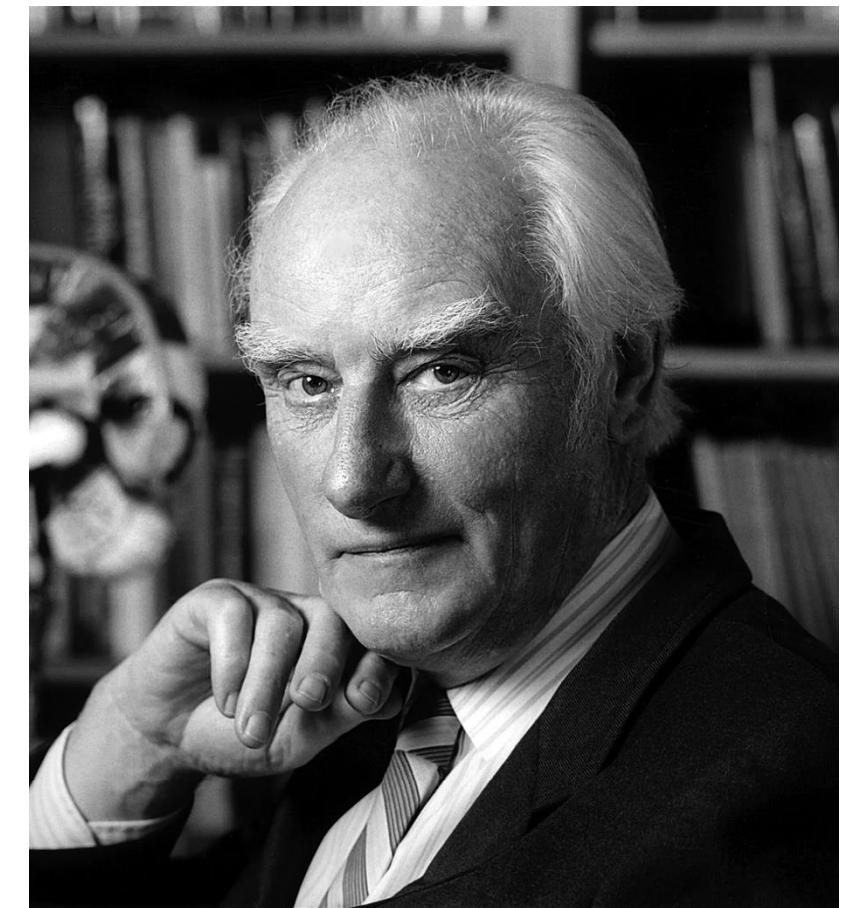
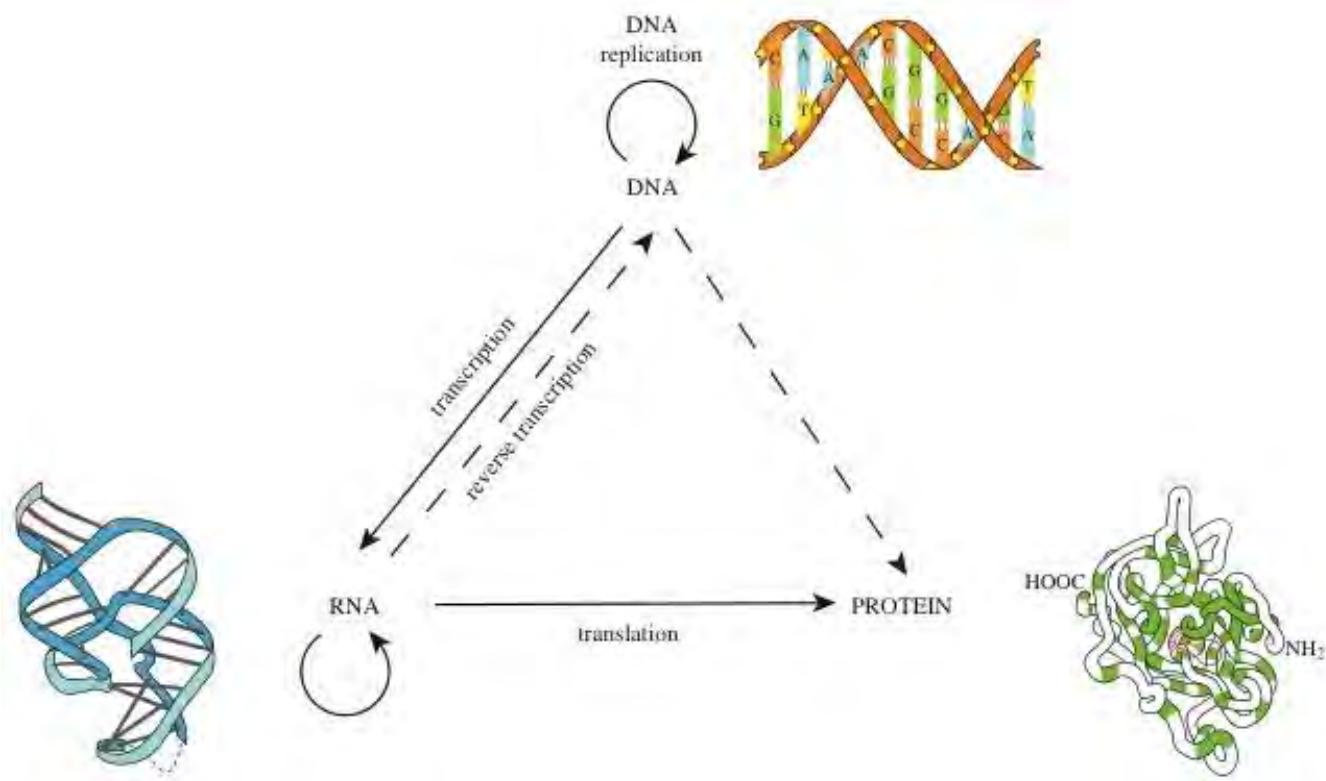


# Транскрипция, процессинг РНК и трансляция

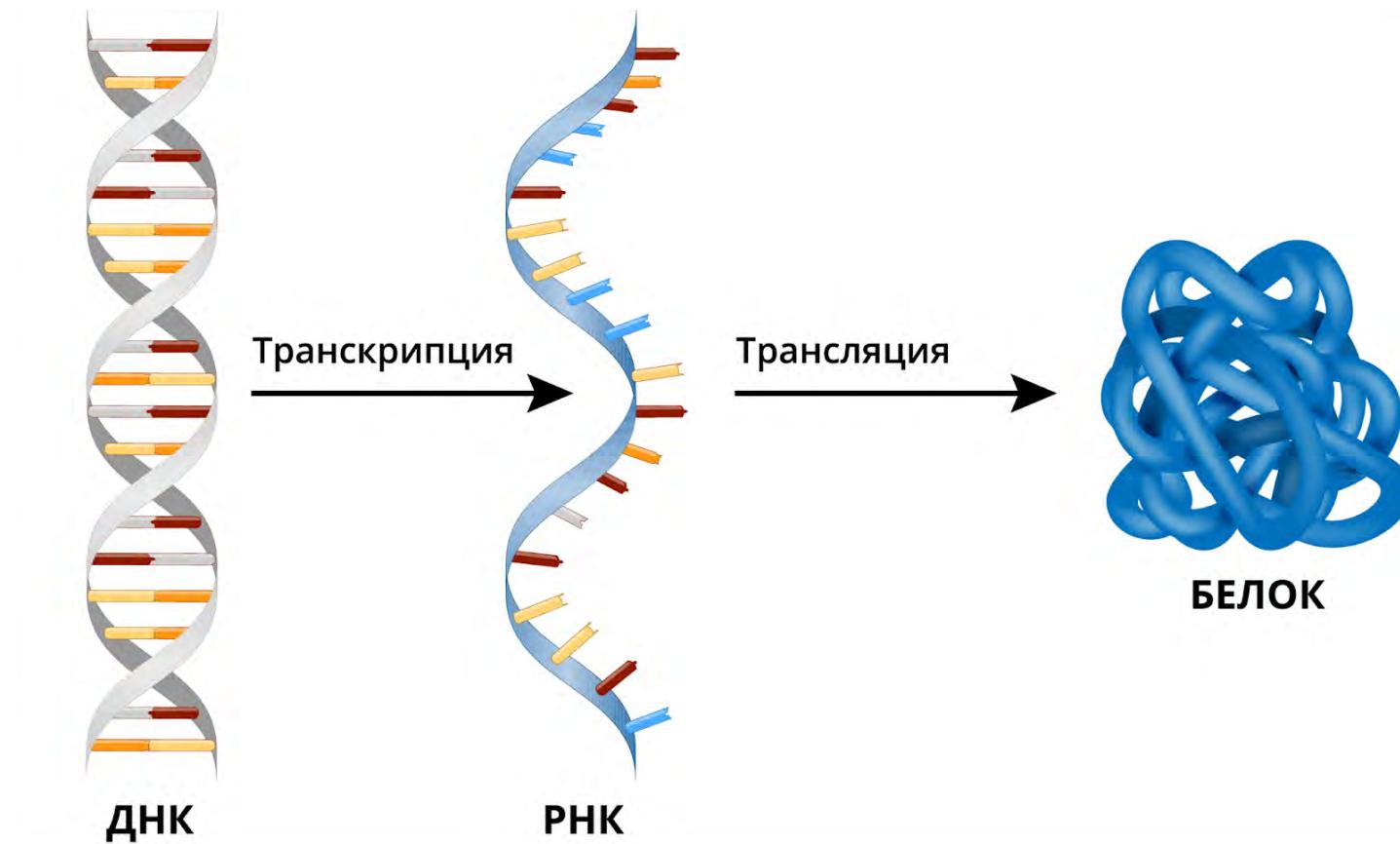
# План лекции

1. Механизм транскрипции
2. Посттранскрипционные модификации РНК
3. Генетический код и его свойства
4. Трансляция белка
5. Посттрансляционные модификации белков
6. Методы анализа транскрипции
7. Методы исследования посттранскрипционных модификаций
8. Методы анализа трансляции белков

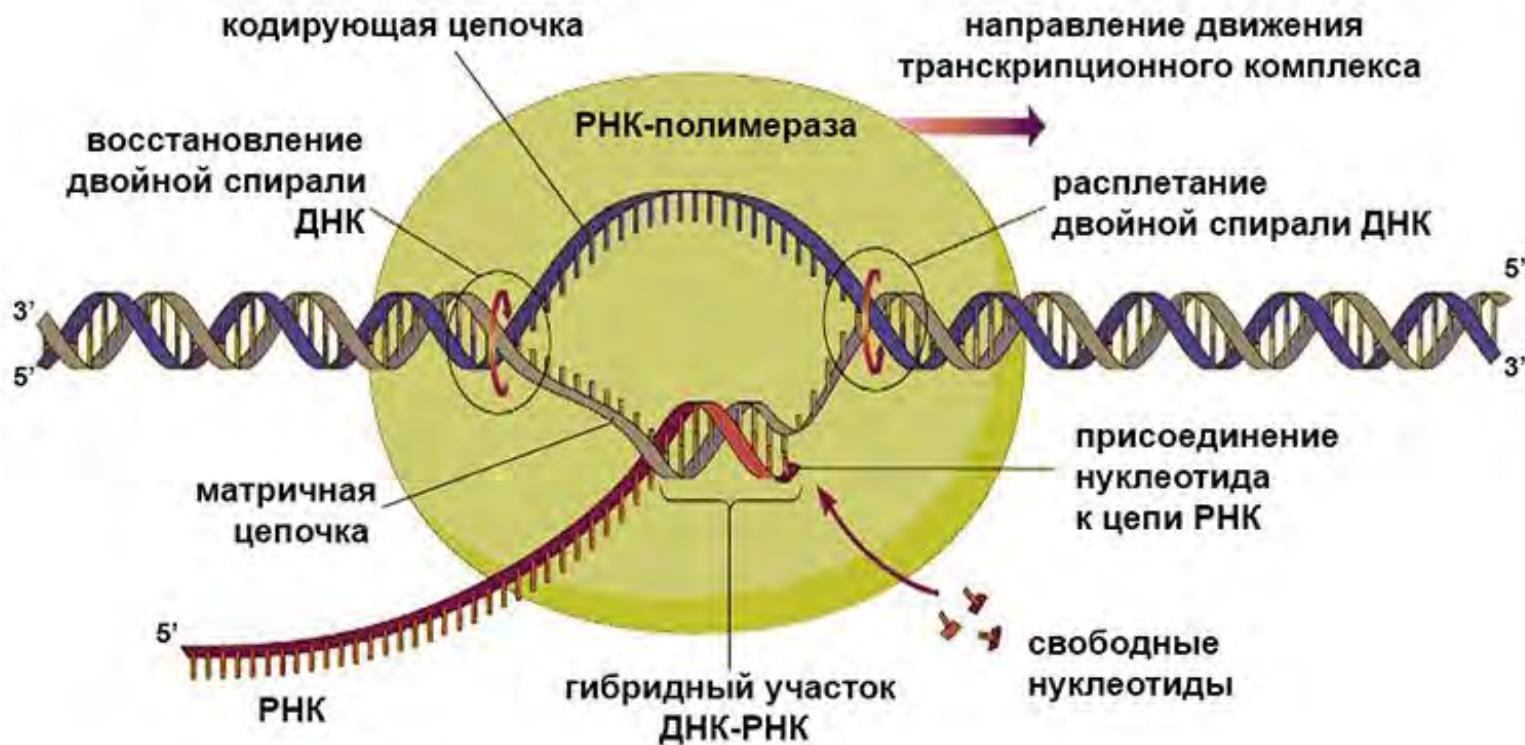
# Центральная догма молекулярной биологии

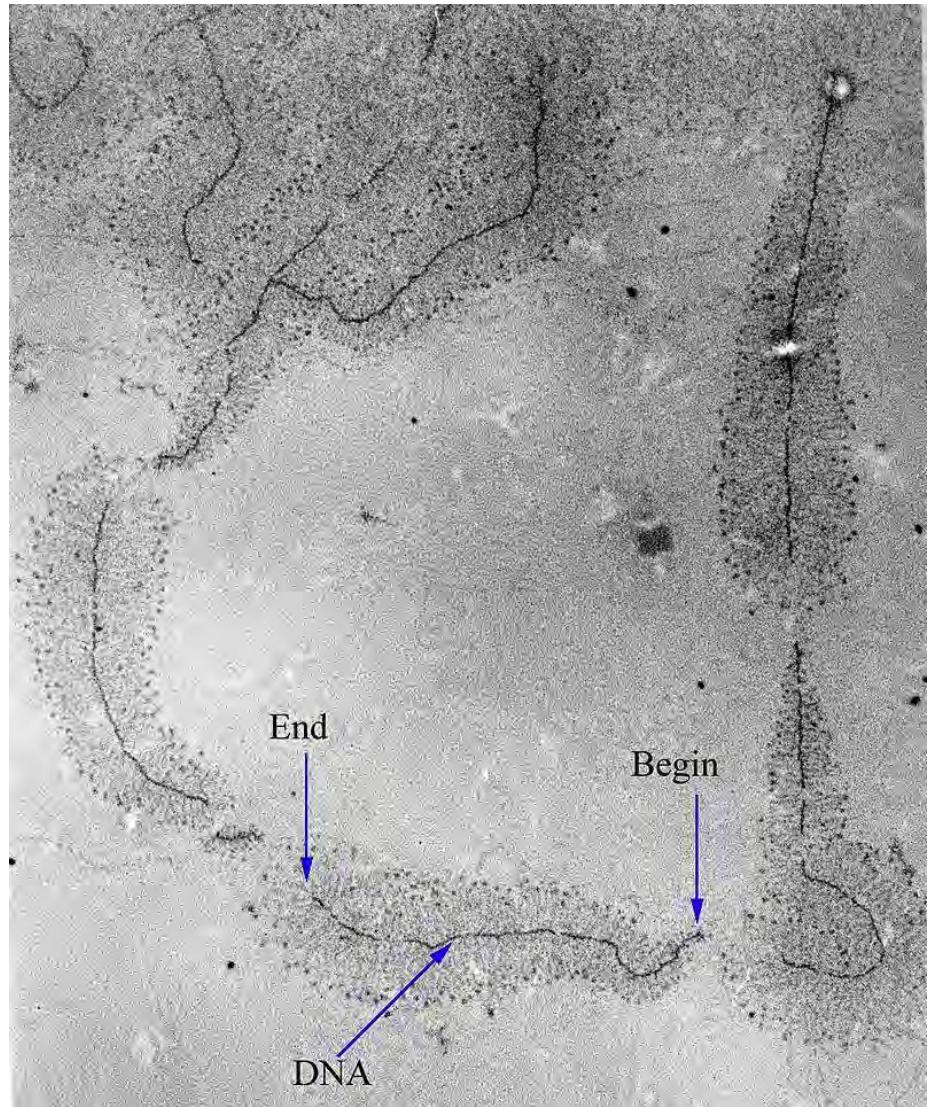


# Транскрипция – синтез РНК на матрице ДНК

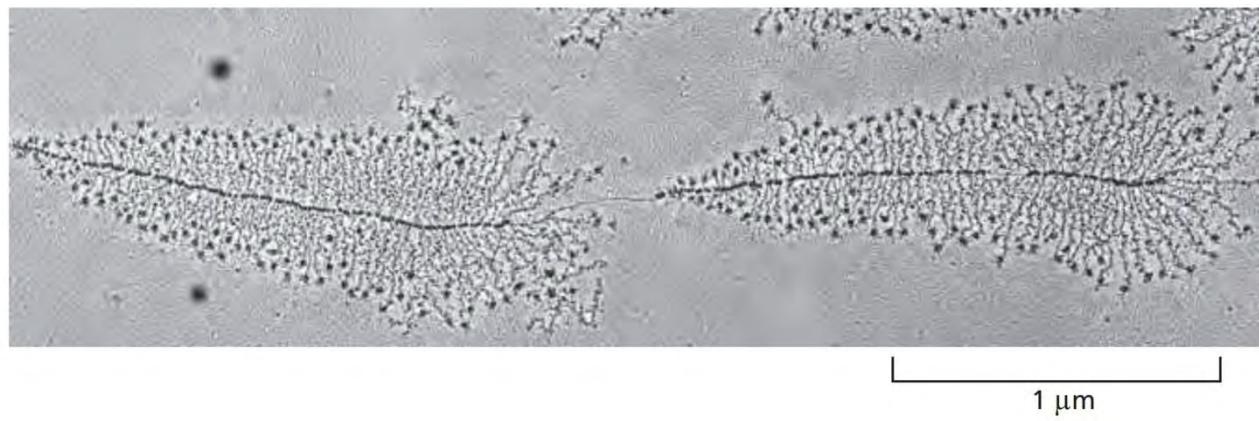


# Транскрипция – синтез РНК на матрице ДНК

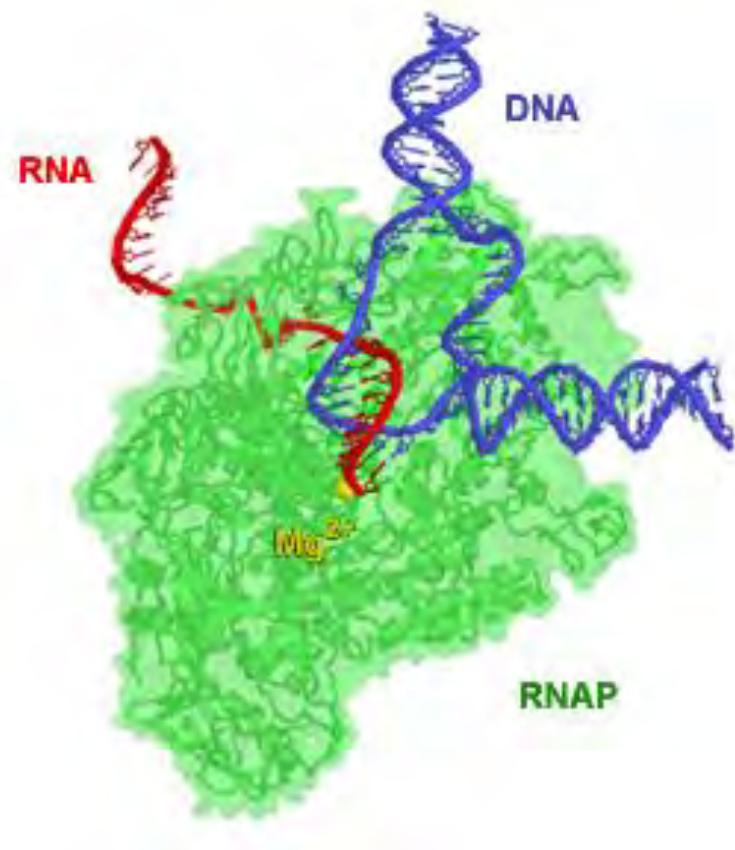
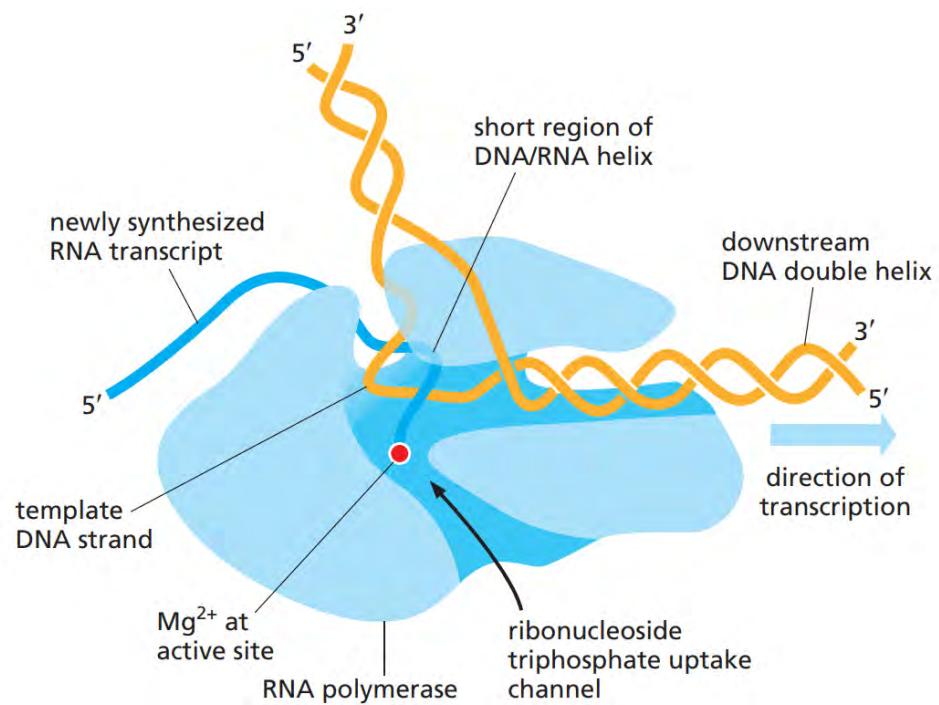




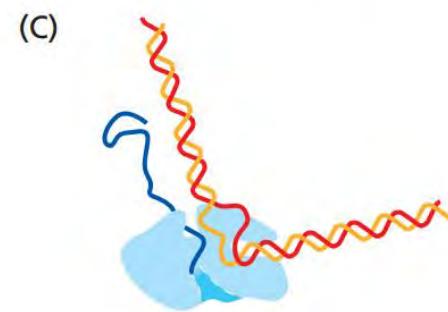
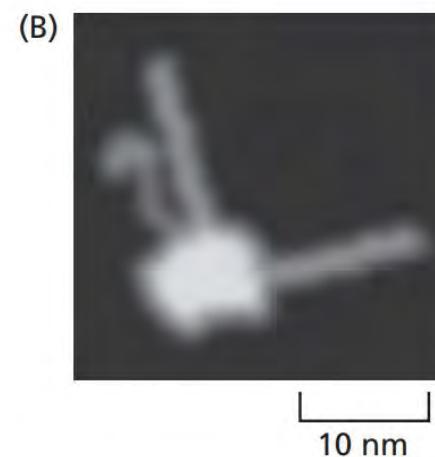
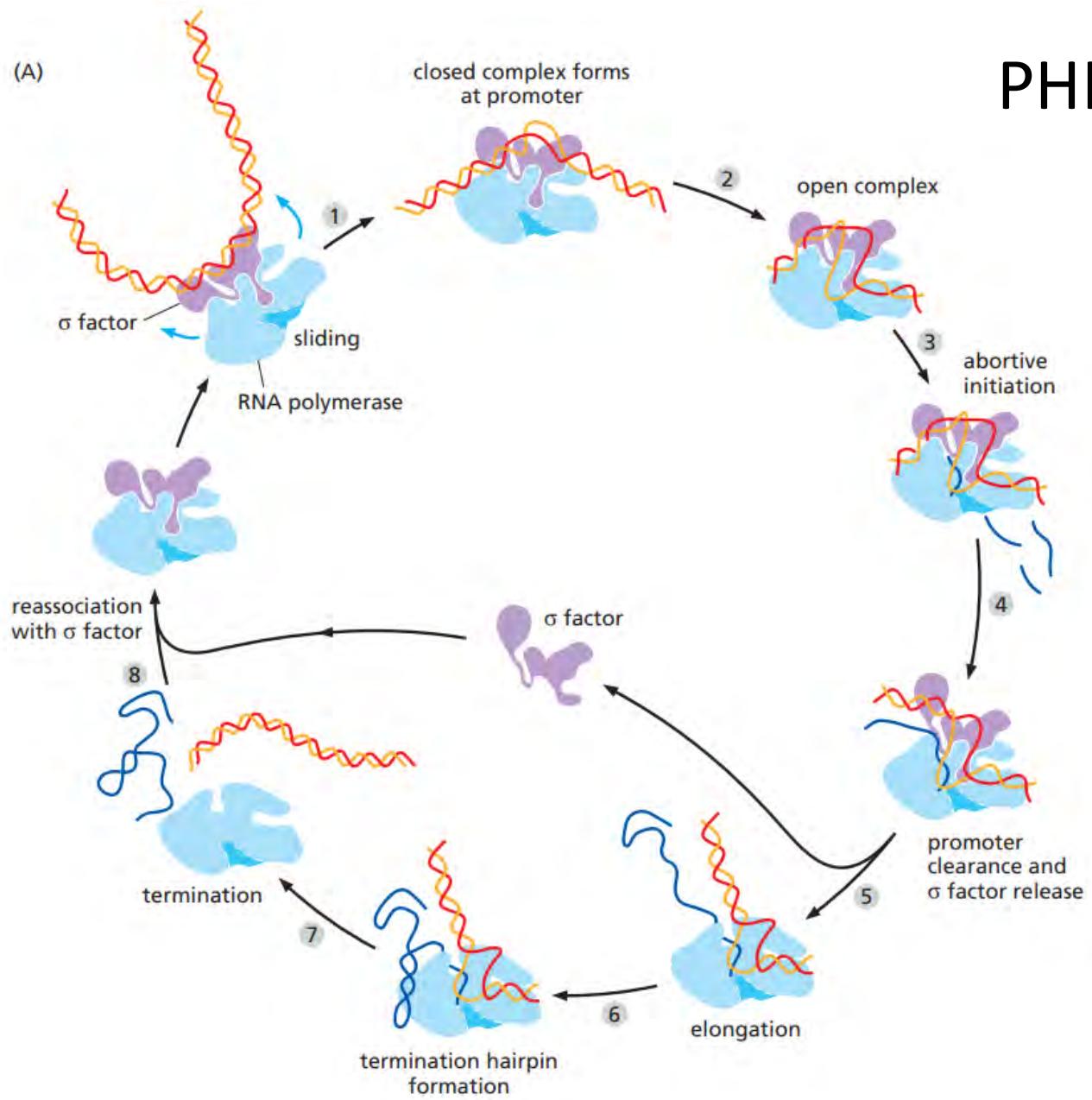
## Транскрипция под электронным микроскопом



# ДНК-зависимая РНК-полимераза



# РНК-полимераза прокариот



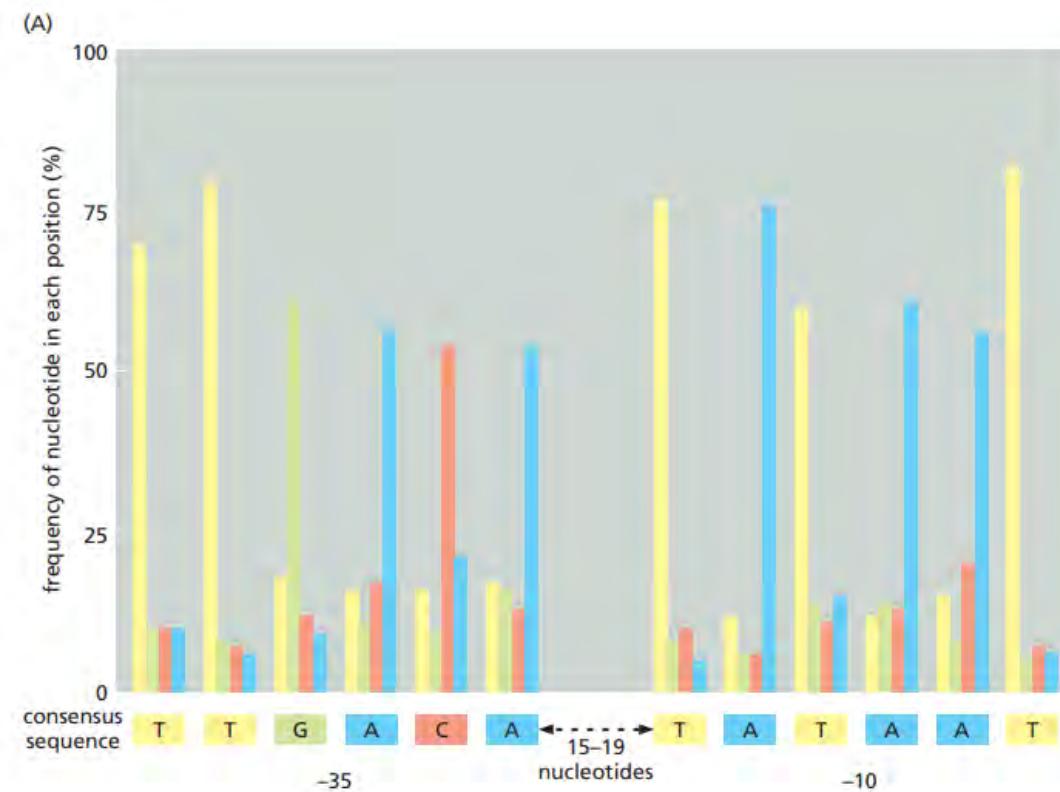
РНК-полимераза производит следующие разновидности РНК:

- **Матричная РНК (мРНК)** — шаблон для синтеза белков в рибосомах.
- **Транспортная РНК (тРНК)**, переносящая аминокислоты к растущей на рибосоме белковой цепочке во время процесса трансляции.
- **Рибосомная РНК (рРНК)**, входящая в состав рибосомы;
- **МикроРНК**, регулирующая активность генов;
- **Катализическая РНК**, обладающая свойствами ферментов.
- **Некодирующая РНК или «РНК-ген»** — большой класс генов, кодирующих РНК, на которых не может быть построено белка.

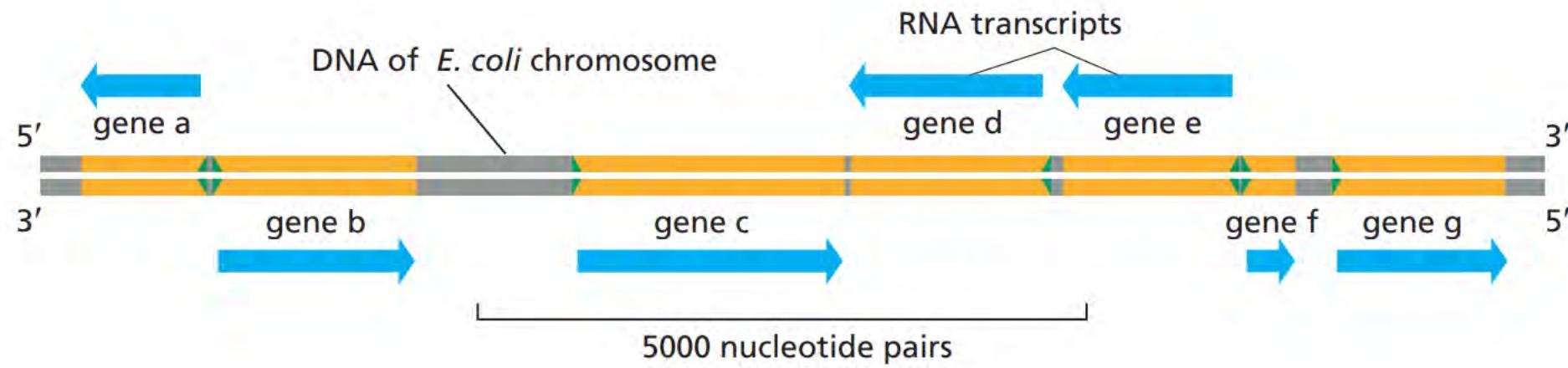
# РНК-полимеразы эукариот

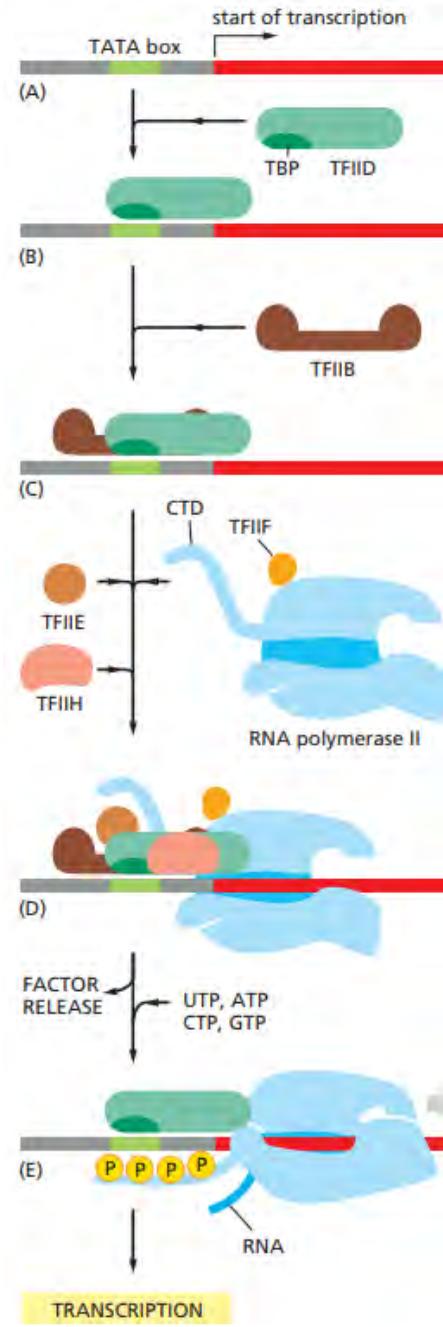
РНК-полимераза	Функция	Оптимальное соотношение $[Mn^{2+}] / [Mg^{2+}]$	Концентрация $\alpha$ -аманитина, полностью подавляющая работу
I	Синтез всех рРНК кроме 5S рРНК	2	нечувствительна
II	Синтез всех мРНК и некоторых сРНК	5	$10^{-8}$ М
III	Синтез всех тРНК, 5S рРНК и некоторых сРНК	2	$10^{-6}$ М

**Промотор — последовательность нуклеотидов ДНК, узнаваемая РНК-полимеразой как стартовая площадка для начала транскрипции.**

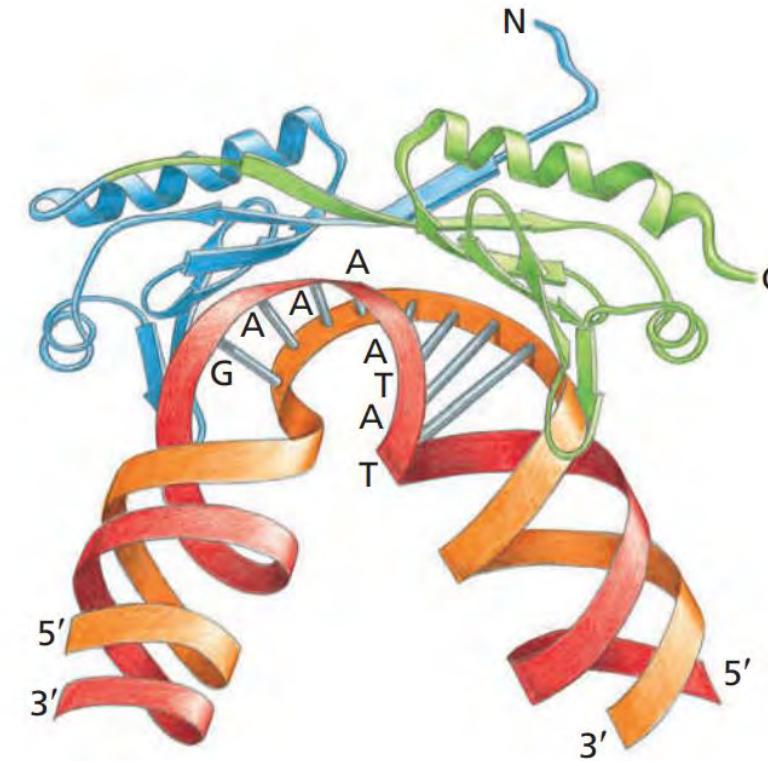


Консенсусные нуклеотидные последовательности для промоторов *E. coli* (на основе сравнения около 300 промоторов).

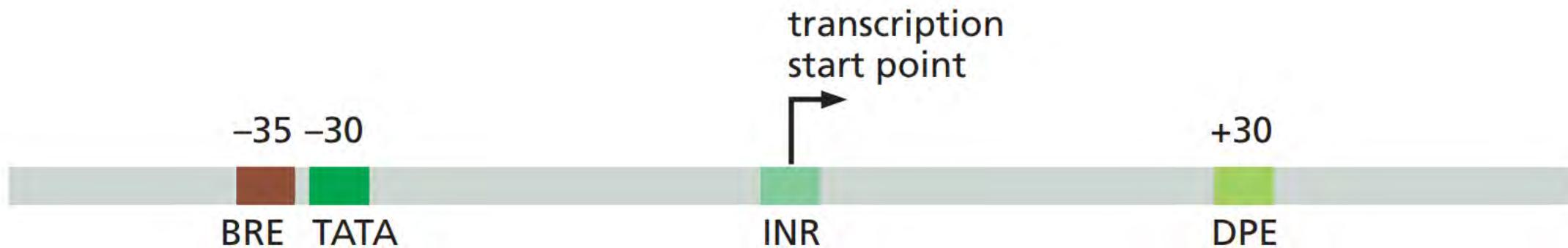




## Универсальные факторы транскрипции у эукариот

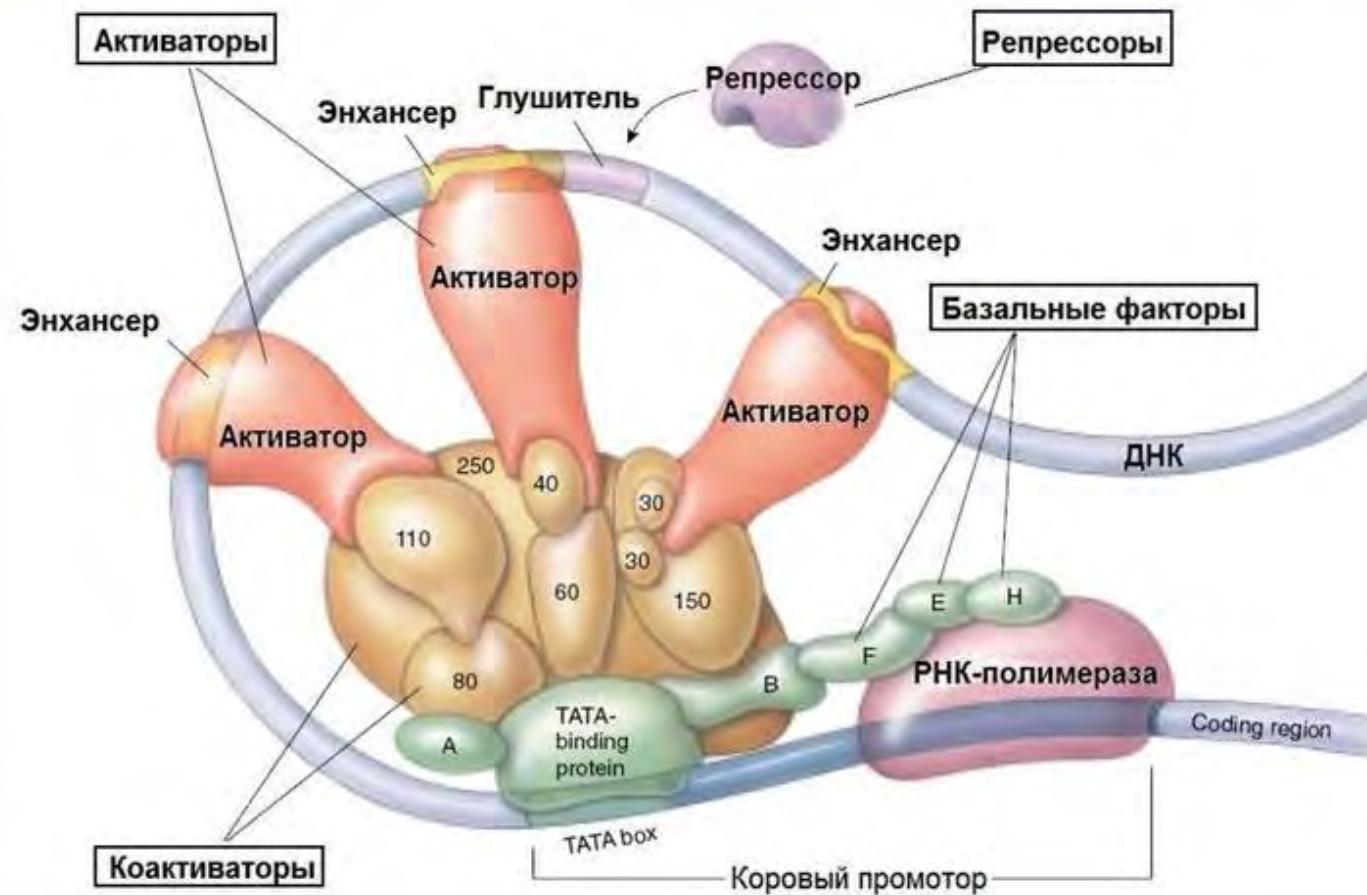


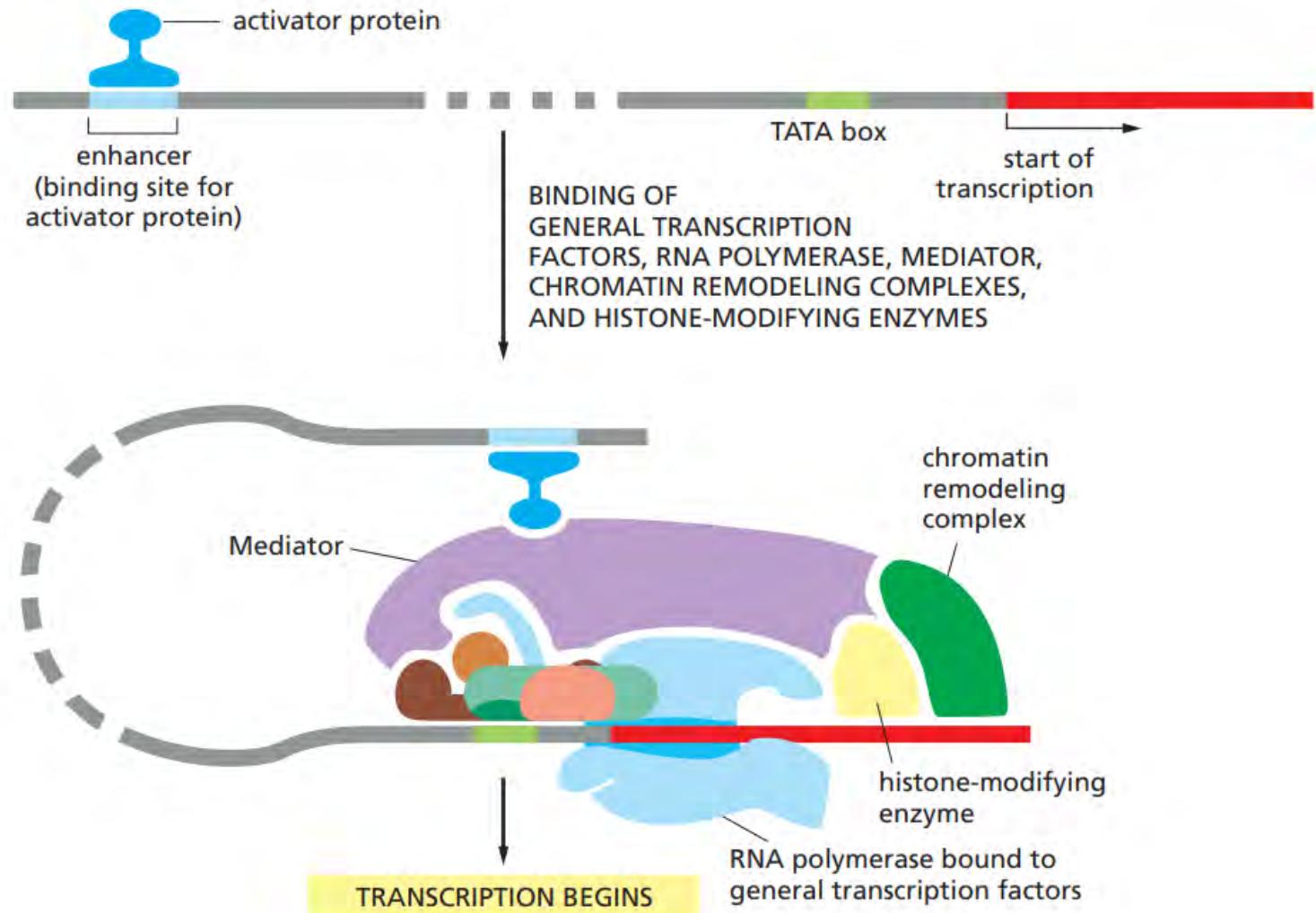
Трехмерная структура ТВР (TATA-binding protein), связанного с ДНК



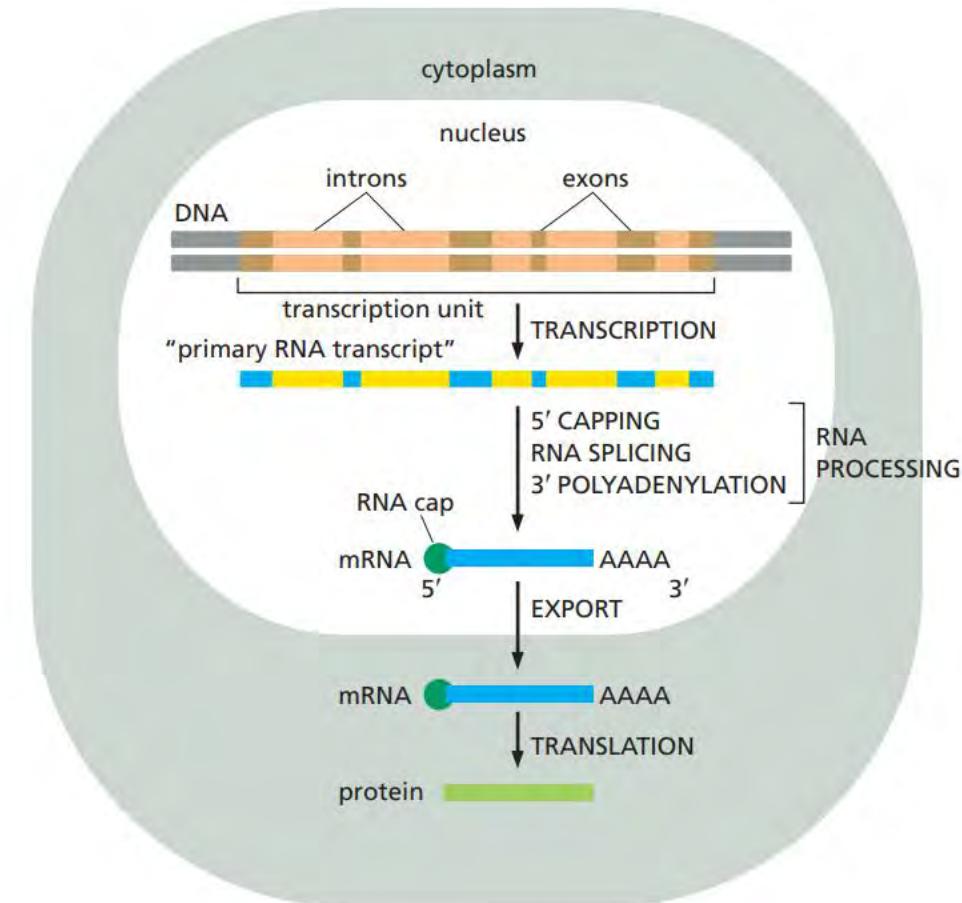
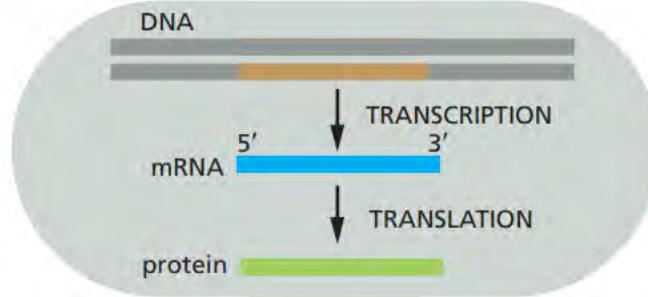
element	consensus sequence	general transcription factor
BRE	G/C G/C G/A C G C C	TFIIB
TATA	T A T A A/T A A/T	TBP subunit of TFIID
INR	C/T C/T A N T/A C/T C/T	TFIID
DPE	A/G G A/T C G T G	TFIID

## Инициация транскрипции у эукариот

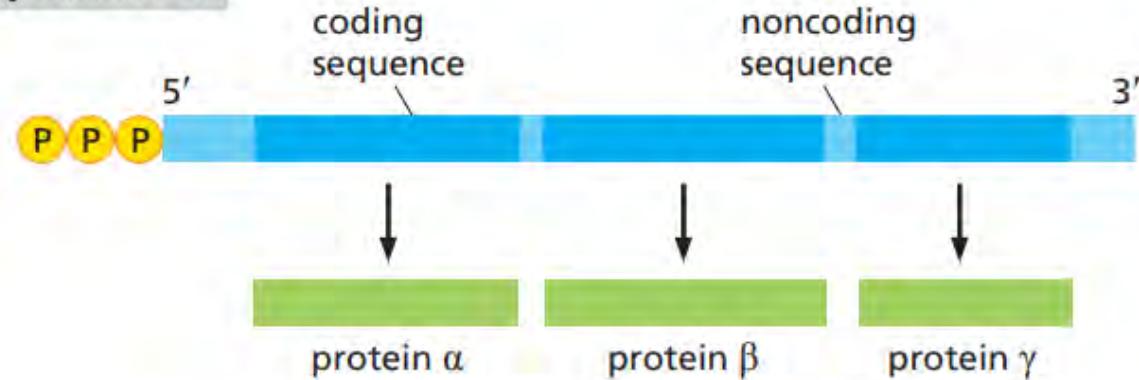




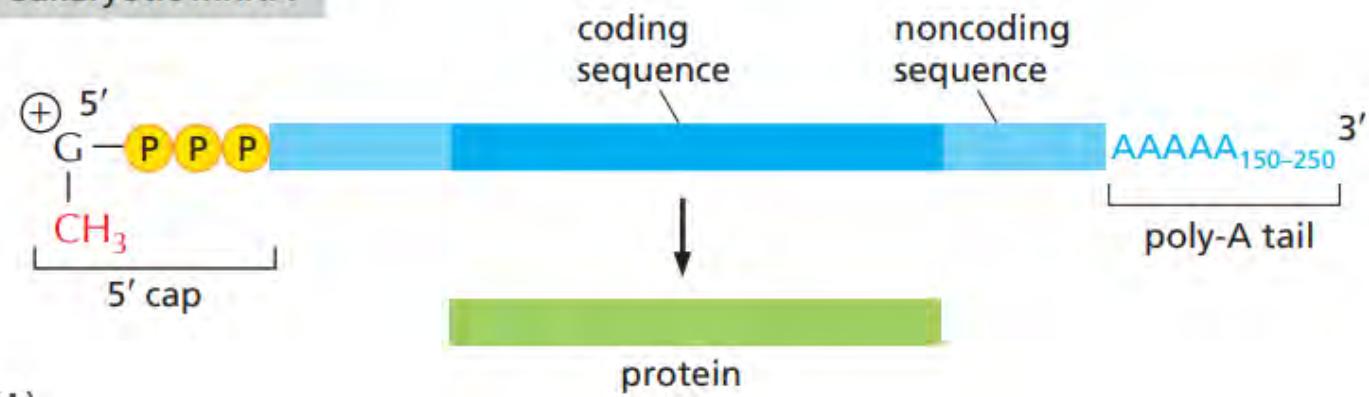
## Транскрипция и трансляция у прокариот и эукариот

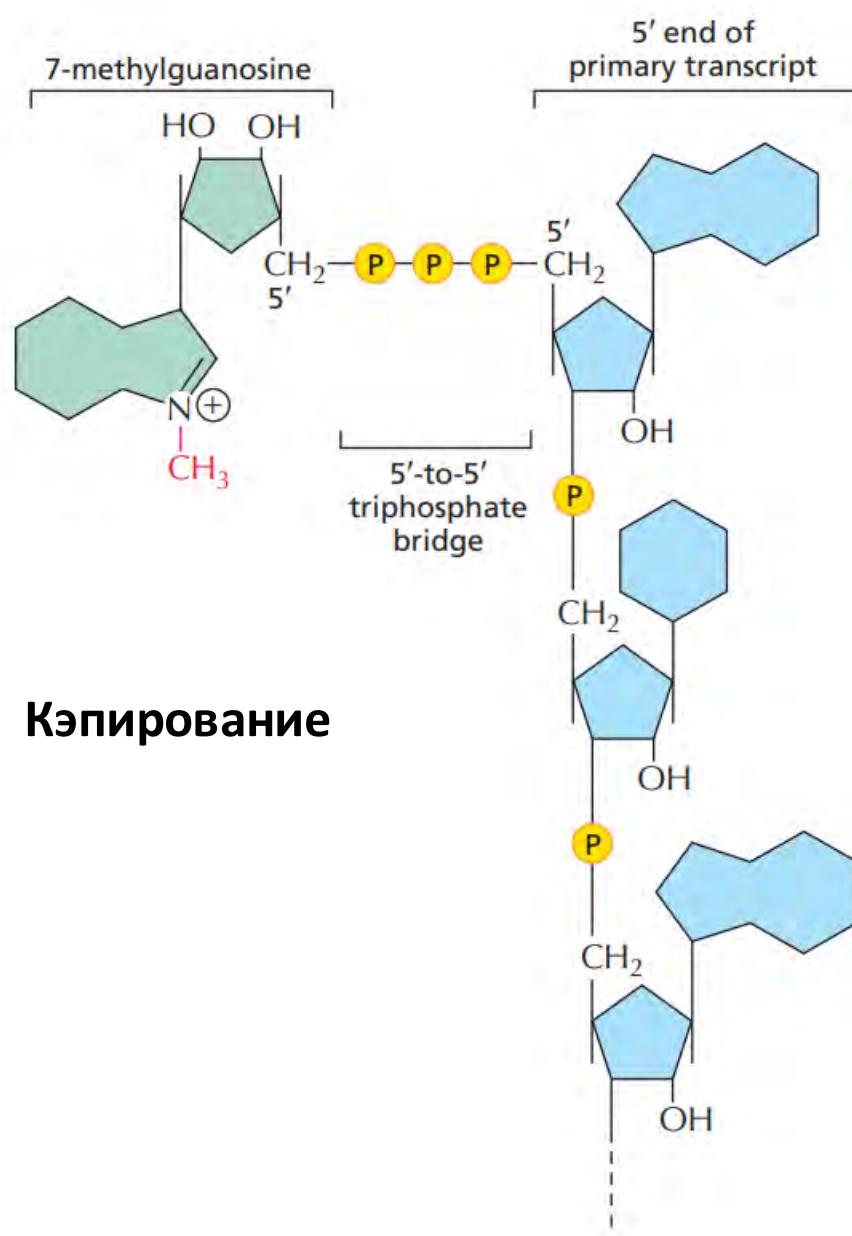
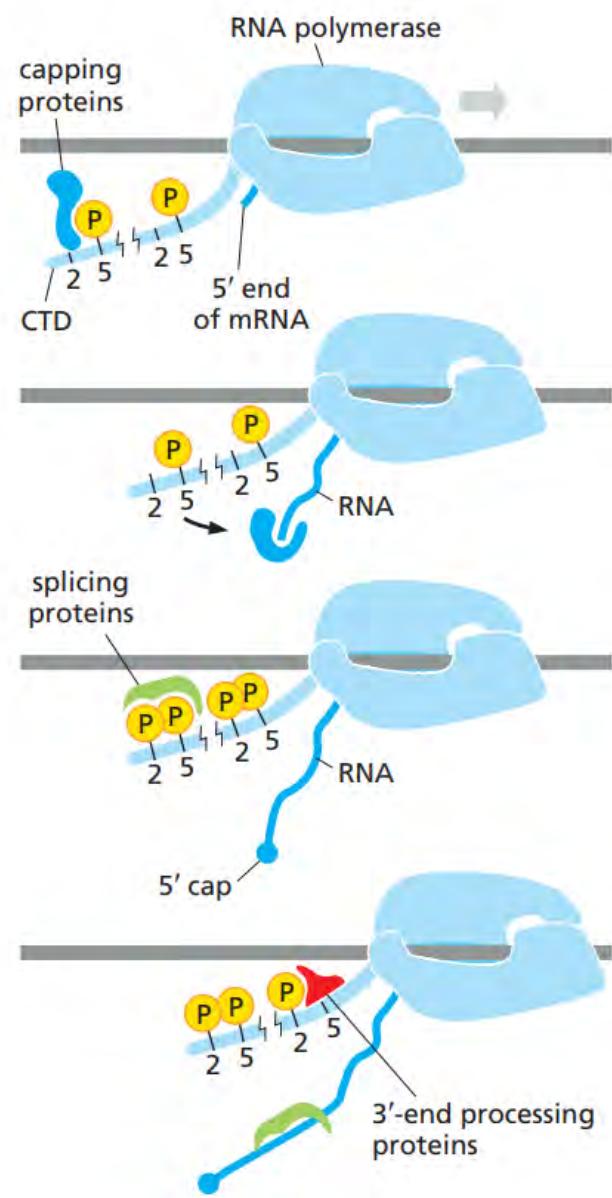


prokaryotic mRNA



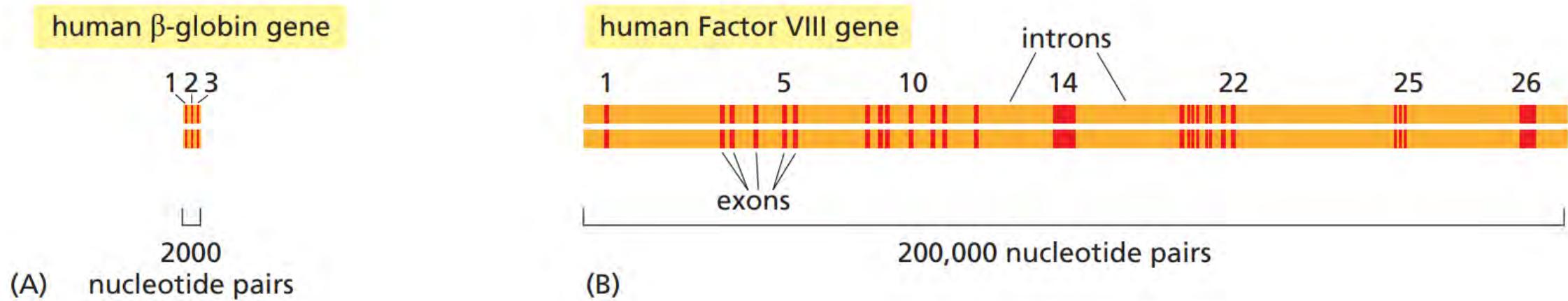
eukaryotic mRNA



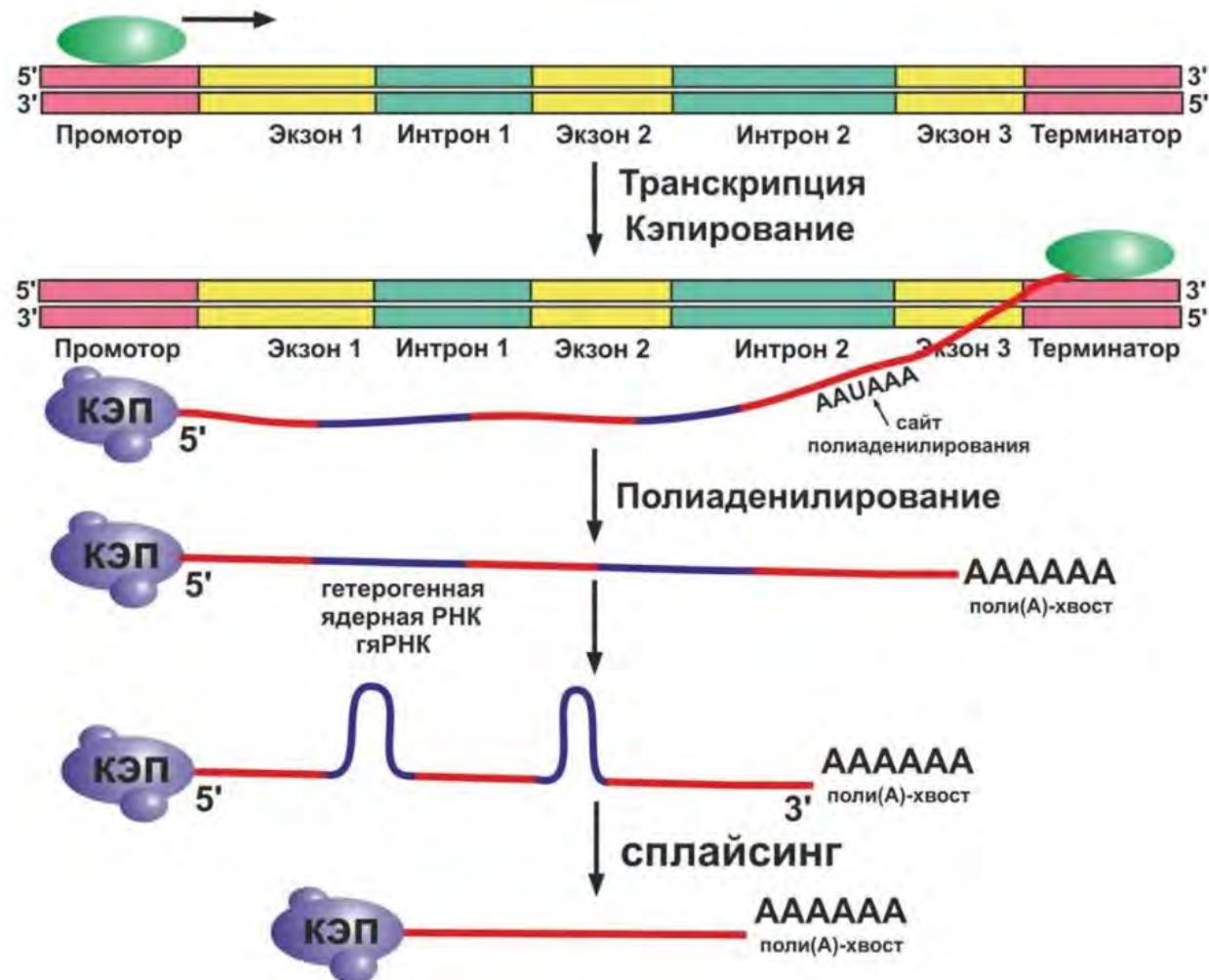


## Кэпирование

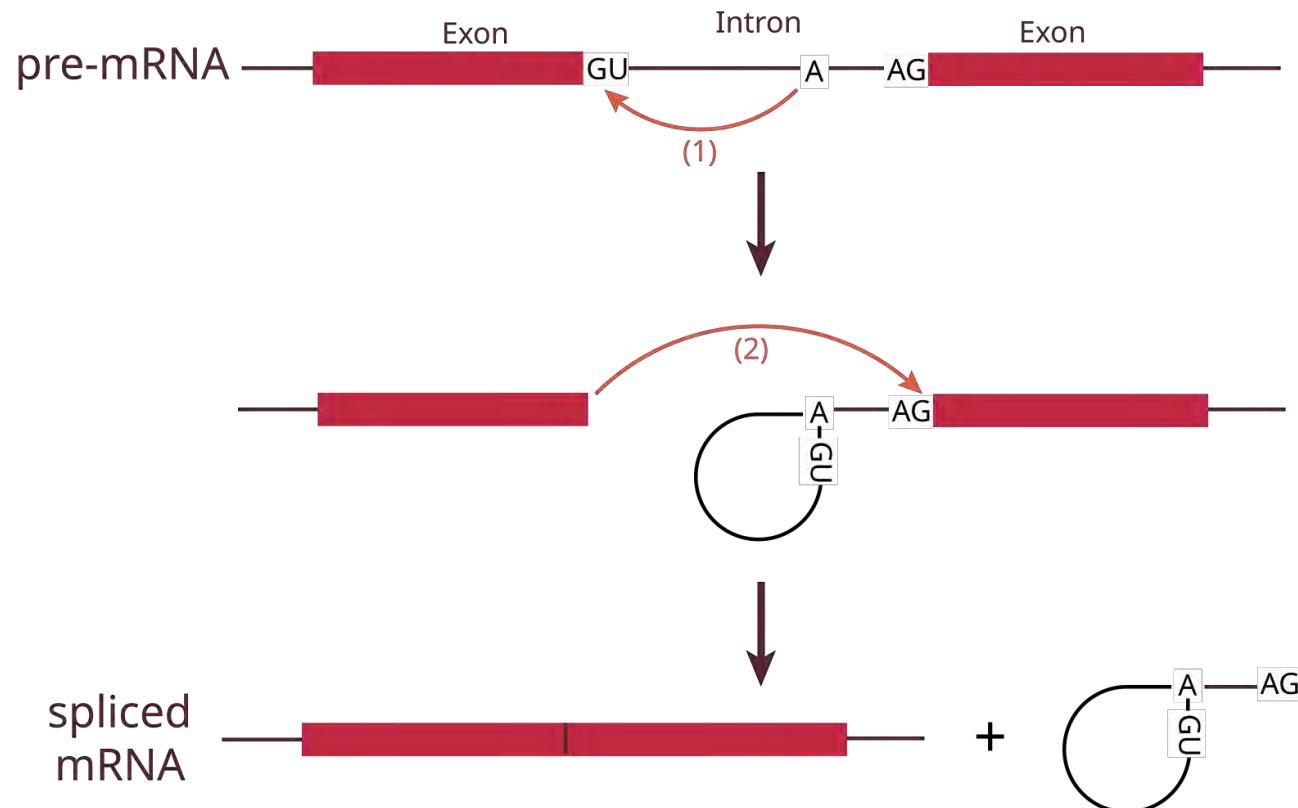
# Структура генов эукариот



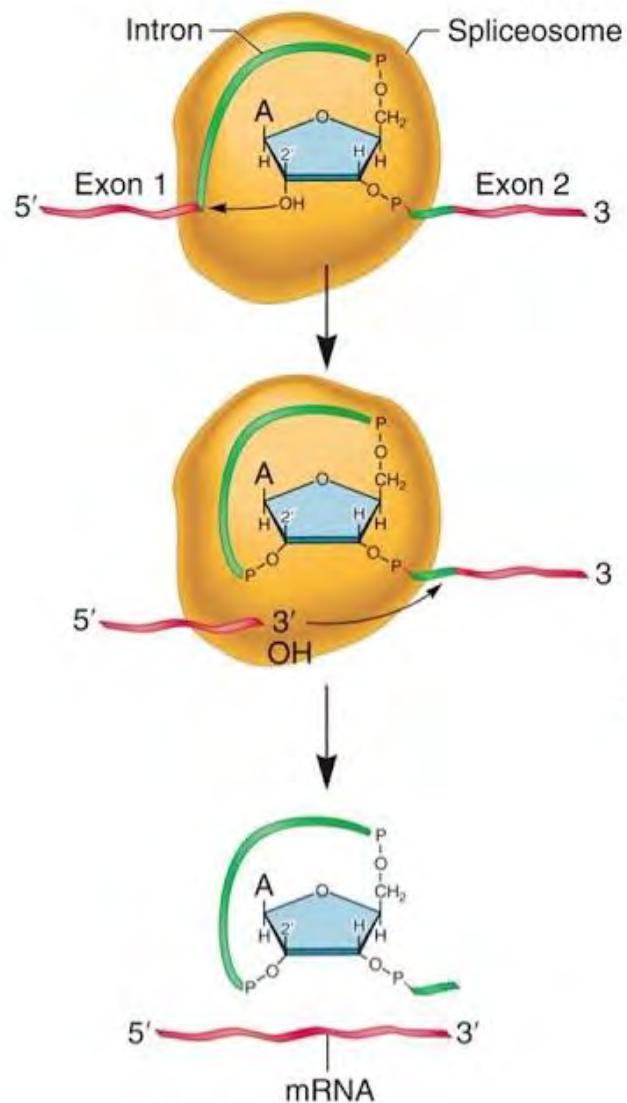
# Процессинг



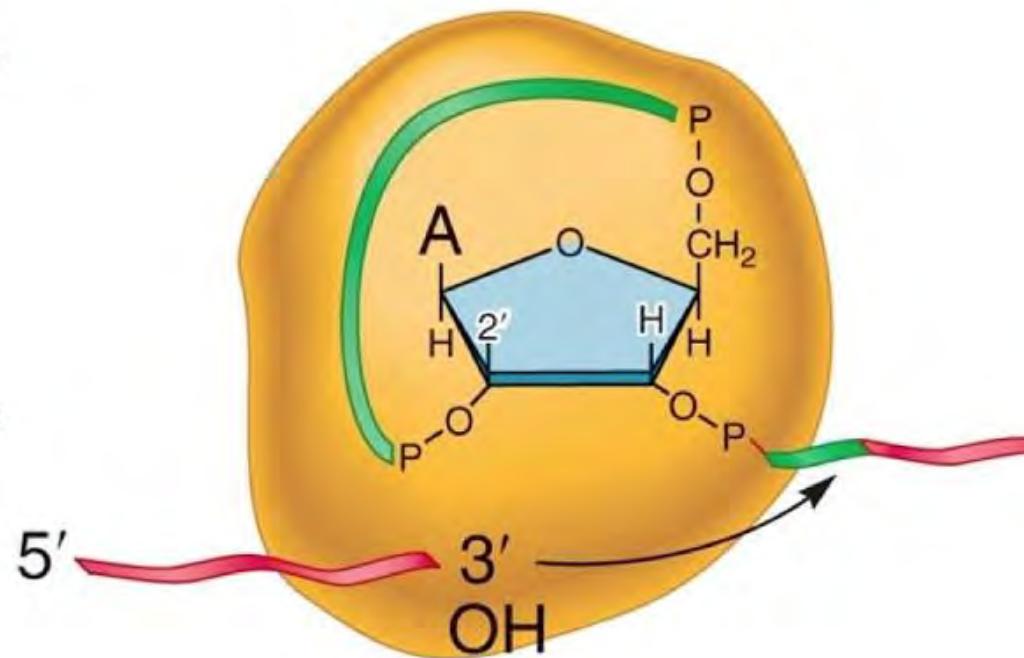
# Спlicing

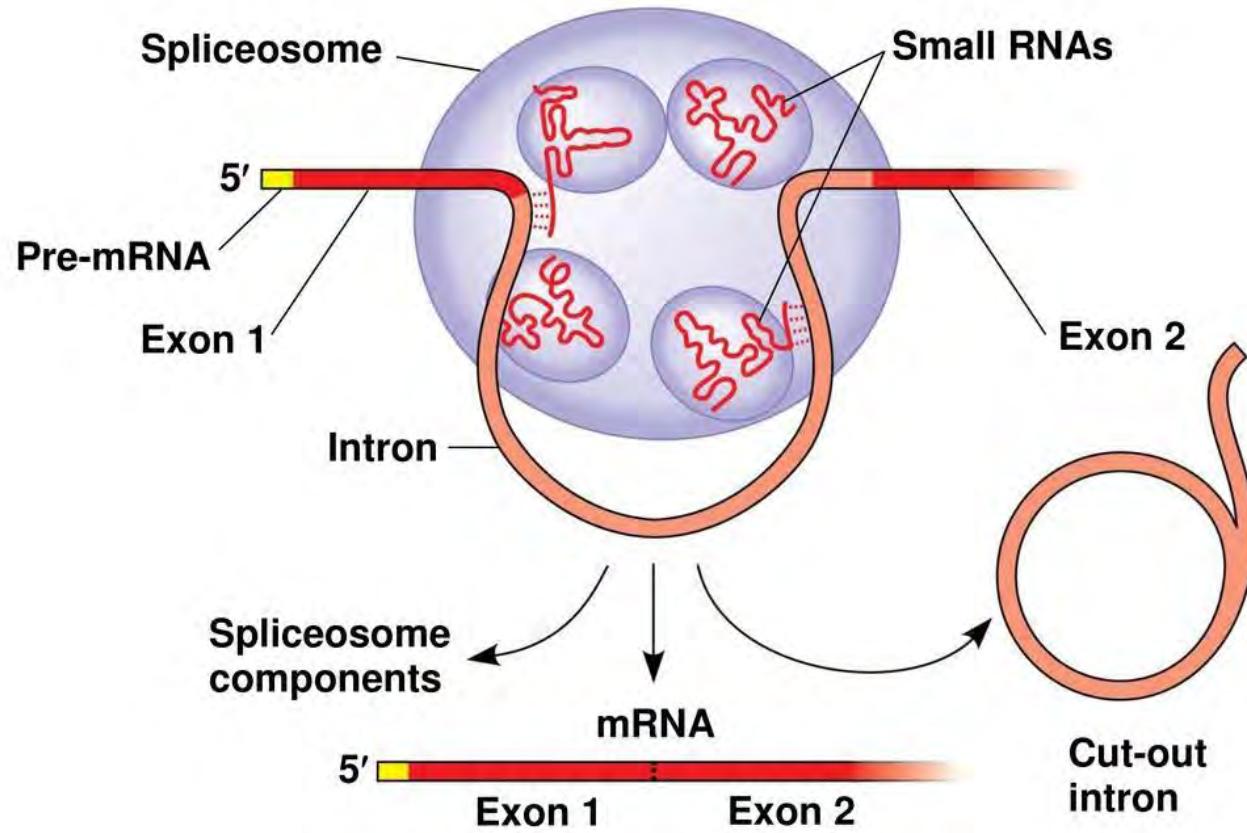


Intron removed via spliceosome  
(very common in eukaryotes)



# Spliceosome

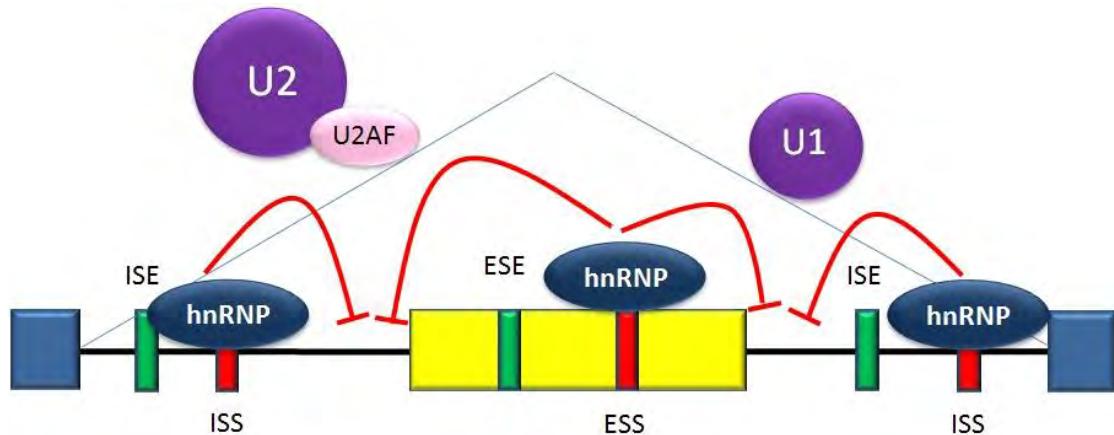




Спайкосома – рибонуклеопротеидный комплекс:

5 мяРНК (U1, U2, U4, U5, U6) + около 100 белков

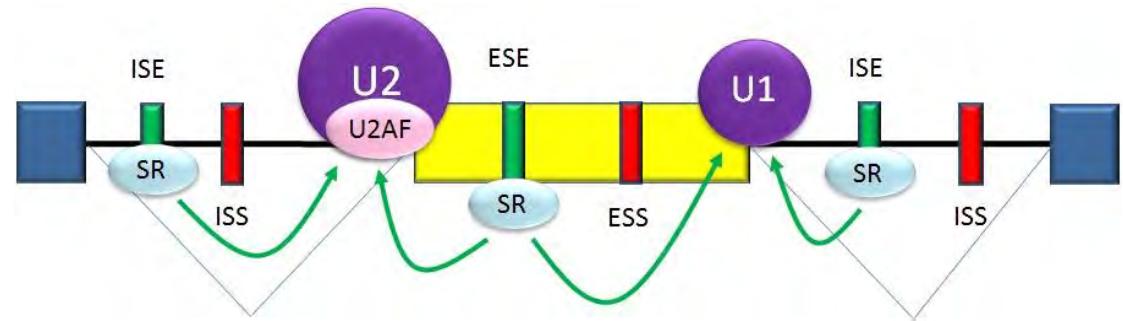
## SR – сплайсинг-регулирующие белки



ESS – экзонный сайленсер сплайсинга

ISS – инtronный сайленсер сплайсинга

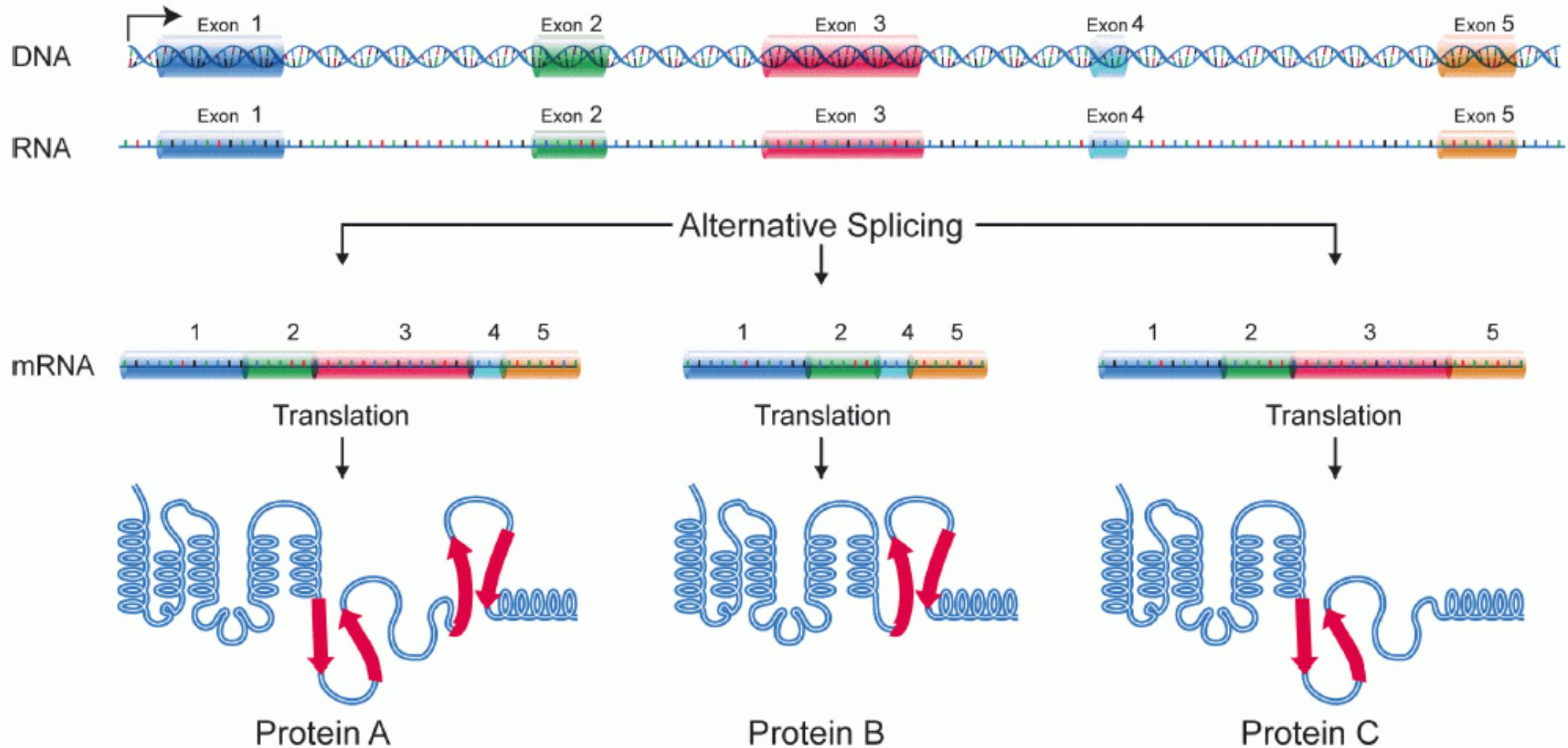
*cis*-регуляторные элементы



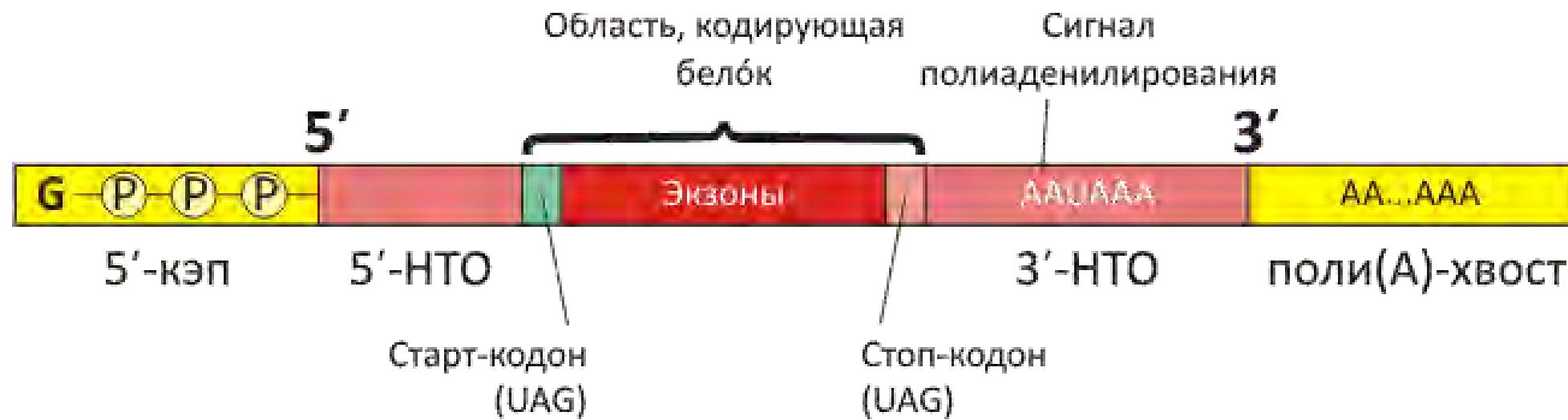
ESE – экзонный энхансер сплайсинга

ISE – инtronный энхансер сплайсинга

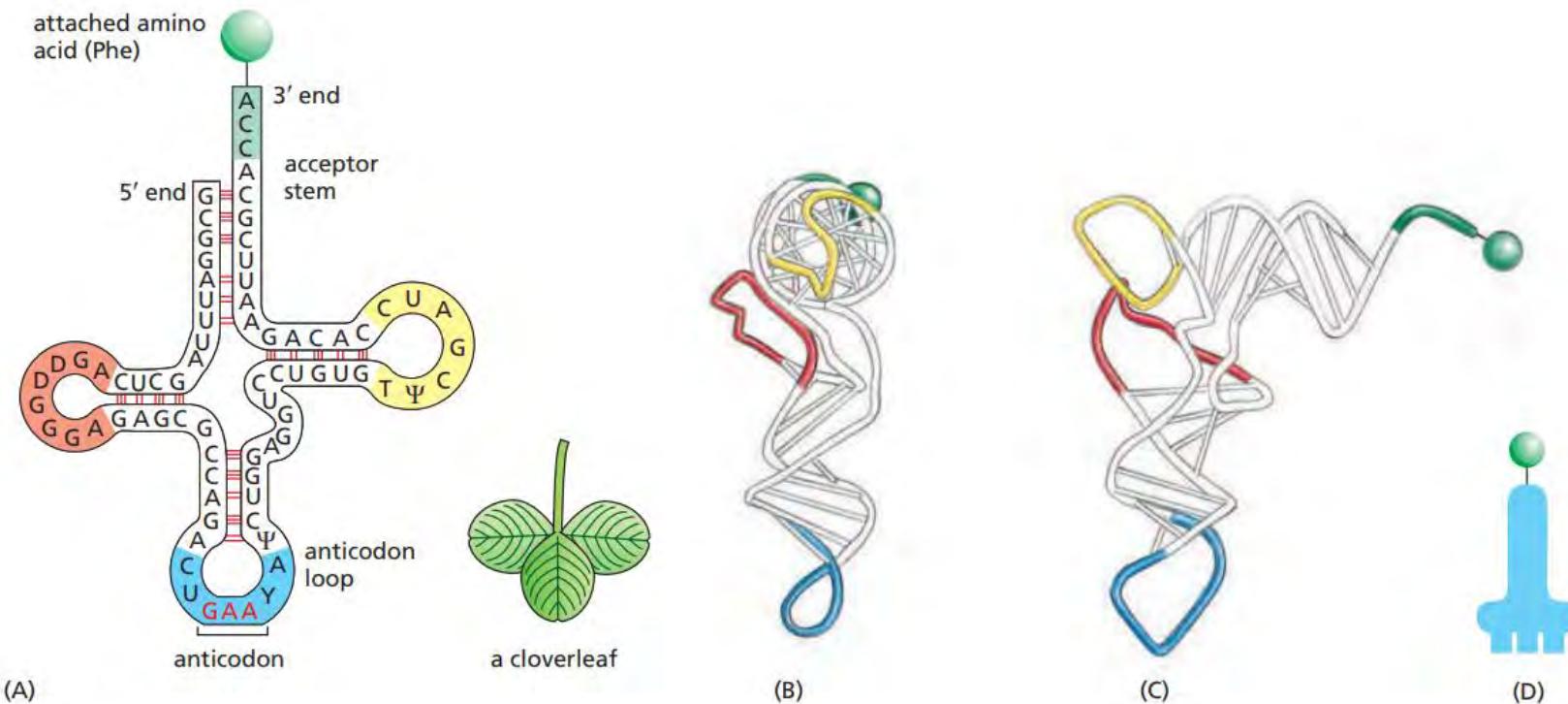
# Альтернативный спlicing



# Зрелая мРНК



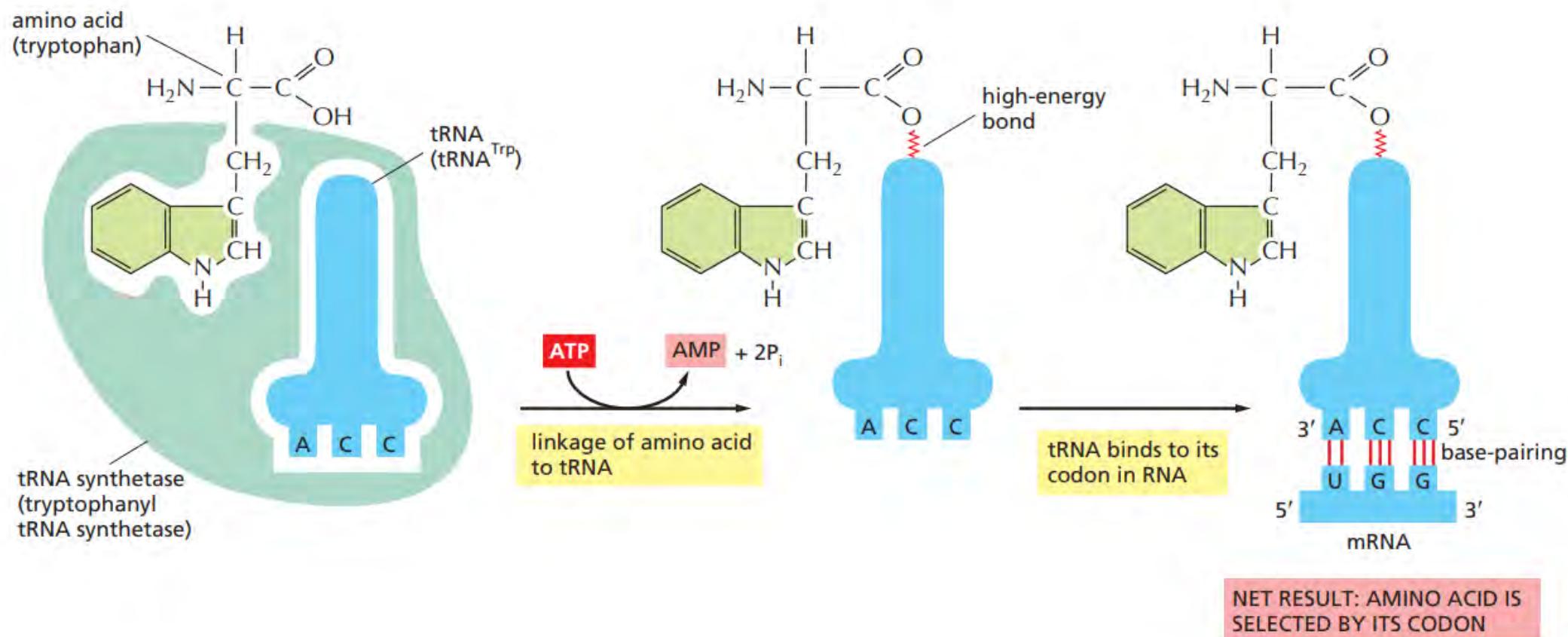
# Структура тРНК



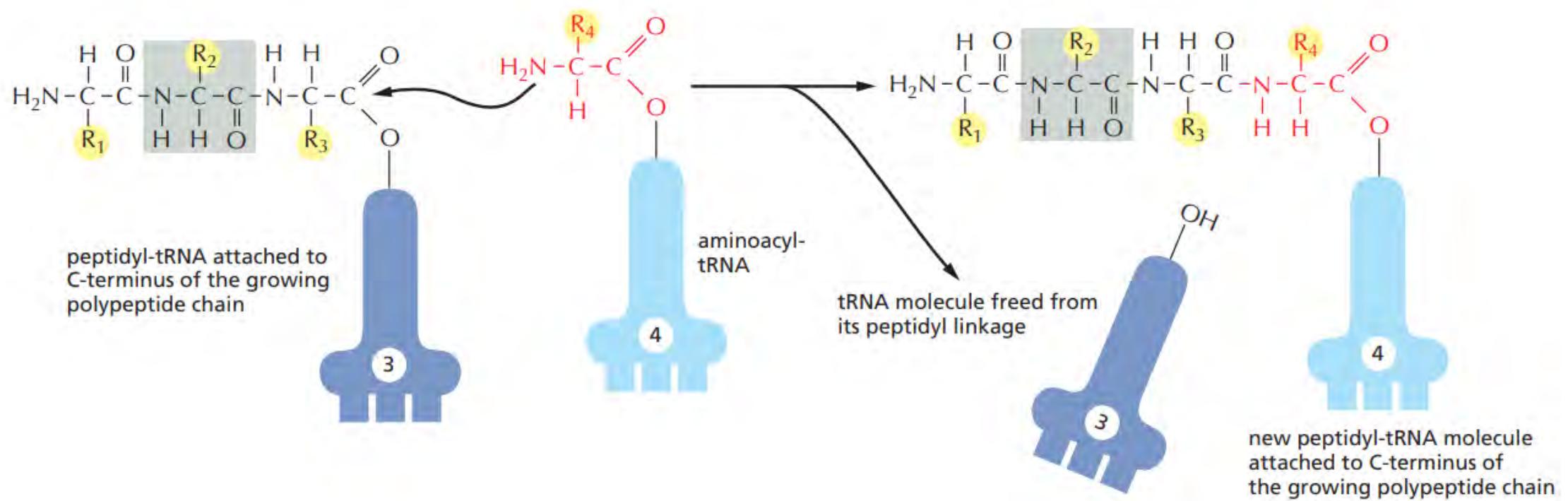
(E)

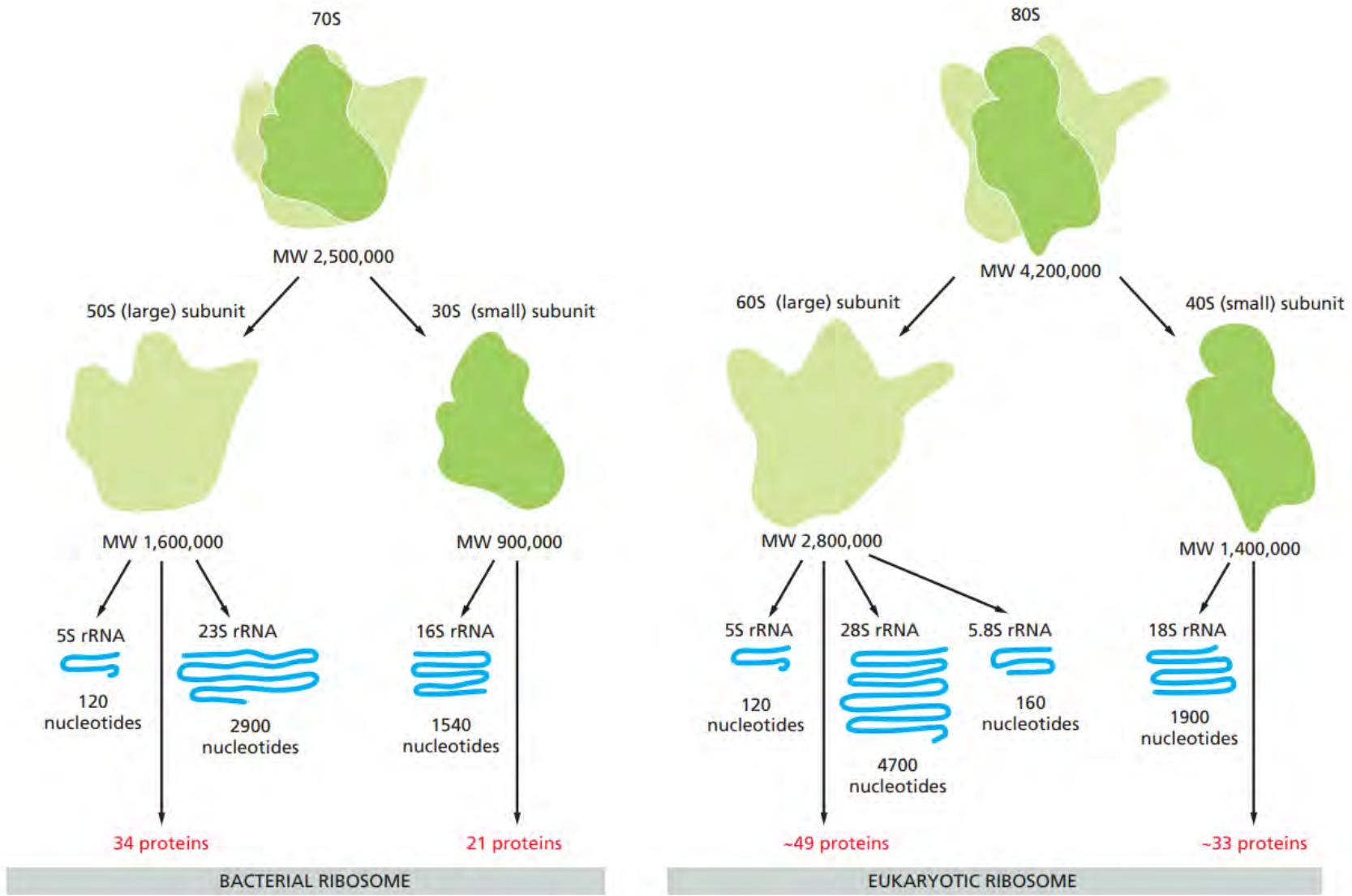
5' GC...AUUAGCU**CAGDDGGGAGAGGCCAGACUGAA**YAΨCUGGAGGUCCUGUGTΨCGAUCCACAGAAUUCGCACCA 3'  
anticodon

# Аминоацил-тРНК-синтетаза



# Образование полипептидной цепи

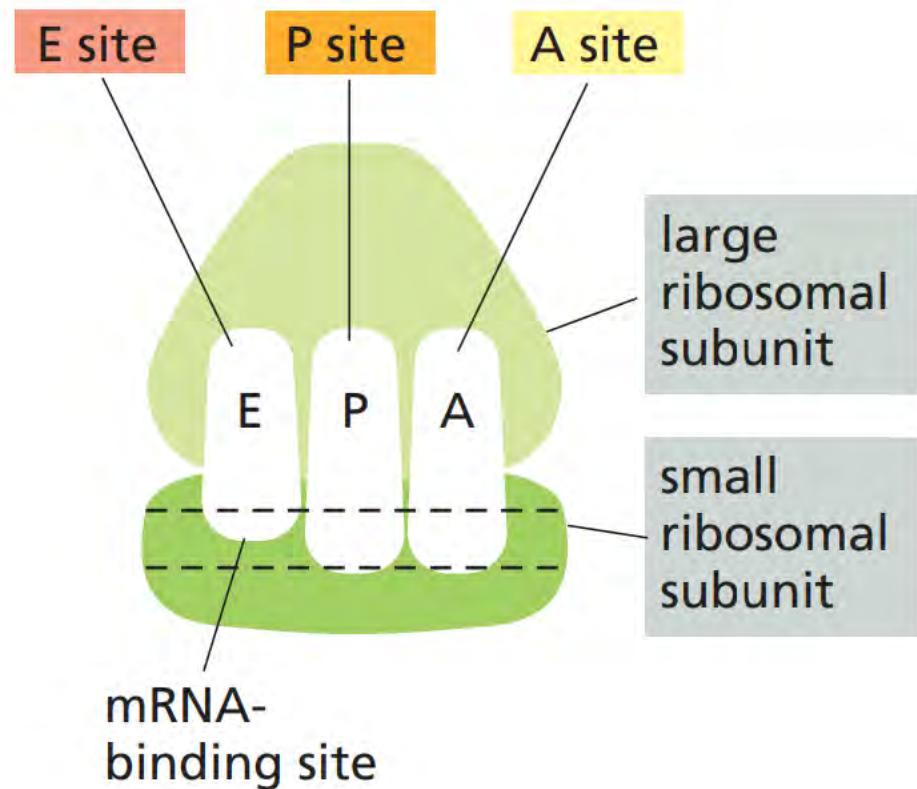




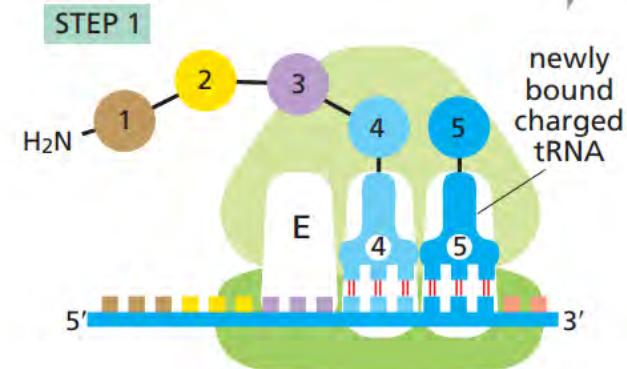
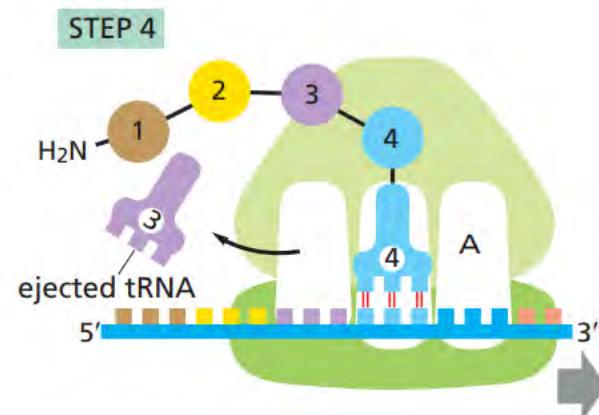
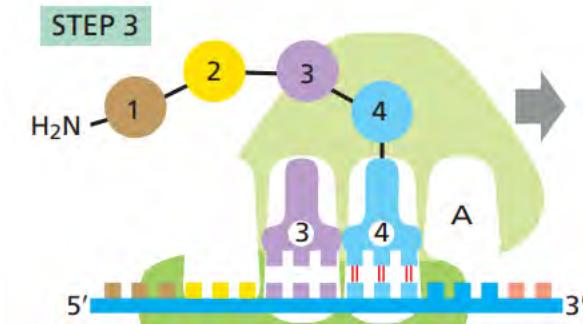
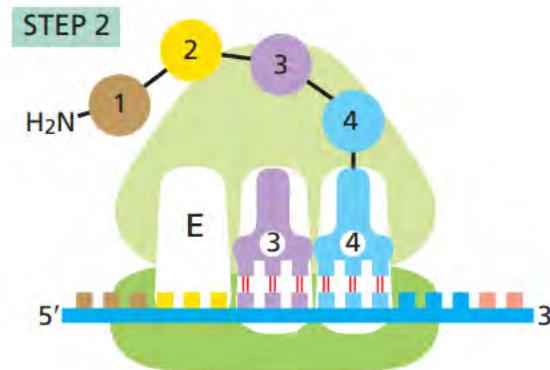
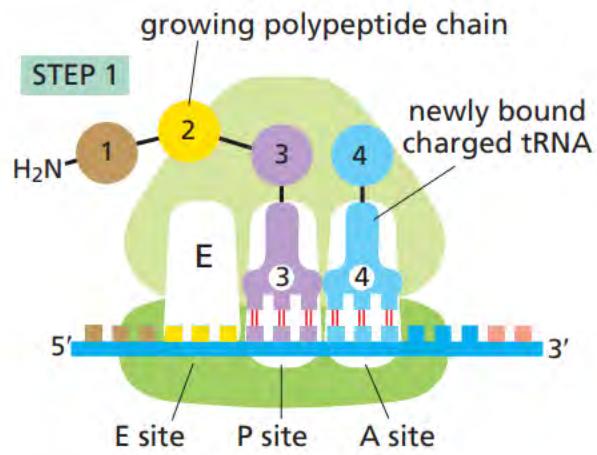
# Структура рибосомы



Синим показана рРНК  
в составе малой субъединицы



# Трансляция

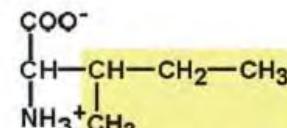
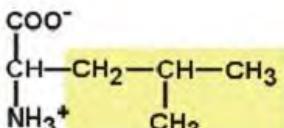
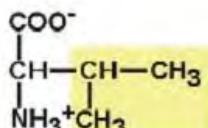
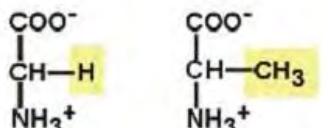
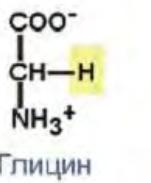


# Свойства генетического кода

1. Универсальность  
(квазиуниверсальность)
2. Триплетность
3. Однозначность
4. Вырожденность
5. Избыточность
6. Неперерывность
7. Наличие стоп-кодонов

Генетический код (иРНК от 5' к 3' концу)

Первое основание	Второе основание				Третье основание
	У	Ц	А	Г	
У	Фен	Сер	Тир	Цис	У
	Фен	Сер	Тир	Цис	Ц
	Лей	Сер	—	—	А
	Лей	Сер	—	Три	Г
Ц	Лей	Про	Гис	Арг	У
	Лей	Про	Гис	Арг	Ц
	Лей	Про	Гли	Арг	А
	Лей	Про	Гли	Арг	Г
А	Иле	Тре	Асн	Сер	У
	Иле	Тре	Асн	Сер	Ц
	Иле	Тре	Лиз	Арг	А
	Мет	Тре	Лиз	Арг	Г
Г	Вал	Ала	Асп	Гли	У
	Вал	Ала	Асп	Гли	Ц
	Вал	Ала	Глу	Гли	А
	Вал	Ала	Глу	Гли	Г



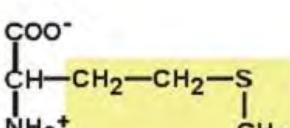
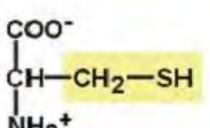
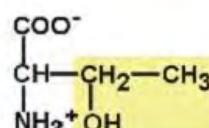
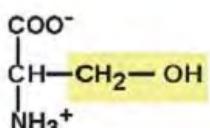
Глицин

Аланин

Валин

Лейцин

Изолейцин

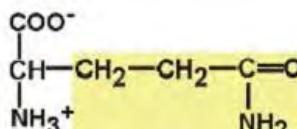


Серин

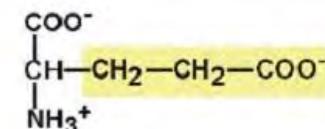
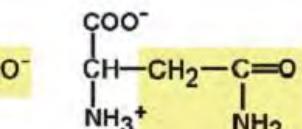
Тreonин

Цистеин

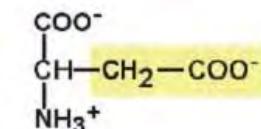
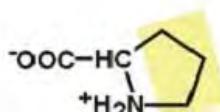
Метионин



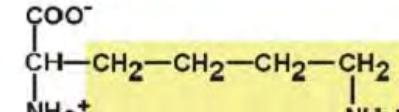
Глутамин

Глутаминовая  
кислота

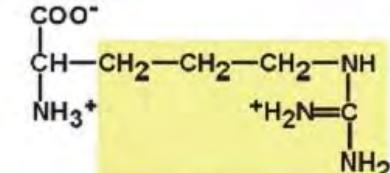
Аспарагин

Аспарагиновая  
кислота

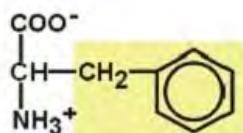
Пролин



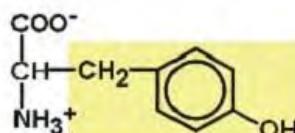
Лизин



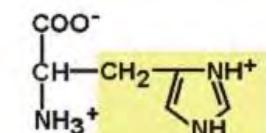
Аргинин



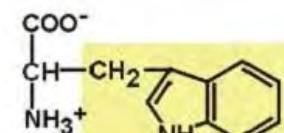
Фенилаланин



Тирозин

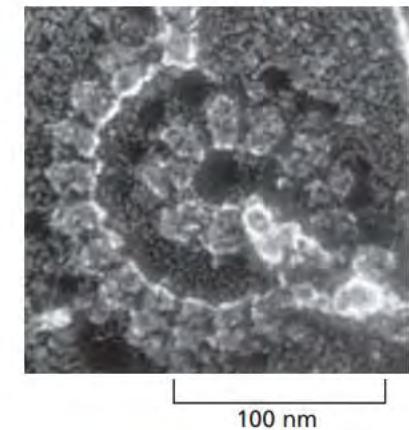
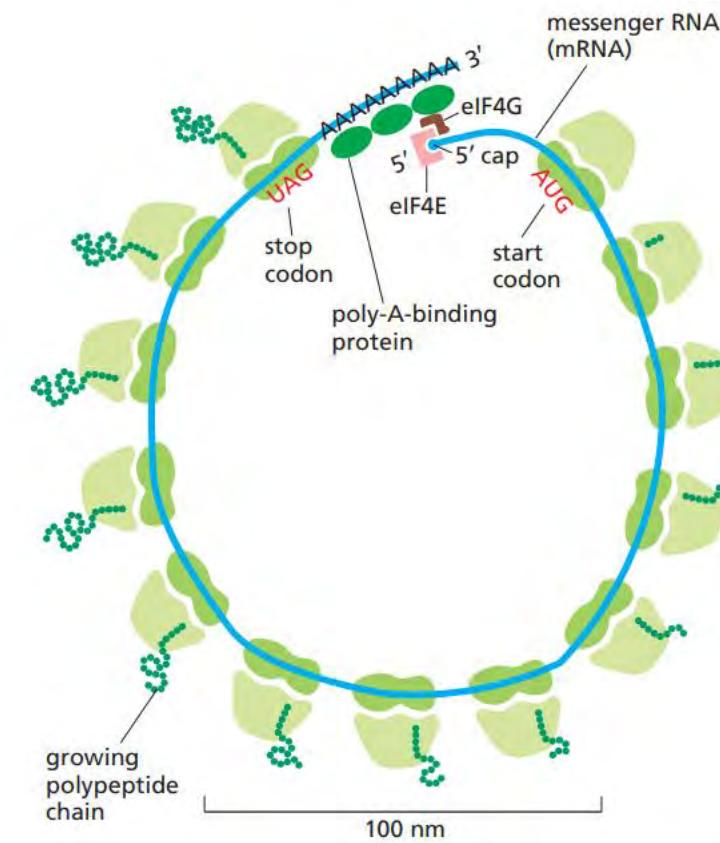
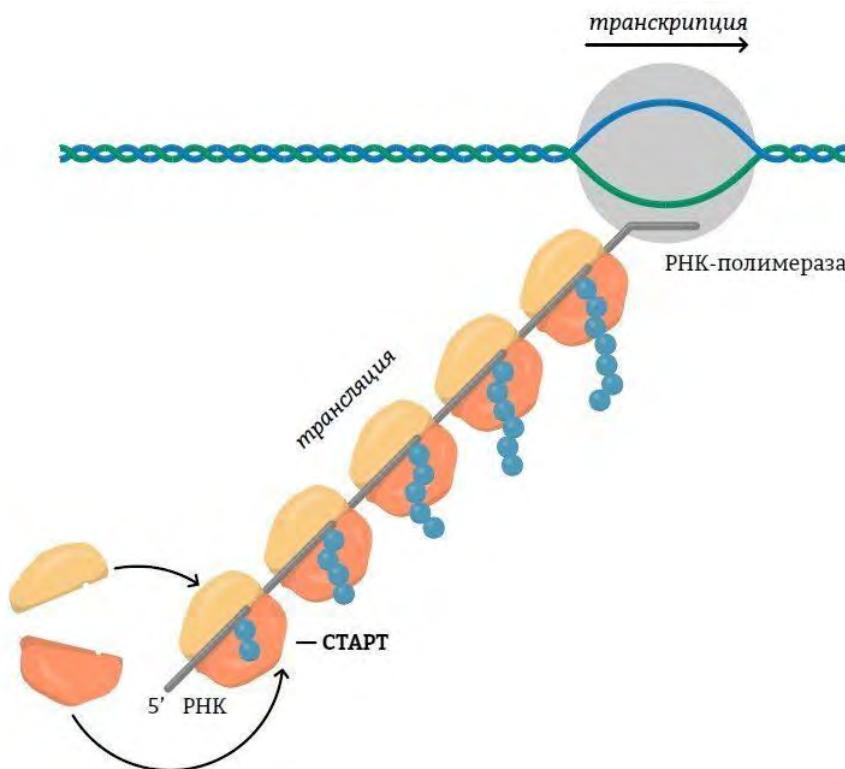


Гистидин

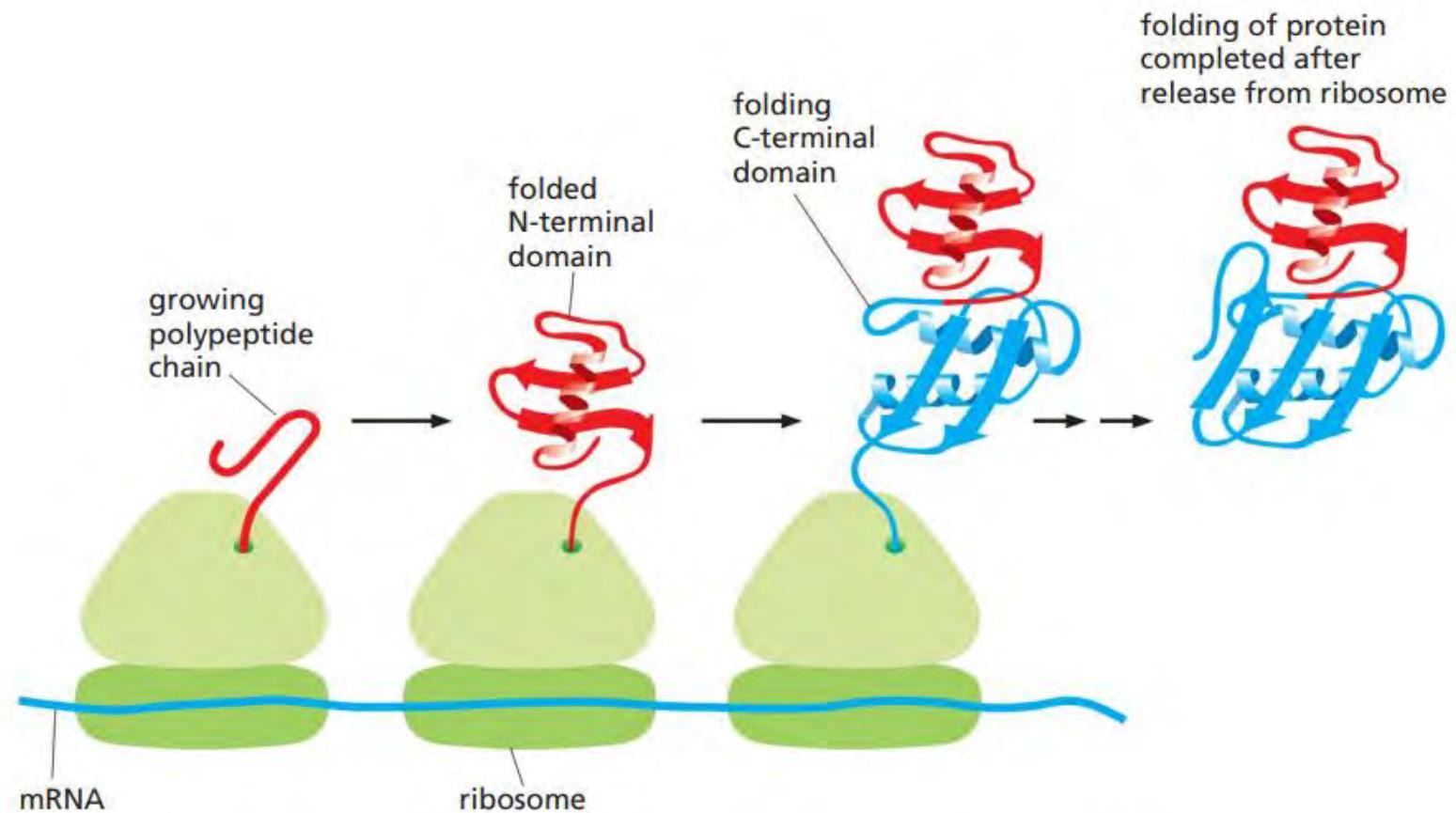


Триптофан

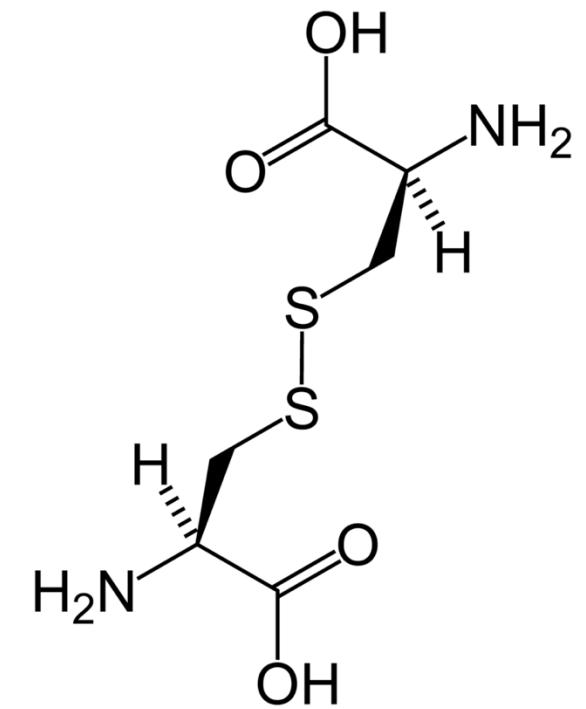
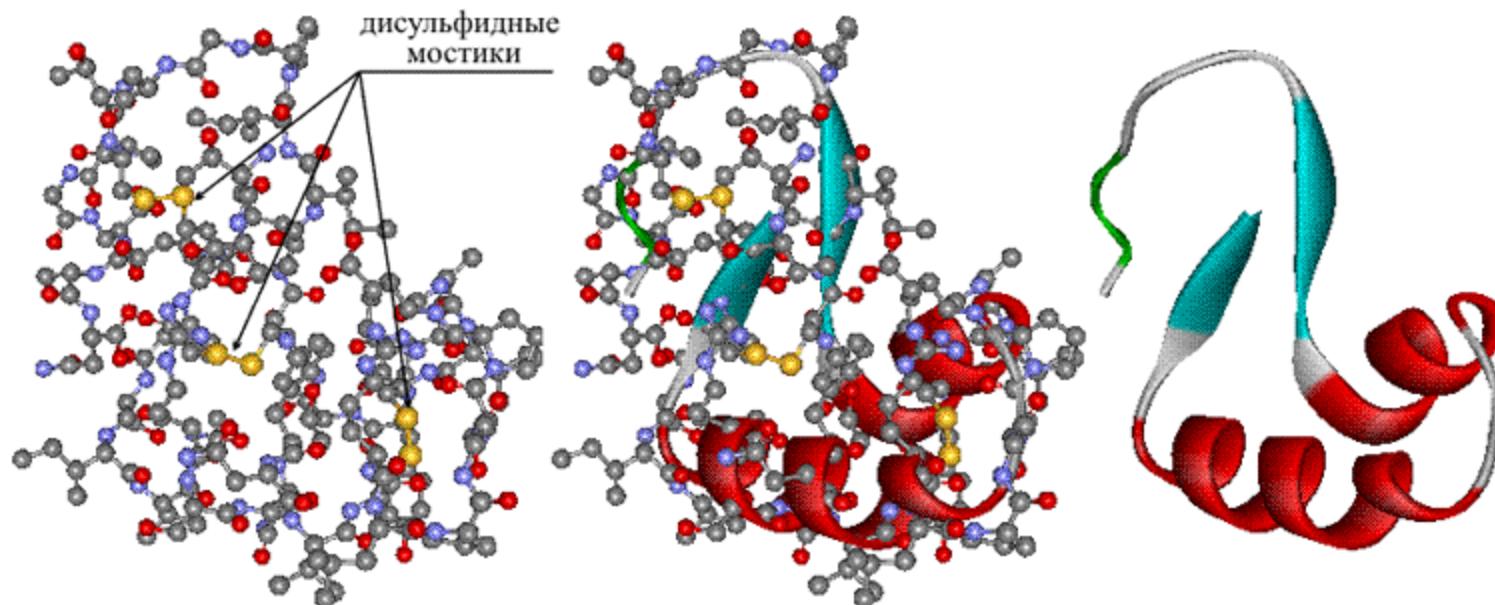
# Полирибосома (полисома)



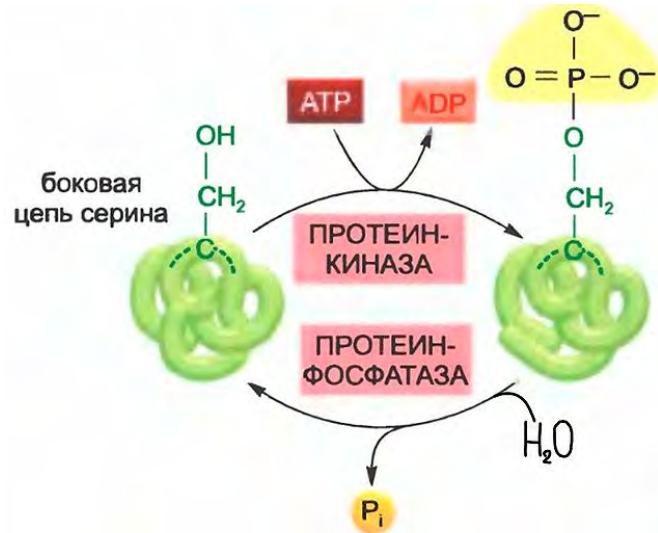
# Котрансляционный фолдинг



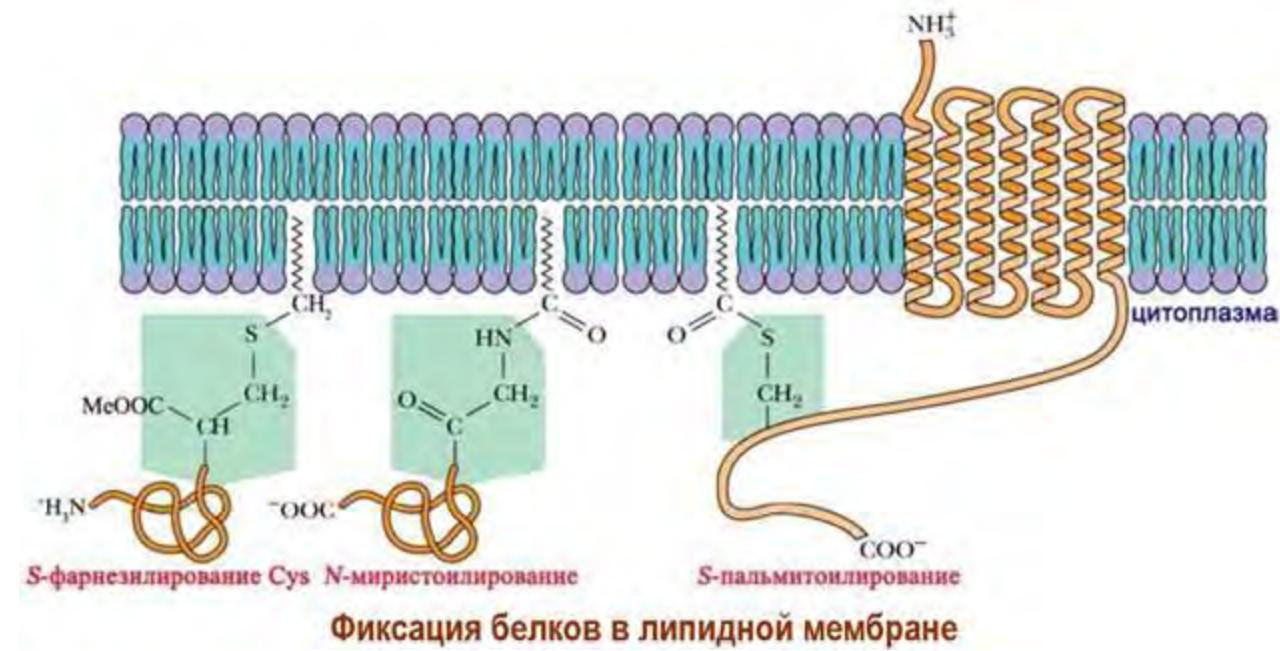
# Посттрансляционные модификации белков



# Посттрансляционные модификации белков

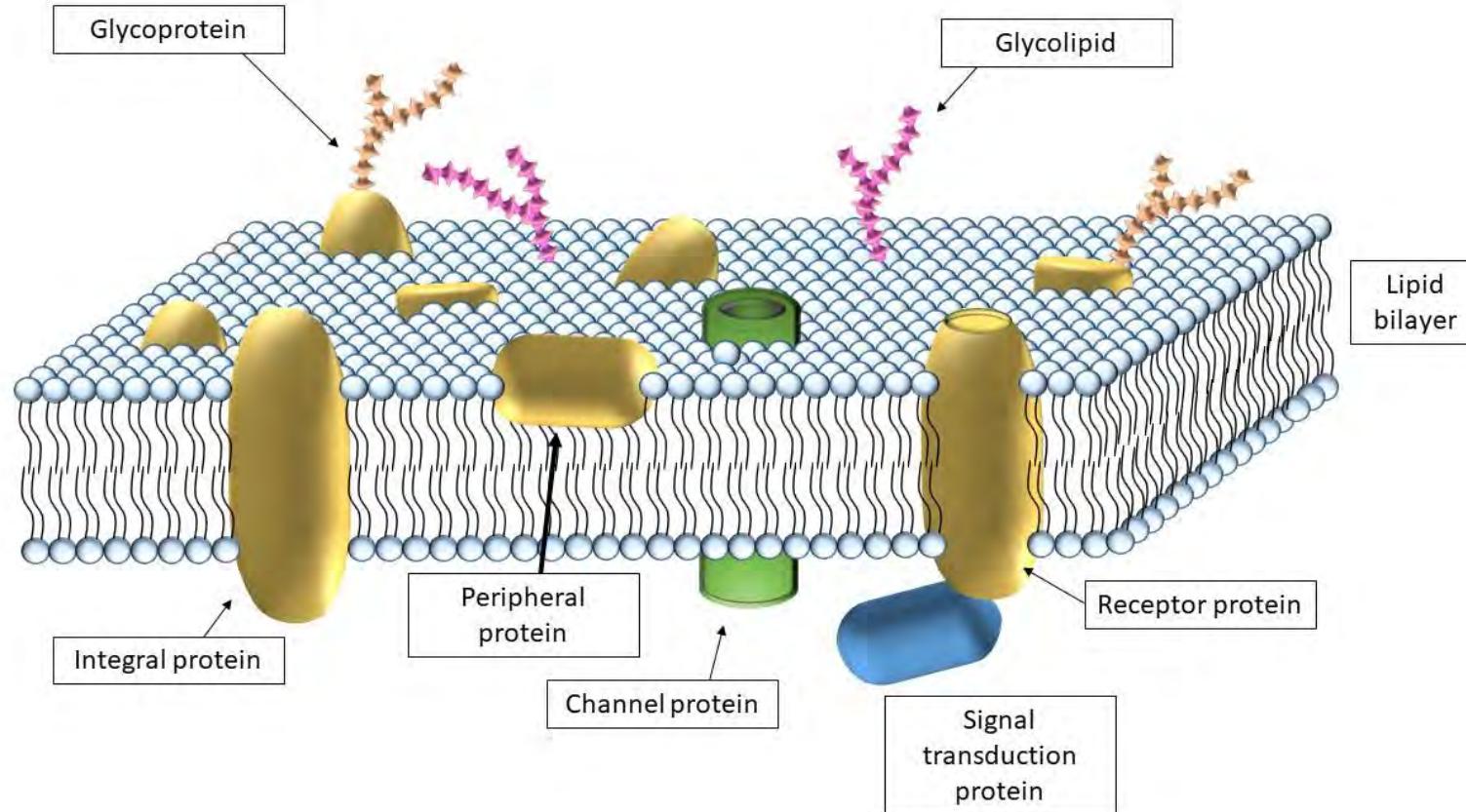


Фосфорилирование

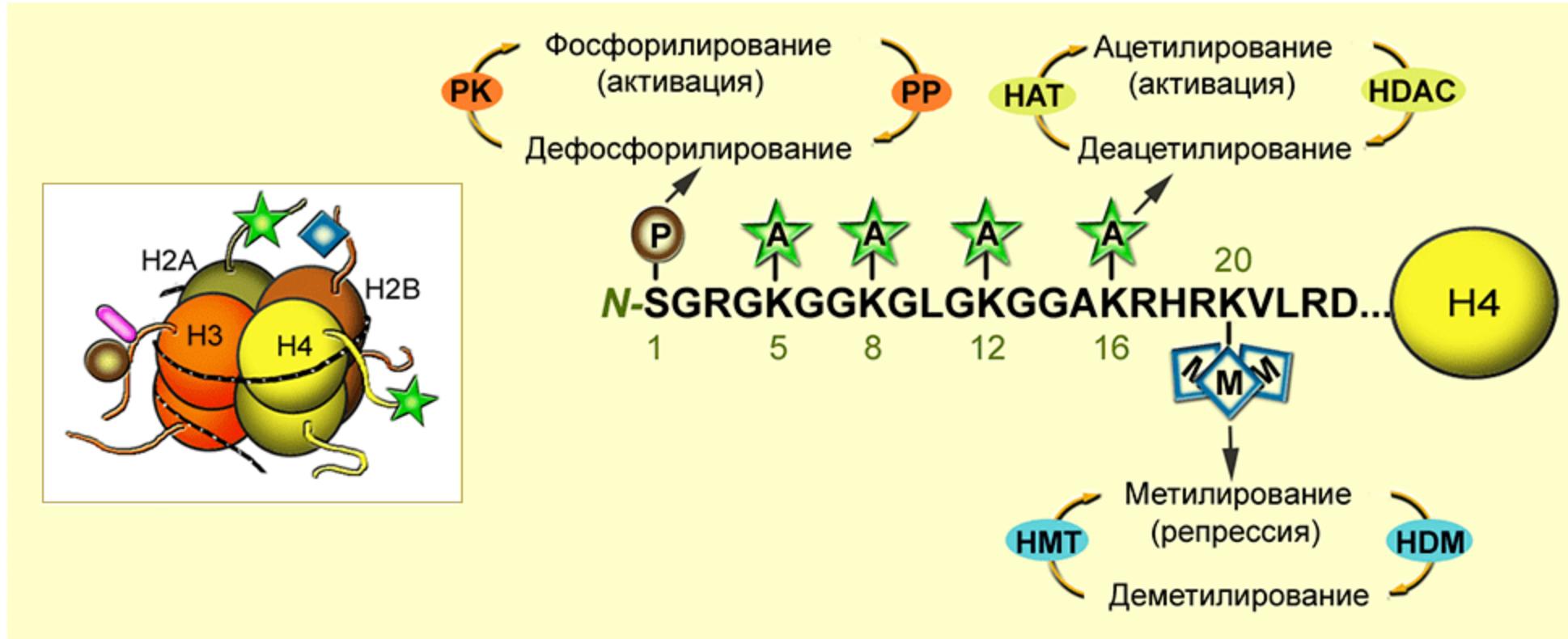


Пренилирование  
(липидирование)

# Посттрансляционные модификации белков



Гликозилирование



Коровье гистоны:

H2A  
H2B x 2  
H3  
H4

Чаще модифицируются аминокислоты N-концевых "хвостиков"

S - серин  
K - лизин

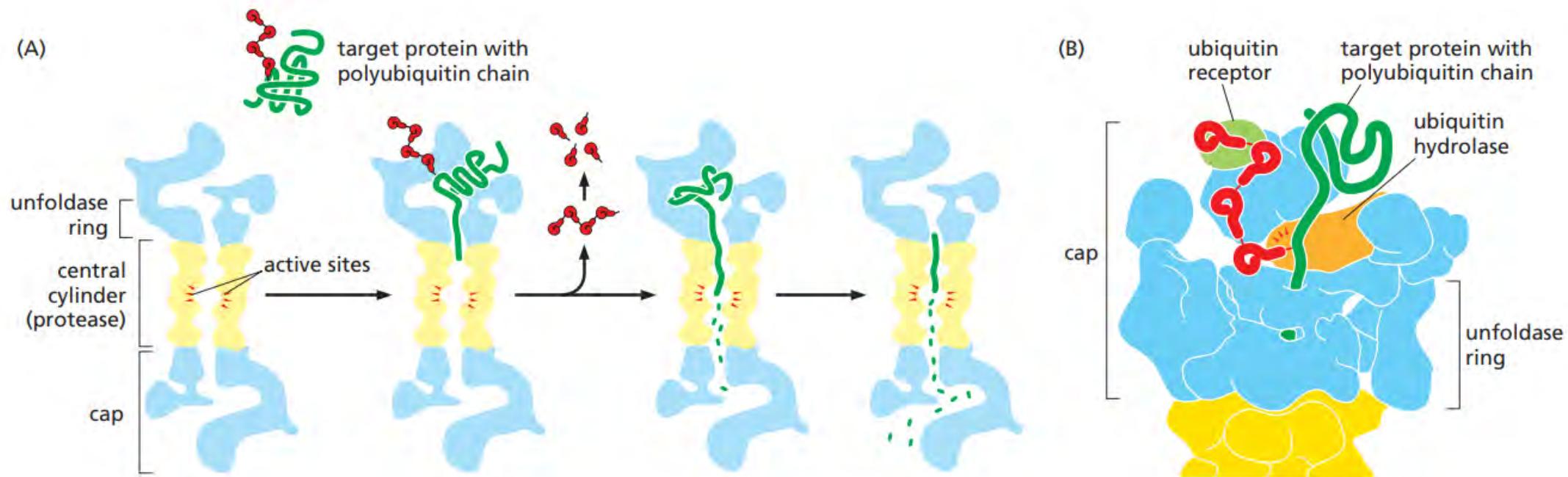
Типы модификаций:

- P - фосфорилирование
- A - ацетилирование
- 3M - метилирование (x 3)
- Ub - убиквитинирование,  
- сумоилирование и др.

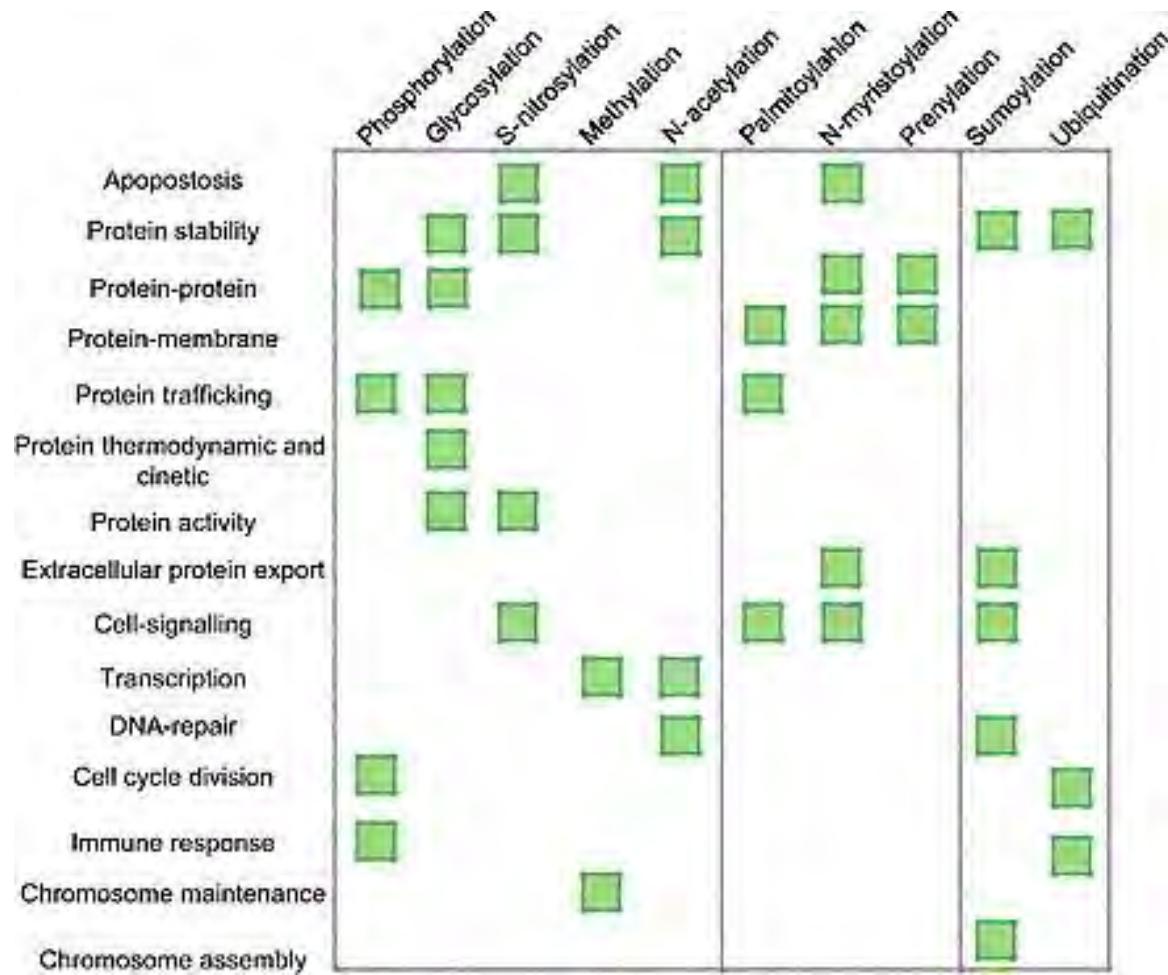
Ферменты:

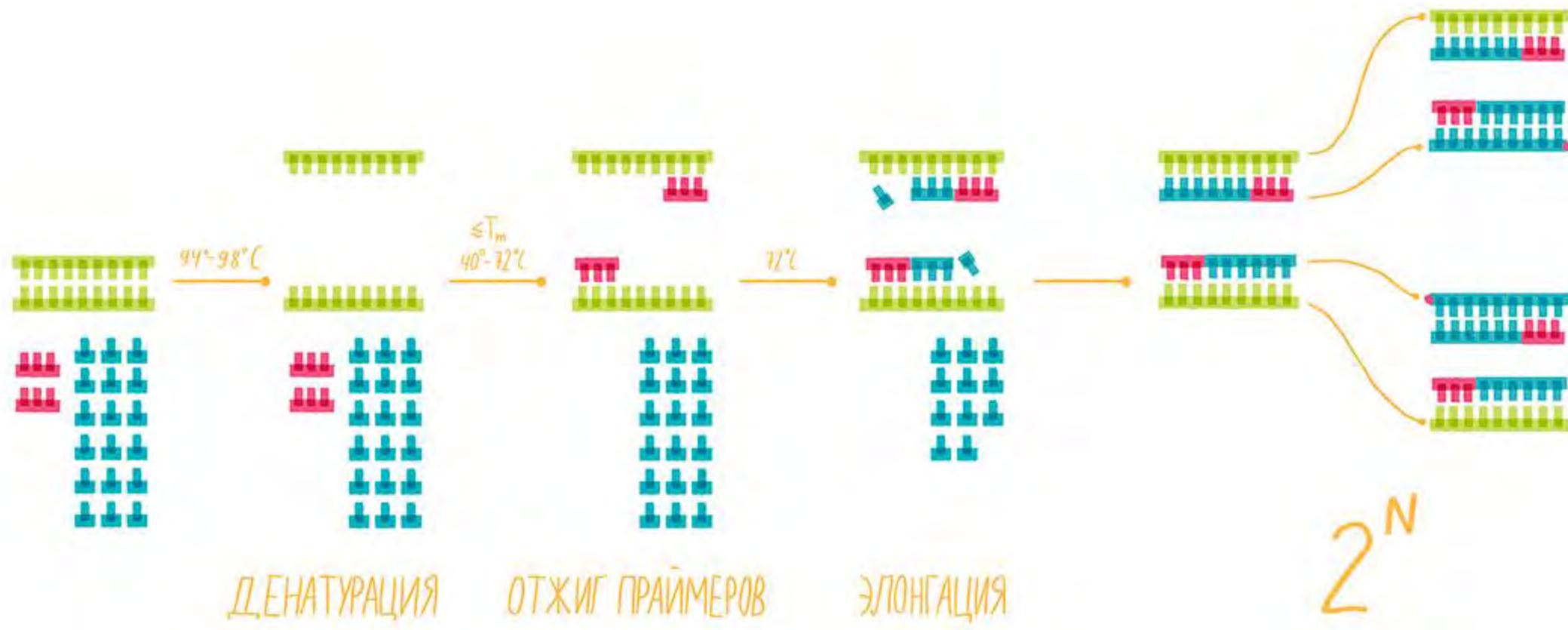
- PK - протеинкиназа
- PP - фосфатаза
- HAT - гистонацетилтрансфераза
- HDAC - гистондеацетилаза
- HMT - гистонметилтрансфераза
- HDM - гистондеметилаза

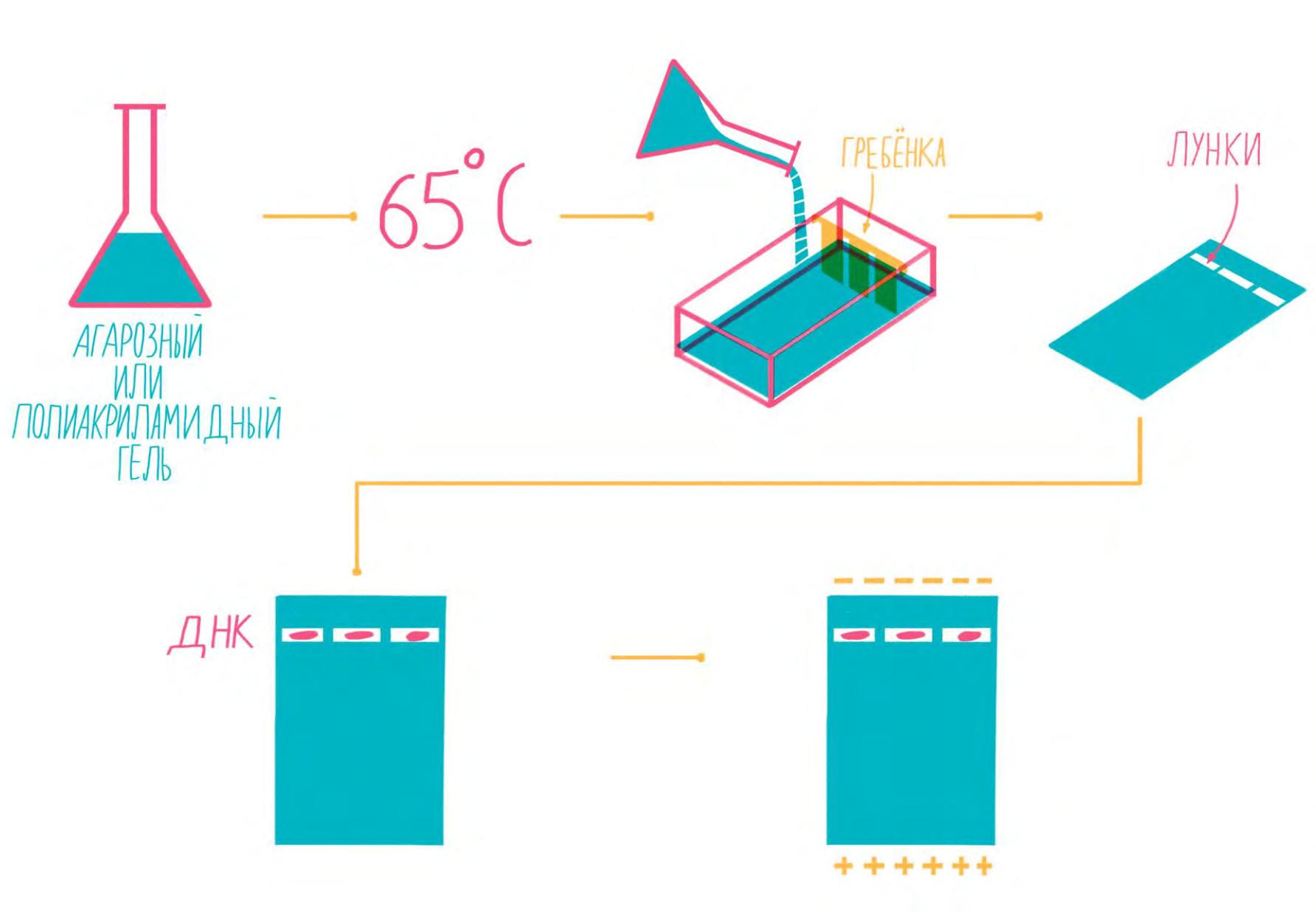
# Расщепление белка протеосомой



## Влияние посттрансляционных модификаций белков на процессы в клетке



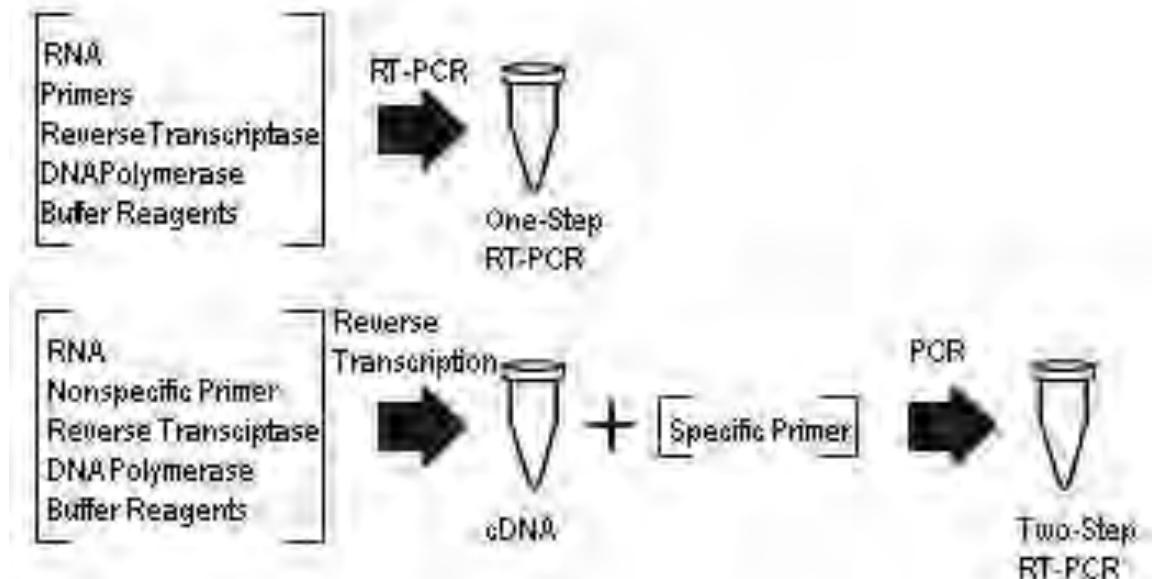
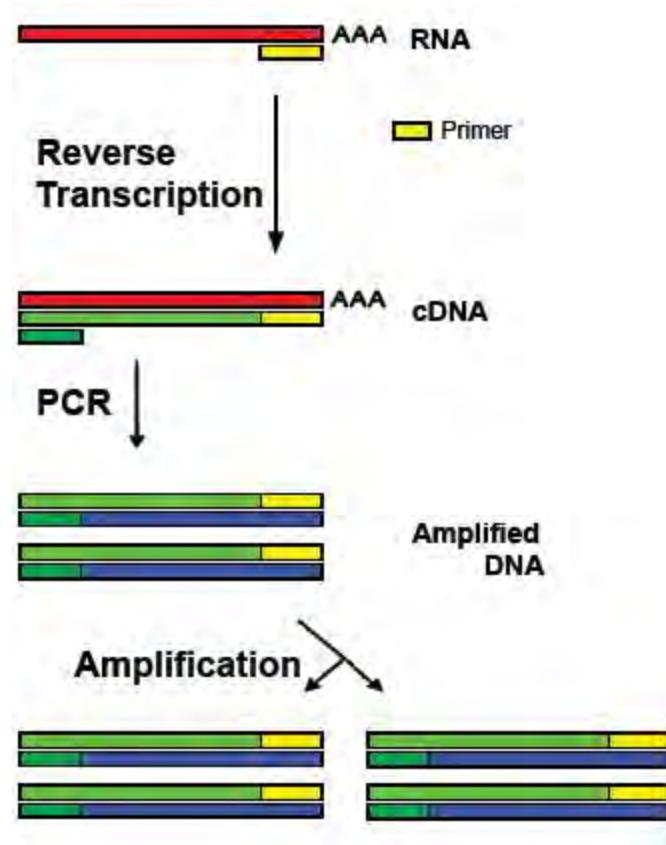




PHK



# ОТ-ПЦР



# Метки для RT-ПЦР

P – репортер

Г –гаситель

Taq-Ман

МОЛЕКУЛЯРНЫЙ МАЯК

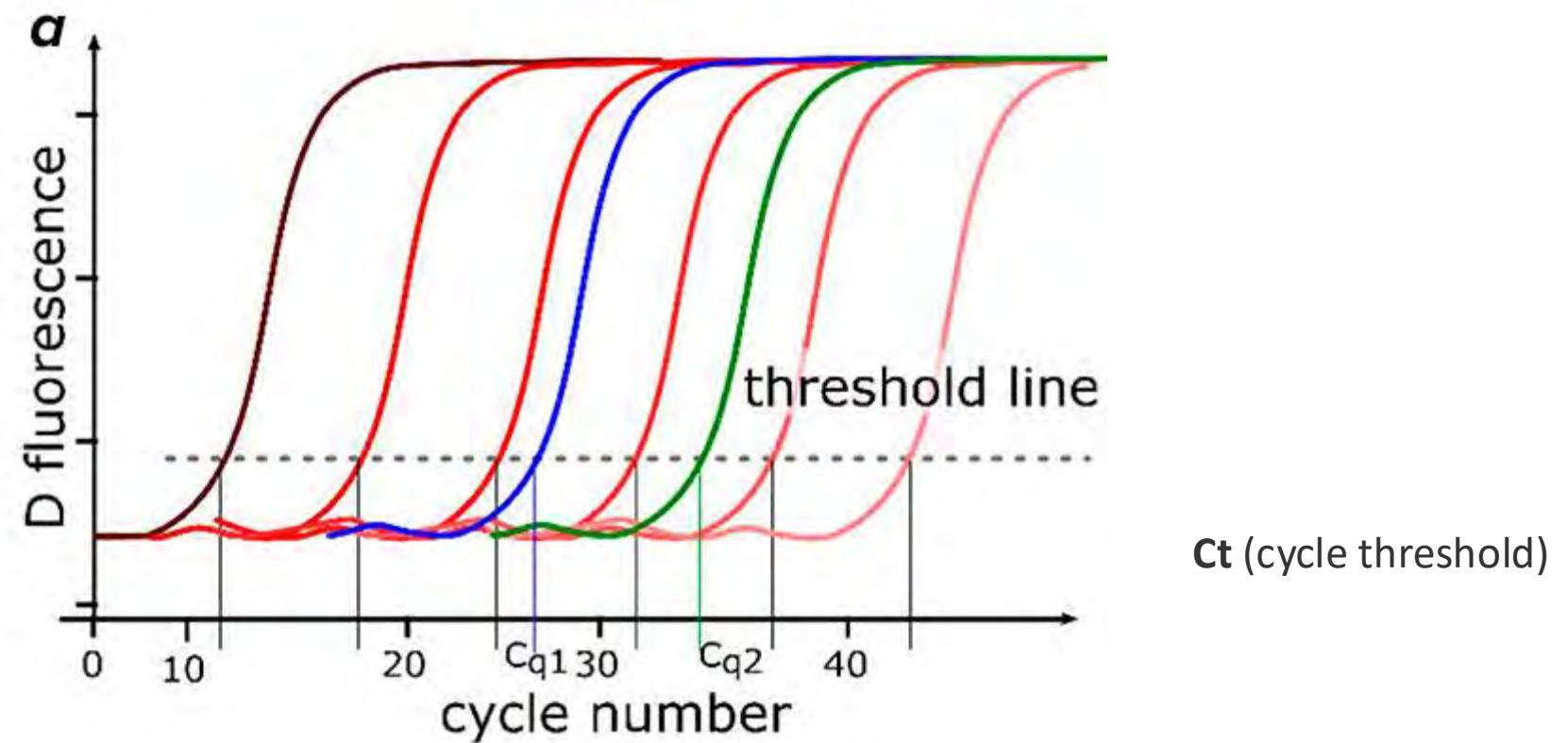
СКОРПИОН

ИНТЕРКАЛИРУЮЩИЙ  
КРАСИТЕЛЬ

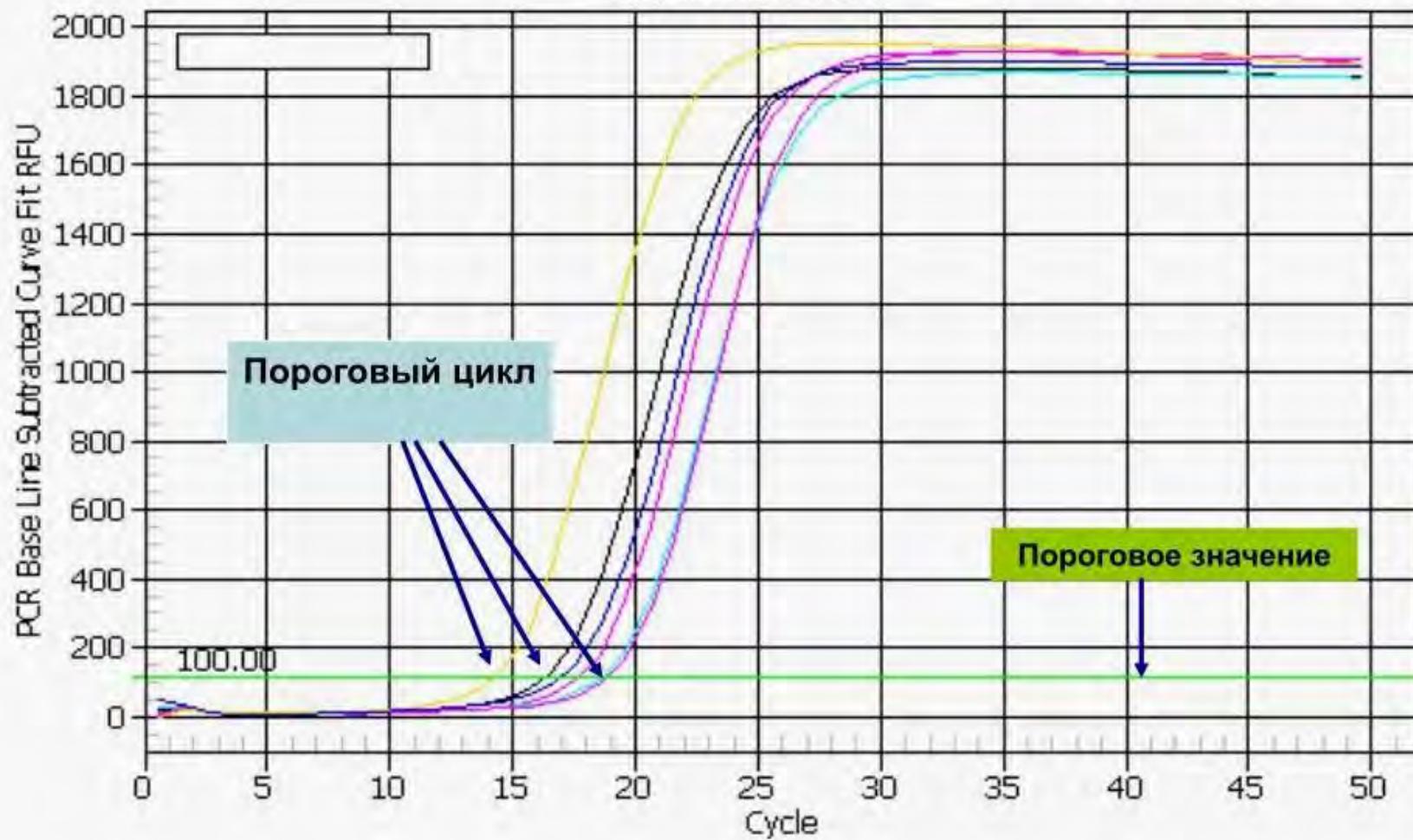
LUX-ПРАЙМЕР

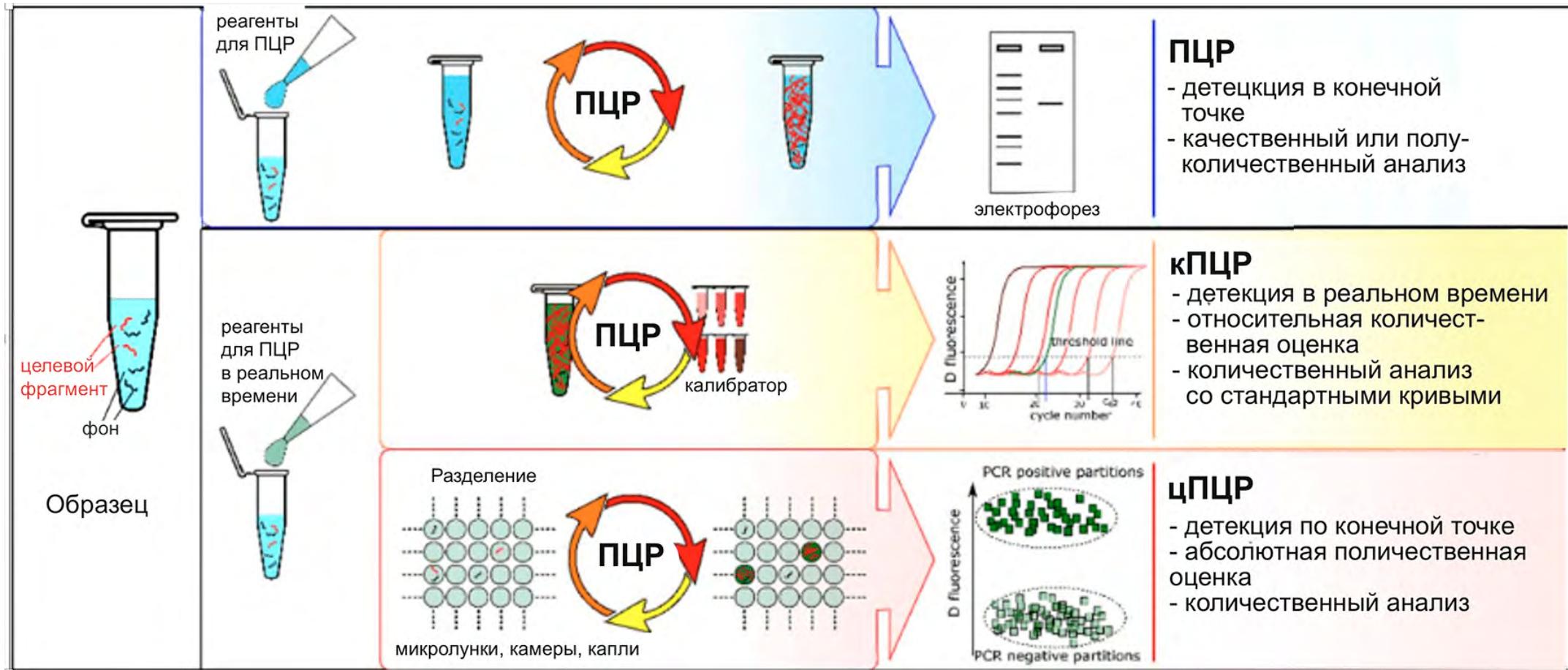
SUNRISE-ПРАЙМЕР

# RT-ПЦР

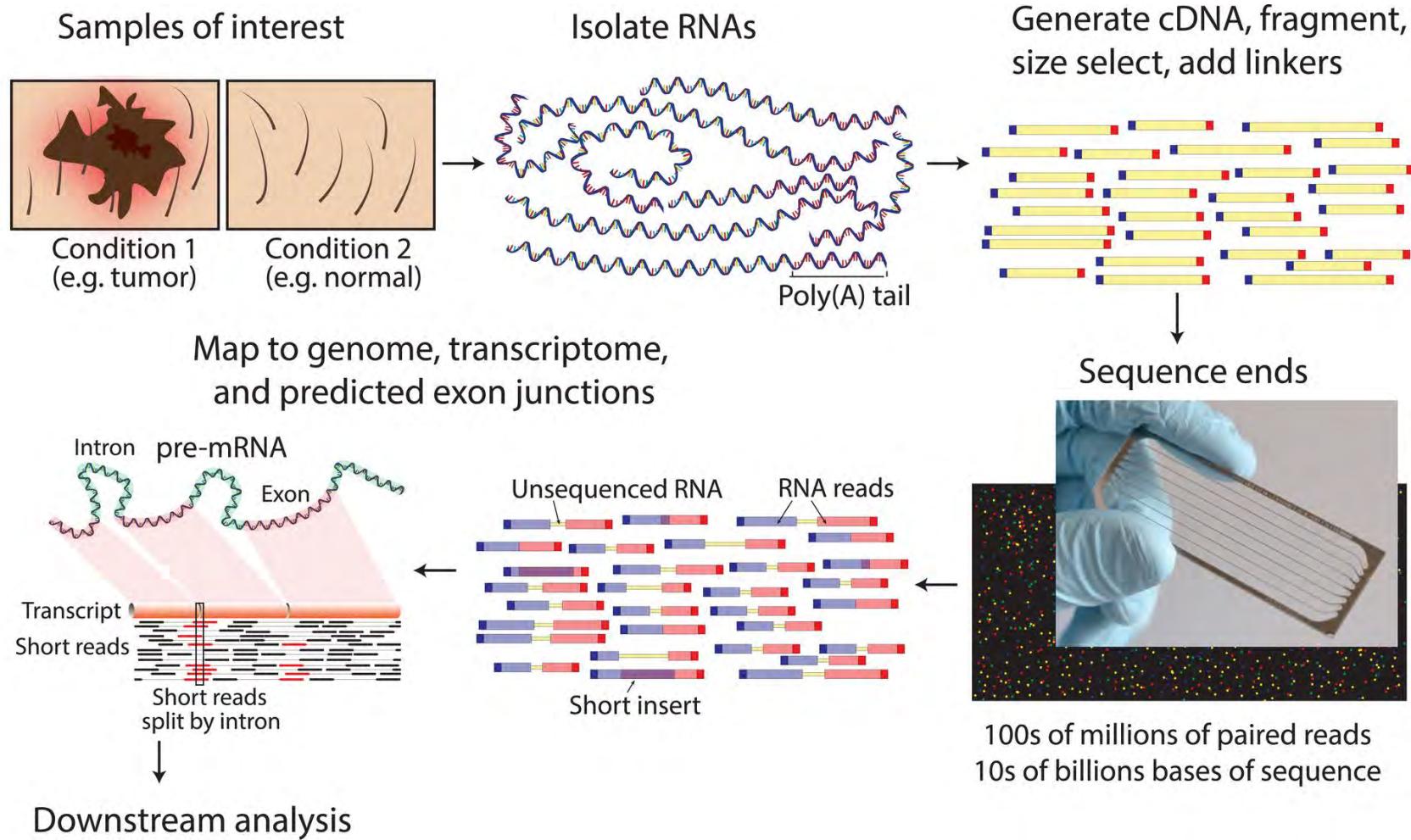


Amplification Chart





# Приименение NGS для исследований РНК



## Транскриптом прокариот:

рРНК – 80%

тРНК – 15 %

мРНК + некодирующие РНК – 5%

# Применение NGS для исследований РНК

Цель исследования	Целевые молекулы для библиотеки	Используемый подход
Уровень экспрессии генов	ПолиА-мРНК	Обогащение полиA-фракции. Часто достаточно коротких прочтений (50-100 п.н.) с 3'-конца
Альтернативный сплайсинг	Экзон-инtronные стыки	Необходимы длинные прочтения (более 300 п.н.) или парноконцевые прочтения (по 100-150 п.н.)
Малые РНК (miRNA, snoRNA, piRNA, sn RNA, tRNA)	Короткая фракция кДНК	Использование 5'-концевого фосфата. Достаточно коротких прочтений (50-100 п.н.)
Антисмыловые и некодирующие РНК	Антисмыловые и некодирующие РНК	Необходимо создание 5'-3'-ориентированной библиотеки кДНК
Транскриптом одной клетки	ПолиА-мРНК	Необходима предварительная амплификация библиотеки кДНК
Транслируемые РНК	РНК на рибосомах	Ферментативное разрушение РНК (кроме участков внутри рибосом)
Двухцепочечные РНК и вторичные структуры на РНК	Двухцепочечные участки РНК	Обработка ацетилирующими агентами

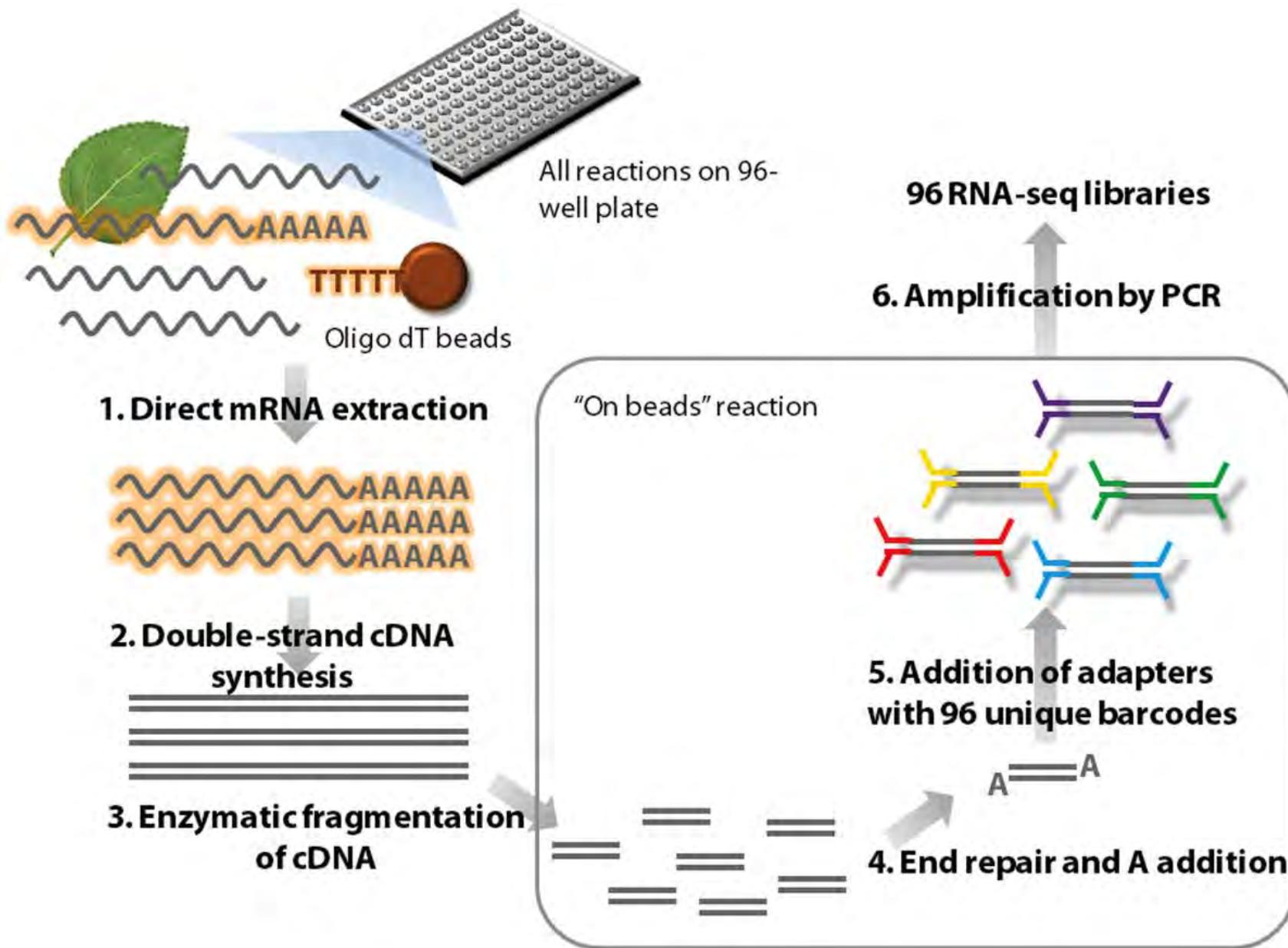
# Общие моменты очистки РНК и синтеза кДНК

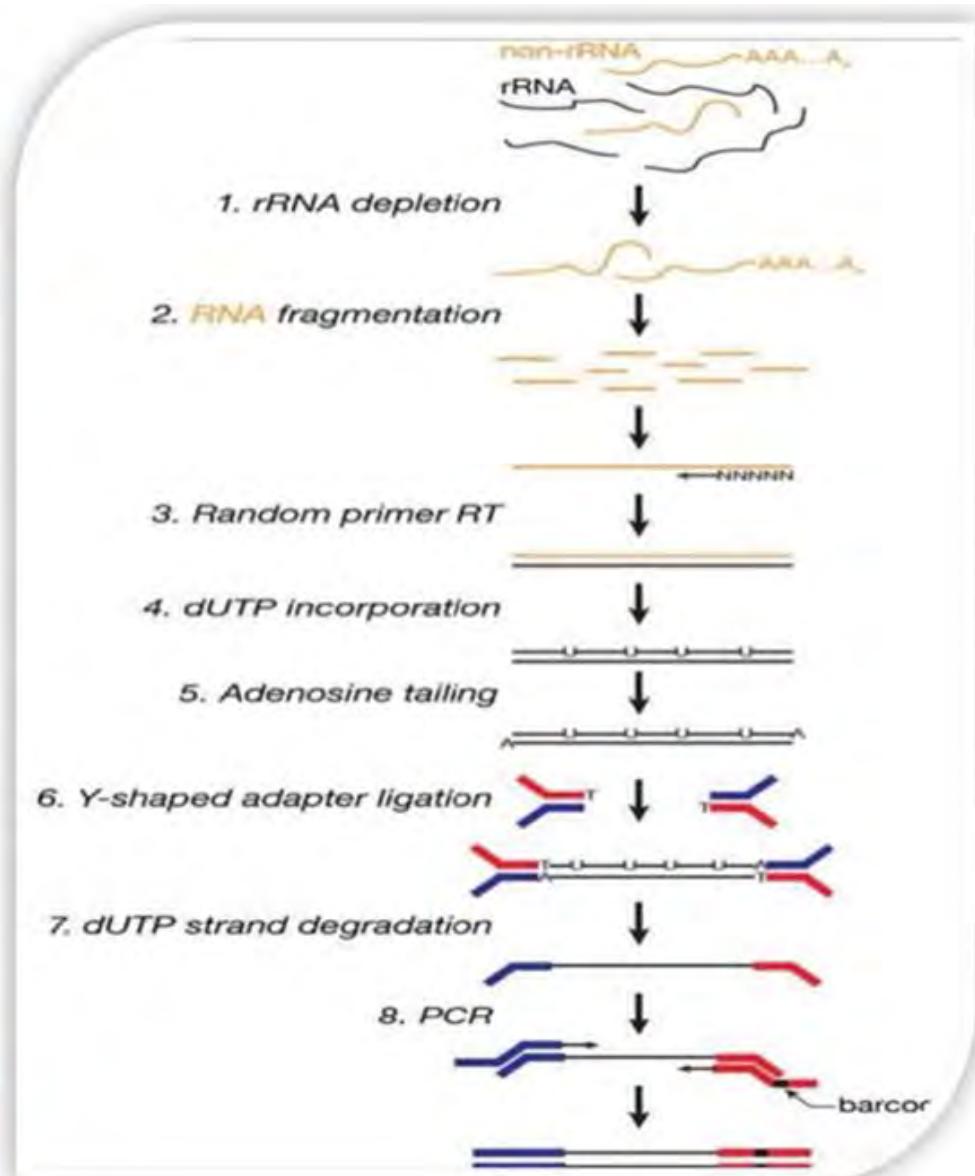
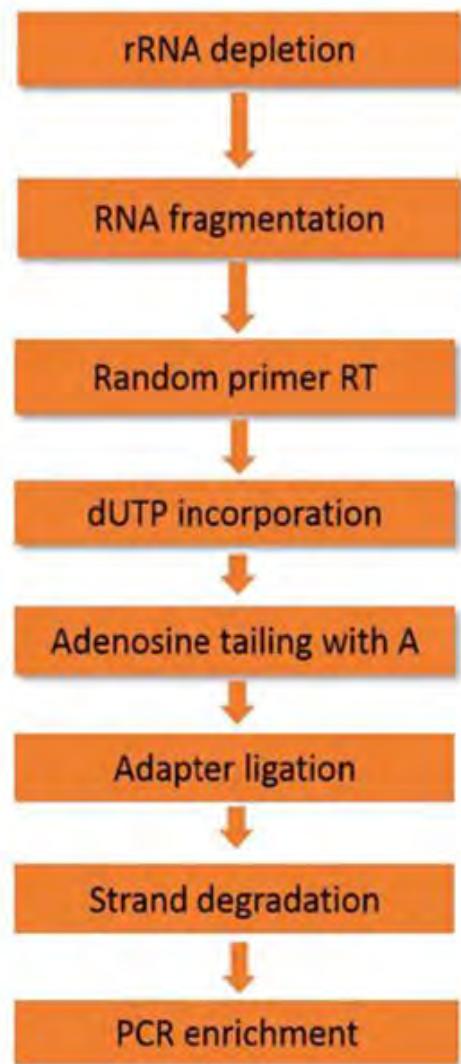
- Финальная очистка препарата ДНКазами («RNase free»!!!)
- Подбор праймеров:
  - самозатравление (без дополнительных праймеров)
  - случайные праймеры (random primers) (6-8 нуклеотидов)\*
  - праймеры с олиго-dT (праймирование одновременно всех мРНК)\*
  - праймеры, специфичные к отдельным транскриптам

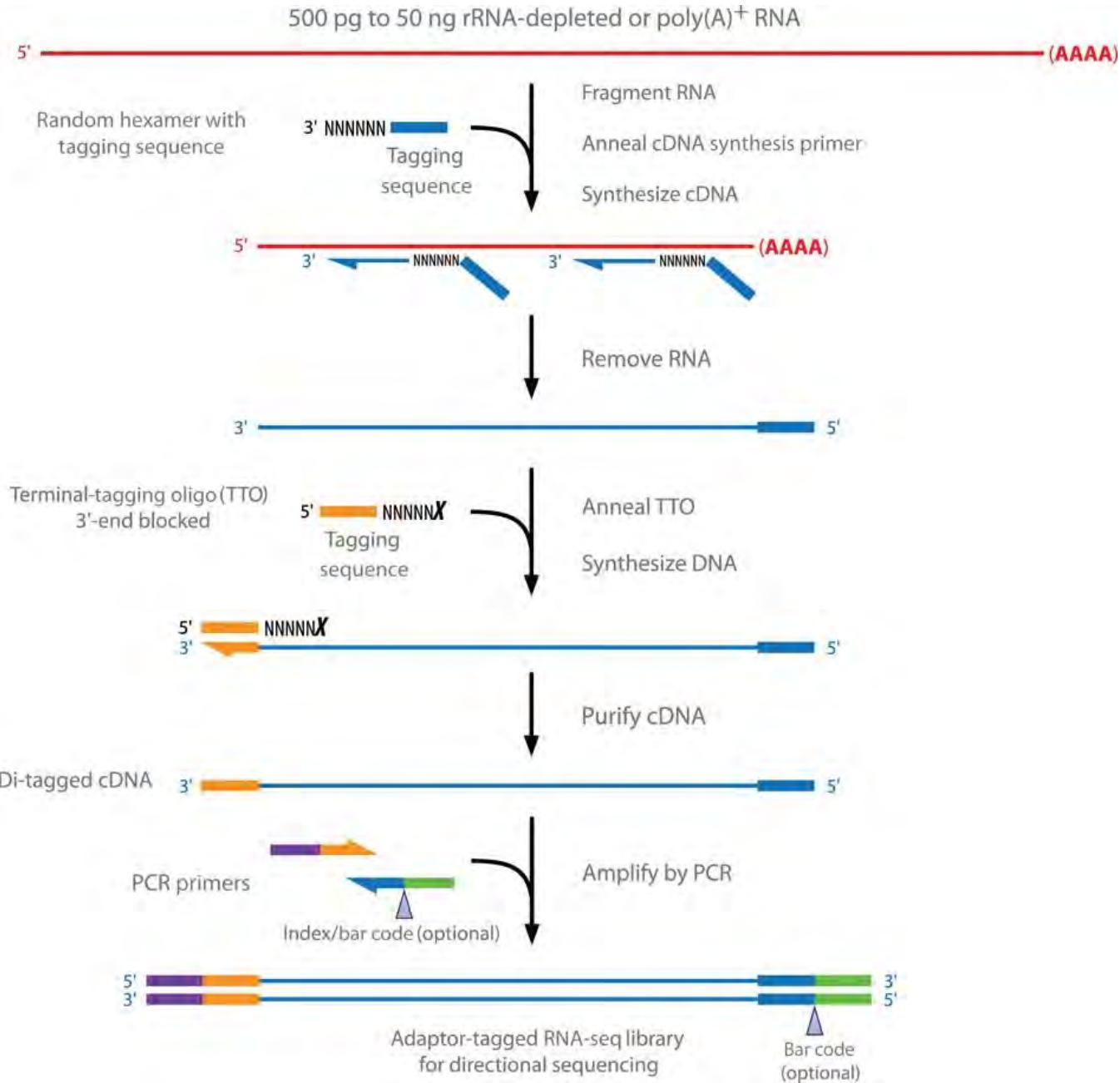
\* Температура отжига ниже, чем температуры, оптимальной для работы термостабильной реакции обратной транскрипции.

# Общие моменты очистки РНК и синтеза кДНК

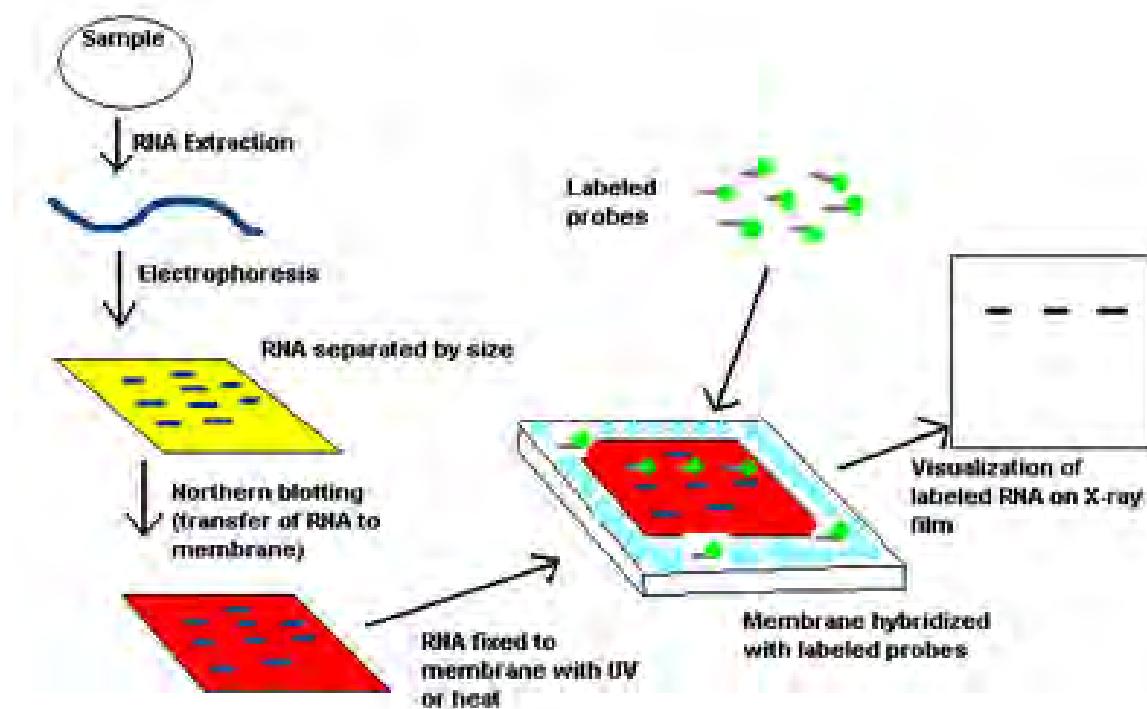
- Ревертазы по сравнению с ДНК-полимеразами чаще допускают ошибки, менее термостабильны (высокая температура требуется для денатурации вторичных структур РНК).



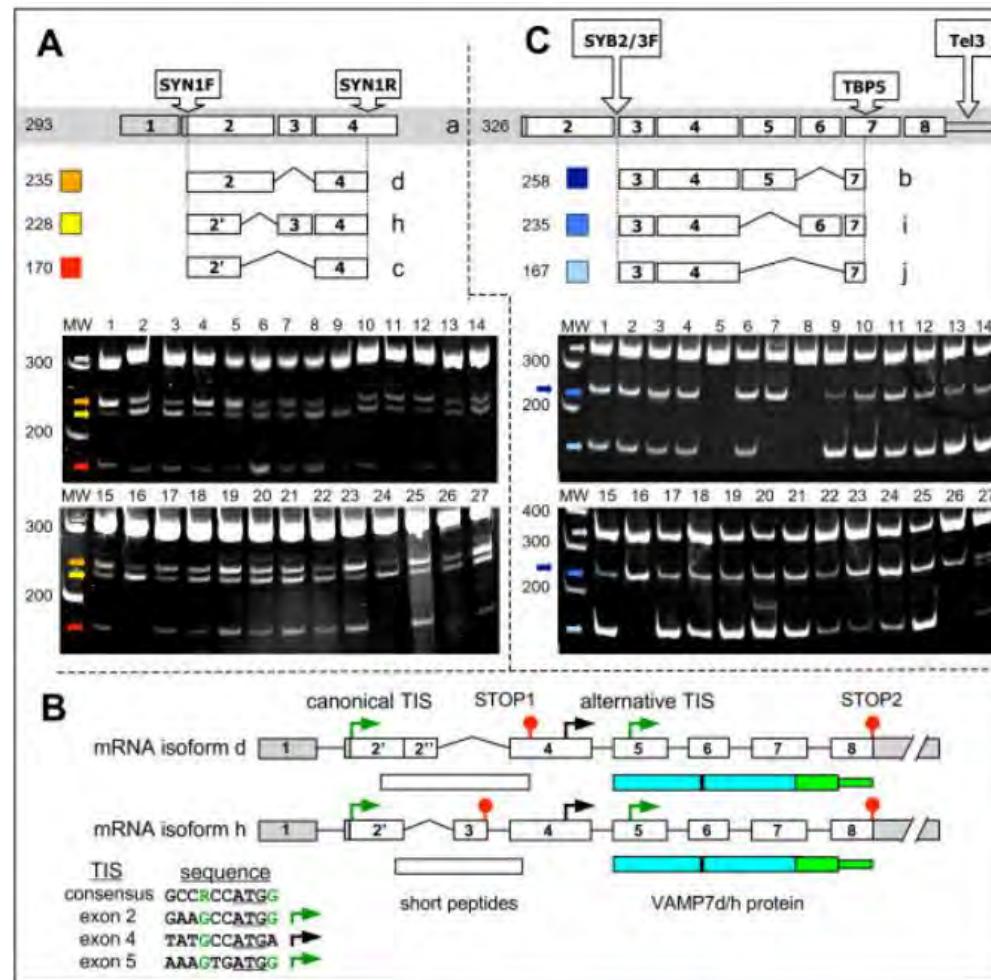




# Northern blot

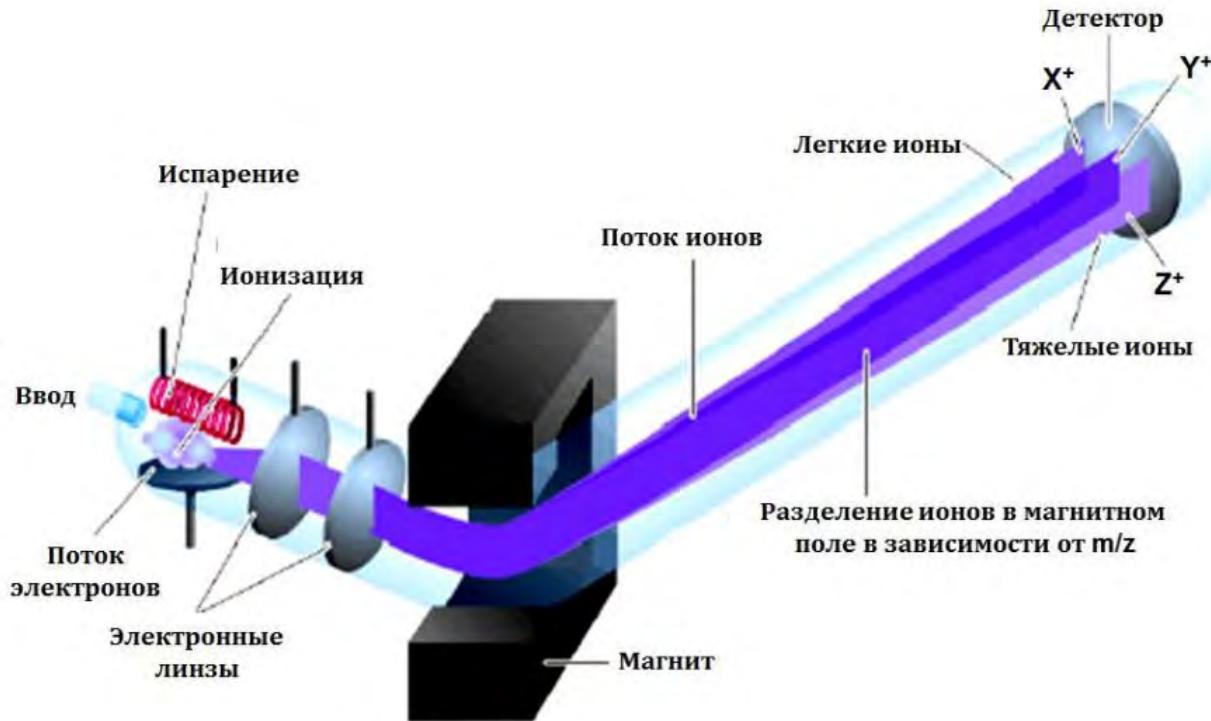


# Исследование альтернативного сплайсинга



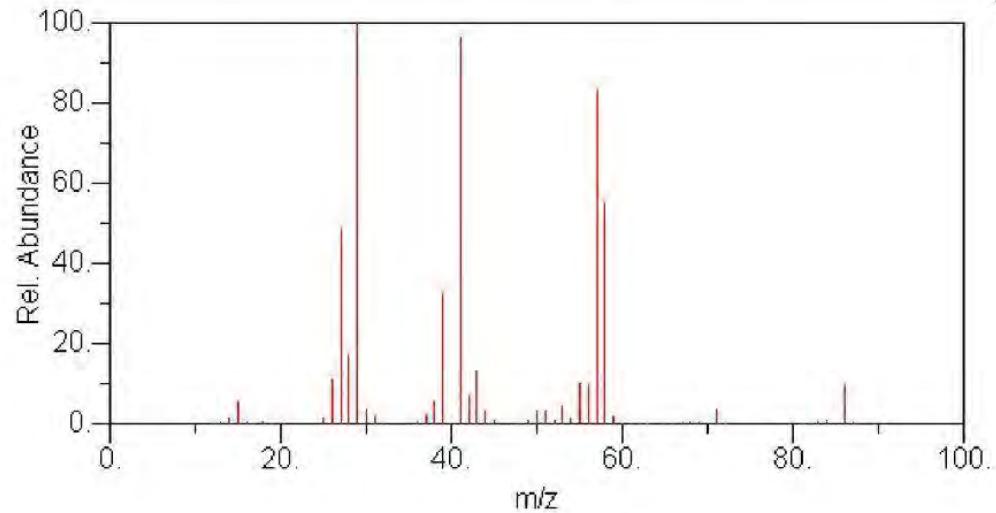


# Масс-спектрометрия



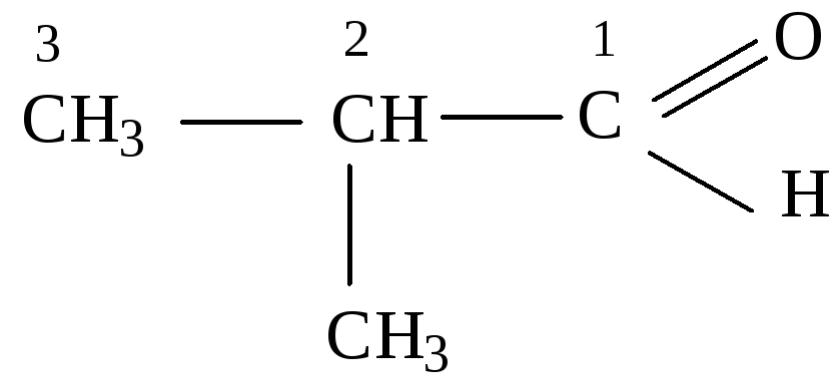
Ионизация компонентов позволяет физически (измеряя интенсивность ионного тока) различить компоненты на основе отношения массы к заряду , производить отдельный подсчёт доли каждого из компонентов (получать *масс-спектр* вещества).

### Масс-спектр метилбутаналя

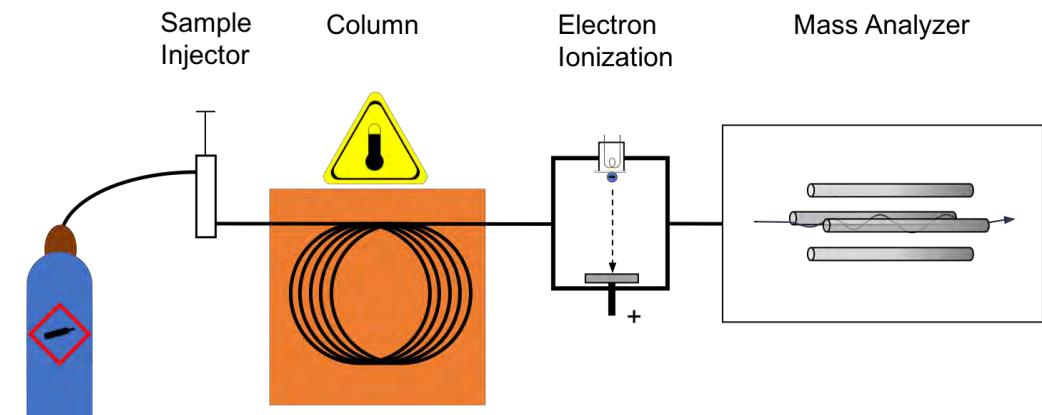


Каждый пик соответствует определенному молекулярному кациону, образующемуся при ионизации.

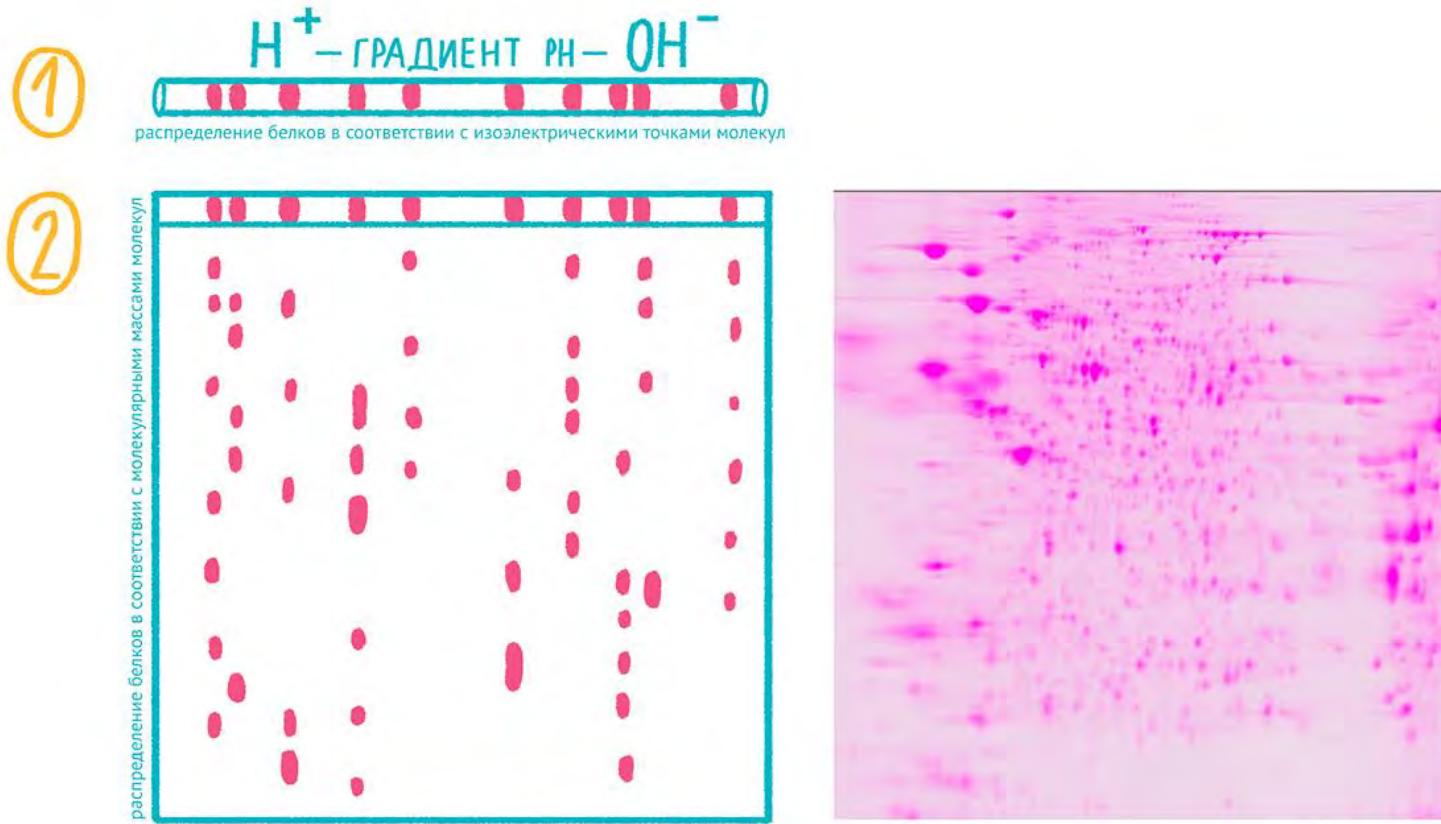
70

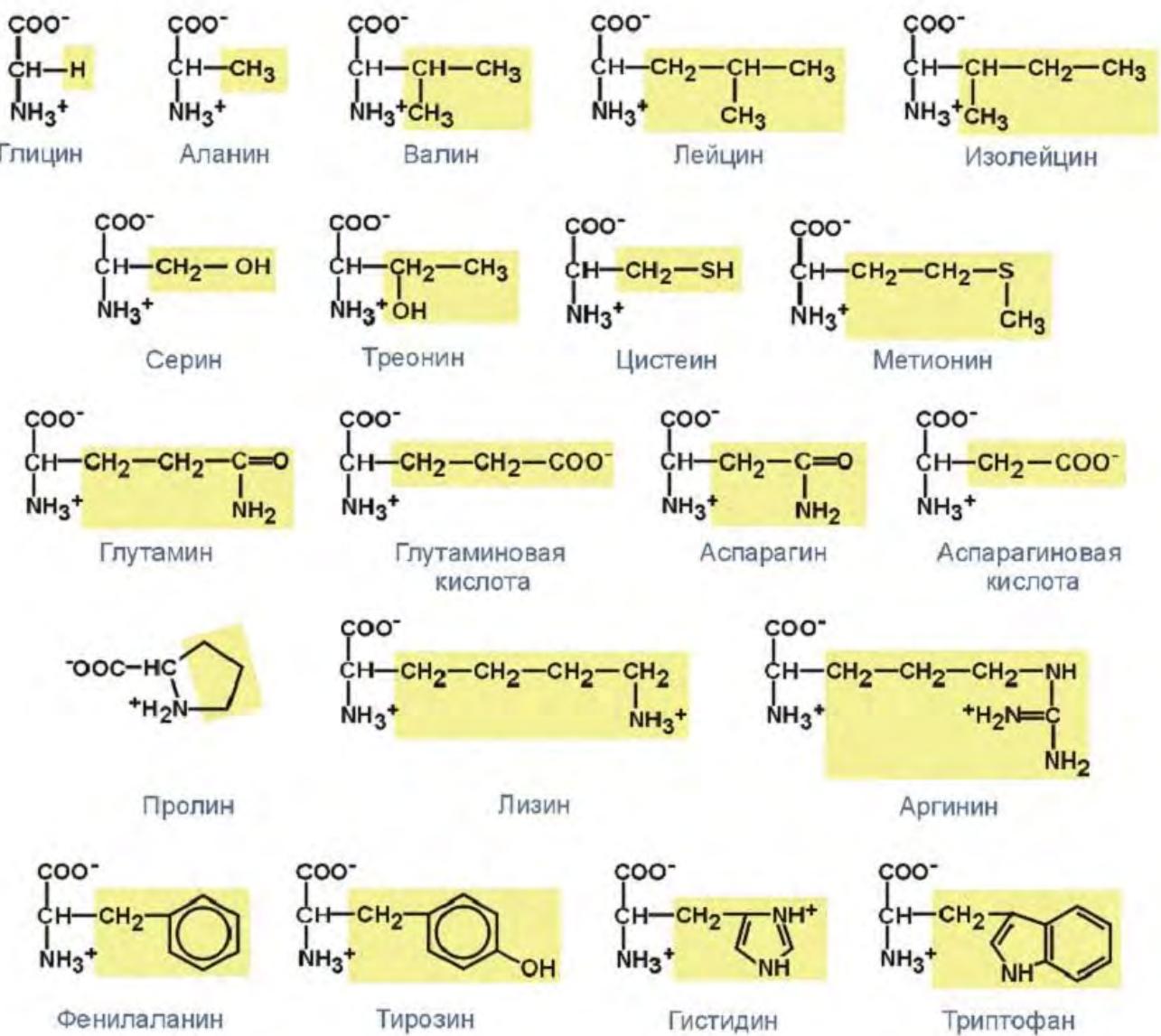


# Хроматография-масс-спектрометрия (Х-МС)

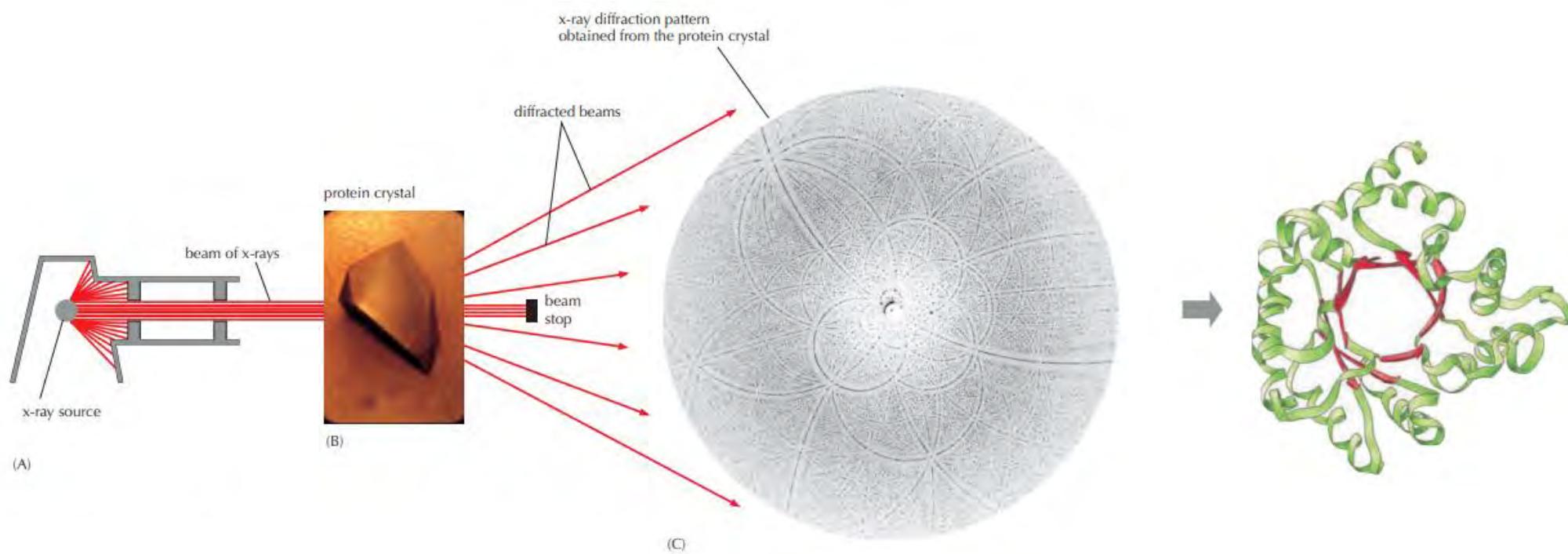


# Двумерный электрофорез белков

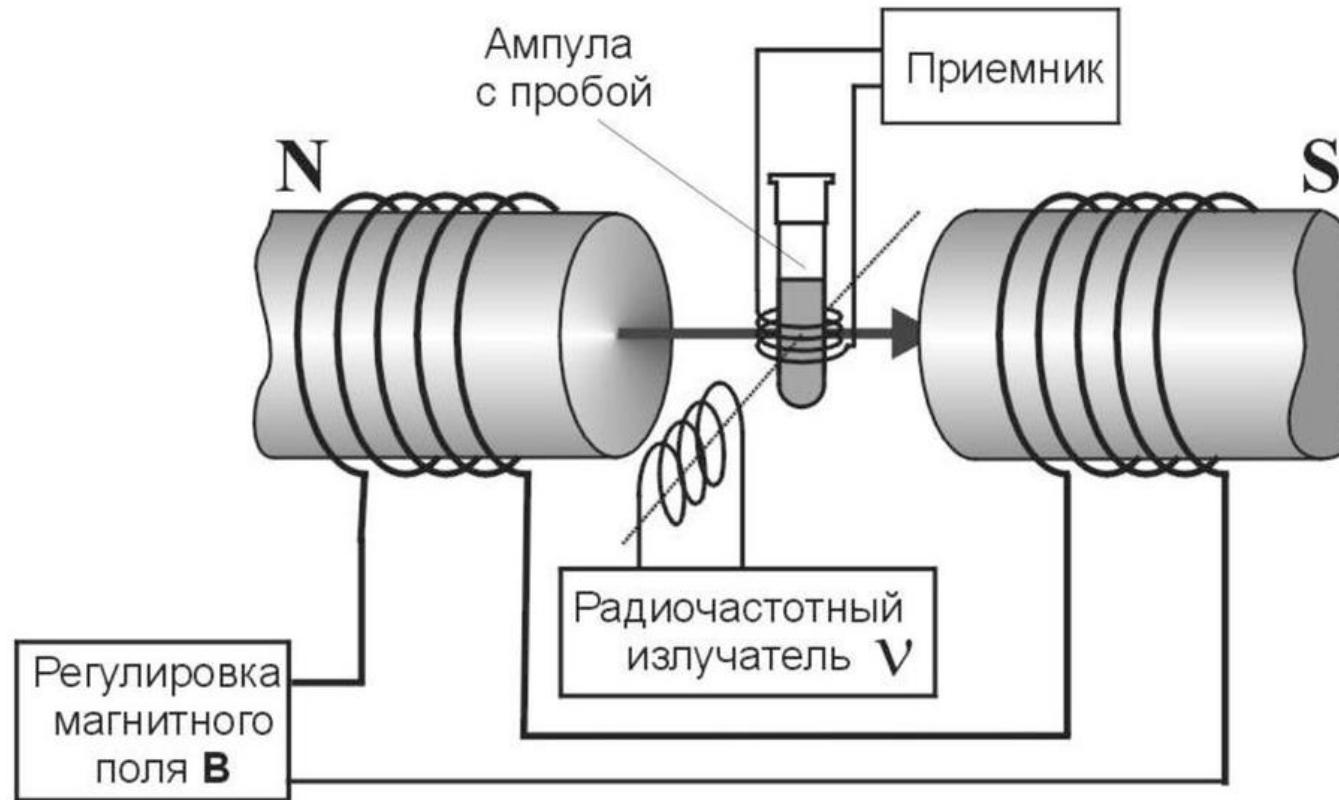




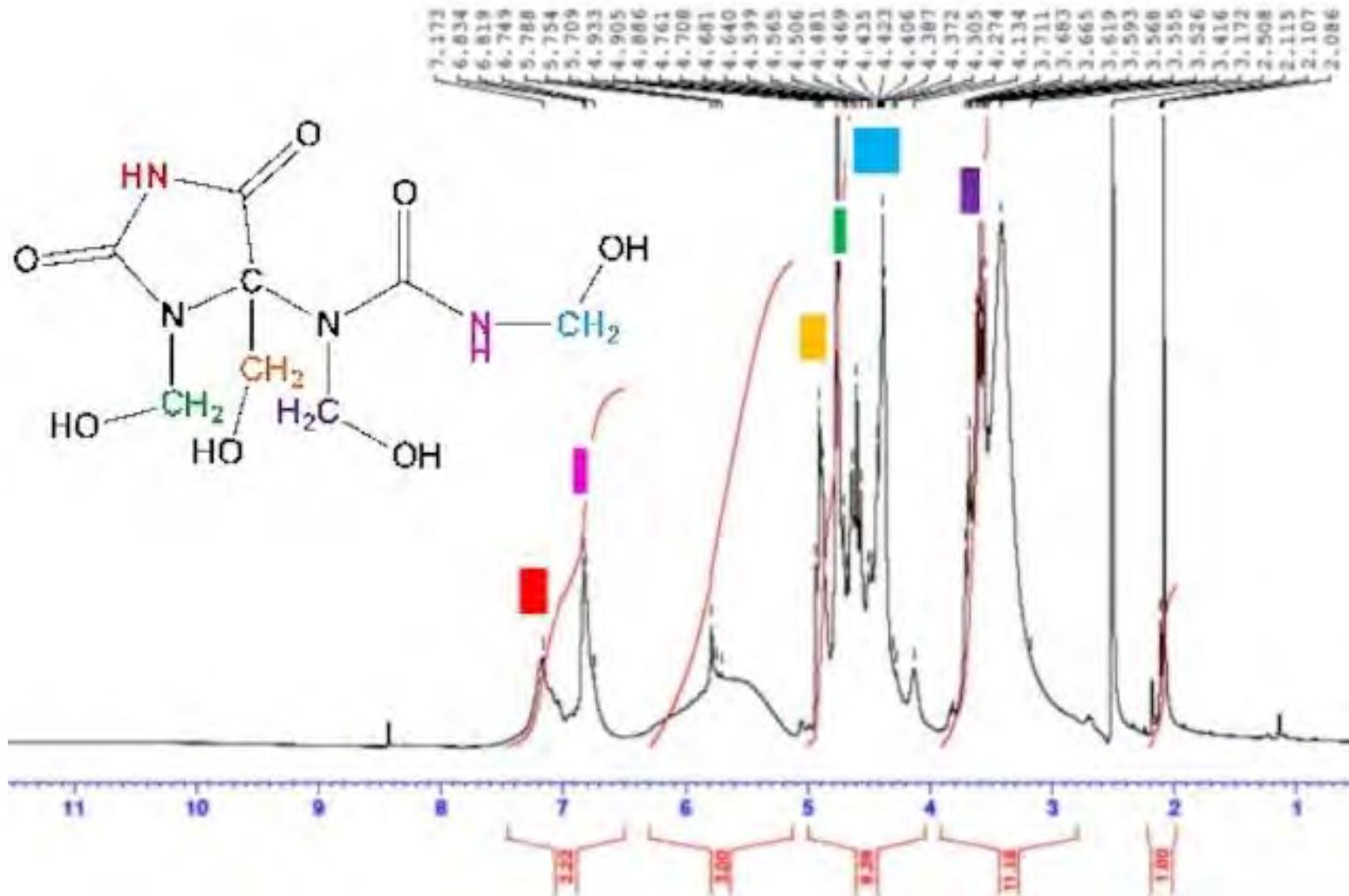
# Рентгеноструктурный анализ белка



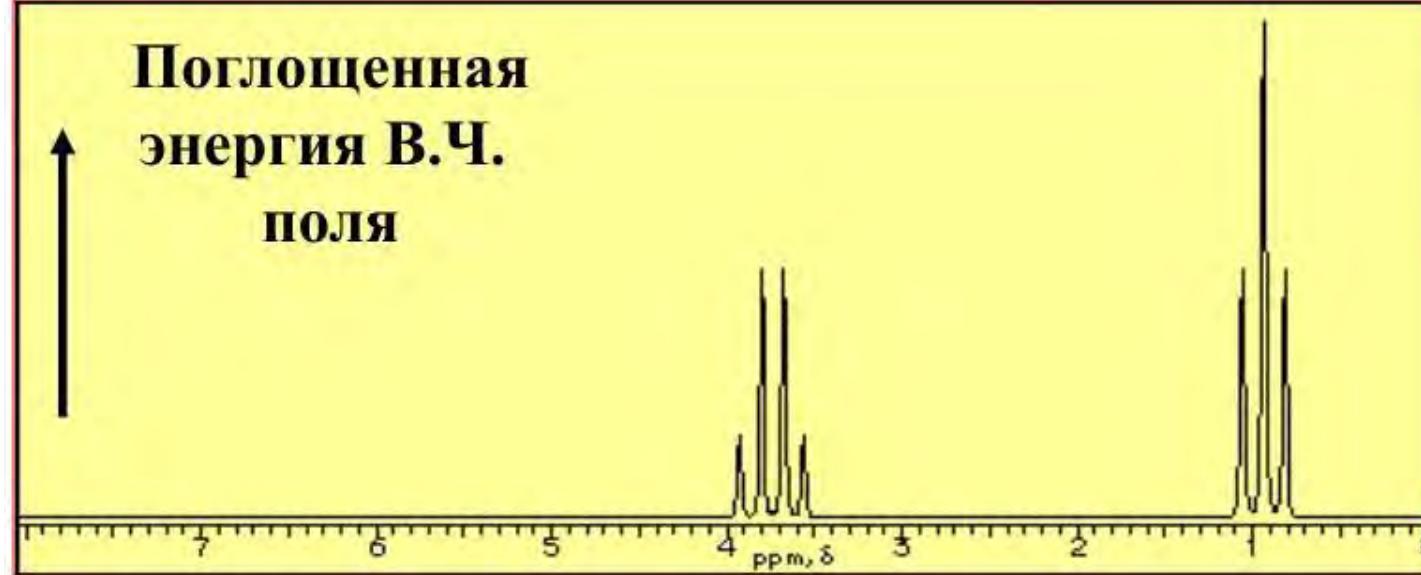
# ЯМР-спектроскопия в изучении белков



Проба исследуемого вещества (в виде жидкости или раствора) помещается в однородное магнитное поле. Затем изменяют напряжённость магнитного поля до тех пор, пока не наступает явление резонанса.

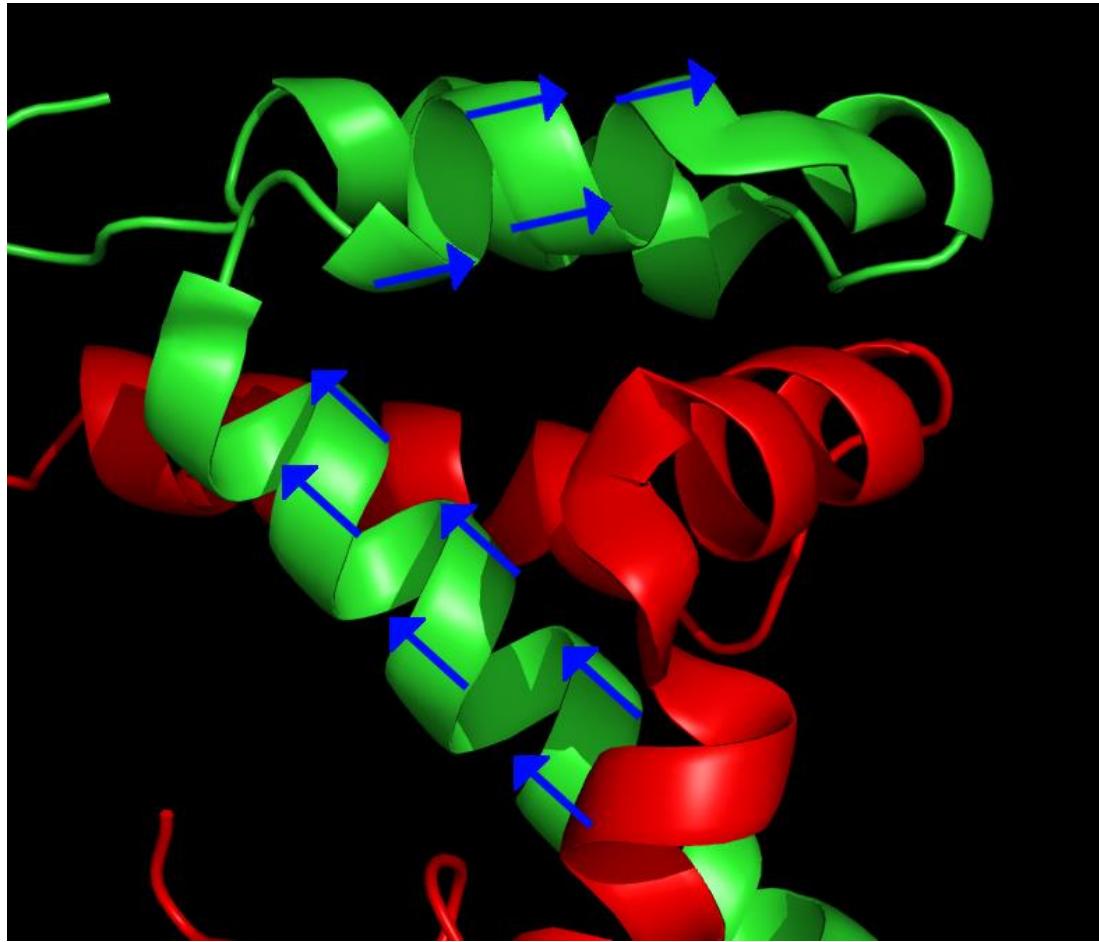


**Поглощенная  
энергия В.Ч.  
поля**



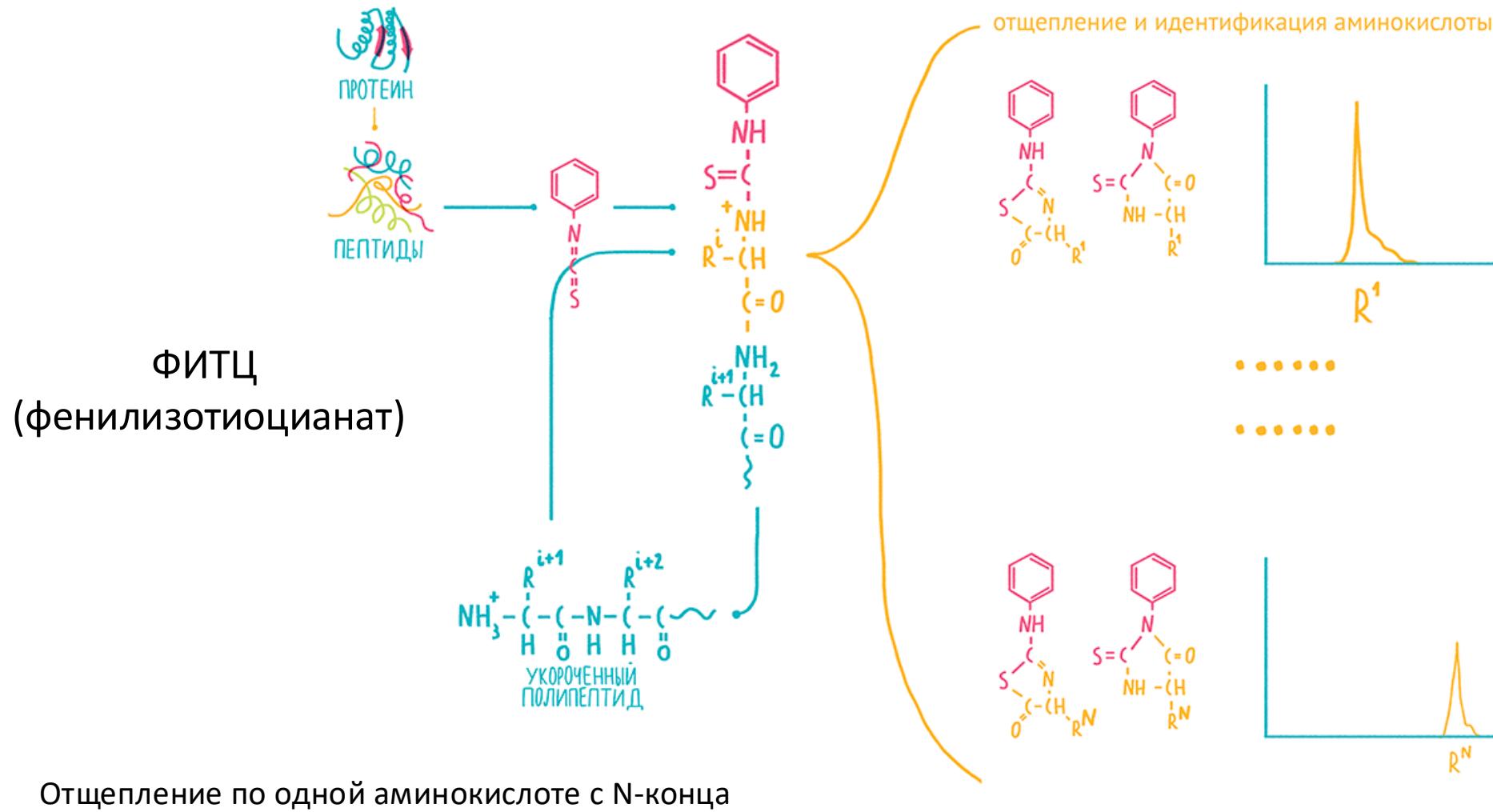
**Частота В.Ч. поля** —————→

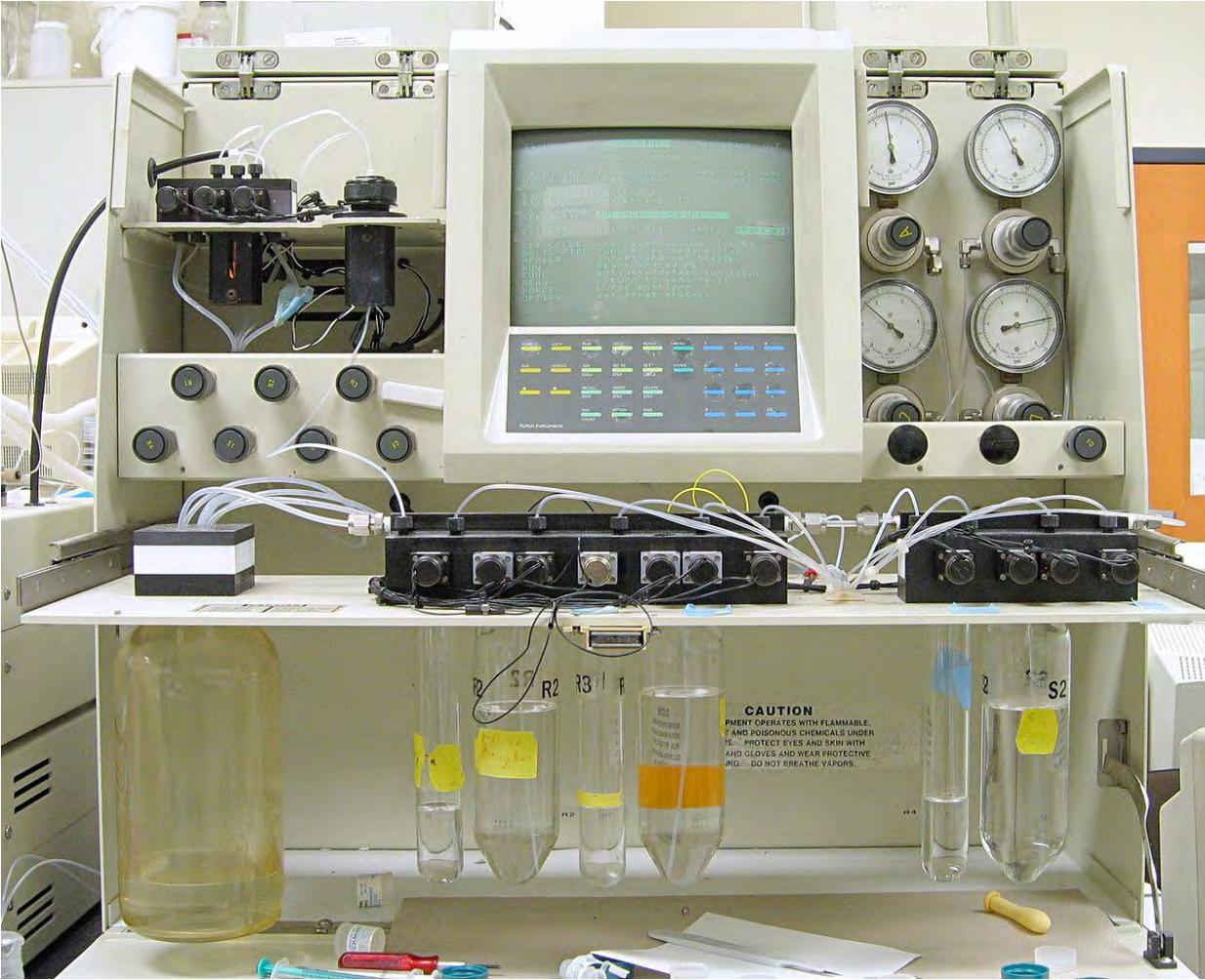
**Спектр  
ЯМР  
 $^1\text{H}$   
в молекуле  
 $\text{CH}_2\text{CH}_3$**



Синие стрелки обозначают ориентацию связи N – H выбранных пептидных связей.  
Определяя ориентацию достаточного количества связей относительно внешнего  
магнитного поля, можно определить структуру белка.

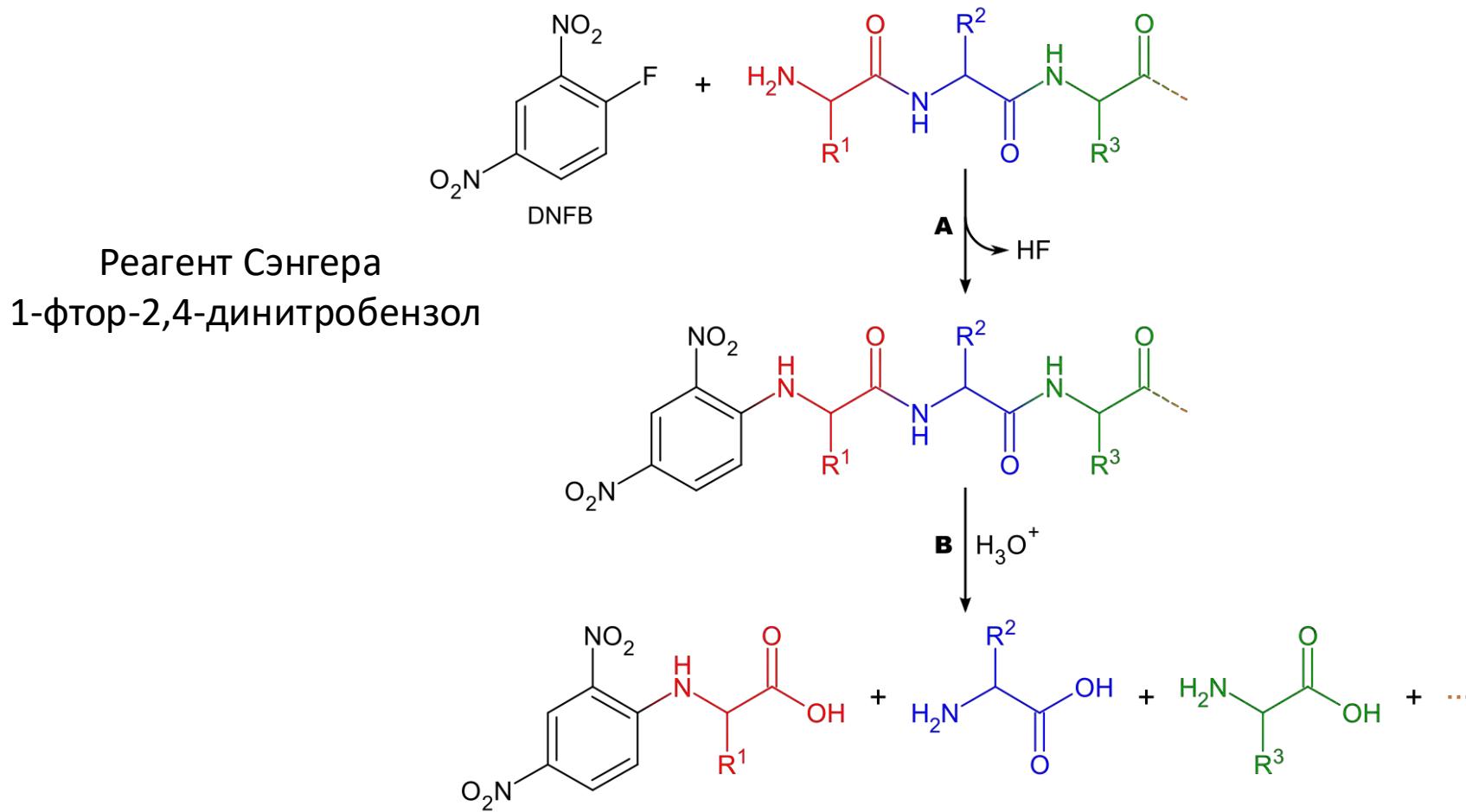
# Секвенирование белка (метод Эдмана)



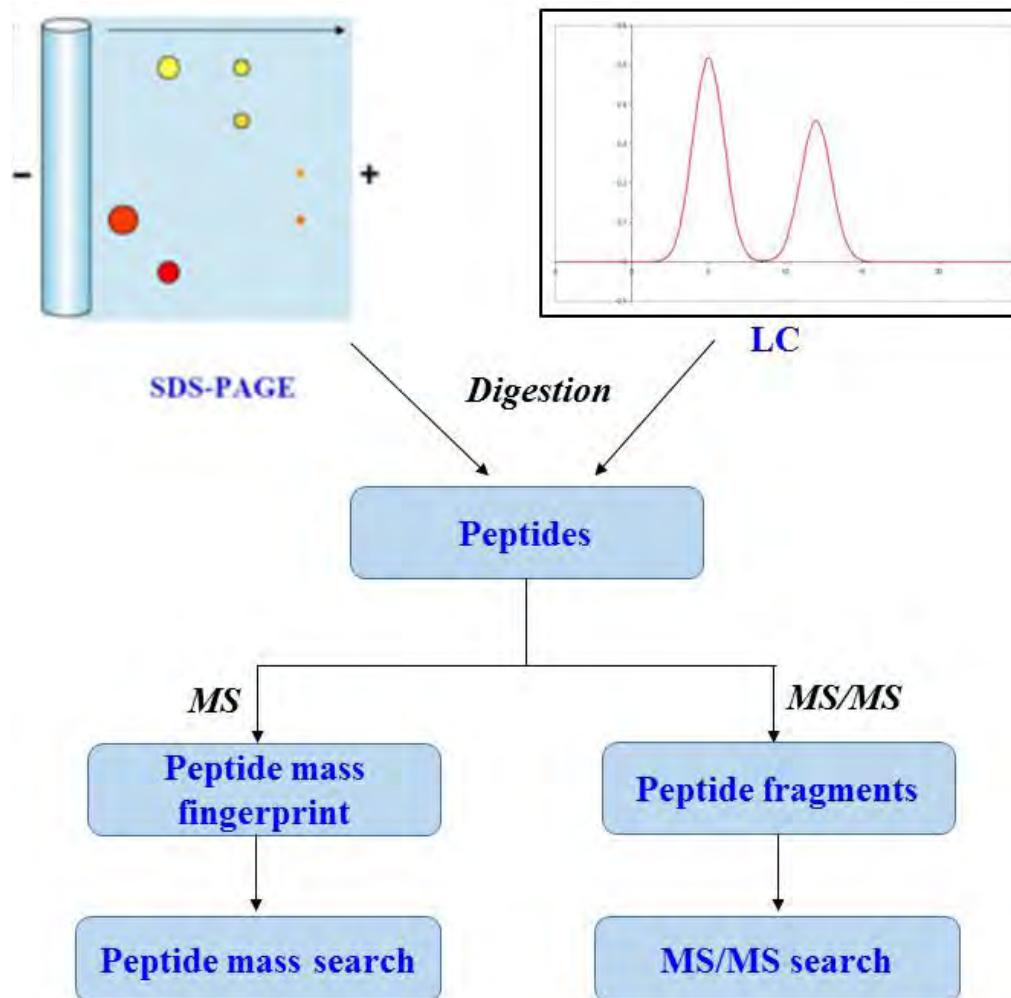


Секвенатор Beckman-Coulter Porton LF3000G (по Эдману)

# Секвенирование белка (метод Сэнгера)

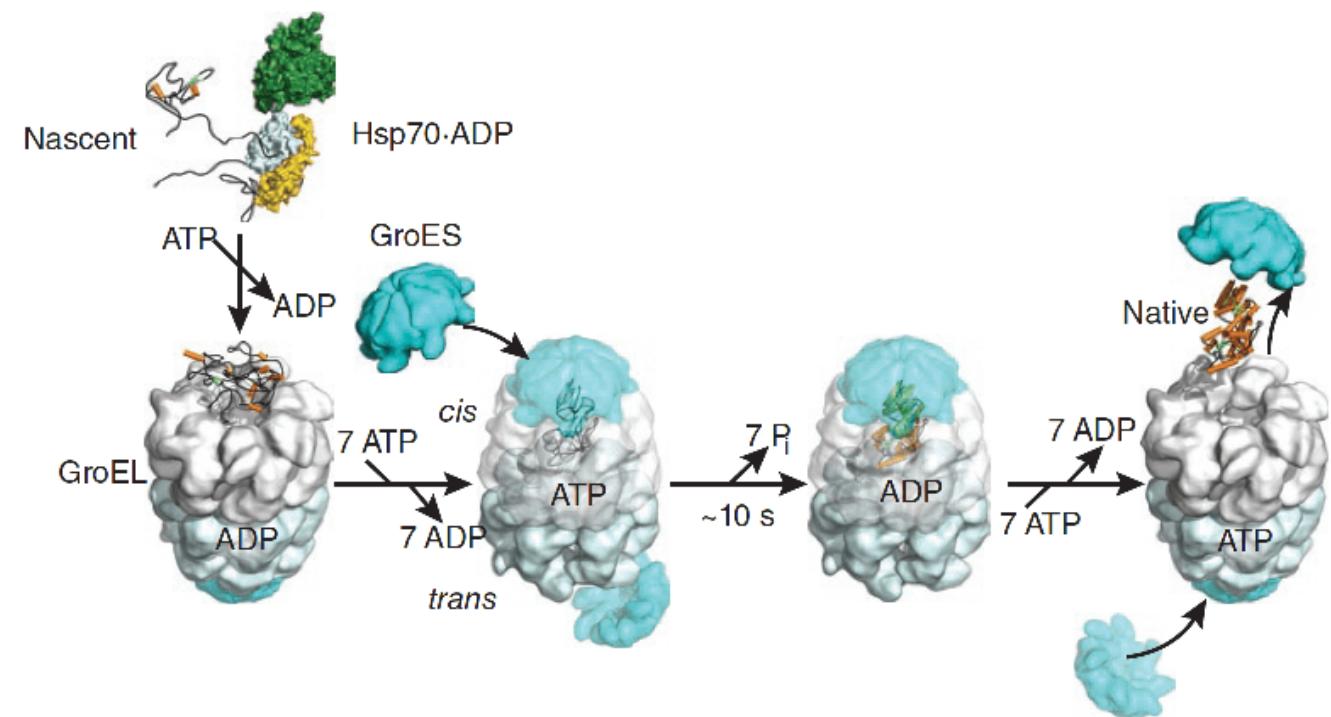
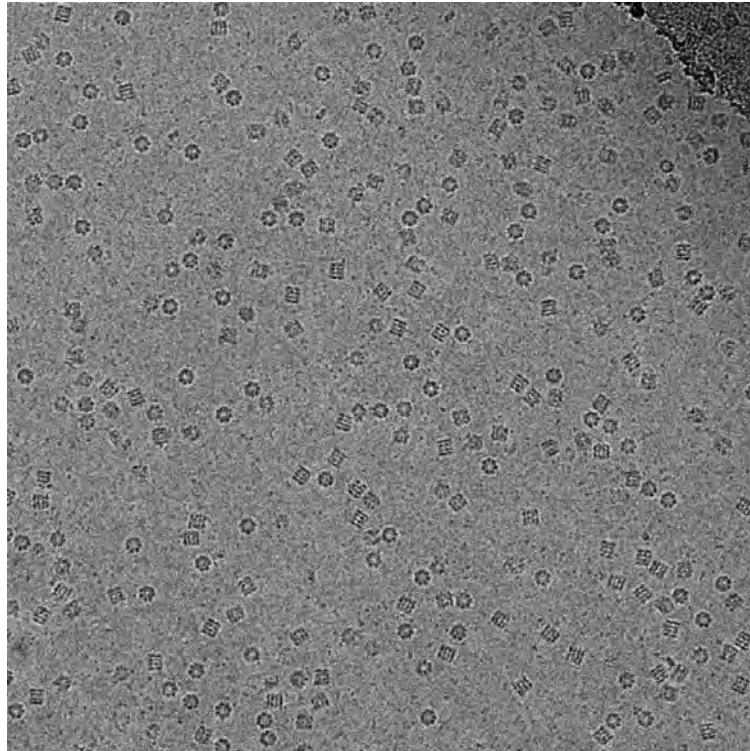


полный кислотный гидролиз динитрофенильного пептида

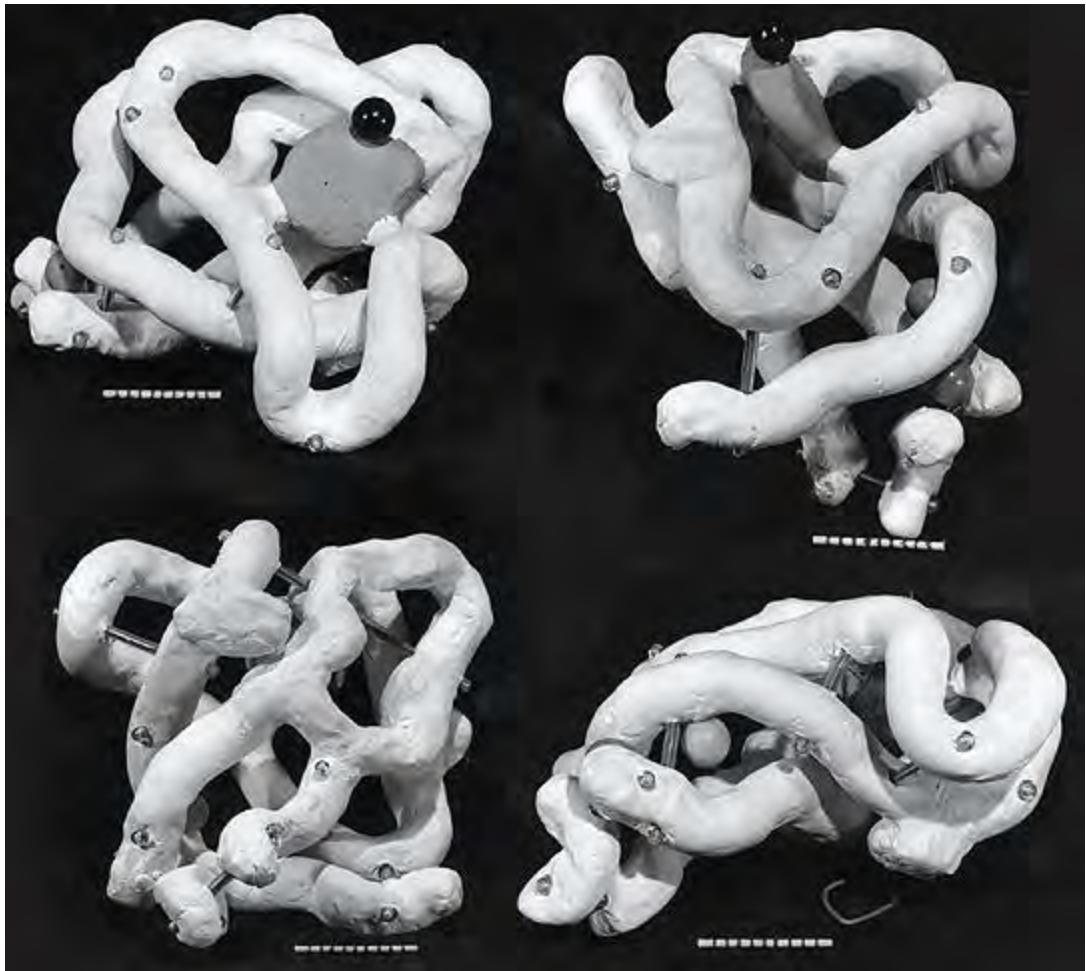
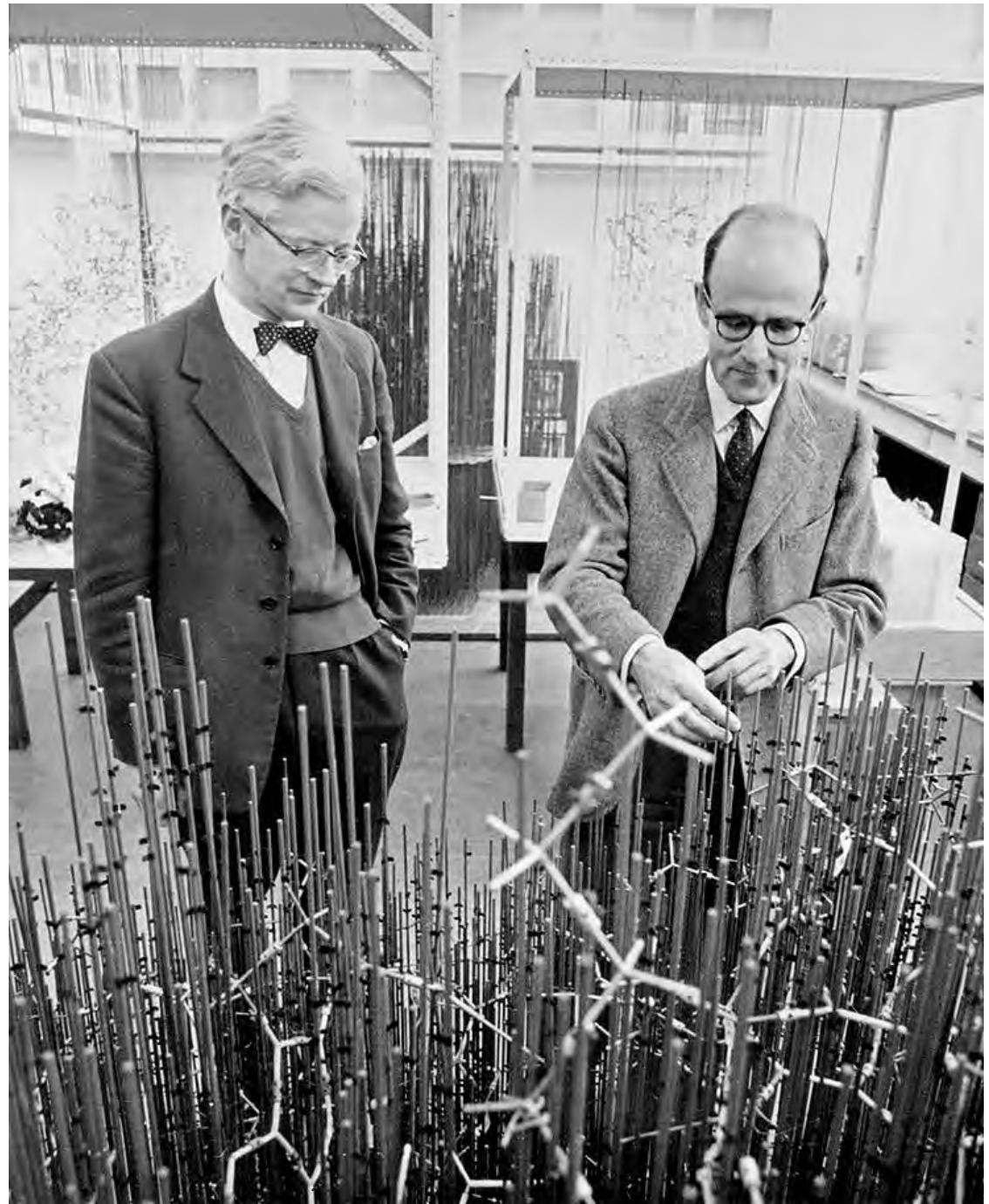


**Массовая дактилоскопия пептидов (PMF)** - идентификации белка, при котором неизвестный интересующий белок сначала расщепляется на более мелкие пептиды, абсолютные массы которых можно точно измерить с помощью масс-спектрометра, такого как MALDI-TOF или ESI-TOF.

# Криоэлектронная микроскопия



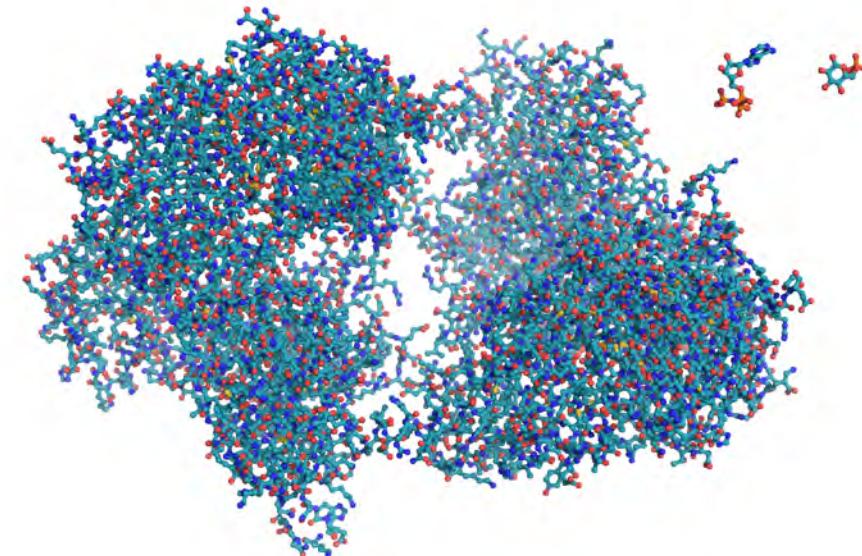
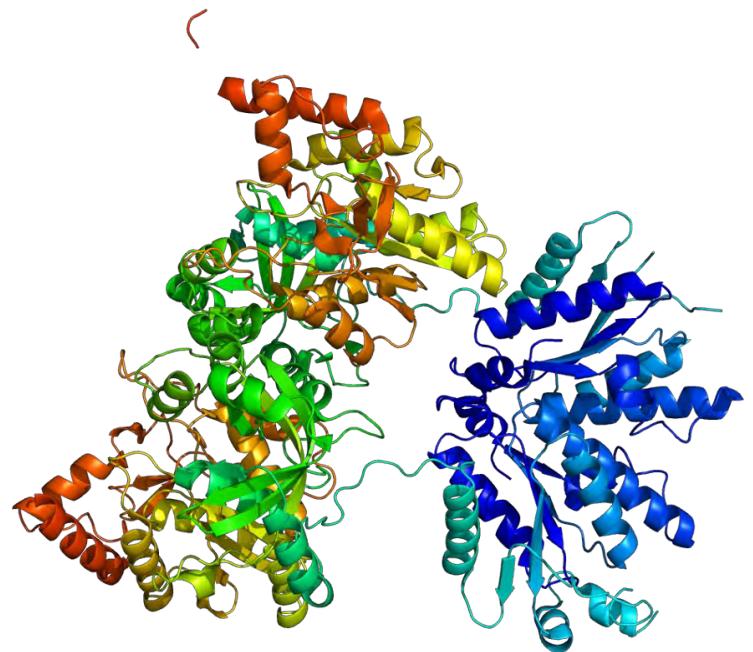
Криоэлектронное изображение белков GroEL,  
сuspendedированных в аморфном льду при  
увеличении в 50 000 раз

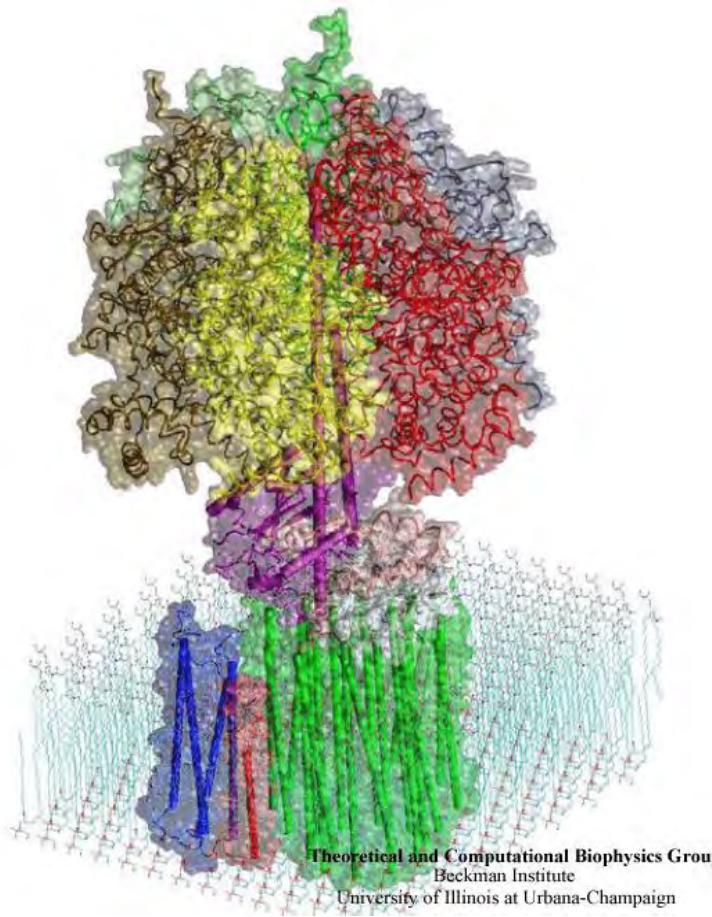


▲  
Первая пространственная структура белковой молекулы. Некрасивая структура миоглобина низкого разрешения из статьи в *Nature* 1958 г.

◀ Джон Кендрю и Макс Перутц демонстрируют пространственную структуру миоглобина, собранную из специального конструктора.

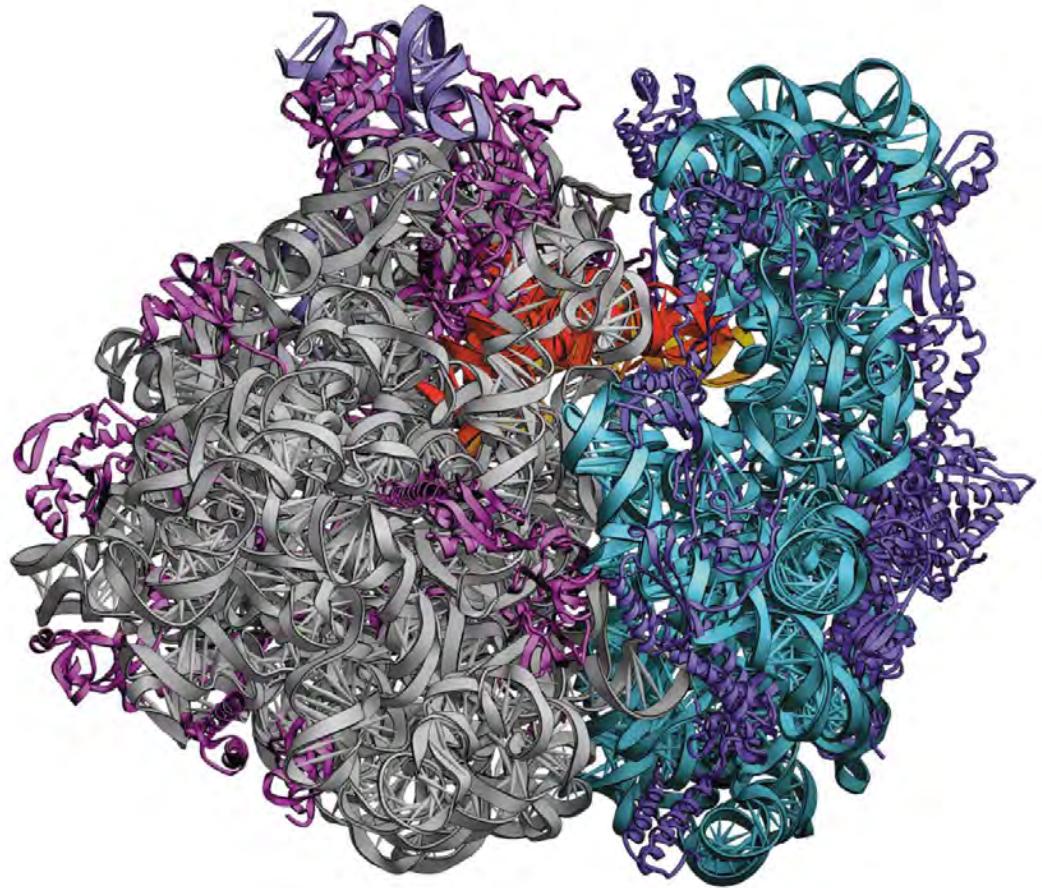
# 3D-моделирование белков





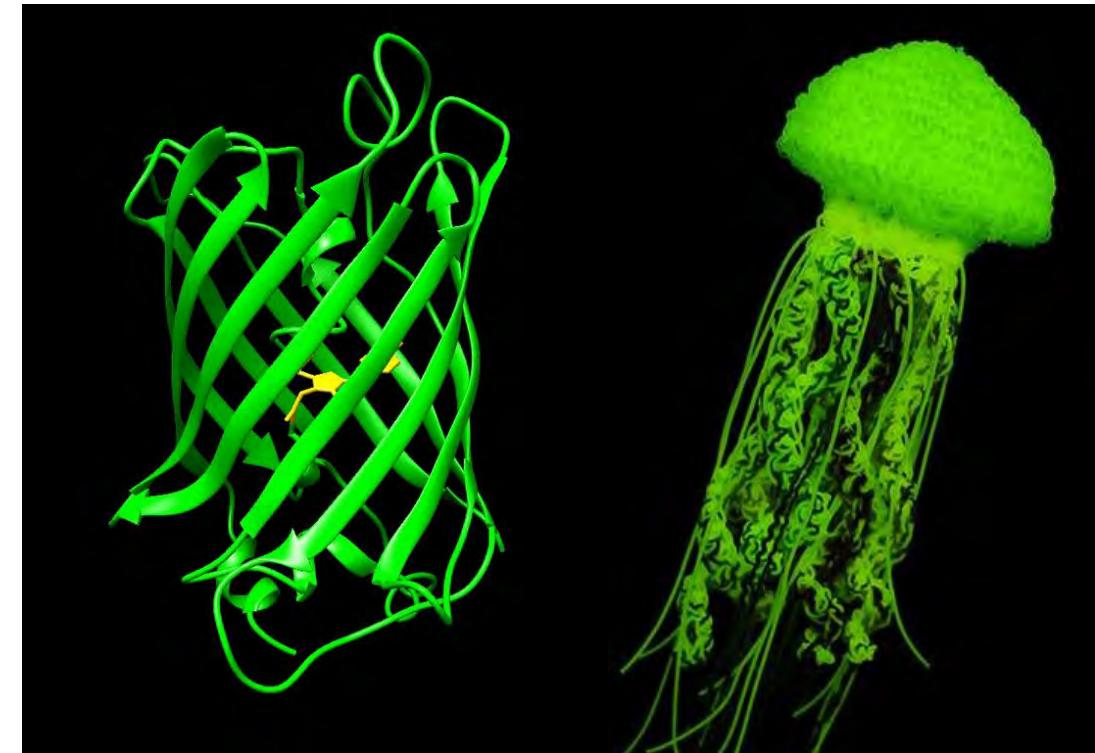
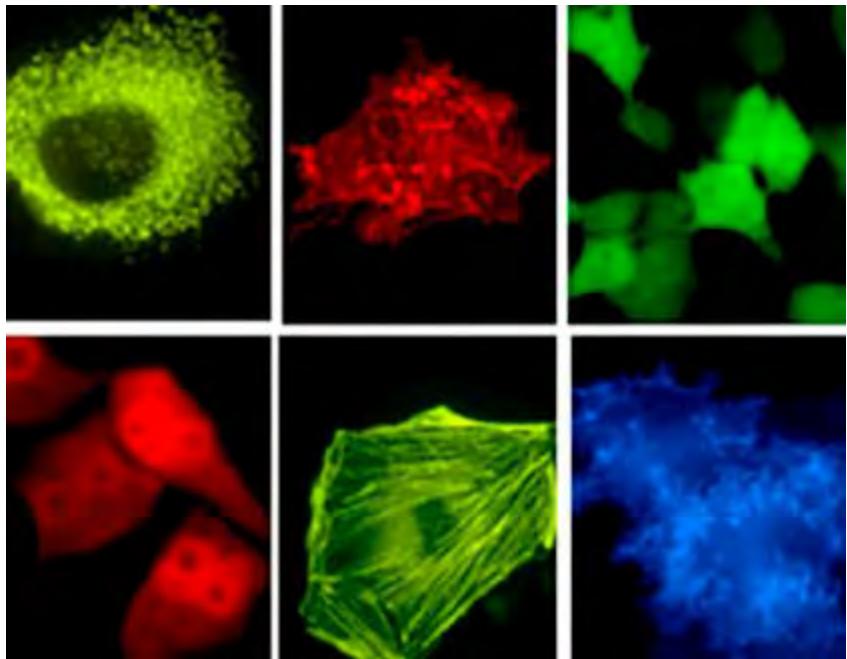
Theoretical and Computational Biophysics Group  
Beckman Institute  
University of Illinois at Urbana-Champaign

**Пространственная структура АТФ-синтазы, полученная в 1994 году.  
Работа удостоилась нобелевской премии.**



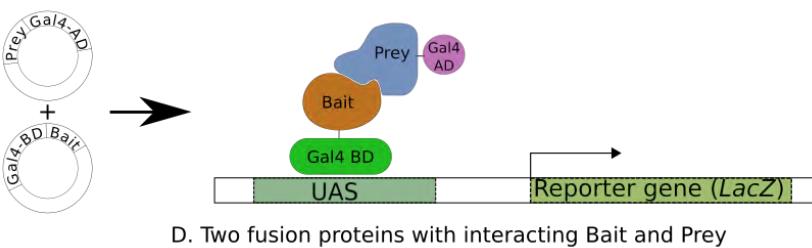
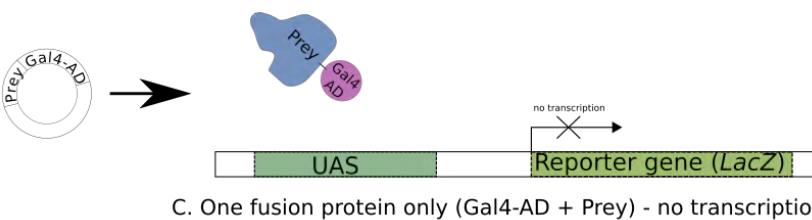
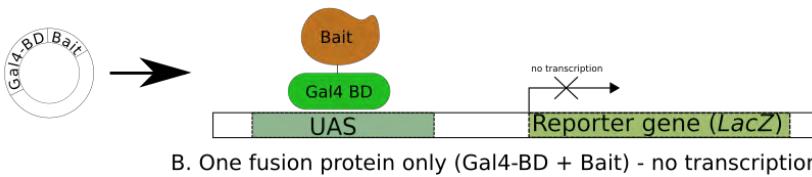
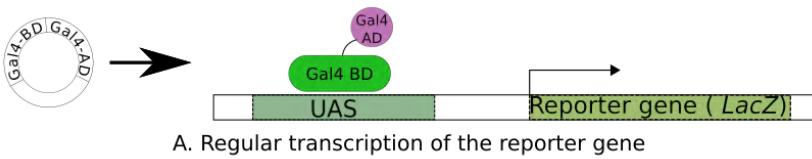
**Пространственная структура рибосомы, полученная в 2000 году.  
Работа удостоилась нобелевской премии.**

# Флуоресцентная микроскопия



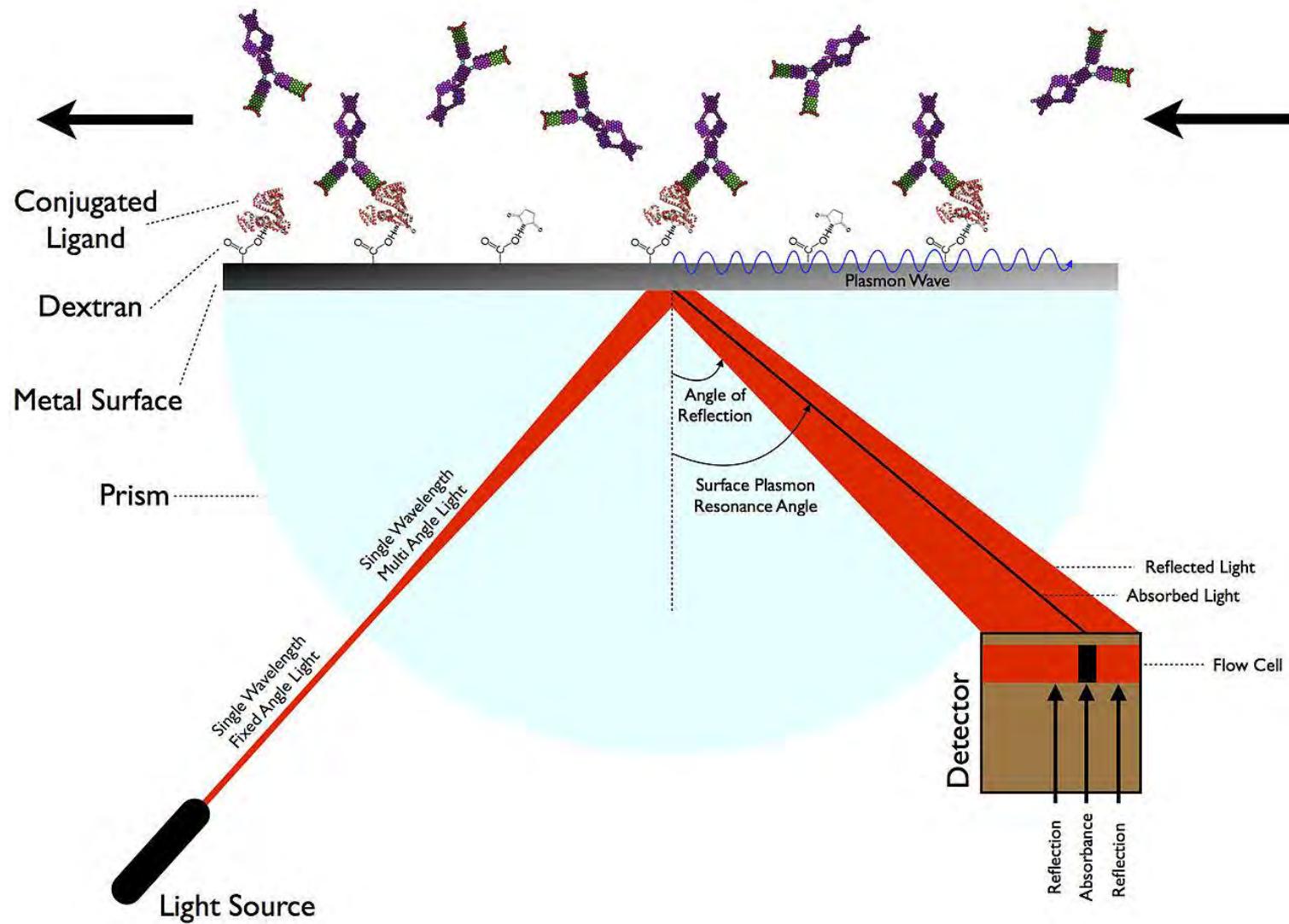
Визуализация различных компонентов  
живых клеток с помощью  
флуоресцентных белков

# Исследование белок-белковых взаимодействий (Yeast Two-Hybrid)

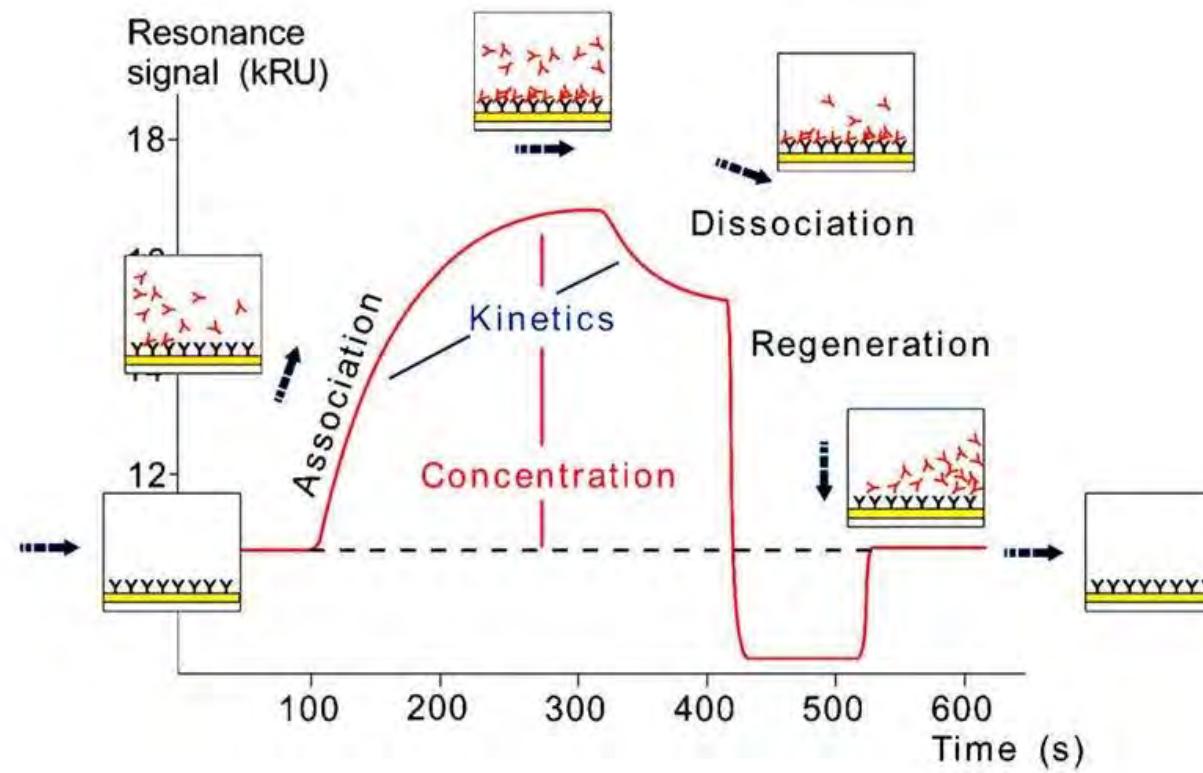


УФ-сшивание с иммунопреципитацией для  
идентификации сайтов связывания РНК белков

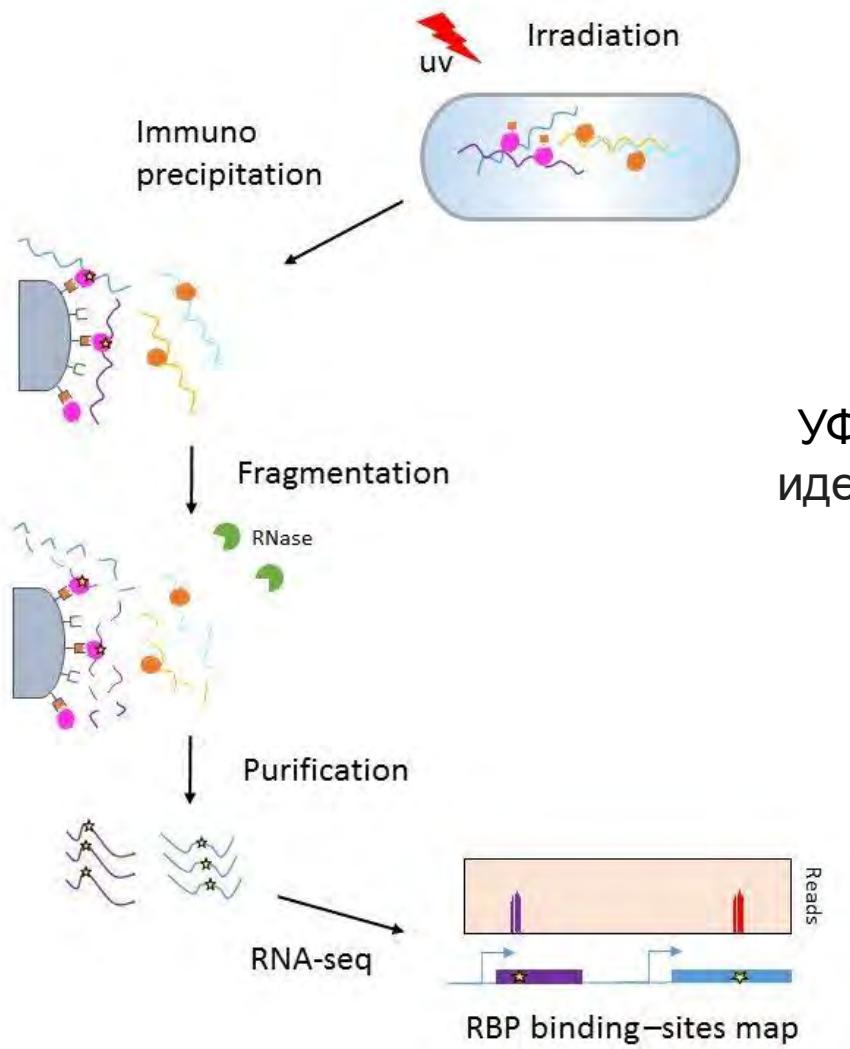
# Поверхностный плазмонный резонанс (SPR)



# Поверхностный плазмонный резонанс (SPR)



## Cross-linking and immunoprecipitation (CLIP, or CLIP-seq)



УФ - «сшивание» с иммунопреципитацией для идентификации сайтов связывания РНК белков

# Cross-linking and immunoprecipitation (CLIP, or CLIP-seq)

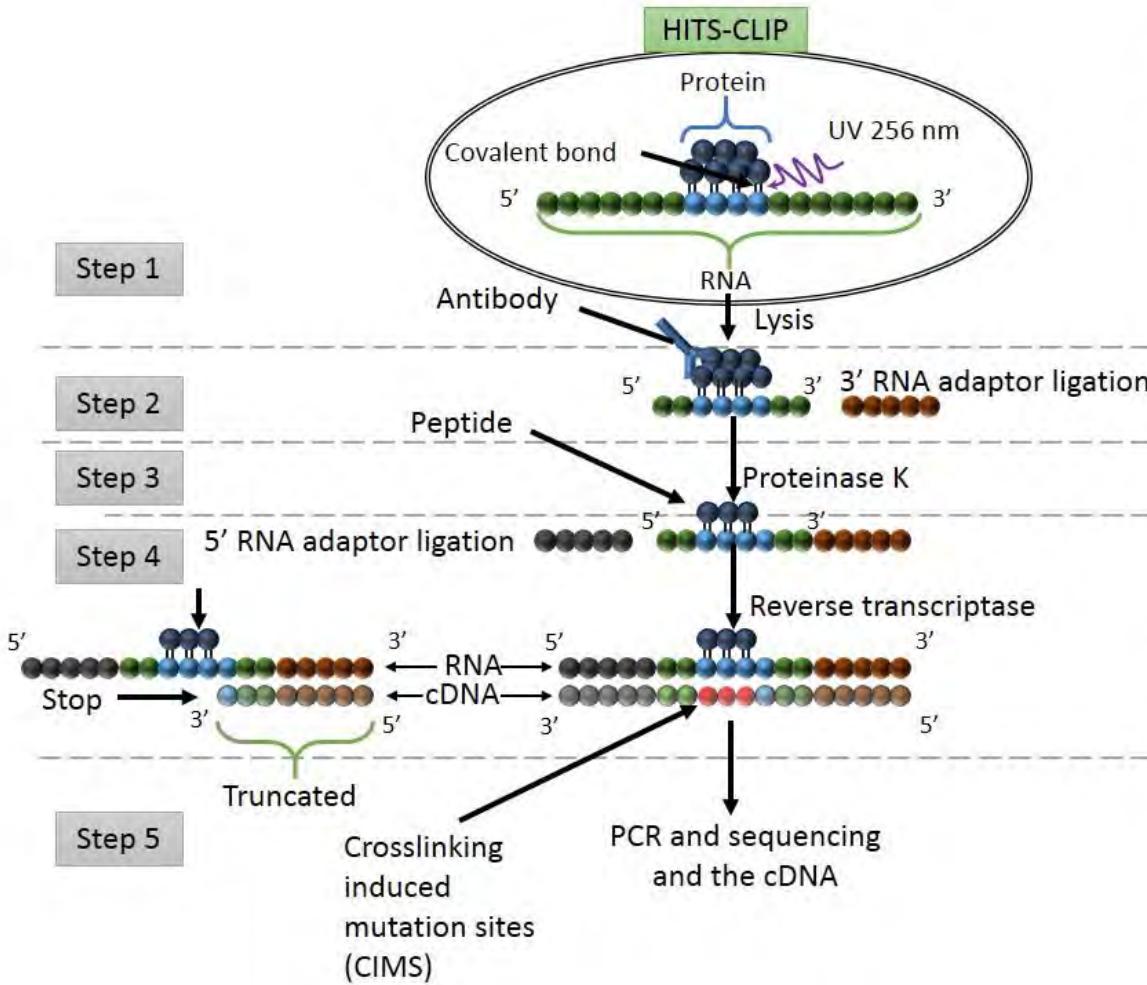


Figure 2: HITS-CLIP

**Step 1**

HITS-CLIP begins with the in-vivo cross-linking of RNA-protein complexes using ultraviolet light. The cell is lysed and the protein of interest is isolated using immunoprecipitation.

**Step 2**

Washing is performed to remove free RNA, and RNA adaptors are ligated at the 3' ends.

**Step 3**

Proteinase K digestion is performed. This leaves a peptide at the cross-link site that modifies the chemical structure of the nucleotide.

**Step 4**

5' RNA adaptors are ligated and cDNA is synthesized using reverse transcription.

**Step 5**

PCR and sequencing of the cDNA.



## Find your protein

UniProtKB • Examples: Insulin, APP, Human, P05067, organism\_id:9606

Advanced | List | Search

UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information. Cite UniProt \*\*

**Proteins**  
UniProt Knowledgebase  
Reviewed (Swiss-Prot) 572,214  
Unreviewed (TrEMBL) 248,266,673

**Species**  
Proteomes  
Protein sets for species with sequenced genomes from across the tree of life

**Protein Clusters**  
UniRef  
Clusters of protein sequences at 100%, 90% & 50% identity

**Sequence archive**  
UniParc  
Non-redundant archive of publicly available protein sequences seen across different databases



VALIDATION ▾ DEPOSITION ▾ DICTIONARIES ▾ DOCUMENTATION ▾ TASK FORCES ▾ DOWNLOADS ▾ STATISTICS ▾ ABOUT ▾ 

**Since 1971, the Protein Data Bank archive (PDB) has served as the single repository of information about the 3D structures of proteins, nucleic acids, and complex assemblies.**

**Celebrating 50 Years of the PDB**  
The Worldwide PDB (wwPDB) organization manages the PDB archive and ensures that the PDB is freely and publicly available to the global community.

**Celebrating 20 Years of the wwPDB Partnership**

**Validate Structure**  
or View validation reports

**Deposit Structure**  
All deposition Resources

**Download Archive**  
Instructions

**Vision and Mission**

**Vision**  
Sustain freely accessible, interoperating Core Archives of structure data and metadata for biological macromolecules as an enduring public good to promote basic and applied research and education across the sciences.

**Mission**

- Manage the wwPDB Core Archives as a public good according to the FAIR Principles.
- Provide expert deposition, validation, biocuration, and remediation services at no charge to Data Depositors worldwide.
- Ensure universal open access to public domain structural biology data with no limitations on usage.
- Develop and promote community-endorsed data standards for archiving and exchange of global structural biology data.

**wwPDB Resources**

**Data Dictionaries**

- Macromolecular Dictionary (PDB/mmCIF)
- Small Molecule Dictionary (CCD)
- Peptide-like antibiotic and inhibitor molecules (BIRD)

**Biocuration**

- Procedures and policies
- Improvements for consistency and accuracy

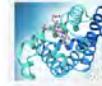
**Community Input: Task Forces and Working Groups**

- Validation Task Forces (X-ray, NMR, 3DEM)
- Small Angle Scattering Task Force
- PDB/mmCIF Working Group
- Hybrid/Integrative Methods Task Force
- Ligand Validation Workshop
- ModelCIF Working Group

**PDB Data Growth & Usage Statistics**

- Depositions: by data center, by year, and by depositor location
- Downloads: by year for all entries

10/22/2024 **Poster Prize Awarded at LACA**  
  
The wwPDB Foundation awarded Zalghum Abbas at the Latin American Crystallographic Association Meeting for *The Role of SEP77 in Magnaporthe oryzae: Structural Insights and Functional Implications*

10/16/2024 **The Nobel Prize in Chemistry 2024**  
  
Congratulations to David Baker, who was recognized for computational protein design, and to Demis Hassabis and John M. Jumper, who were recognized for protein

## Status

Reviewed (Swiss-Prot) (9)

 Unreviewed (TrEMBL)  
(20)

## Popular organisms

A. thaliana (10)

## Taxonomy

Filter by taxonomy

## Group by

Taxonomy

Keywords

Gene Ontology

Enzyme Class

## Proteins with

3D structure (3)

Activity regulation (3)

Alternative products  
(isoforms) (3)

Alternative splicing (3)

Beta strand (3)

## UniProtKB 29 results or search "ahk3" as a Gene Name or Strain

 Tools ▾ Download (29) Add View: Cards  Table Customize columns Share ▾

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
Q9C5U1	AHK3_ARATH	Histidine kinase 3 [...]	AHK3, ORE12, At1g27320, F17L21.11	Arabidopsis thaliana (Mouse-ear cress)	1,036 AA
Q9SAZ5	AHP3_ARATH	Histidine-containing phosphotransfer protein 3	AHP3, ATHP2, At5g39340, MUL8.2	Arabidopsis thaliana (Mouse-ear cress)	155 AA
Q8L9T7	AHP5_ARATH	Histidine-containing phosphotransfer protein 5	AHP5, At1g03430, F21B7.5	Arabidopsis thaliana (Mouse-ear cress)	157 AA
Q9C5U0	AHK4_ARATH	Histidine kinase 4 [...]	AHK4, CRE1, RAW1, WOL, At2g01830, T23K3.2	Arabidopsis thaliana (Mouse-ear cress)	1,080 AA
Q8RY18	TRF1A_ARATH	TNF receptor-associated factor homolog 1a [...]	TRAF1A, MUSE14, At5g43560, K9D7.6	Arabidopsis thaliana (Mouse-ear cress)	1,055 AA
Q9ZNV9	AHP1_ARATH	Histidine-containing phosphotransfer protein 1	AHP1, ATHP3, At3g21510, MIL23.8	Arabidopsis thaliana (Mouse-ear cress)	154 AA
Q9ZNV8	AHP2_ARATH	Histidine-containing phosphotransfer protein 2	AHP2, ATHP1, At3g29350, MUO10.6	Arabidopsis thaliana (Mouse-ear cress)	156 AA
Q9C5U2	AHK2_ARATH	Histidine kinase 2 [...]	AHK2, At5g35750, MXH1.16	Arabidopsis thaliana (Mouse-ear cress)	1,176 AA
Q9ZWJ9	ARR2_ARATH	Two-component response regulator ARR2	ARR2, ARP5, At4g16110, dl4095w, FCAALL.297	Arabidopsis thaliana (Mouse-ear cress)	664 AA
A0A2H9ZW06	A0A2H9ZW06_9ASPA	histidine kinase [...]	AHK3, AXF42_Ash020182	Apostasia shenzhenica	998 AA
A0A9Q0FQK4	A0A9Q0FQK4_9ROSI	histidine kinase [...]	AHK3, Tsubulata_026526	Turnera subulata	1,038 AA

Feedback

Help