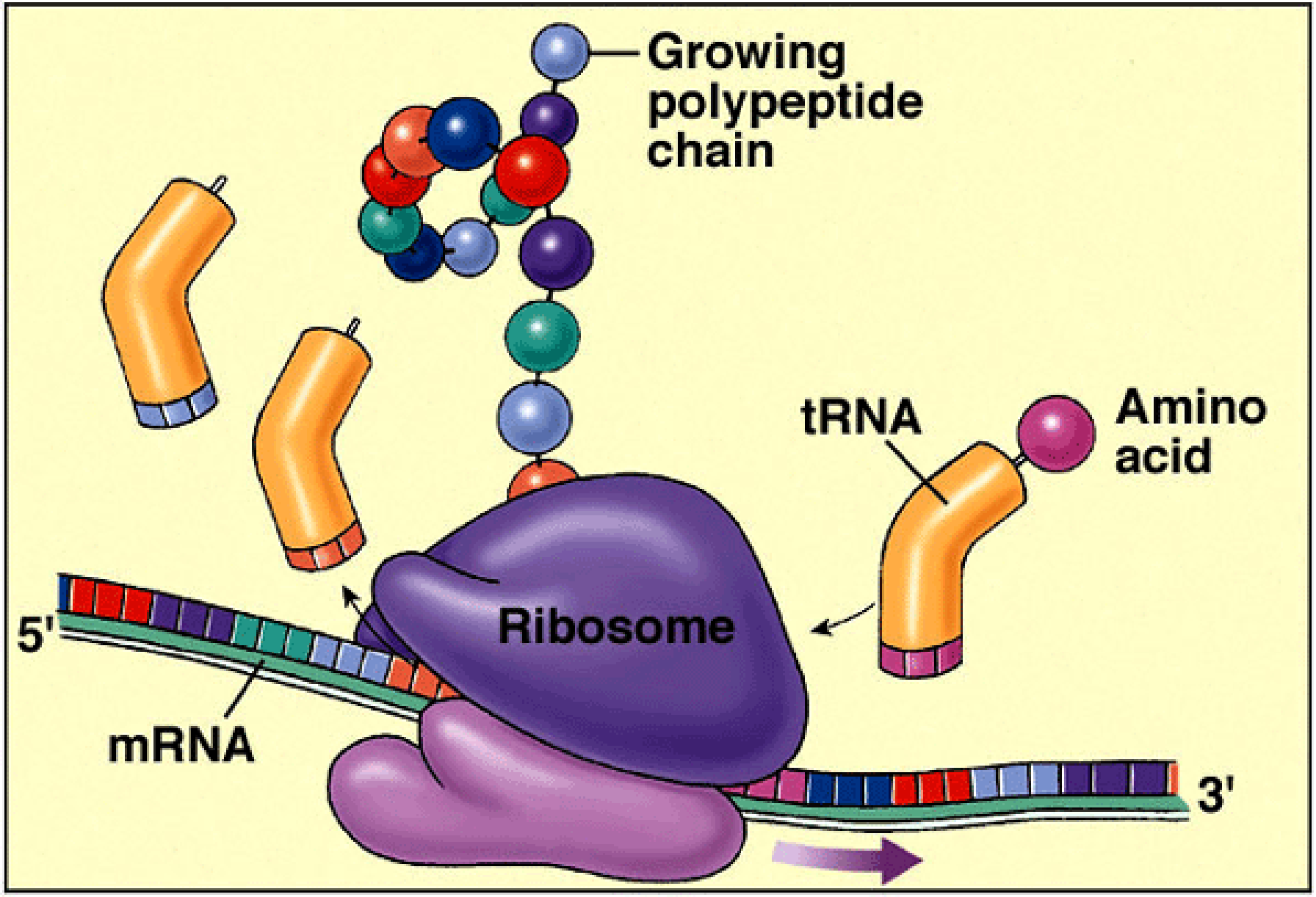


Трансляция (биосинтез белка)

- Особенности генетического кода
- **Основные этапы трансляции:**
- Активация аминокислот
  (особенности структуры тРНК, синтез аминоацил-тРНК);
- Инициация
  (структура рибосом, белковые факторы, синтез первой пептидной связи);
- Элонгация;
- Терминация;
- Процессинг белка. Пост-трансляционная модификация – добавление липидов, сахаров, фосфатов, сульфатов и.т.д.

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

Translation



ОСОБЕННОСТИ ГЕНЕТИЧЕСКОГО КОДА

- триплетность
- неперекрываемость
- вырожденность

Второе основание

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

5’ – A T G C C T A G G T A C C T A T G A – 3’

3’ – T A C G G A T C C A T G G A T A C T – 5’

DNA

Транскрипция

5’ – A U G C C U A G G U A C C U A U G A – 3’

mRNA

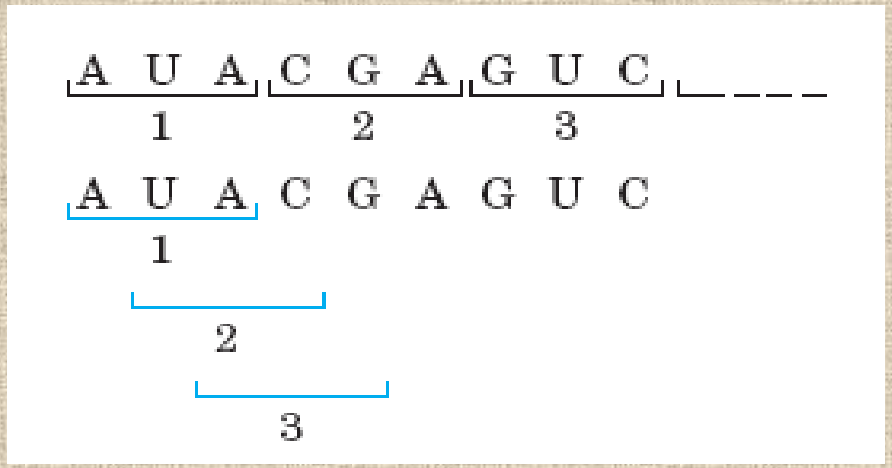
5’ – A U G C C U A G G U A C C U A U G A – 3’

Трансляция

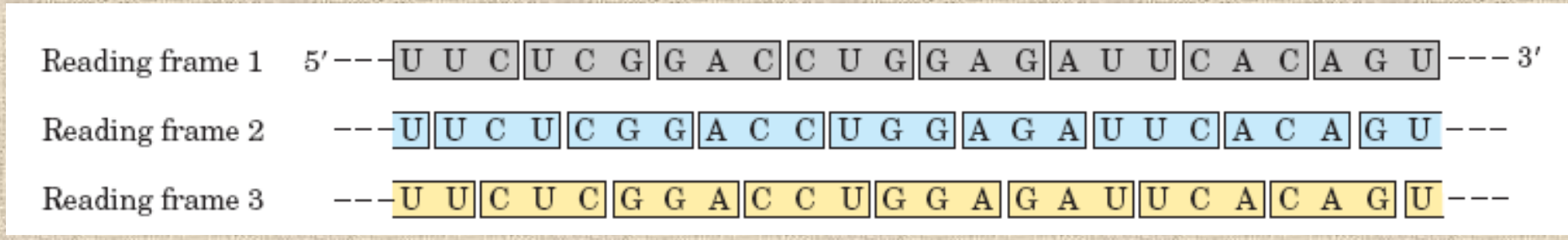
N-Met - Pro - Arg - Tyr - Leu-*C*

Белок

Неперекрывающийся код



Перекрывающийся код



The genetic code arranged according to the pattern of degeneracy

Amino Acids with one Codon

AUG	UGG
Met	Trp

Amino Acids with two codons

AAA AAG	AAC AAU	CAA CAG	CAC CAU	GAA GAG	GAC GAU	UAC UAU	UGC UGU	UUC UUU
Lys	Asn	Gln	His	Glu	Asp	Tyr	Cys	Phe

Amino acid with three codons

AUA AUC AUU
Ile

Amino Acids with four codons

ACA ACC ACG ACU	CCA CCC CCG CCU	GCA GCC GCG GCU	GGA GGC GGG GGU	GUA GUA GUG GUU
Thr	Pro	Ala	Gly	Val

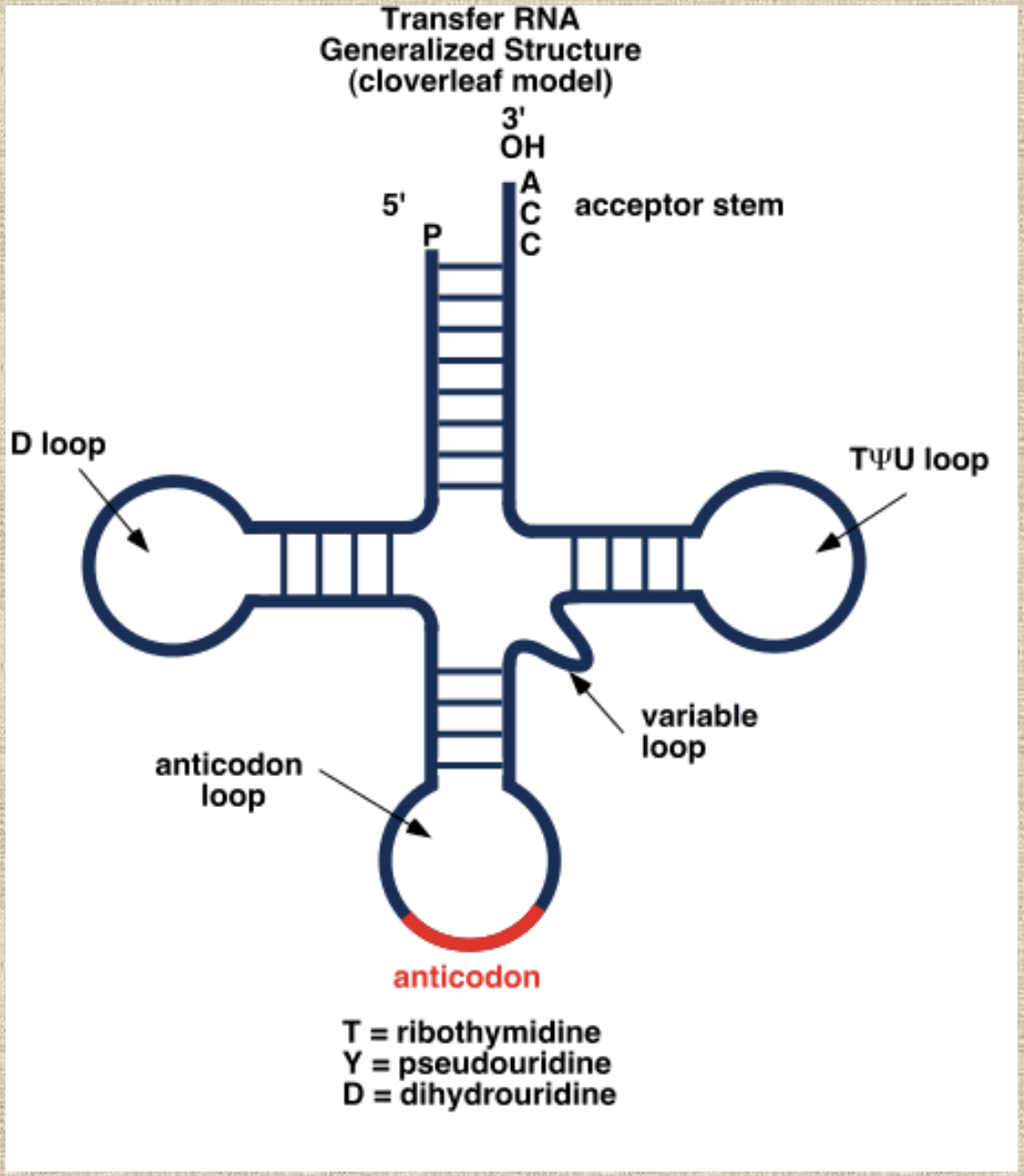
Aminio Acid with four codons

CGA CGC CGG CGU AGA AGG	GUA CUC CUG CUU UUA UUG	UCA UCC UCG UCU AGC AGU
Arg	Leu	Ser

ВЫРОЖДЕННОСТЬ ГЕНЕТИЧЕСКОГО КОДА

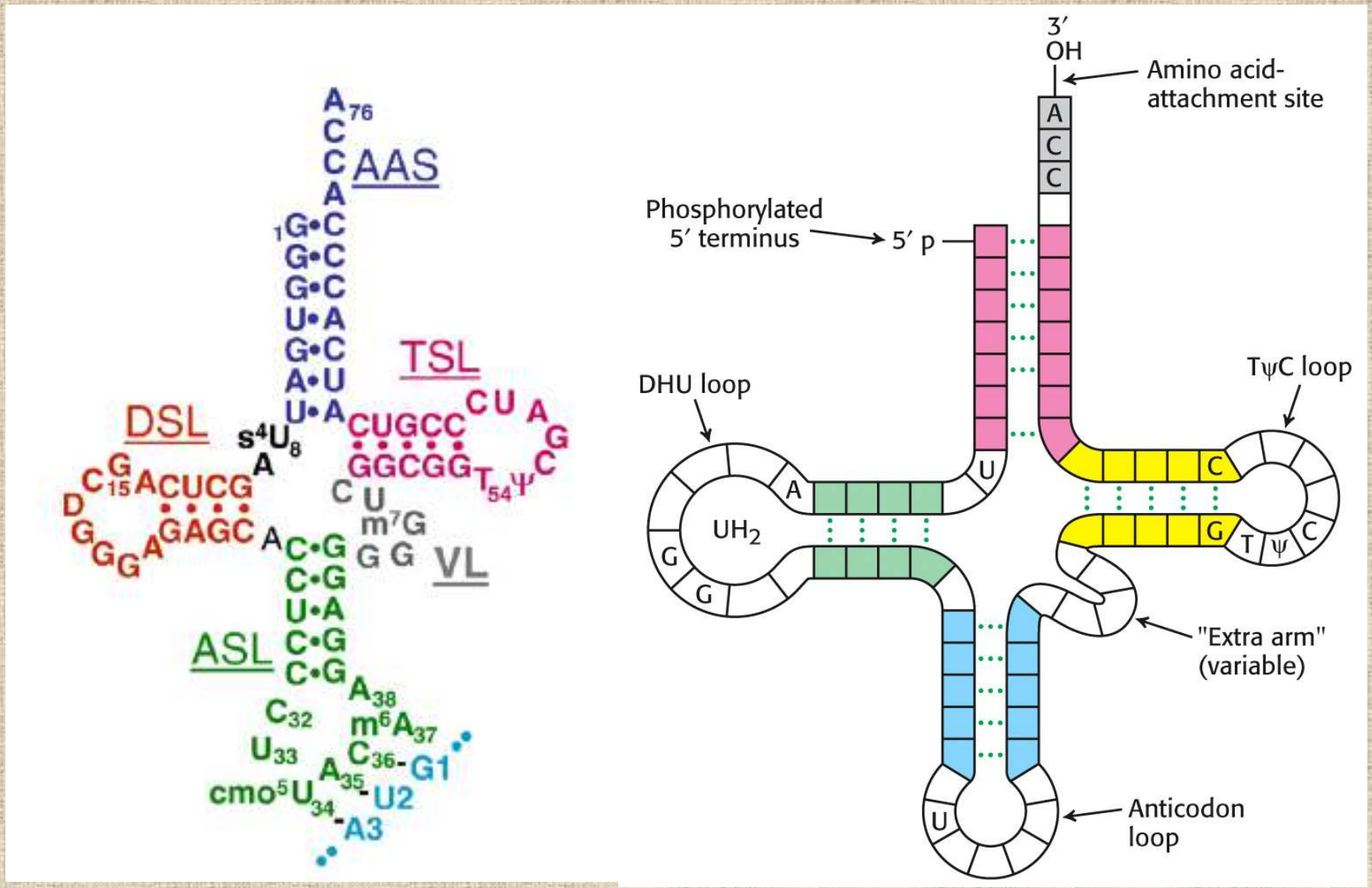
The genetic code

- consists of 64 triplet codons (A, G, C, U) $4^3 = 64$
- all codons are used in protein synthesis
  - 20 amino acids
  - 3 termination (stop) codons: UAA, UAG, UGA
- AUG (methionine) is the start codon (also used internally)
- multiple codons for a single amino acid = degeneracy


Transfer RNA
Generalized Structure
(cloverleaf model)
3'
OH
A
C
C
acceptor stem
5'
P
D loop
TΨU loop
variable
loop
anticodon
loop
anticodon
T = ribothymidine
Y = pseudouridine
D = dihydrouridine

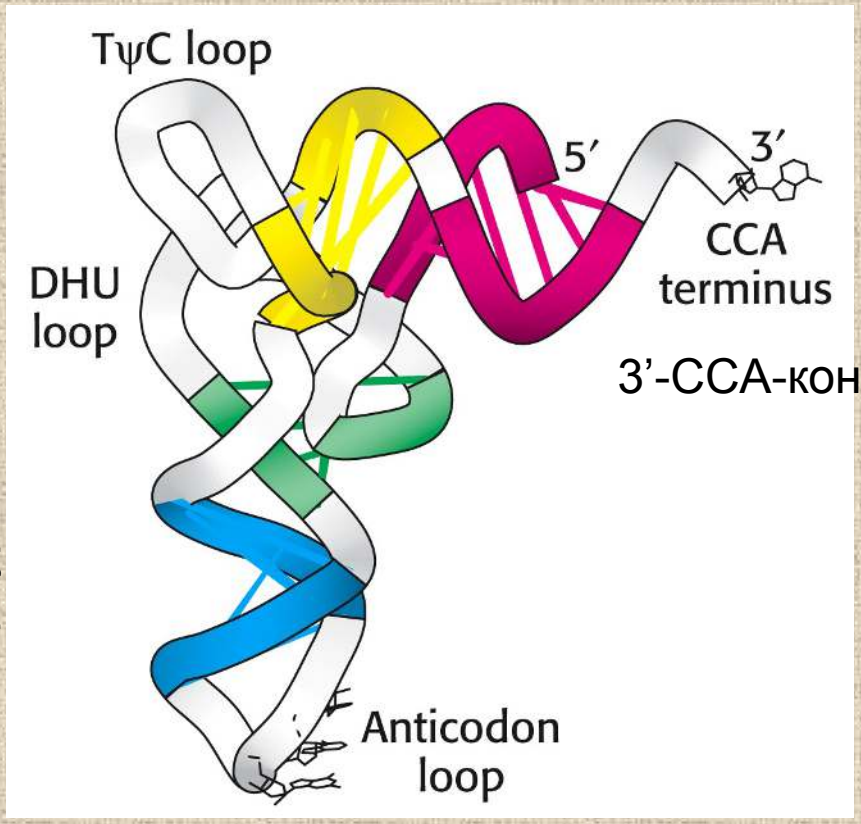
тРНК

Вторичная структура тРНК



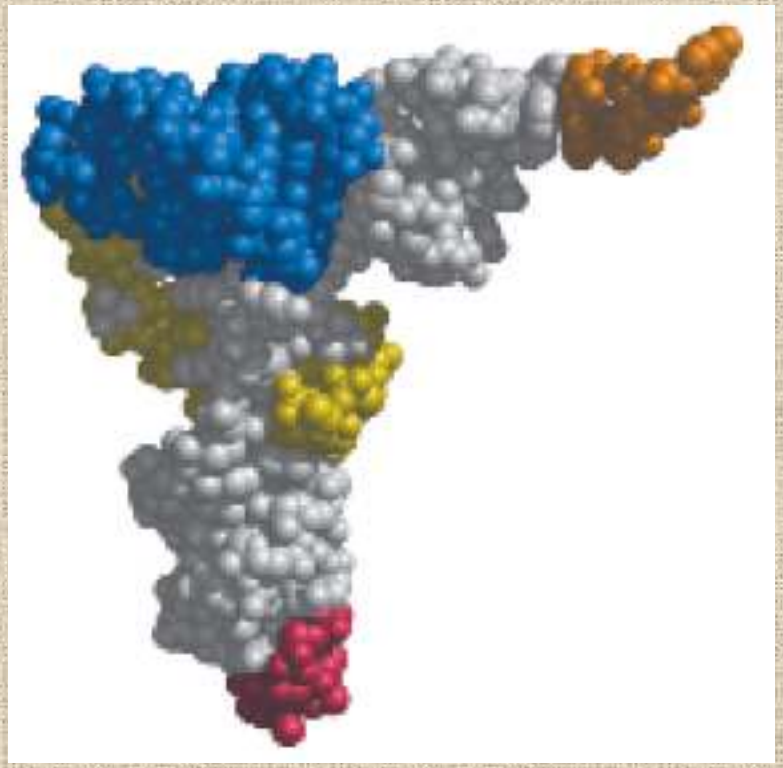
Третичная структура тРНК

Псевдоуридиновая петля



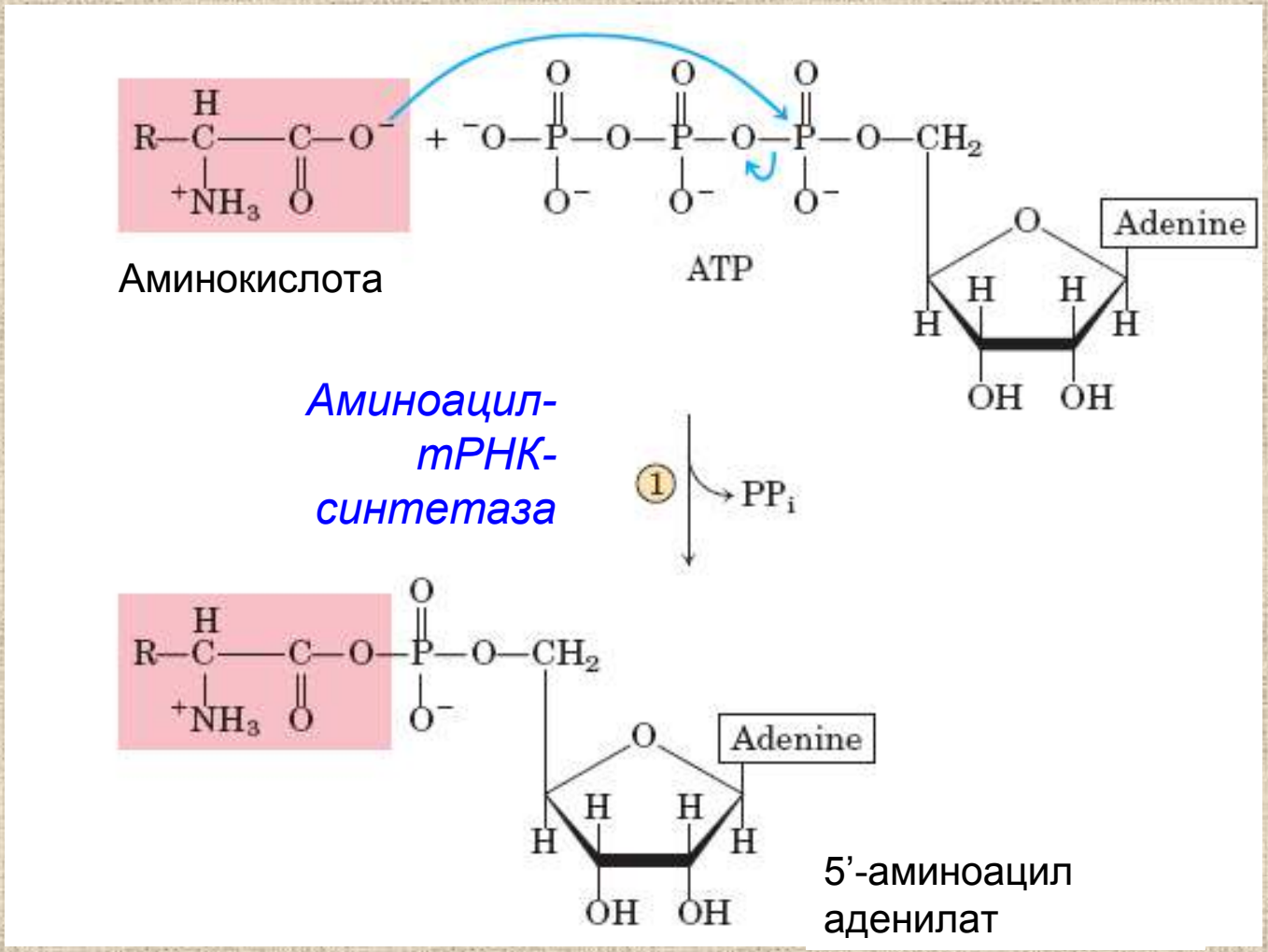
Дигидроуридиновая петля

3’-CCA-конец



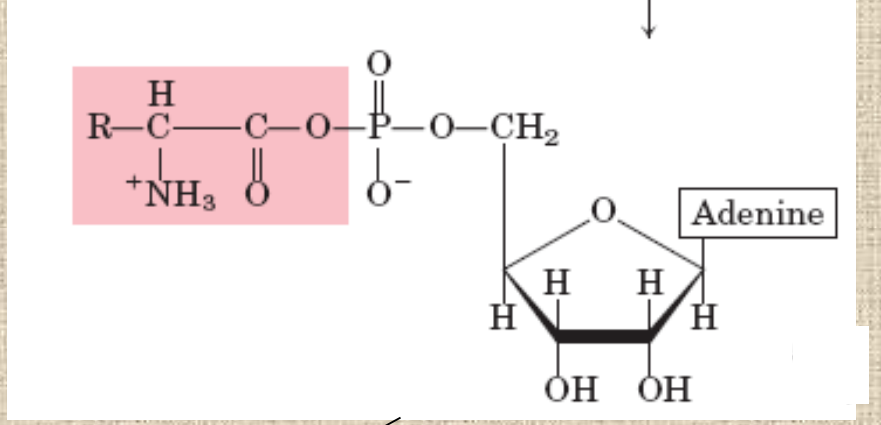
Антикодоновая петля

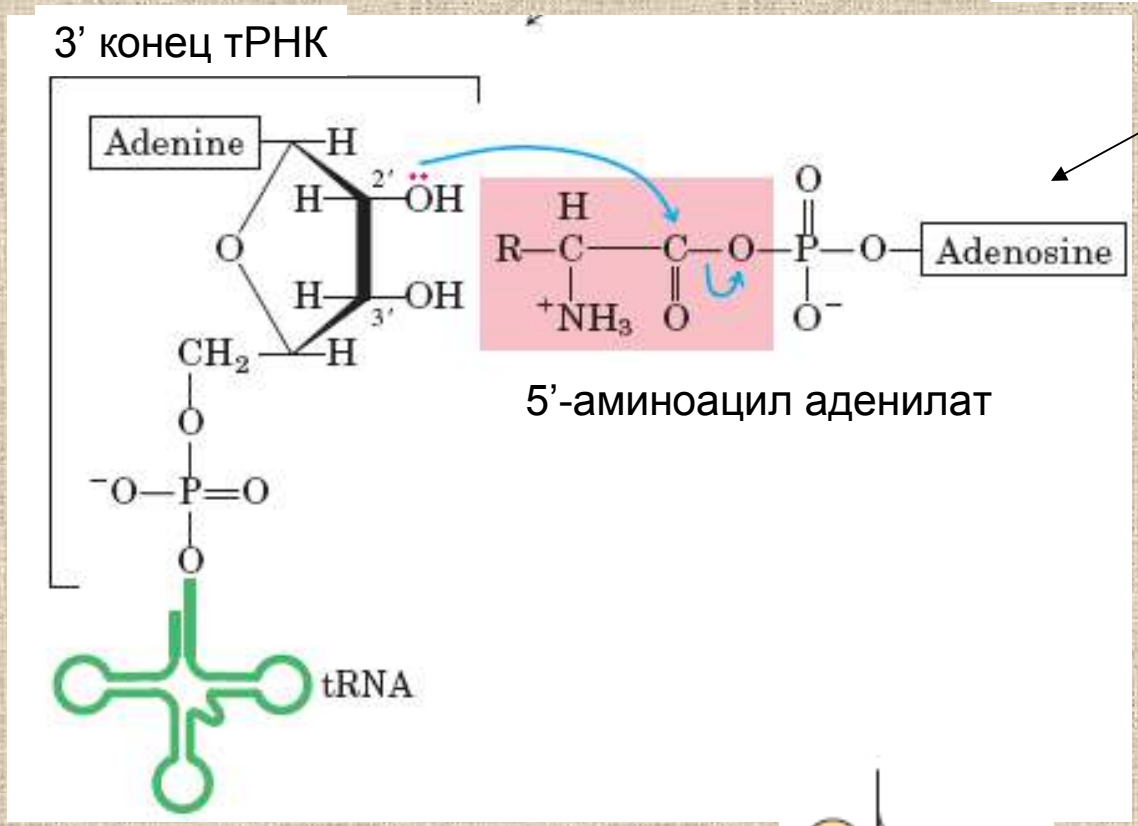
1. Активация аминокислот. Синтез аминоациладенилата



2.Аминоацилирование тРНК

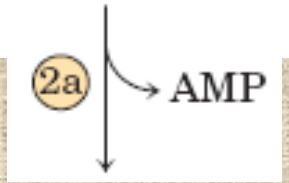
ааатРНК-синтетазы 1 класса

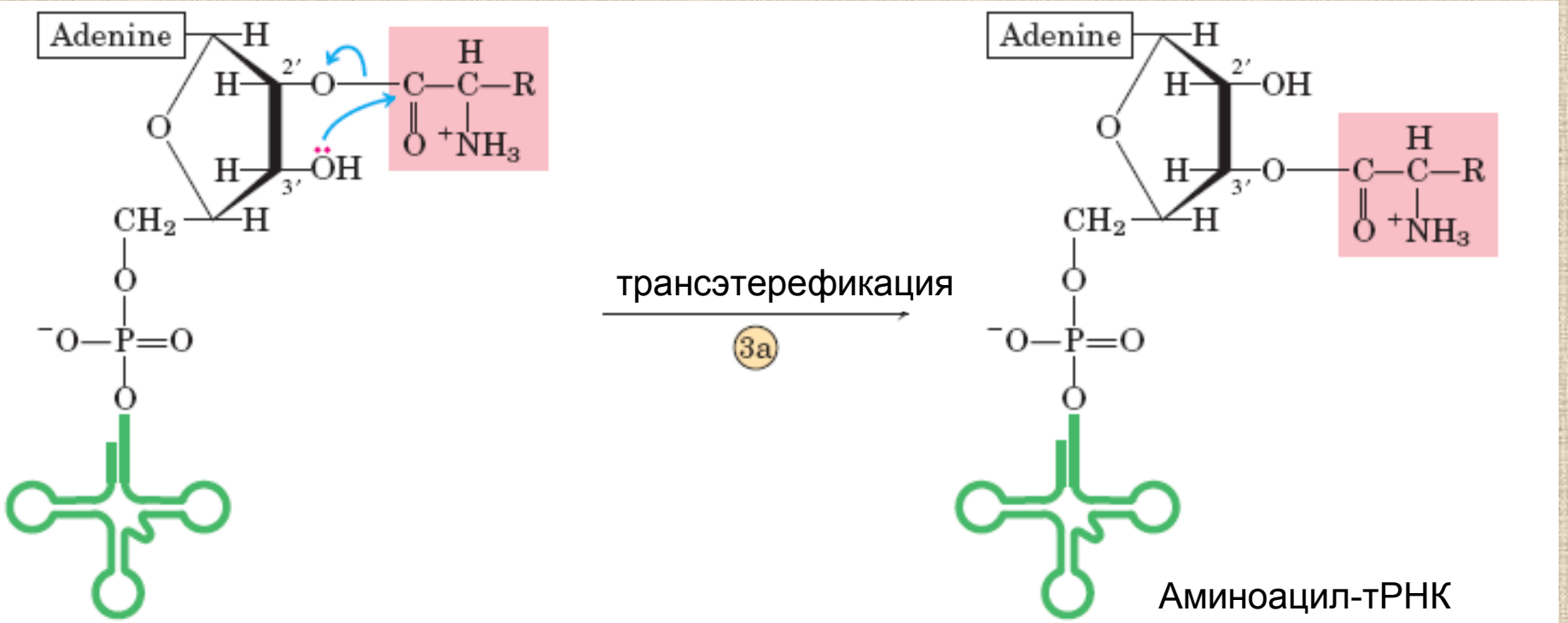




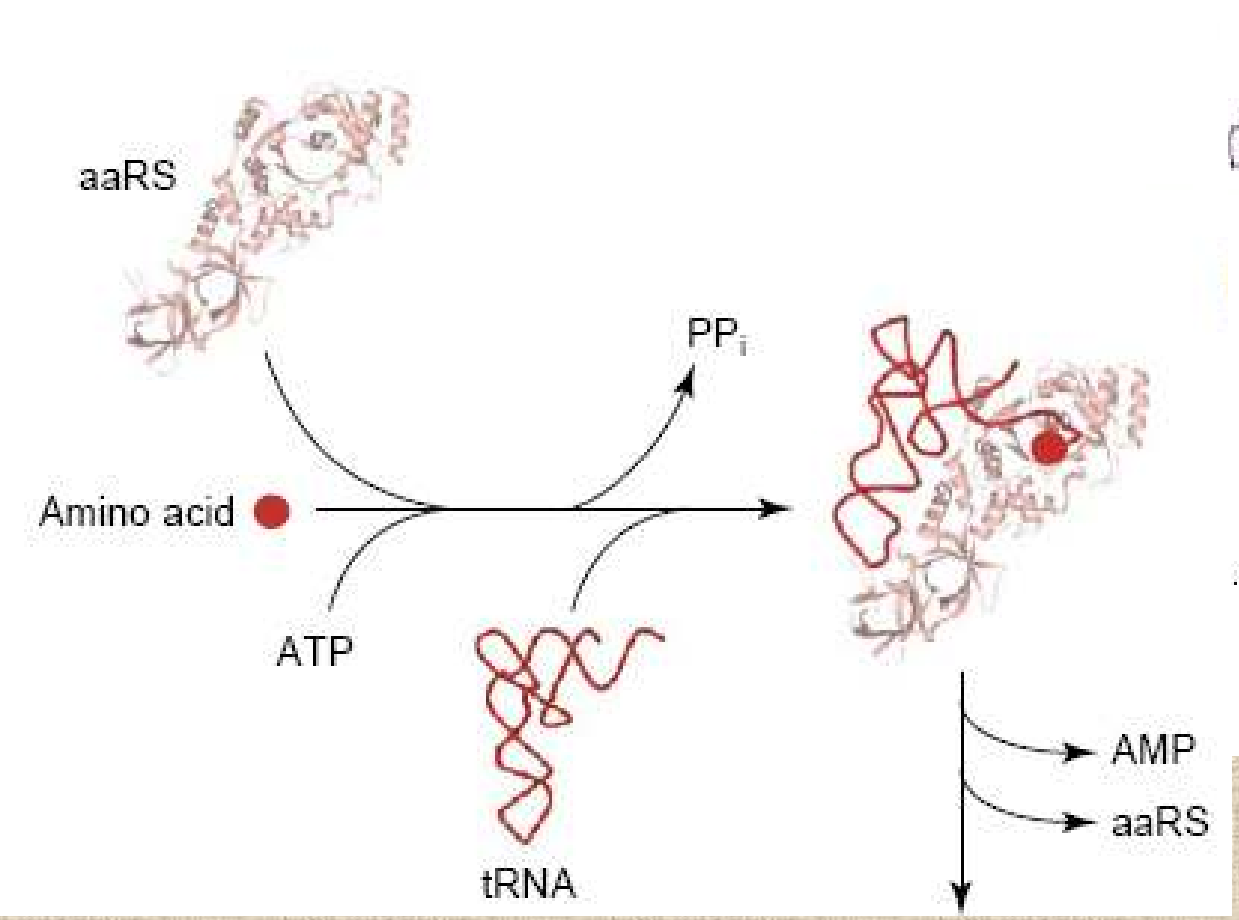
5’-аминоацил аденилат

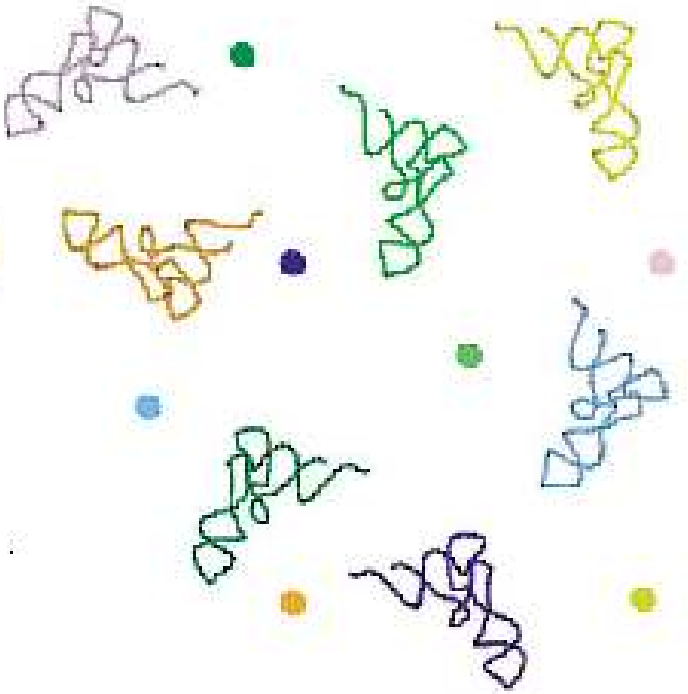
Аминоацил-тРНК-синтетаза 1 класса

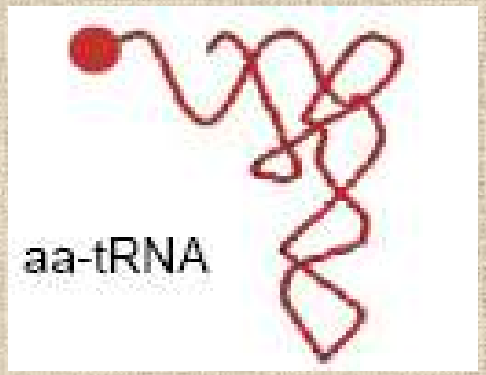




Class I		*Class II*	
Arg	Leu	Ala	Lys
Cys	Met	Asn	Phe
Gln	Trp	Asp	Pro
Glu	Tyr	Gly	Ser
Ile	Val	His	Thr

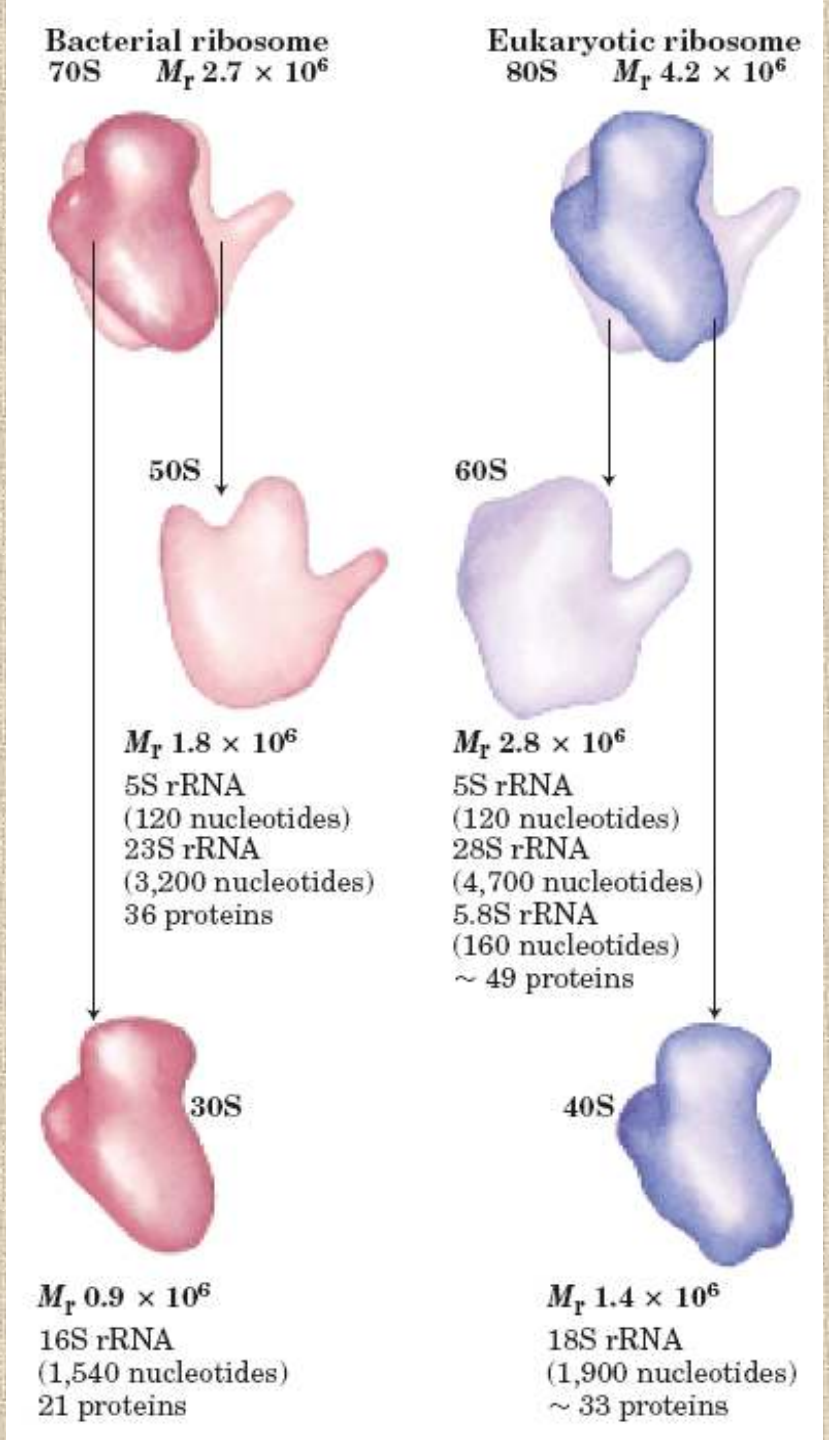

aaRS
PPi
Amino acid
ATP
tRNA
AMP
aaRS



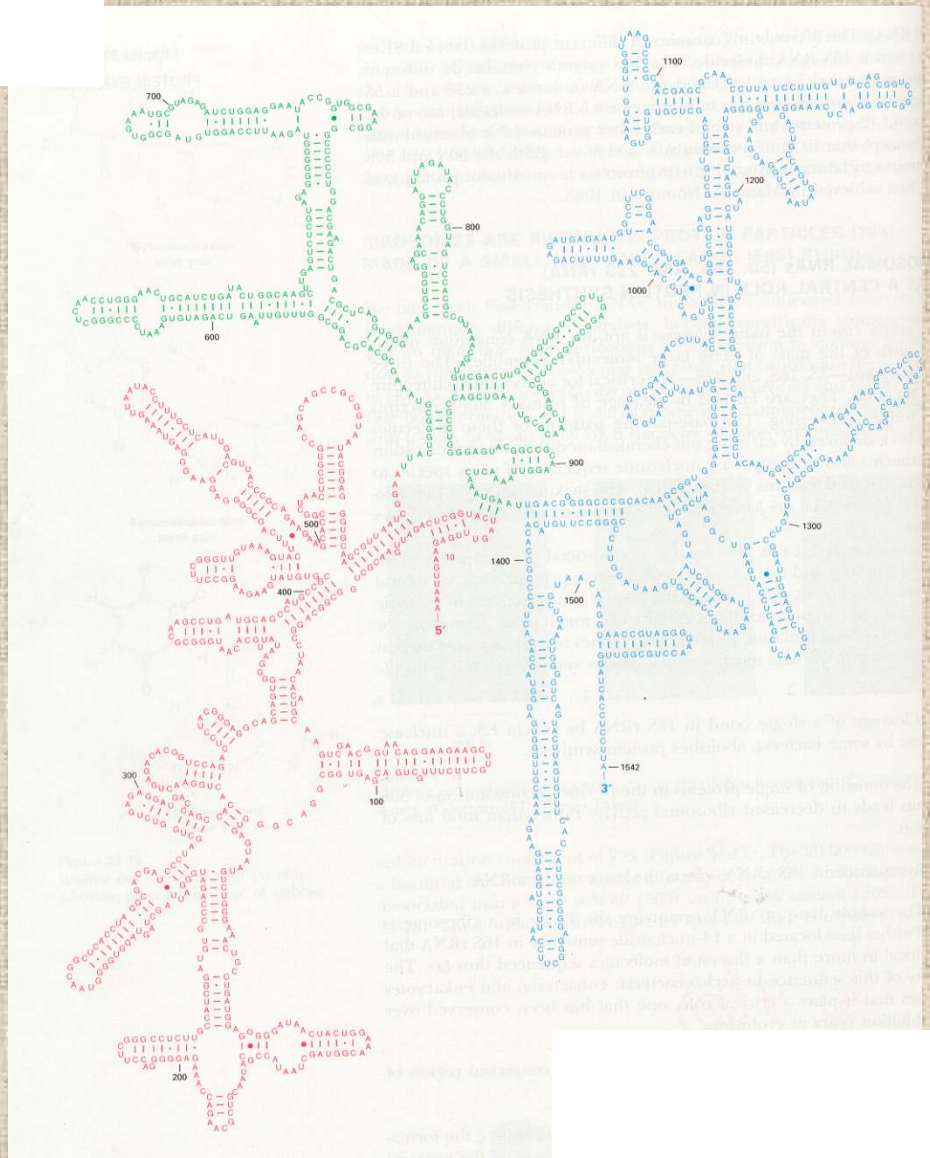

aa-tRNA

СТРУКТУРА РИБОСОМ

Субъединицы состоят из рРНК и белков, соединяются только для трансляции



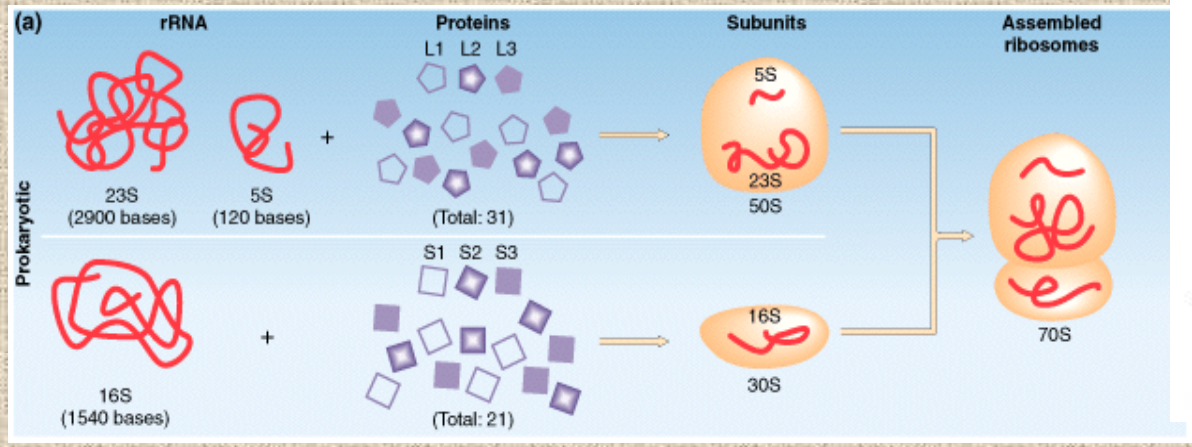
Компоненты рибосомы

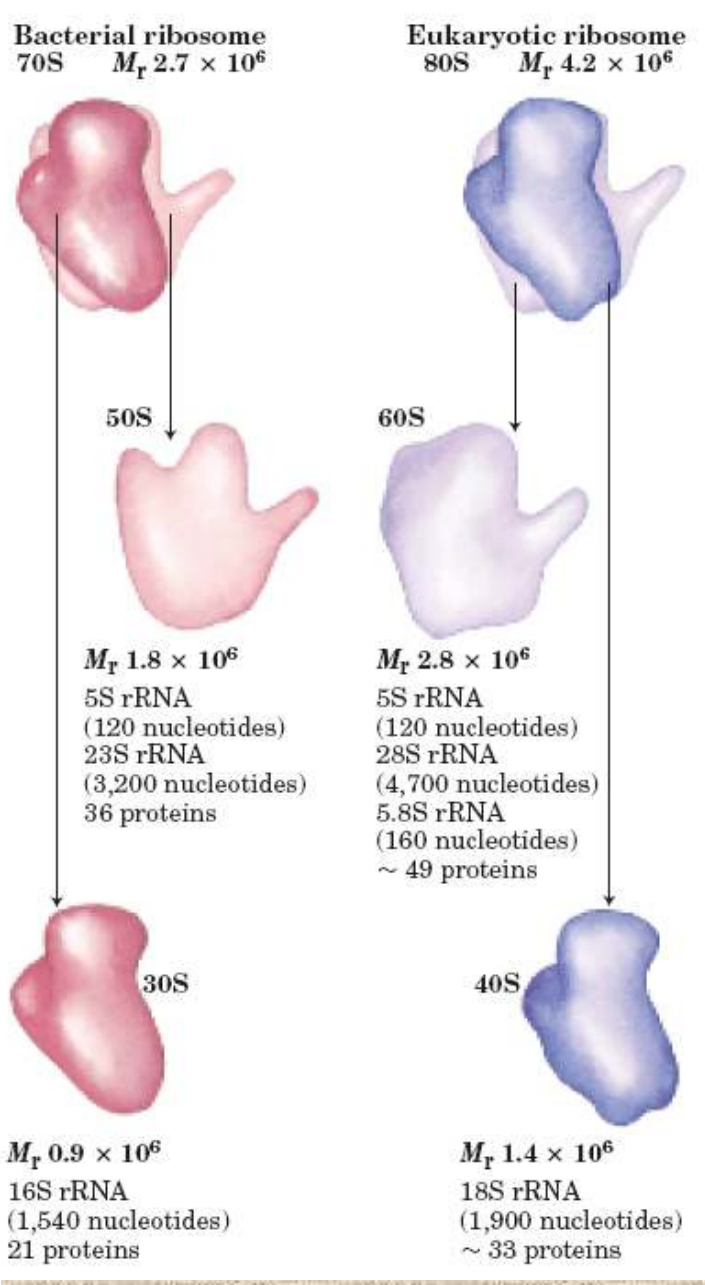


- Малая субъединица
- **16 S рРНК у прокариот** и 18S рРНК у эукариот
- Ее последовательность используется для установления эволюционного родства между организмами
- + белки

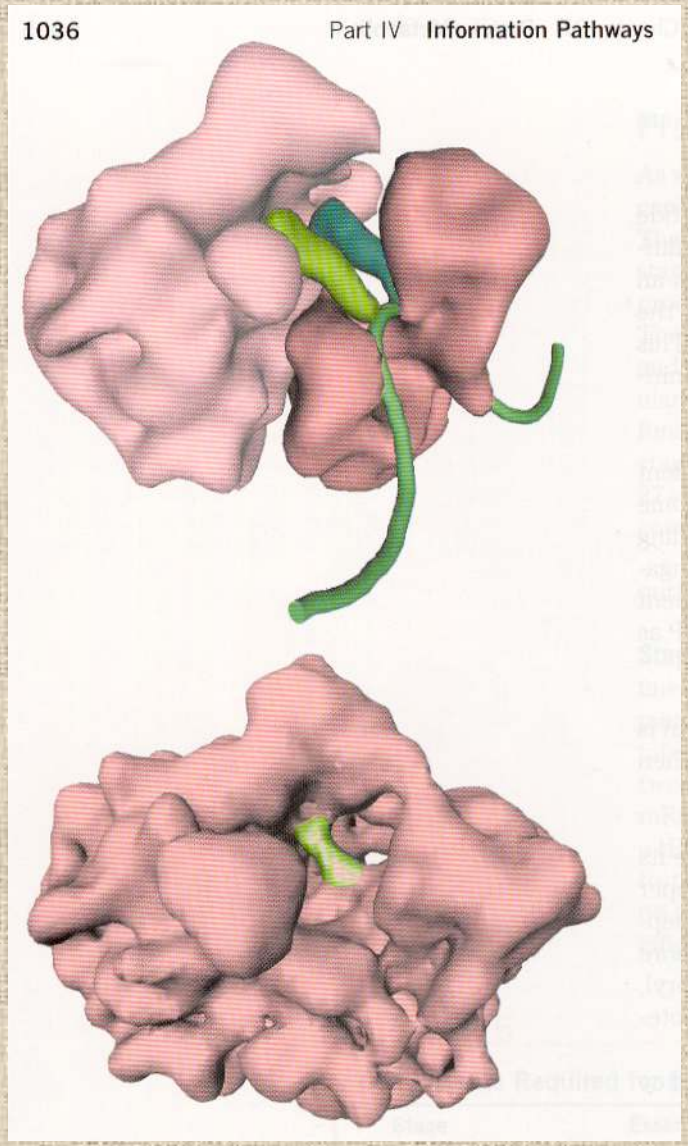
Компоненты рибосомы II

- Большая субъединица (на примере прокариот)
  - 23 S rRNA
  - 5 S rRNA
  - 26 белков





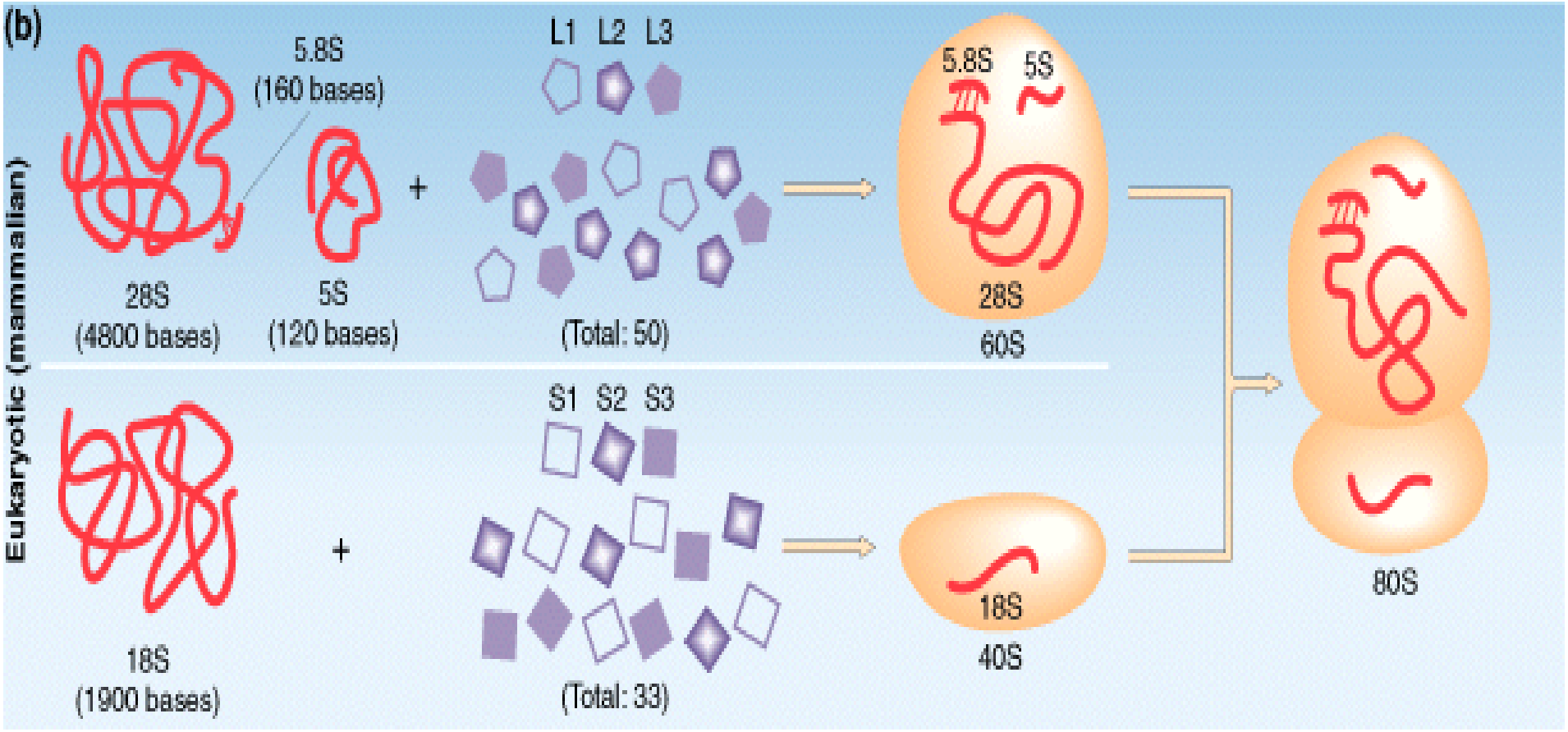
Бороздка между субъединицами



- Место связи анти-кодон-кодон
- Пептидная цепь выходит оттуда тоже

Эукариотические рибосомы

- 40S + 60S
  - **18S**, 28S + 5.8S rRNA +5S

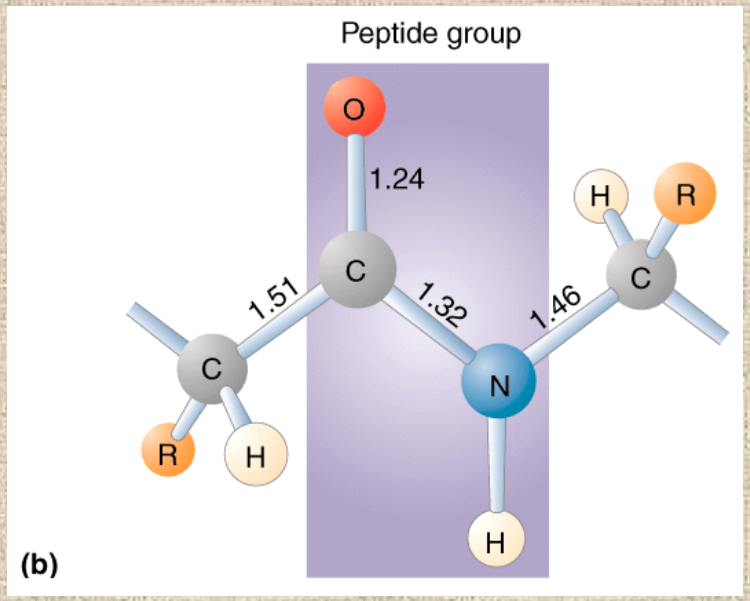


Участки связывания тРНК на рибосоме

- **A-сайт**: узнавание аминоацил-тРНК;
  позиционирование аминоацила.

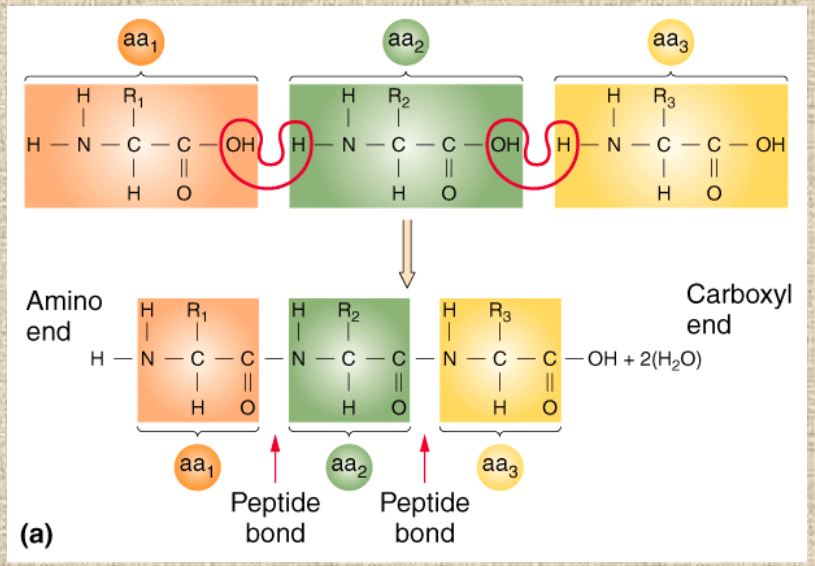
- **P-сайт:** связывание инициаторной fMet-тРНК;
  связывание пептидил-тРНК
  при транслокации.

- **E-сайт:** выход деацилированной тРНК с
  50S субчастицы.



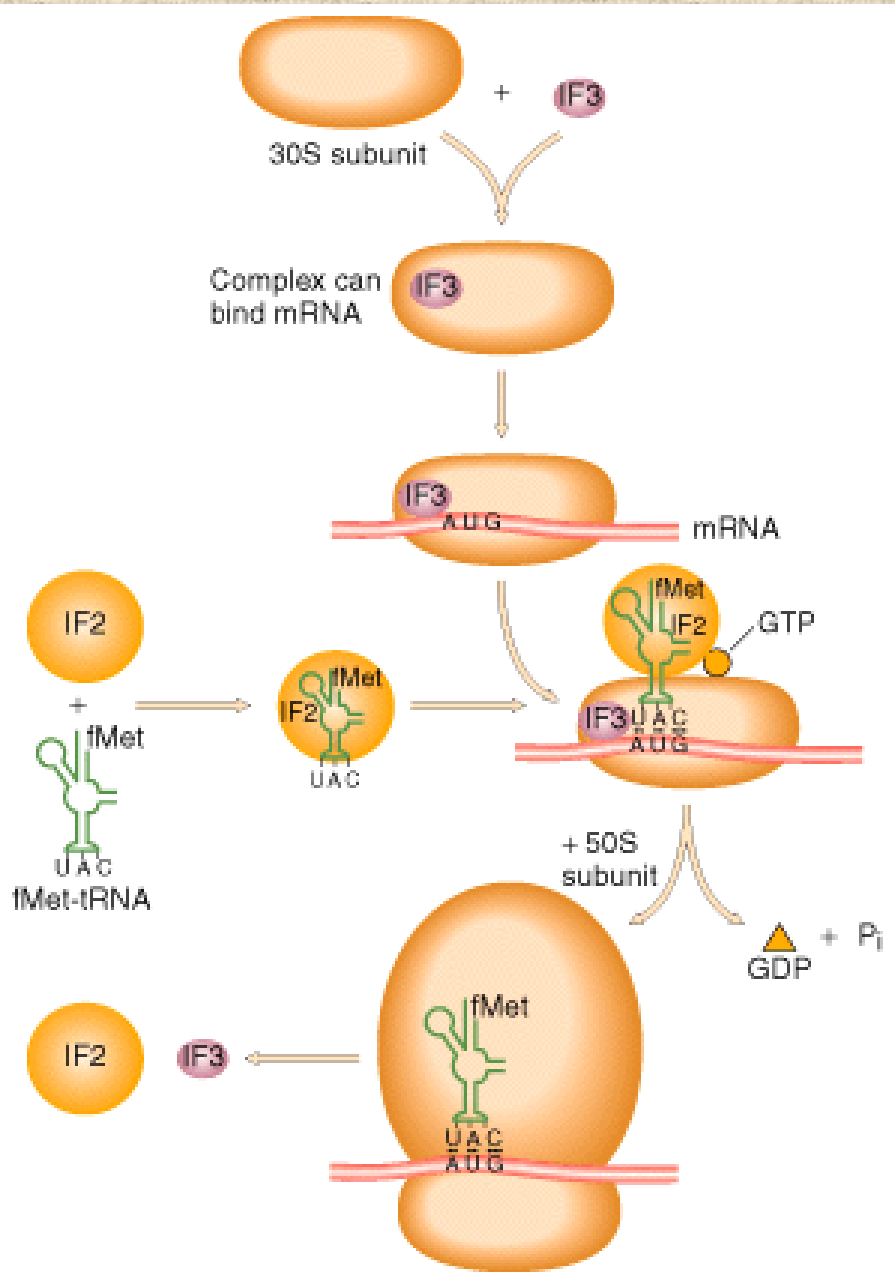
Пептидная связь

- Костяк белка
- Функция белка определяется боковыми цепями аминокислот его составляющих



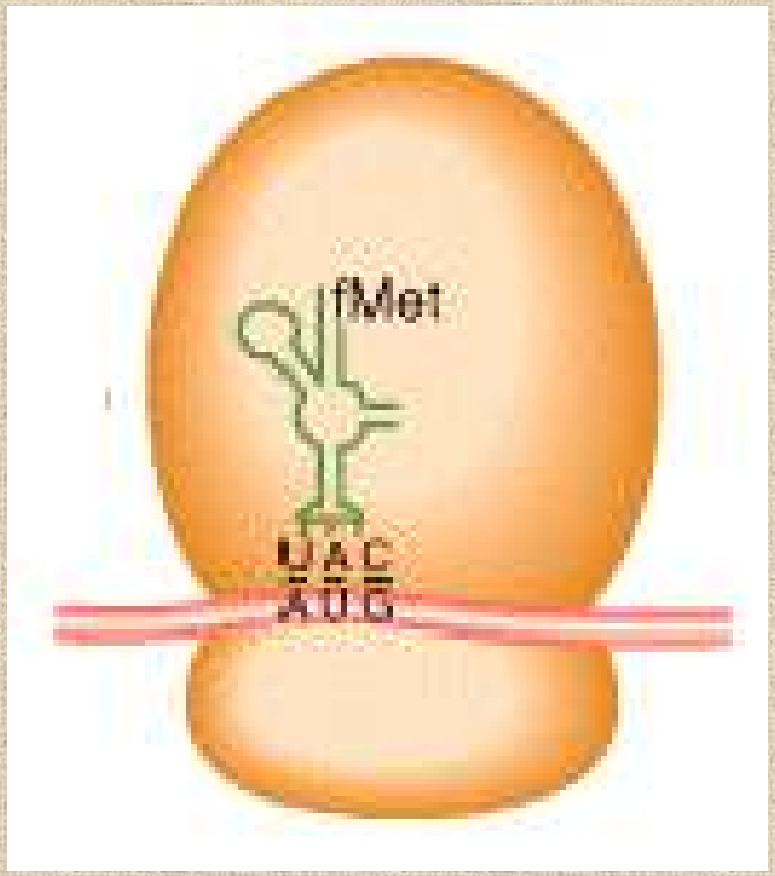
Инициация

- Сборка рибосомы
  - 70 S из 30S и 50S субъединиц у прокариот
  - 40S + 60S = 80S у эукариот
    - Требует факторов инициации
    - Требует (ф)метионил-тРНК

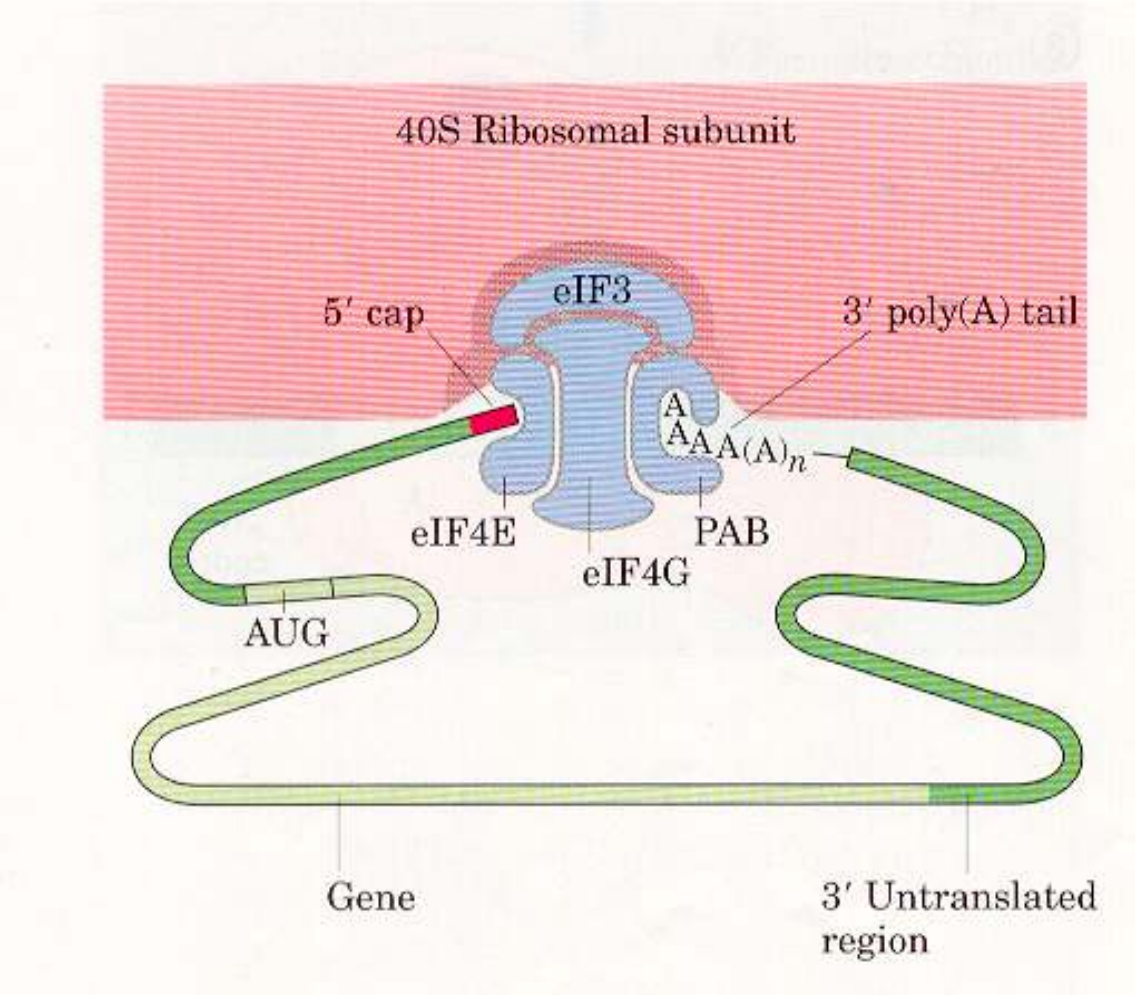


Комплекс инициации у прокариот

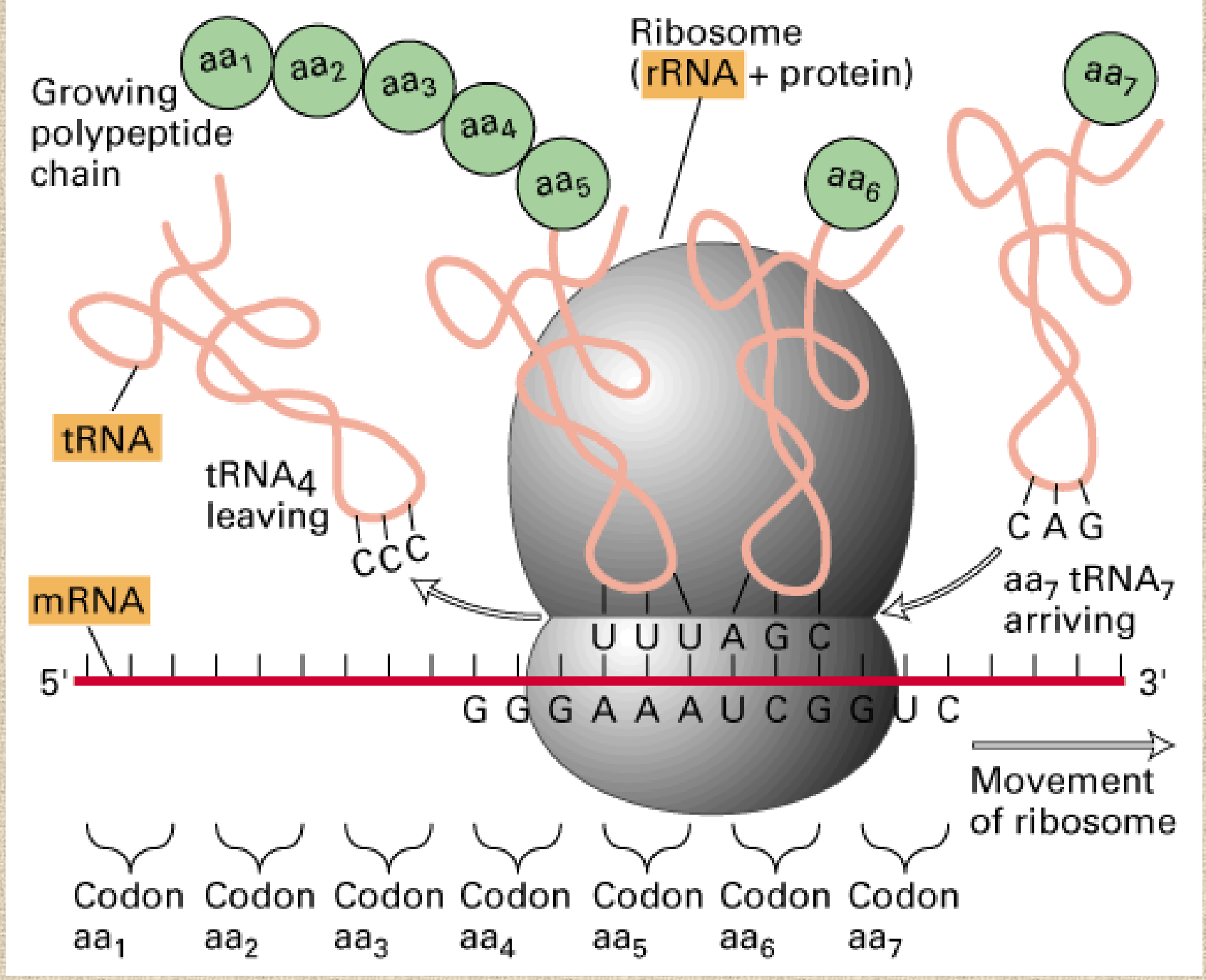
- Требуются факторы инициации
- Последовательность **Шайна-Дальгарно** (Shine-Dalgarno sequence,) — была описана австралийскими учеными Джоном Шайном и Линн Дальгарно и является сайтом связывания рибосом на молекуле мРНК прокариот, обычно на расстоянии около 10 нуклеотидов до стартового кодона AUG



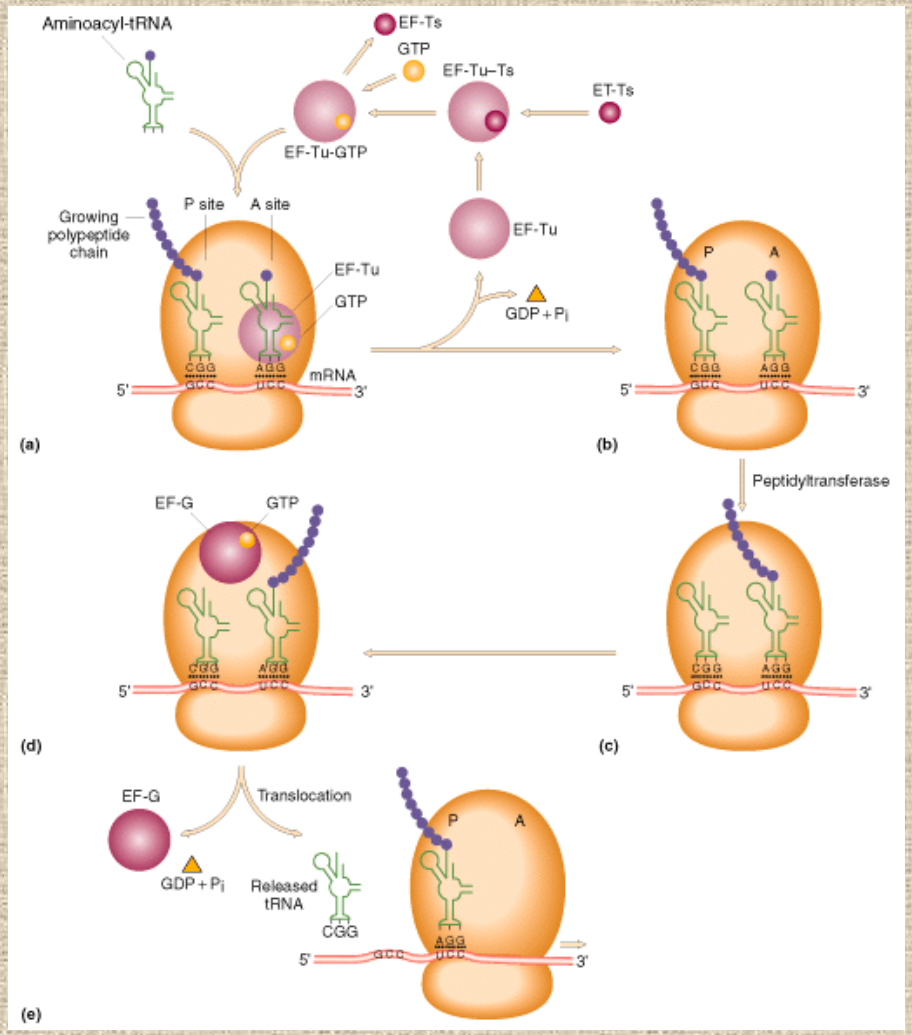
Эукариотическая инициация



- Эукариотичеккая мРНК имеет специальные белки на 5 ’ и 3 ’ конце (связываются с 5’ CAP и 3’ poly A)
- Первый AUG после CAP – начало трансляции


Growing polypeptide chain
aa1
aa2
aa3
aa4
aa5
Ribosome (rRNA + protein)
aa6
aa7
tRNA
tRNA4 leaving
C C C
C A G
aa7 tRNA7 arriving
mRNA
5'
3'
U U U A G C
G G G A A A U C G G U C
Movement of ribosome
Codon aa1
Codon aa2
Codon aa3
Codon aa4
Codon aa5
Codon aa6
Codon aa7

Элонгация и терминация



- Элонгация
  - Аминоацилированная тРНК устанавливается на рибосоме в месте кодона
  - Пептид переносится на эту тРНК
  - Цикл повторяется
- Терминация
  - Стоп кодон соединяется со специальным белком вместо тРНК

Терминация трансляции

PLAY

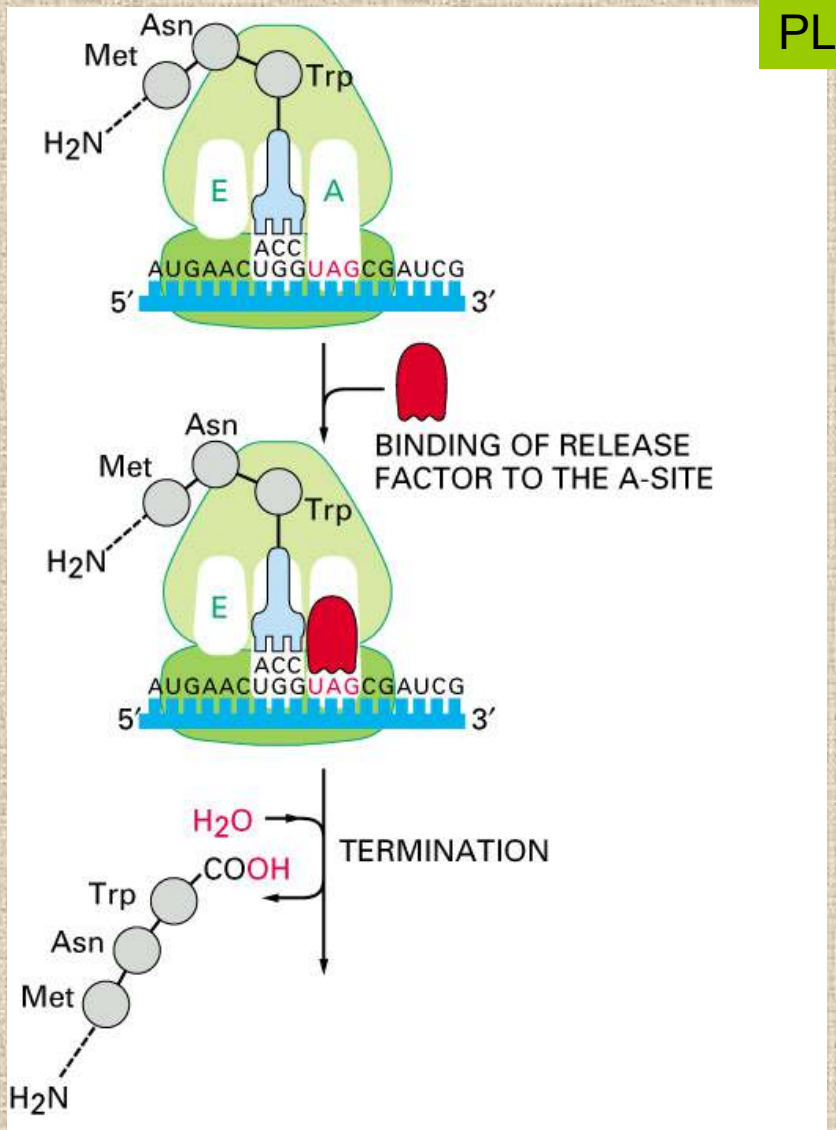


Figure 6–73 part 1 of 2. Molecular Biology of the Cell, 4th Edition.

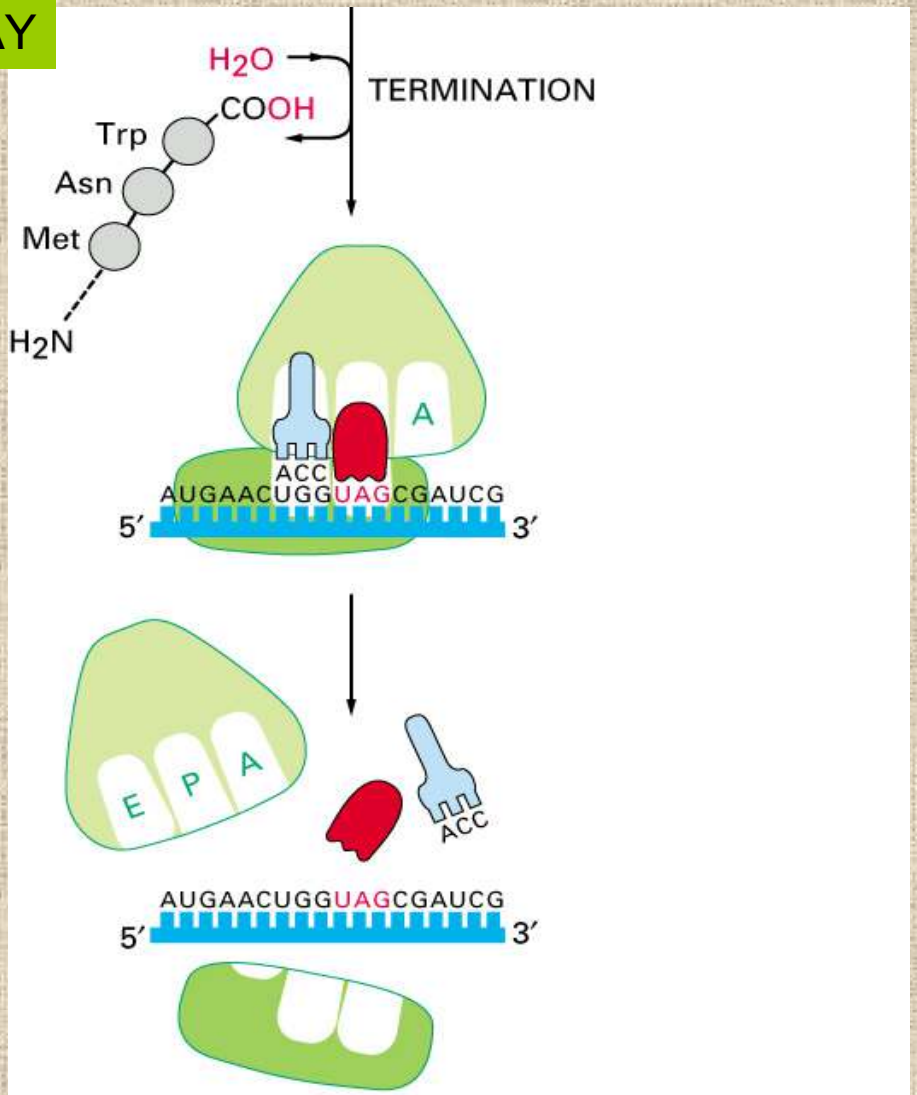
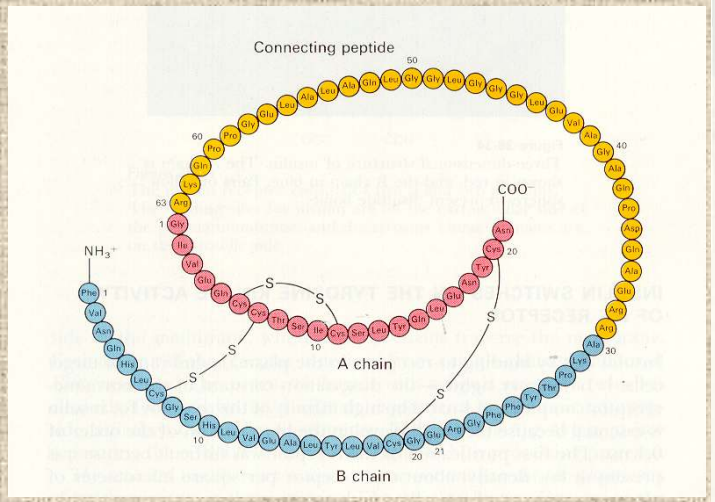
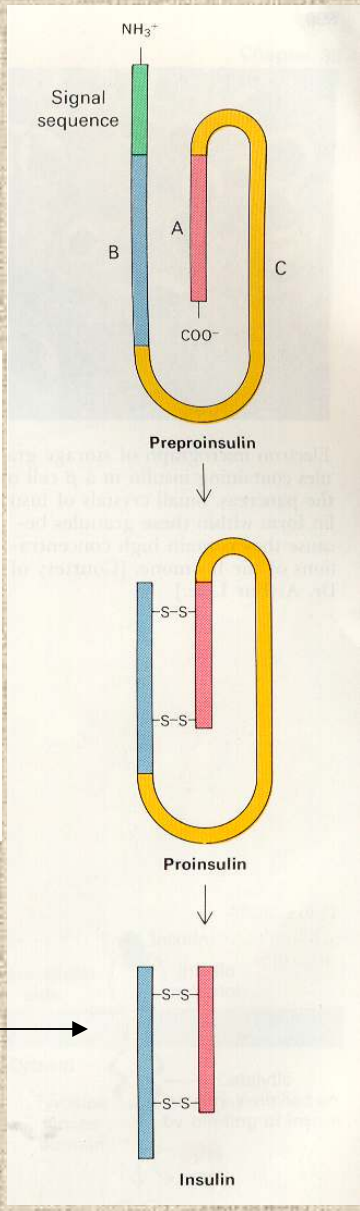


Figure 6–73 part 2 of 2. Molecular Biology of the Cell, 4th Edition.

Процессинг

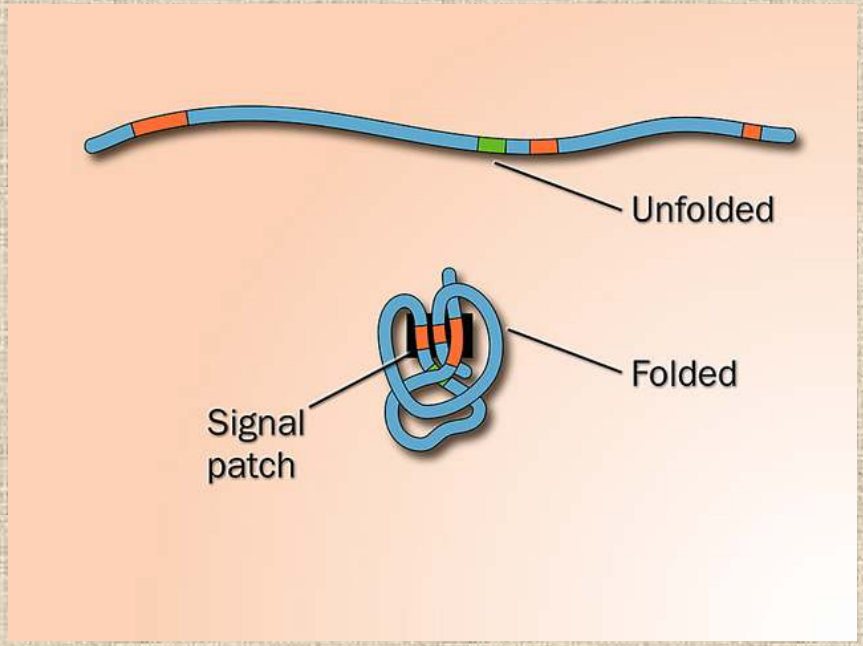
- Начинается с момента синтеза белка
- Требует специальных ферментов
- Может происходить разрезание пептидной цепи с образованием конечного продукта (например, инсулин)



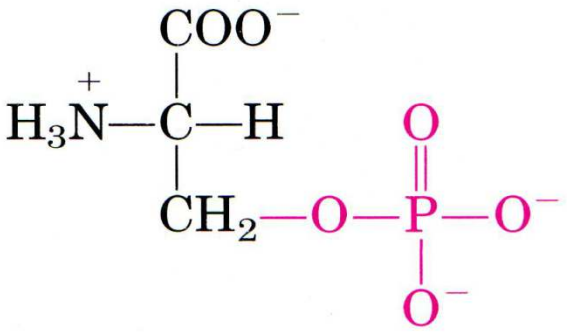


Посттрансляционные модификации белков

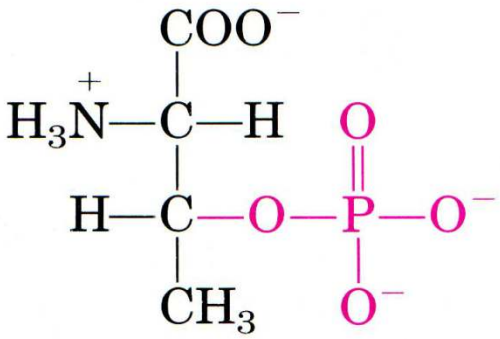
- Формирование третичной структуры – фолдинг белка
- Удаление сигнальной последовательности
- Модификации N- и C-концов (удаление fMet, N-ацетилирование)
- Ковалентные модификации аминокислотных остатков



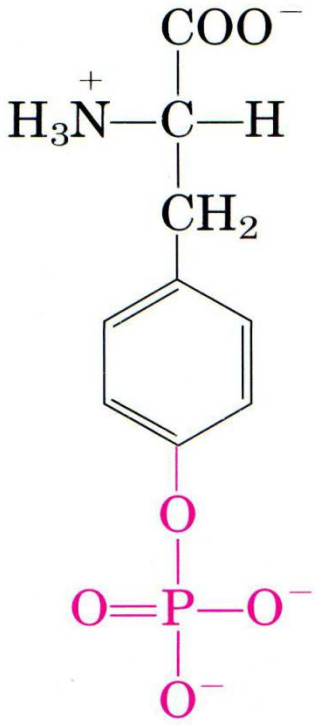
Ковалентные модификации аминокислотных остатков



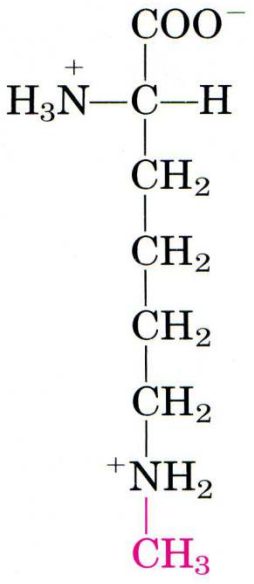
Phosphoserine



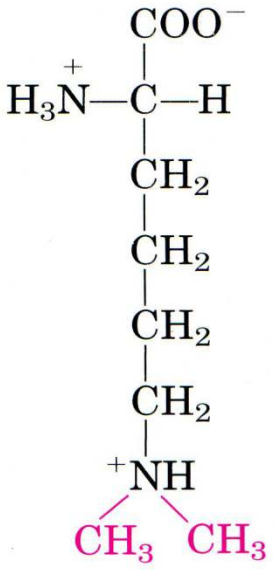
Phosphothreonine



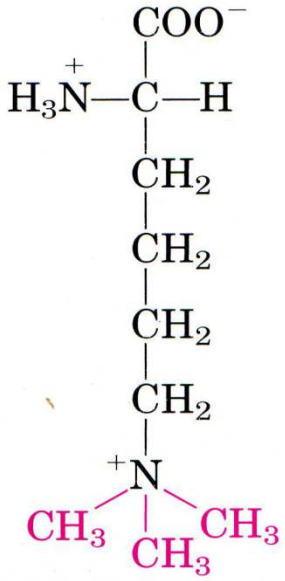
Phosphotyrosine



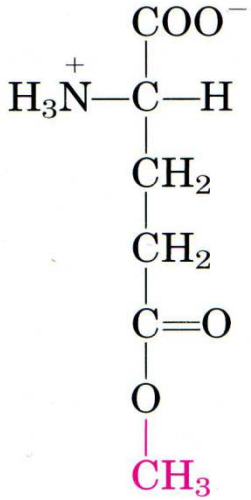
Methyllysine



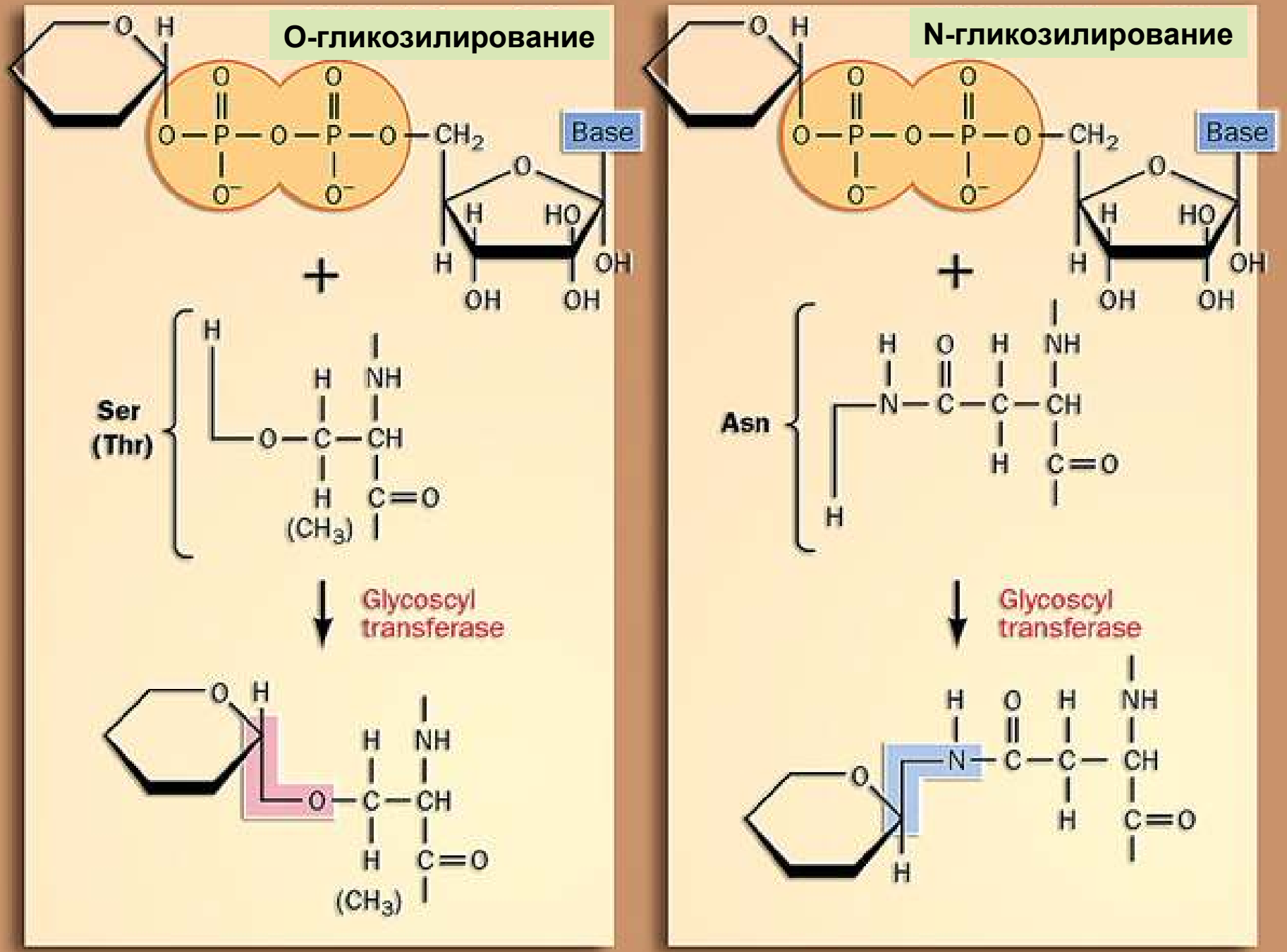
Dimethyllysine



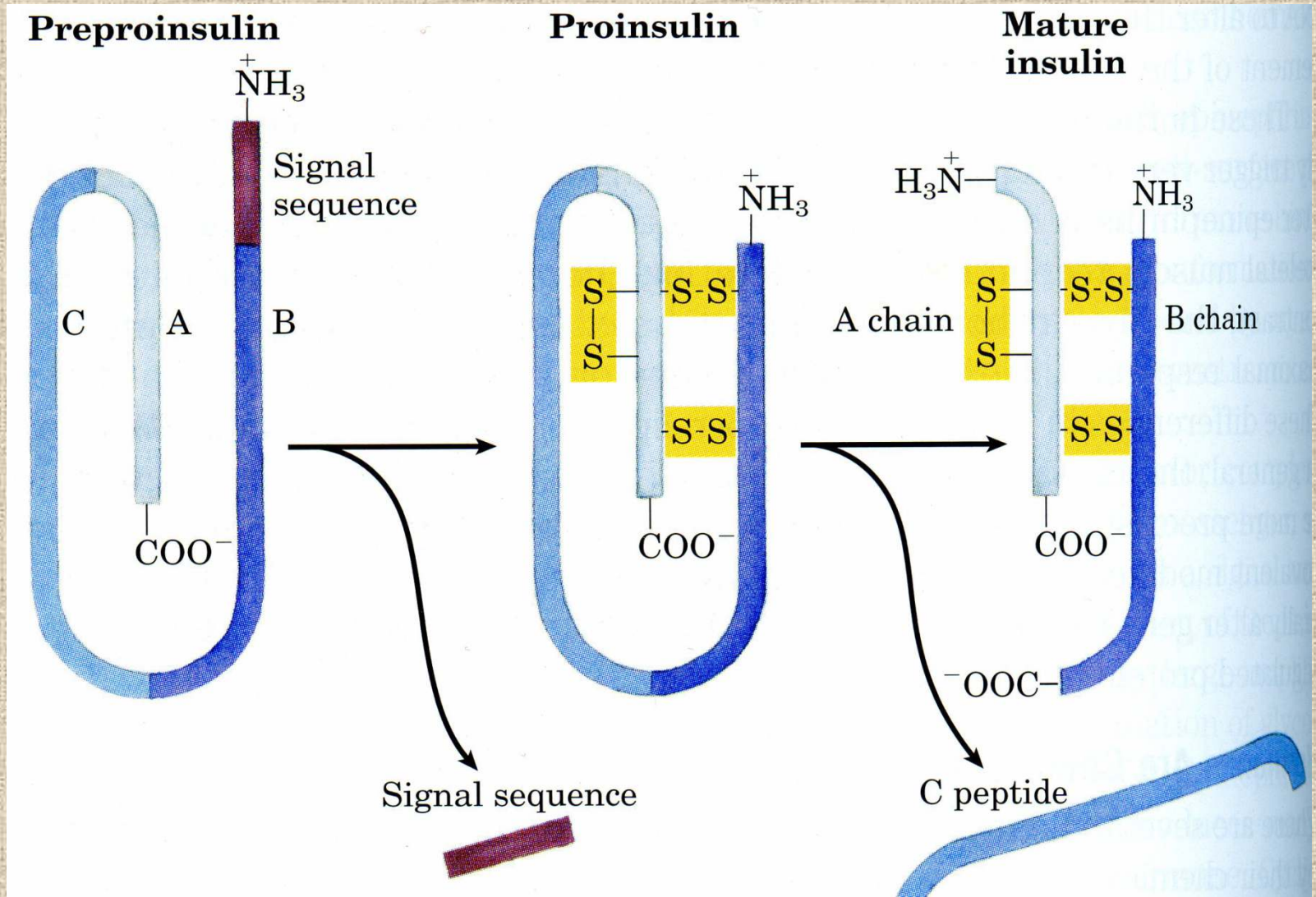
Trimethyllysine



Methylglutamate


O-гликозилирование
N-гликозилирование
Base
Ser
(Thr)
Asn
Glycoscyl
transferase
Glycoscyl
transferase

Ограниченный протеолиз; добавление дисульфидных связей



Основные этапы трансляции

АКТИВАЦИЯ АМИНОКИСЛОТ → **ИНИЦИАЦИЯ** → **ЭЛОНГАЦИЯ** → **ТЕРМИНАЦИЯ**

тРНК	мРНК; инициирующий кодон	мРНК	Терминальный кодон в мРНК
Аминокислоты	30S субъединица; 50S субъединица	Рибосома 70S	Белковые факторы терминации (RF)
ATP, Mg^{2+}	Белковые факторы инициации (IF-1, 2, 3)	Белковые факторы элонгации (Tu,Ts, G)	ATP
Амиоацил-тРНК-синтетазы	fMet-тРНК	Аминоацил-тРНК	
	GTP, Mg^{2+}	GTP	

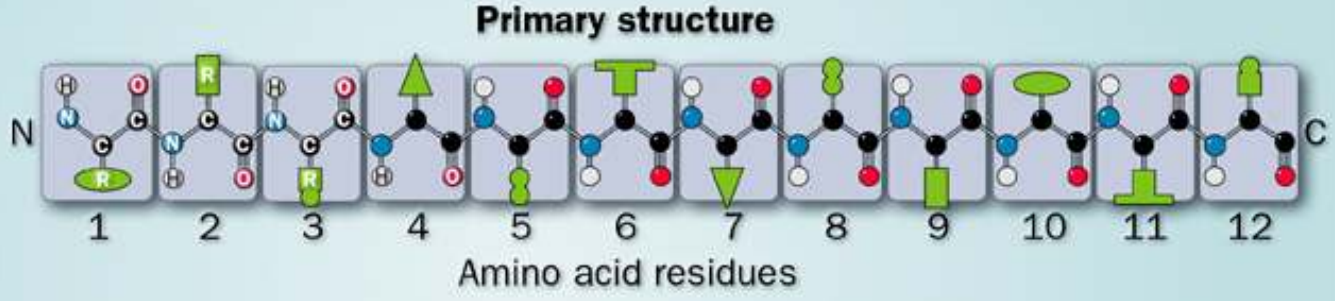
ПРОЦЕССИНГ БЕЛКА

(фолдинг, ограниченный протеолиз, ковалентные модификации)

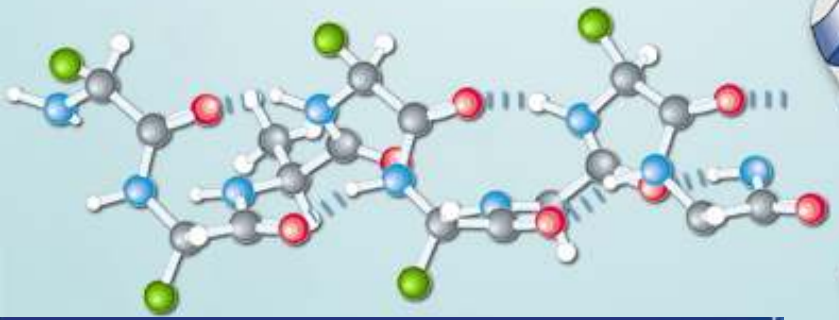
Уровни структурной организации белков

- Первичный – последовательность АК
- Вторичный – альфа-спираль или бета-слой
- Третичный – скручивание вторичной структуры в глобулу
- Четвертичный – соединение несколький полипептидных цепей в особую пространственную структуру

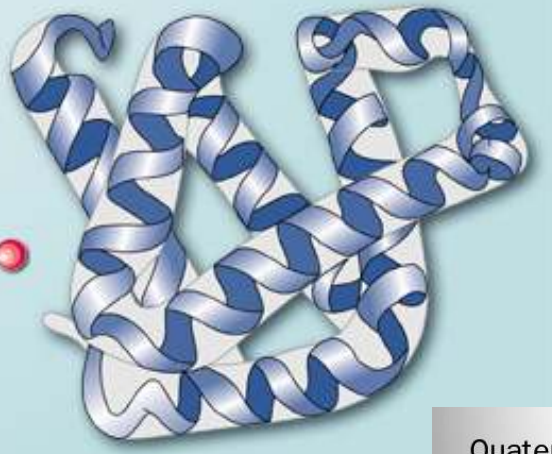
ИЕРАРХИЯ БЕЛКОВОЙ СТРУКТУРЫ

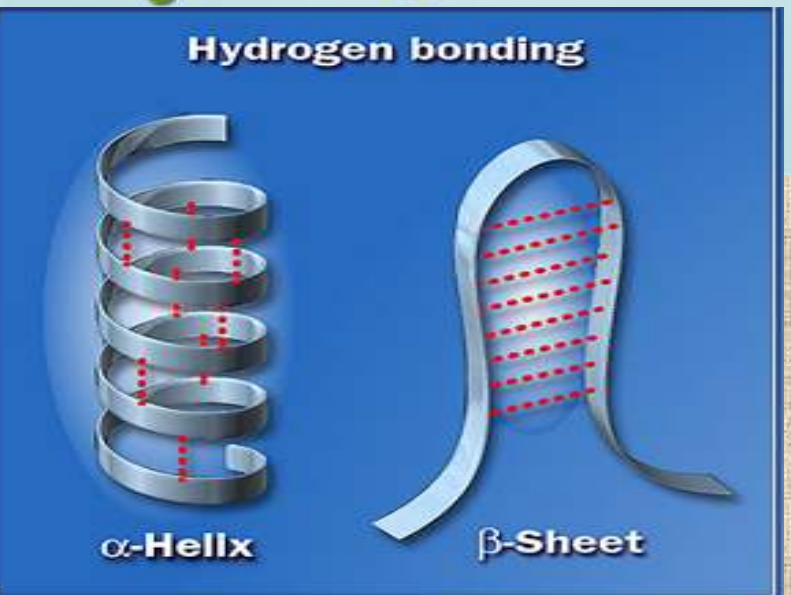


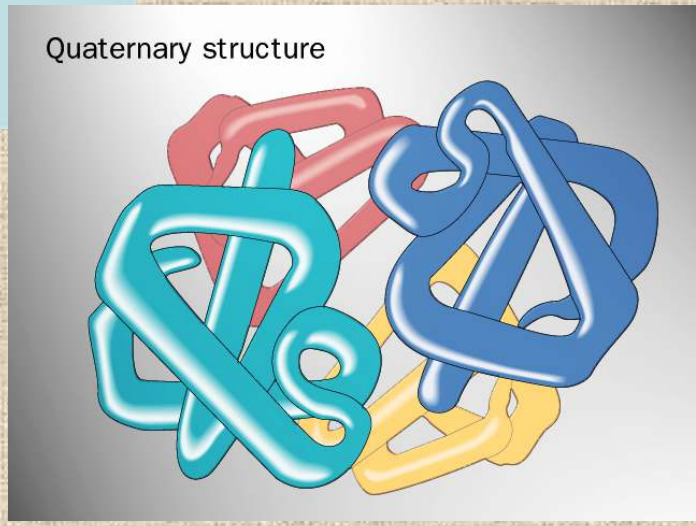
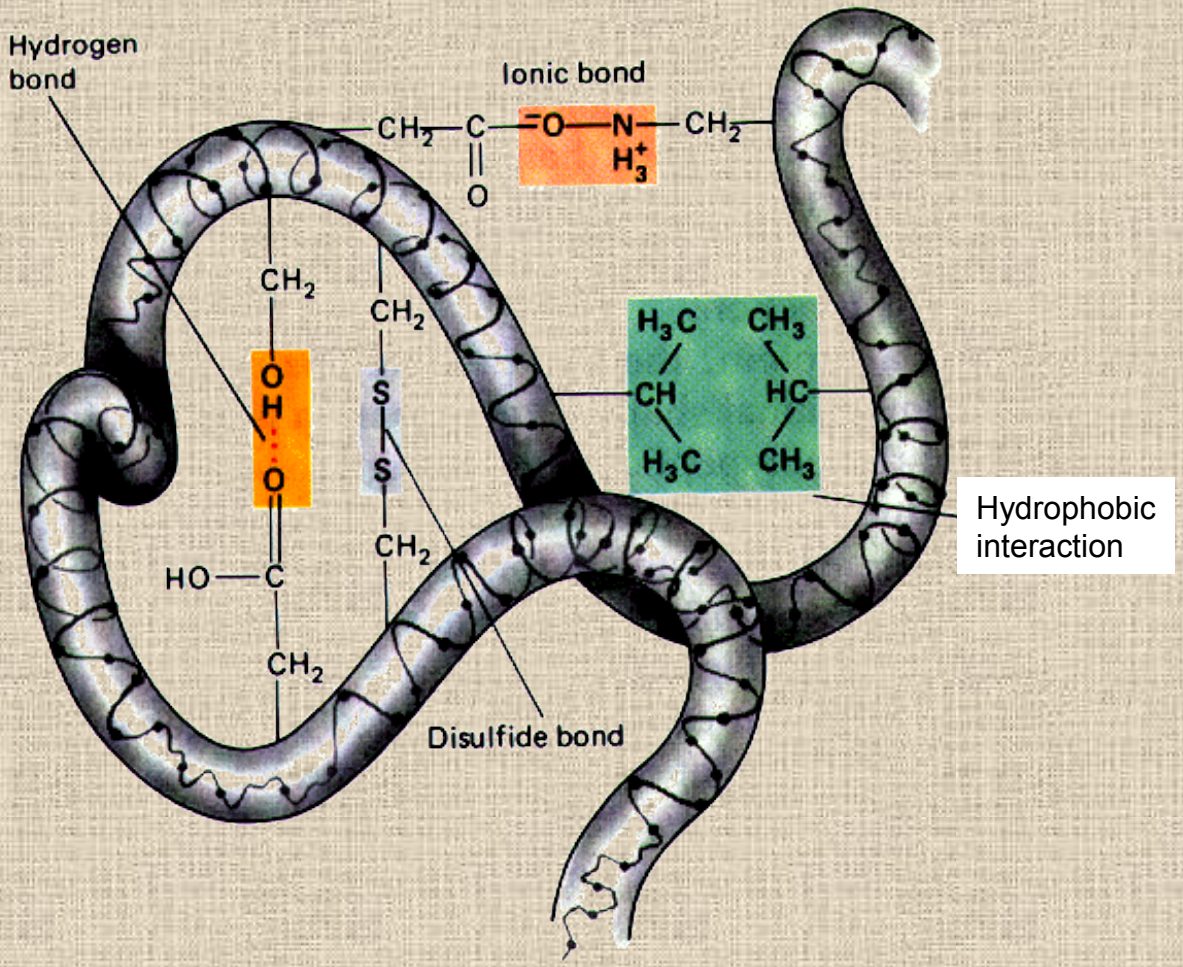
Secondary structure



Tertiary structure







Примеры нековалентных взаимодействий между аминокислотными остатками в третичной структуре