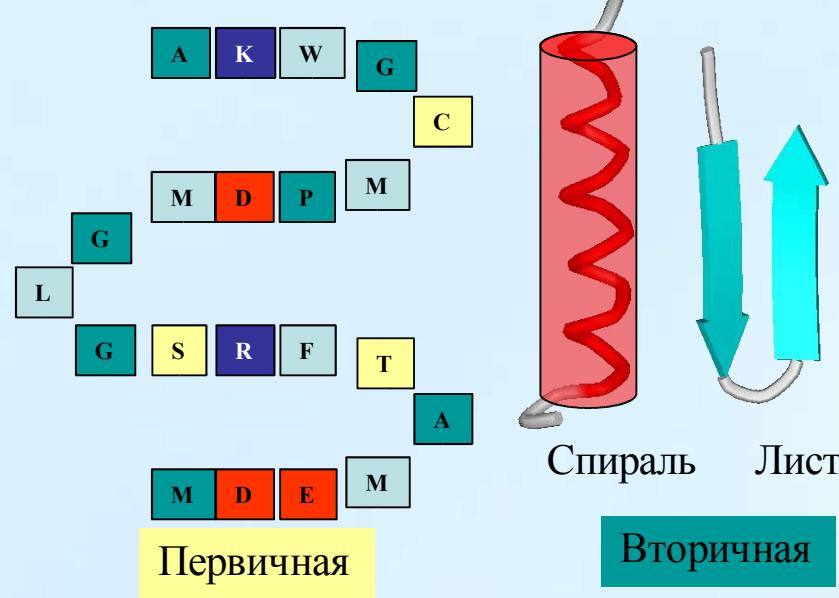


# **Молекулярный дизайн, компьютерное моделирование и анализ структуры и эволюции функциональных сайтов глобулярных белков**

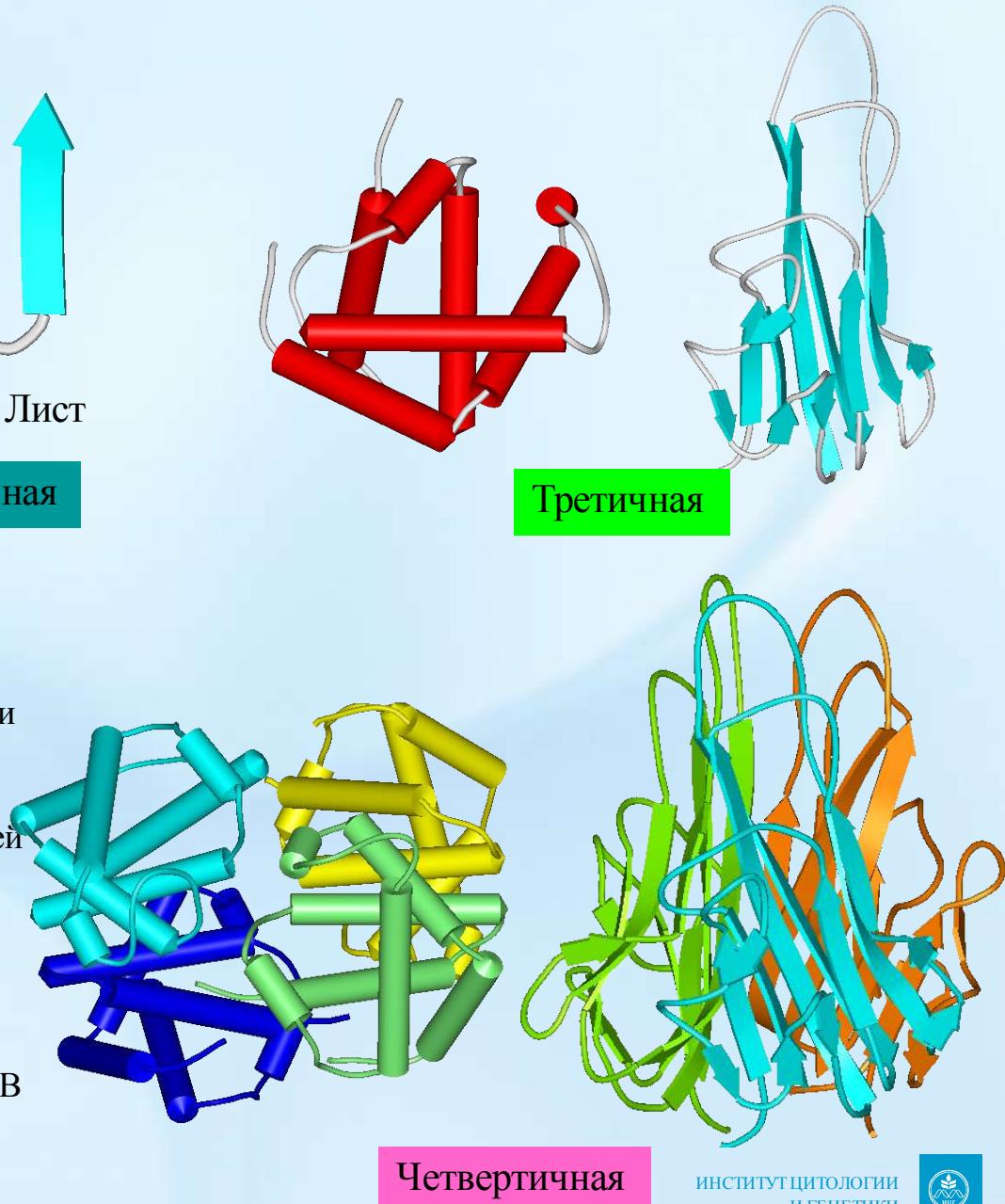
**Институт Цитологии и Генетики СО РАН  
Лаборатория Теоретической Генетики  
с.н.с., к.б.н. В.А. Иванисенко**



# Различные уровни организации белковой структуры



Первичная структура белковой цепи представляет собой аминокислотную последовательность. Вторичная структура состоит из регулярной организации спиралей и листов. Третичная структура является трехмерной нативной конформацией полипептидной цепи, как правило, включающей компактную упаковку элементов вторичной структуры. Четвертичная структура представляет собой комплекс нескольких полипептидных цепей (субъединиц), формирующих функциональную биомолекулярную структуру. (Гемоглобин PDB ID 1HBR; ФНО-альфа человека PDB ID 5tsw).

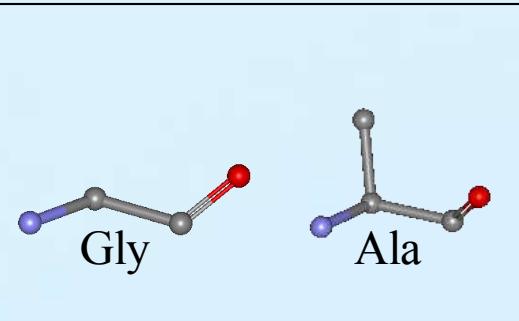


Четвертичная

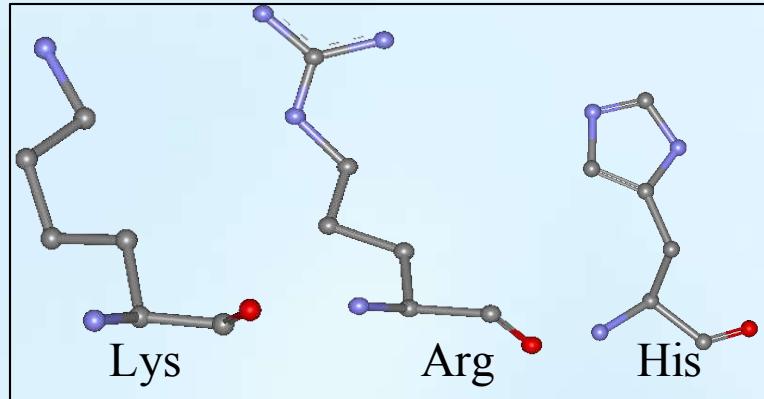
ИНСТИТУТ ЦИТОЛОГИИ  
И ГЕНЕТИКИ



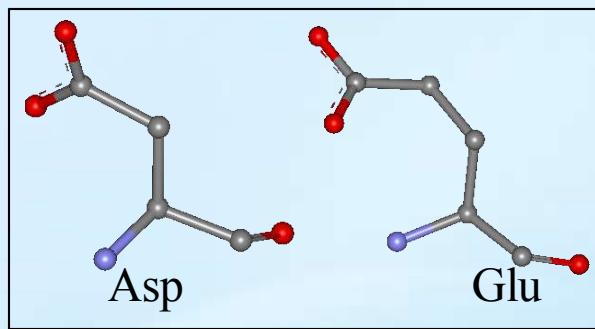
## Малые



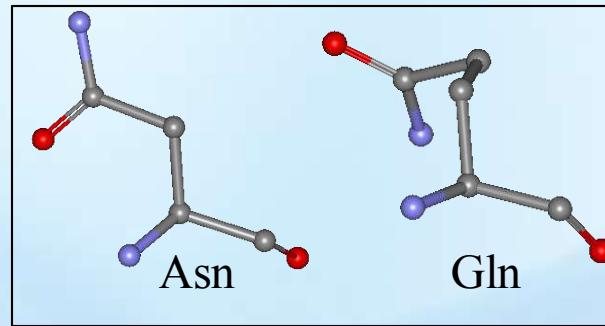
## Положительно заряженные



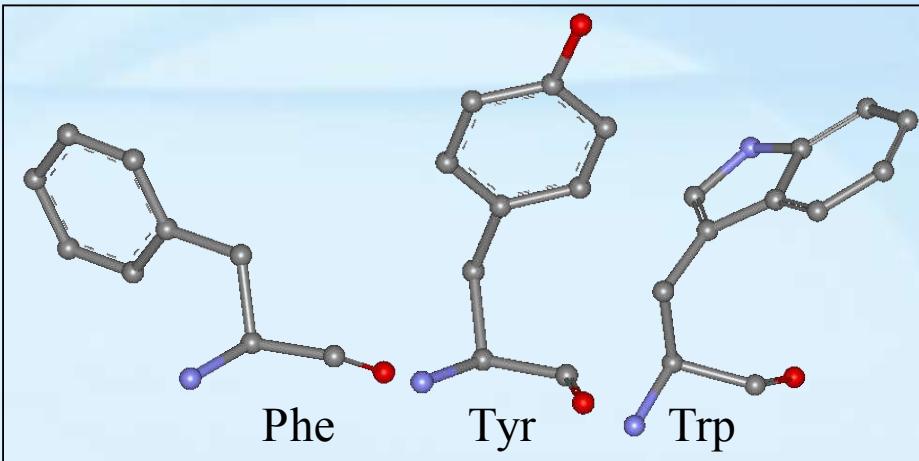
## Отрицательно заряженные



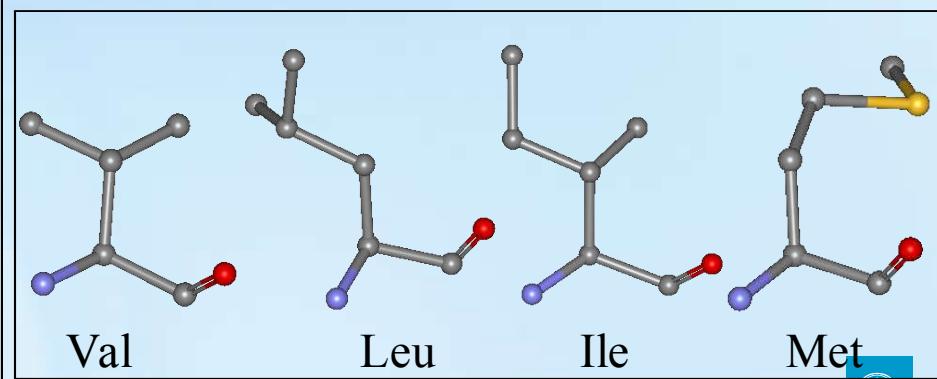
## Полярные



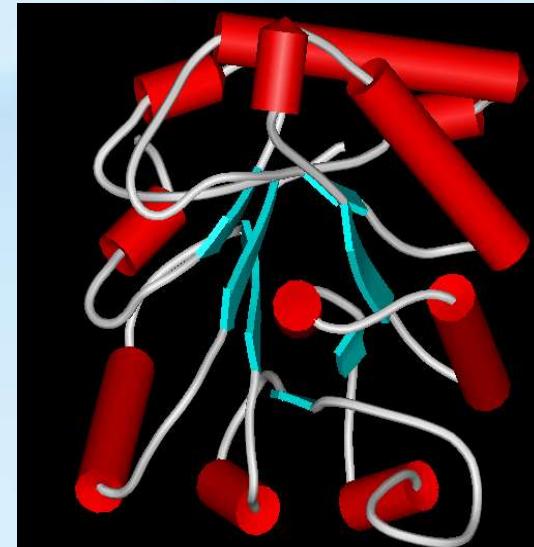
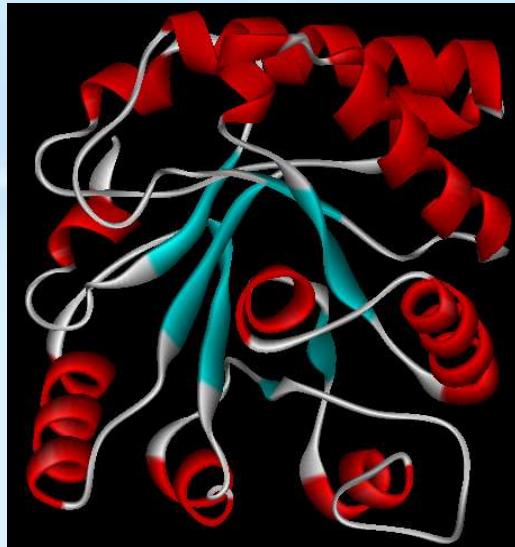
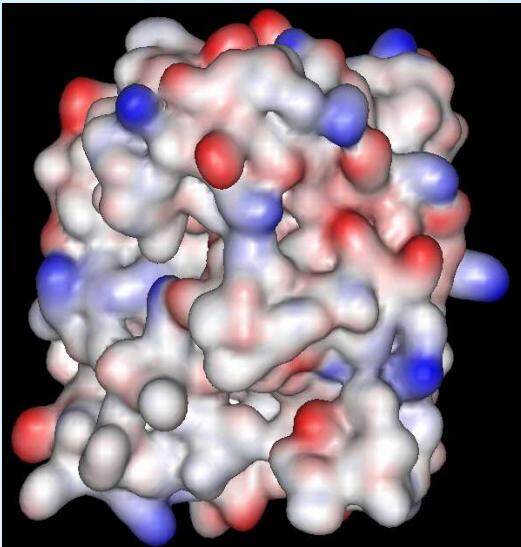
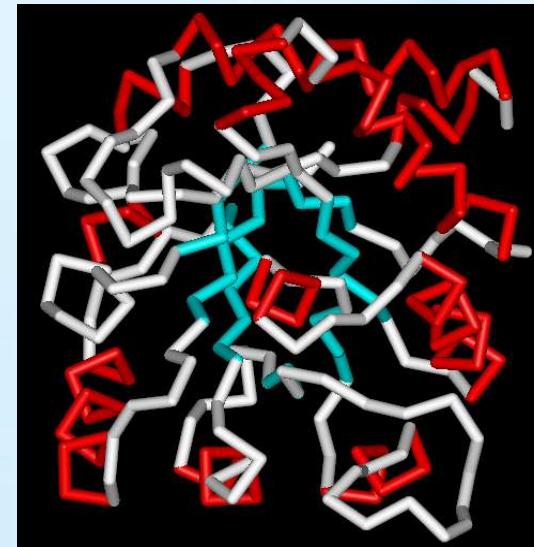
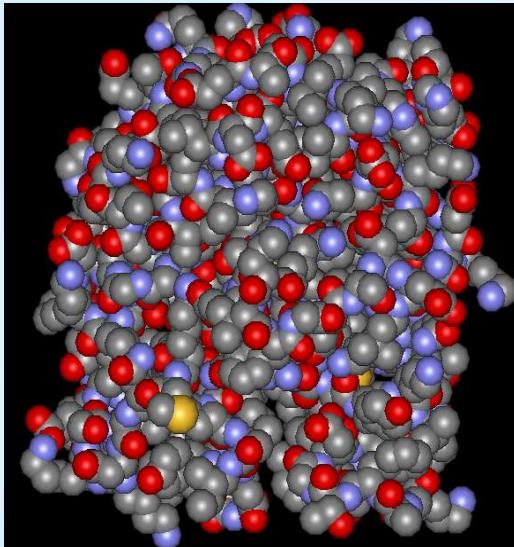
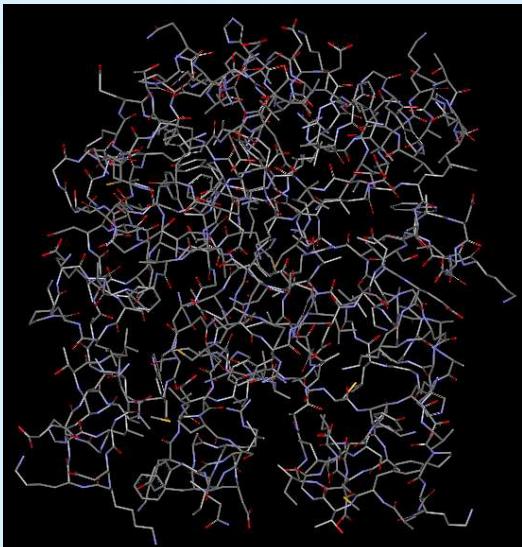
## Ароматические гидрофобные



## Алифатические гидрофобные



Визуализация пространственных структур белков позволяет решать многие задачи по их структурно-функциональной организации.  
Наиболее распространенными являются следующие модели визуализации.



**В настоящее время секвенированы геномы более 230  
организмов**

**В ближайшие годы ожидается секвенирование еще примерно  
1000 геномов различных организмов**

**В базе данных SWISS-PROT содержится более 150 000  
функционально аннотированных белковых  
последовательностей**

**В базе данных TREMBL содержится более 1 560 000 белковых  
последовательностей, требующих функциональную аннотацию**



# **Задачи компьютерной протеомики и функциональной геномики**

## **Функциональная аннотация белков**

**Биохимическая функция**

**Клеточная функция**

## **Молекулярный дизайн**

**Модифицированные белки с улучшенными медико-биологическими свойствами**

**Модифицированные белки с приобретением новой функции**

**Искусственные белки**

## **Поиск фармакологических мишней**

**Блокирование функции определенных генов и белков**

**Возобновление поврежденной функции**

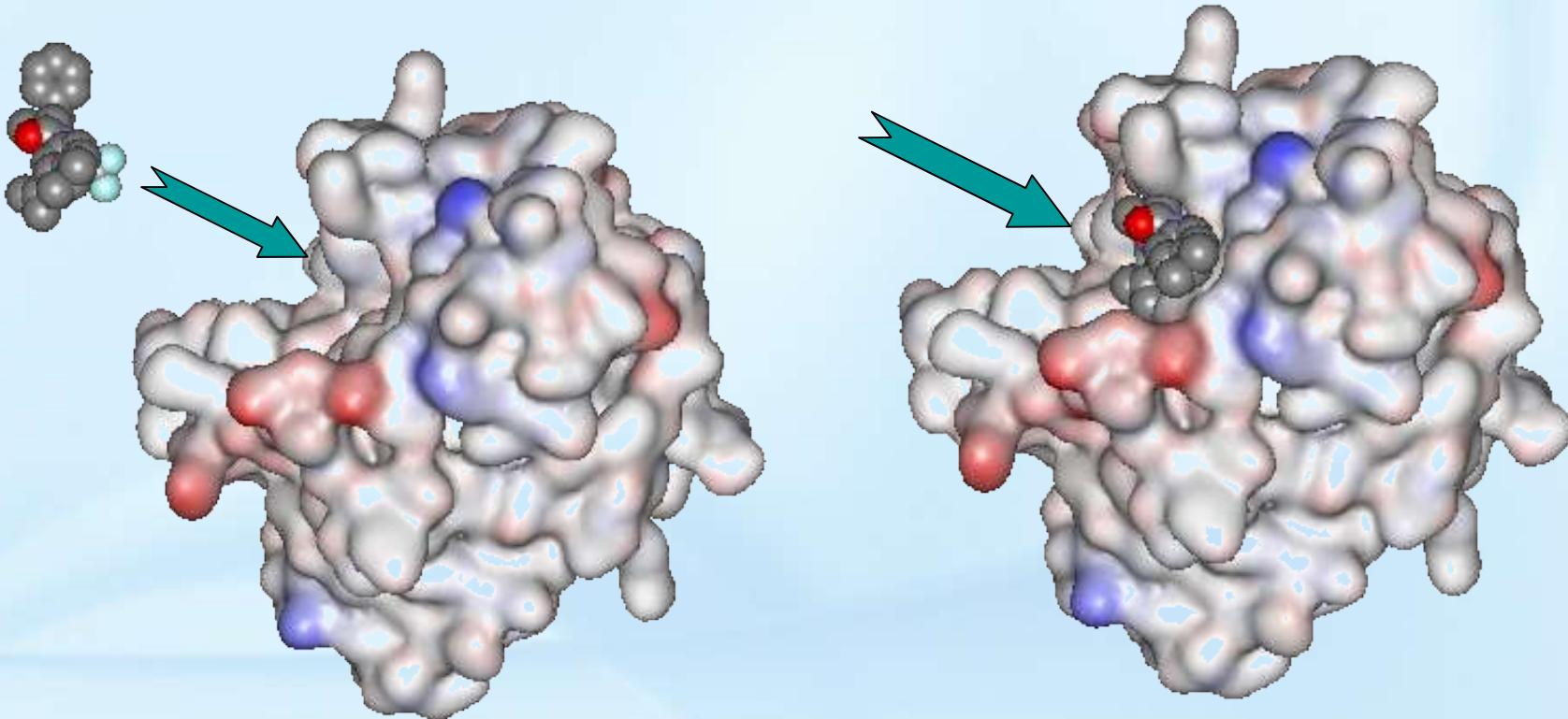
## **Молекулярная эволюция**

**Предковые последовательности**

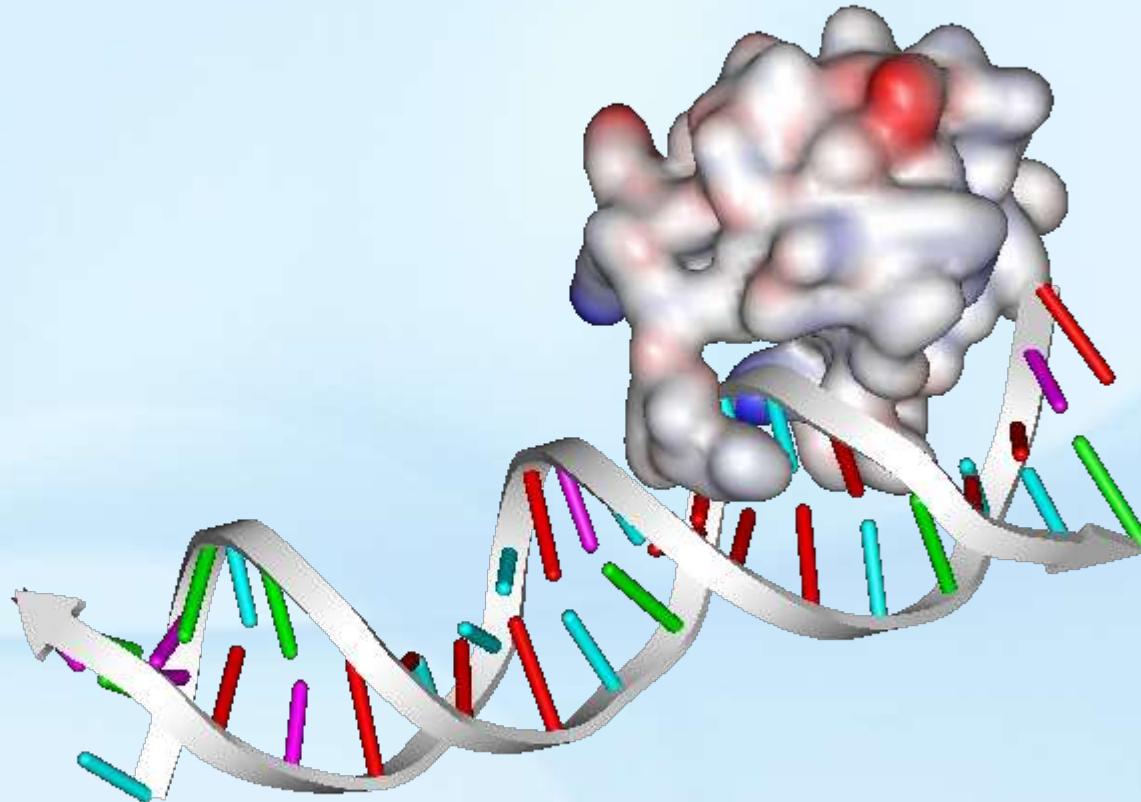
**Молекулярные механизмы эволюционного процесса**



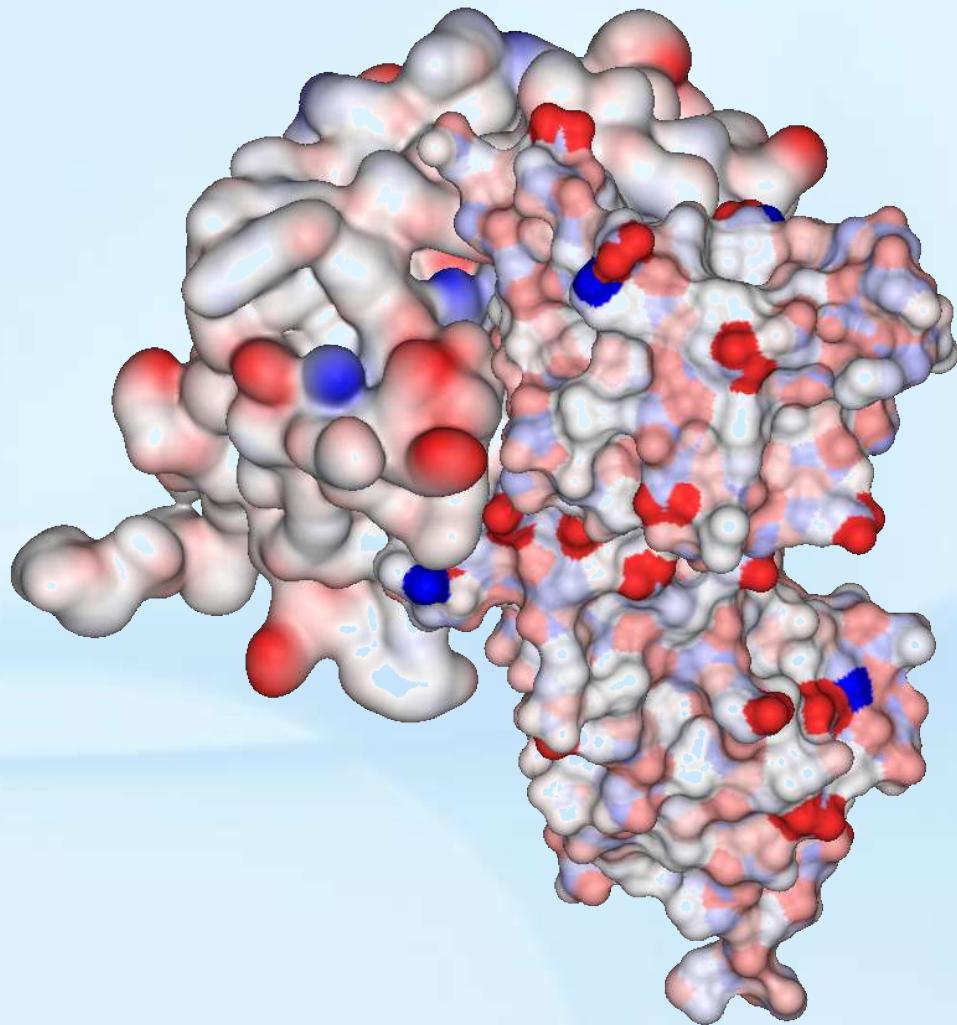
## Докинг белков с малыми молекулами



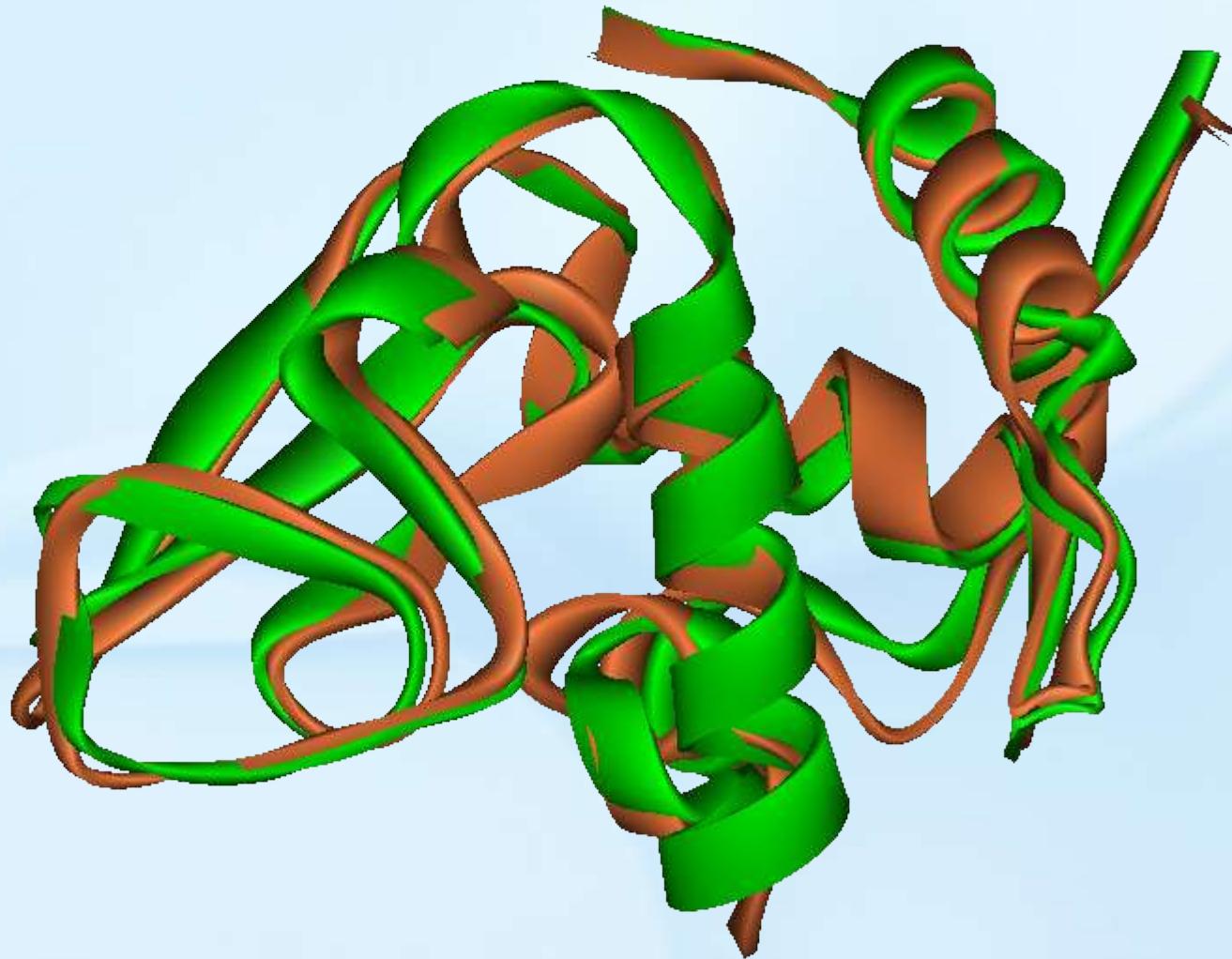
# Докинг белков с ДНК



# Белок-белковые взаимодействия



# Структурное выравнивание



# Множественное выравнивание последовательностей

POLG_TBEVS	GLFGKGGSIVTCVKAASCCEAKKKKATGHVYDANKIVVYT
O10383/281-581	GLFGKGGSIVACVKAACEAKKKKATGYVYDANKIVVYT
Q82951/281-581	GLFGKGGRIVACAKFSCETKKKATGAVVYDWNKITVYR
POLG_POWVL/279-579	GFFFGKGGSIVACAKFECEFAKKAVGHVYDSTKIKTYW
POLG_YEFV1/286-579	GLFGKGGSIVACAKFTCAK-----SMSLFEVVDQTKIQYV
POLG_MVEW/293-591	GLFGKGGSIDTCAKFTCSN-----SAAGRLILPEDIKYE
O09754/1-298	GLFGKGGSIDTCAKFECTS-----KALIGRTIQLOPENIKYE
POLG_KUNJM/291-589	GLFGKGGSIDTCAKFRACST-----KATGRTIILKENIKYE
POLG_STEVM/289-587	GLFGKGGSIDTCAKFTCKN-----KATGKTILRERENIKYE
O10246/1-296	GLFGKGGSLITCAKFKCWT-----KLEGKIVQYENLKYI
POLG_DEN3/281-574	GLFGKGSSLVTCAKFQCL-----SIEGKVWQHENNLKYUT
O09234/281-576	GLFGKGGGIVTCAMFRCKK-----NMEGKVWQPENLEXYT
POLG_DEN4/280-575	GLFGKGGGVVTCAKFSCSG-----KITGNLVRIENLEYUT

паттерны

Регулярные выражения, описывающие последовательность аминокислот в сайте.

Например, для N-glycosylation site паттерн имеет вид  $N\text{-}\{P\}\text{-}[ST]\text{-}\{P\}$ .

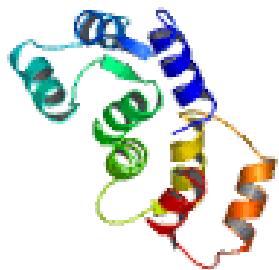
весовые матрицы

**В случае правильно выравненных аминокислотных последовательностей совмещаются вторичные и третичные структуры белков.**

	10	20	30	40	50
RSF	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓
1sbp	k d i q LLNVS YDP TREL YEQYNKAF S AHWK q e t g d n VV IDQSHG g s g KQATSVIN - - - g i	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑
RSF	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓
1wod	- - g k I TVFAA - - - AS L TNAMQD I ATQFKKEKGVDVVSSFA - - s s s TLARQIEA - - - g a	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑
RSF	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓
1mrp	- - - D I TVYNG - - - q HKEAATAVAKAFEQE TG I KVTLNSG - - KSEQLAGQLKEE g d k T	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑
RSF	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓
1pot	- - NTLYFYNW - - - TEYVPPG - - LLEQFTKE TG I KV I YSTY - E s N E T M Y A K L K T - Y K d g	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑
RSF	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓
1atg	- - - ELKVVT A - - - TNFLG TLEQLAGQF AKQTGH AVV I S SG - - S S GPV Y A Q I VN - - - GA	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑	↓↓↓↓↓↓↓↓	↑↑↑↑↑↑↑↑



# Молекулярная динамика



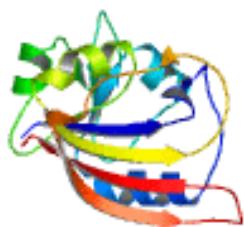
**Calmodulin**



**HIV-RT**



**SERCA pump**



**DHFR**

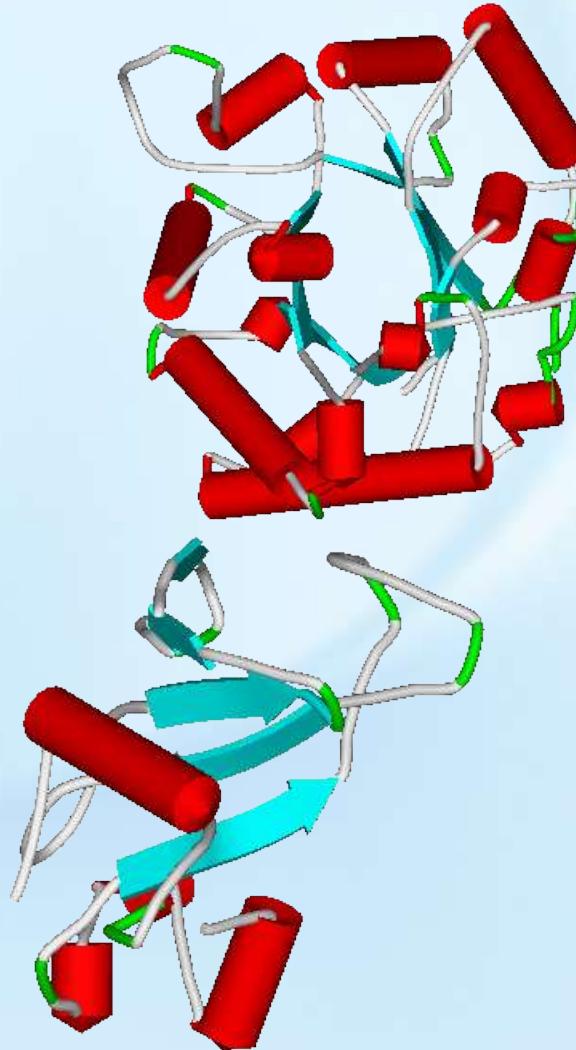
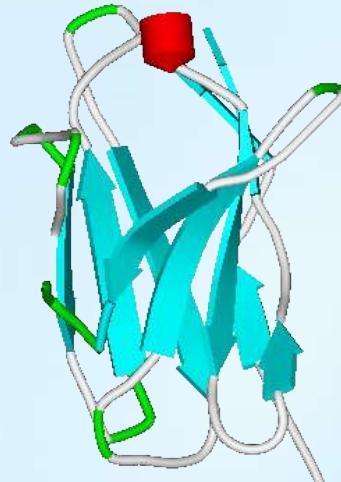
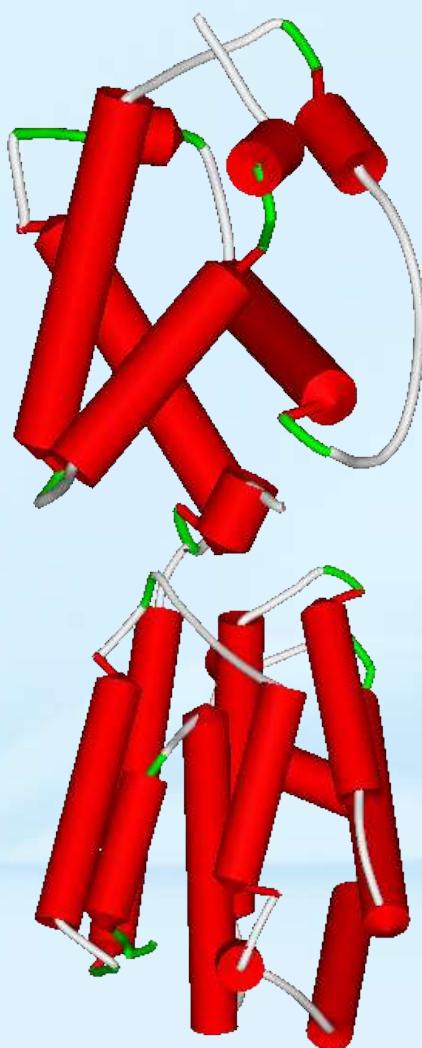


**HIV-protease**



**SOD**

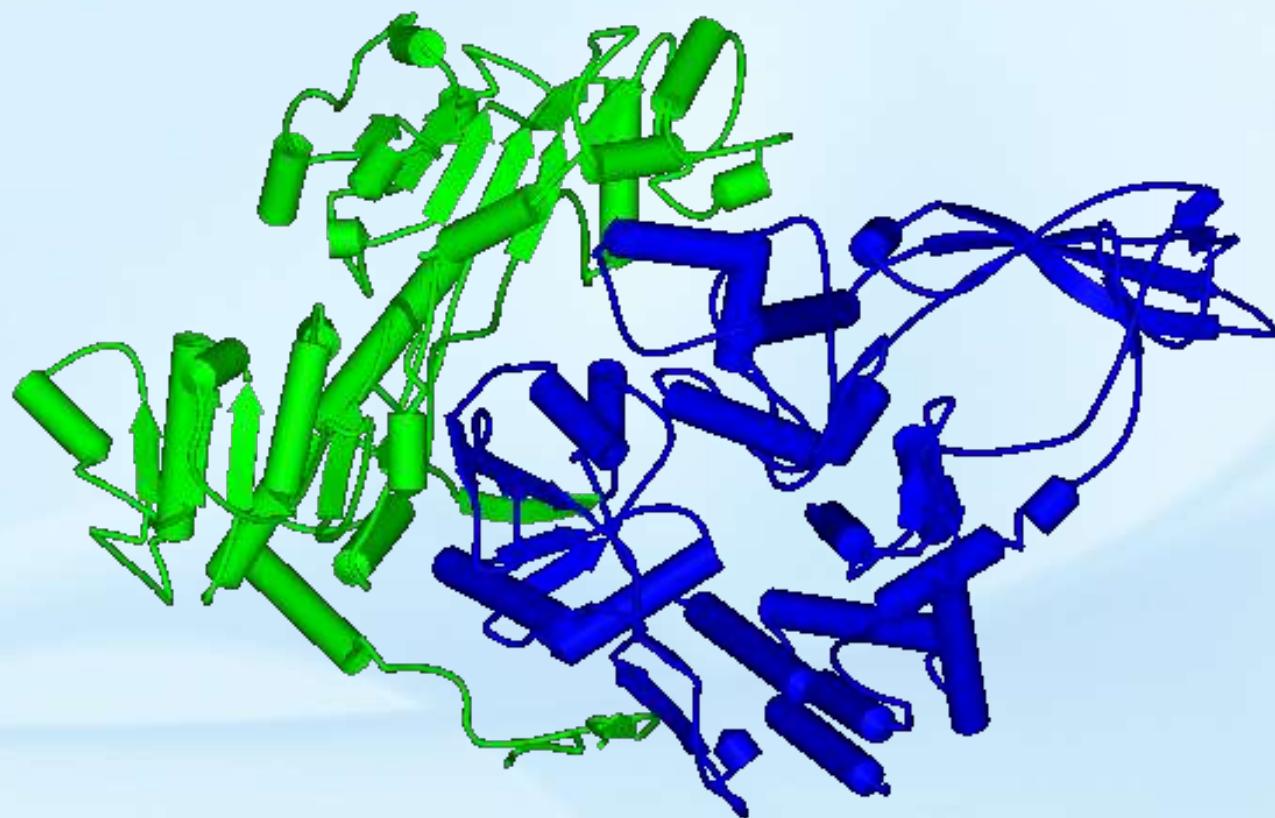
## Предсказание третичной структуры белков



В базе данных структурной классификации 3D структур белковых доменов SCOP приведено 800 типов укладки (фолдов) из 54745 белковых доменов.  
Разнообразие такое, что на 54745 доменов, представленных в 20619 белков из PDB приходится 800 типов укладок, указанных в SCOP.



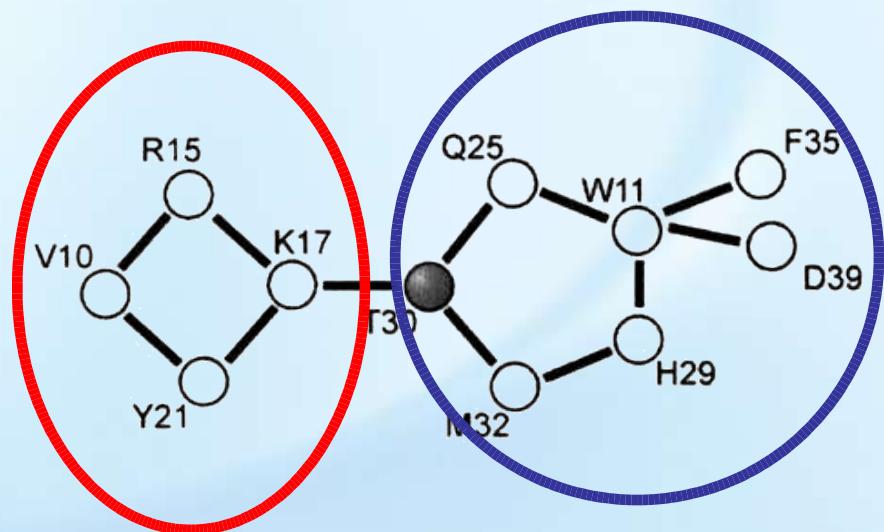
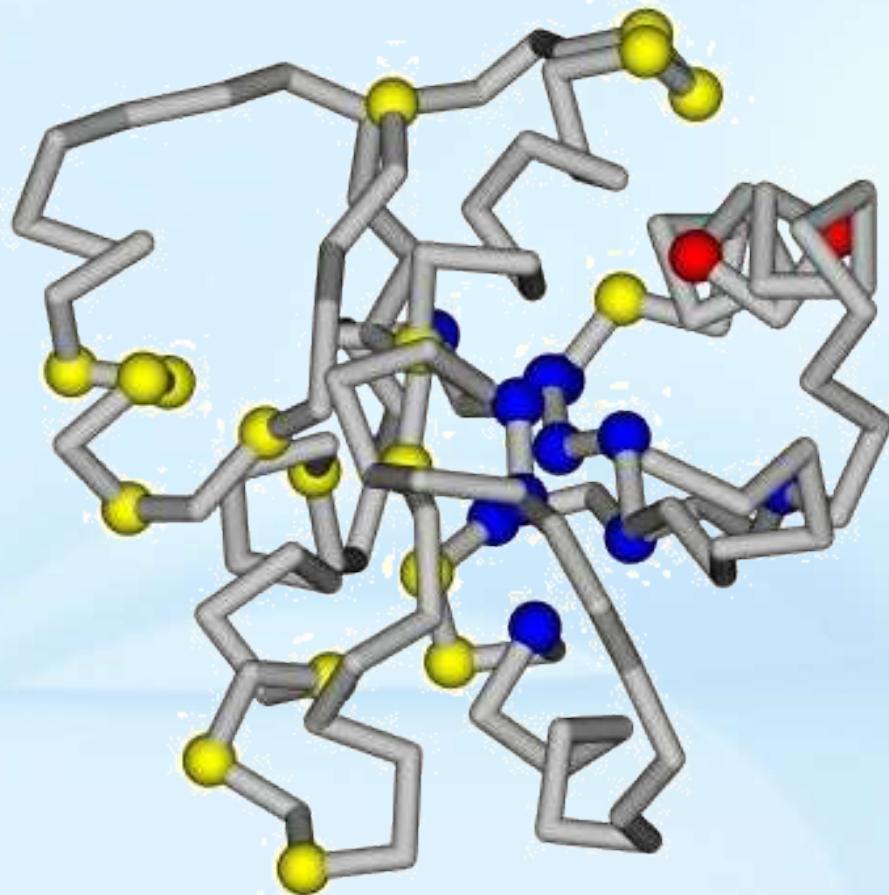
## Определение доменной структуры белков



Обратная гираза (reverse gyrase) (PDB Id 1GL9). Зеленым выделен домен, выполняющий функцию геликазы, синим – топоизомеразы. Организм: ARCHAEOGLOBUS FULgidus



# Поиск кластеров аминокислотных остатков, обладающих определенными физико-химическими или эволюционными свойствами



# **Мишени и механизмы действия лекарств**

## **РЕПЛИКАЦИЯ/ТРАНСКРИПЦИЯ ДНК**

ДНК – алкилирующие агенты, minor groove binders, intercalating agents, wrong substrates (trojan horses)

## **ФУНКЦИИ РНК**

РНК – интерференция (подавление трансляции и ускорение деградации РНК)

## **СТРУКТУРА И ФУНКЦИЯ БЕЛКОВ**

Процесс фолдинга после трансляции - ингибиторы белкового фолдинга

Процесс создания биологически активных модулей (units) для полисубъединичных белков - ингибиторы белкового фолдинга  
Активные центры ферментов - обратимые и необратимые ингибиторы  
Транспортеры - uptake inhibitors

## **УРОВЕНЬ ГЕННОЙ СЕТИ И ПУТЕЙ ПЕРЕДАЧИ СИГНАЛА**

Рецепторы – агонисты и антагонисты

## **СУПРАМОЛЕКУЛЯРНЫЕ СТРУКТУРЫ**

Ионные каналы – блокаторы и активаторы



# ГЕНОМНЫЕ ПРОЕКТЫ: ЧЕЛОВЕК

Характеристики генома человека:

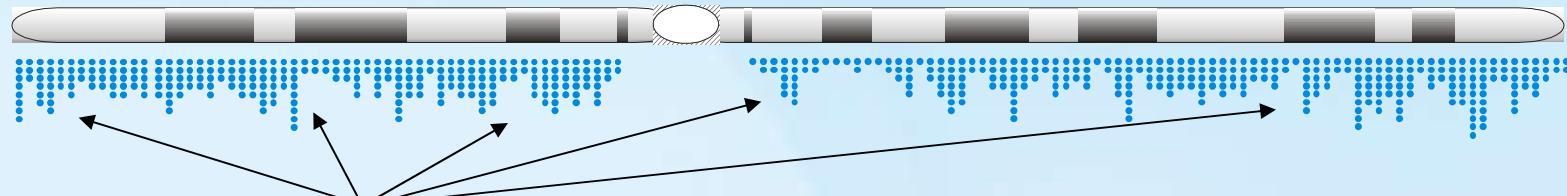
размер ~ 3.3 миллиарда п.о. ATGCCCGGGTTTAATGCGTCAGTGACTGCACA...

количество генов > 35000

цитогенетическая карта

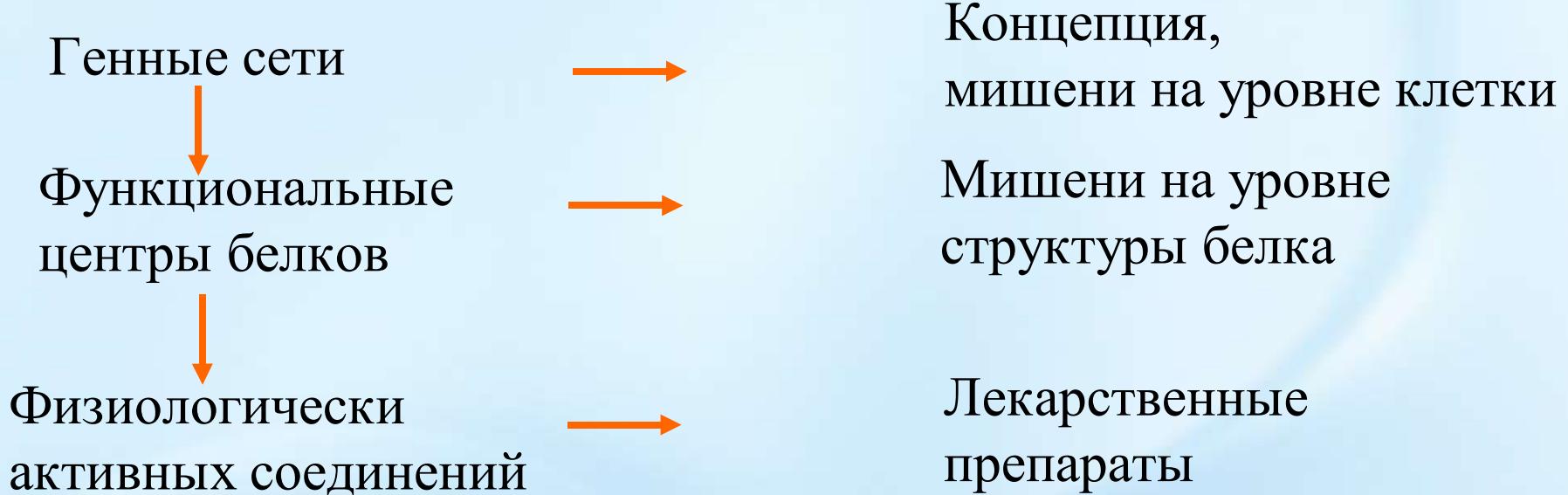
физическая карта:

локализация генов

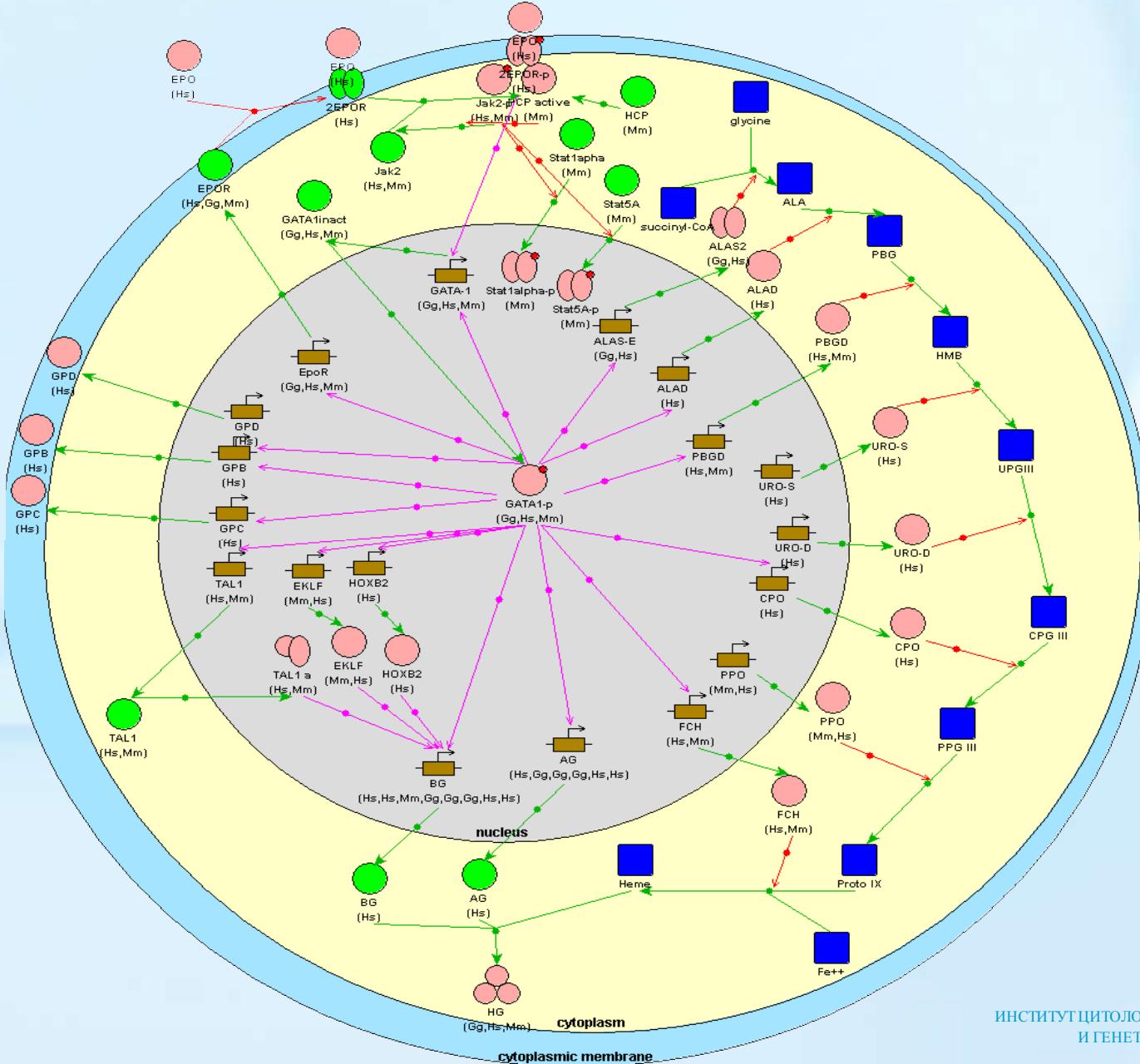


генетические полиморфизмы: более 2.5 миллиона  
вариантов!!!!



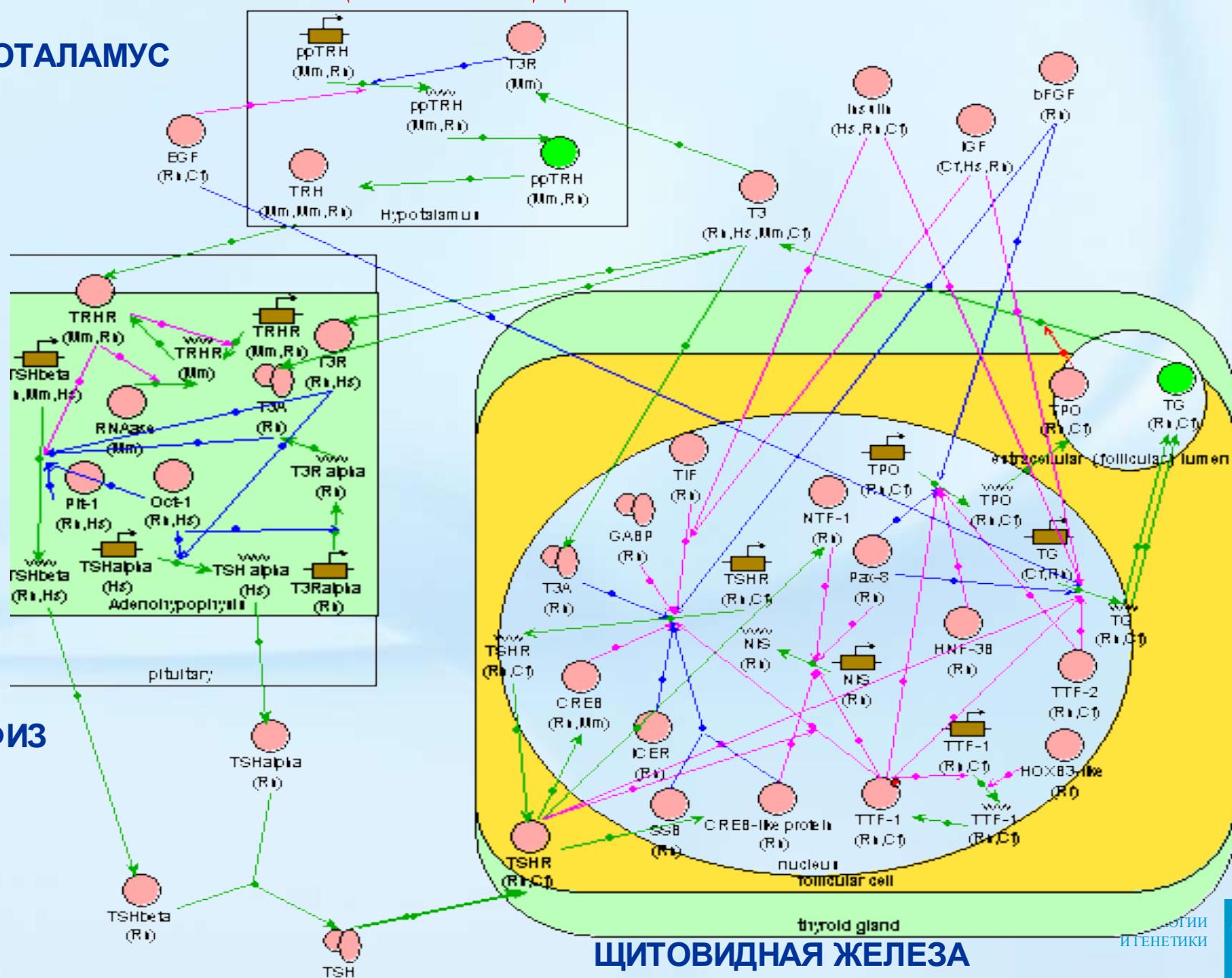


# ФРАГМЕНТ ГЕННОЙ СЕТИ СОЗРЕВАНИЯ И ДИФФЕРЕНЦИРОВКИ ЭРИТРОЦИТА



# ФРАГМЕНТЫ ГЕННОЙ СЕТИ РЕГУЛЯЦИИ ФУНКЦИИ ЩИТОВИДНОЙ ЖЕЛЕЗЫ

## ГИПОТАЛАМУС

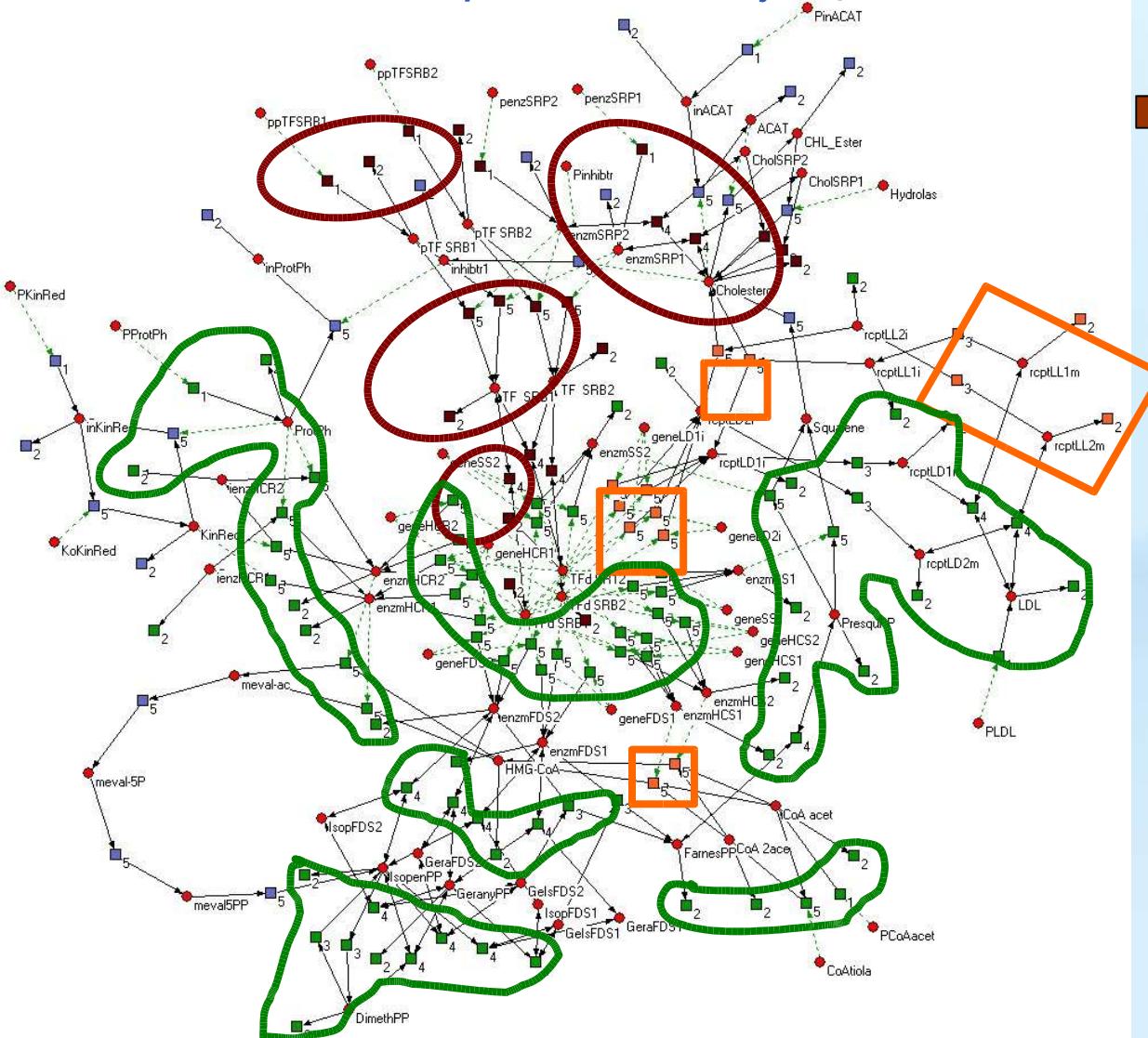


## **ЩИТОВИДНАЯ ЖЕЛЕЗА**

# ЛИМИТИРУЮЩИЕ ПРОЦЕССЫ В ГЕННОЙ СЕТИ

## регуляции биосинтеза холестерина в клетке

Генная сеть с указанием чувствительности стационарного содержания свободного холестерина в клетке к мутационным изменениям параметров

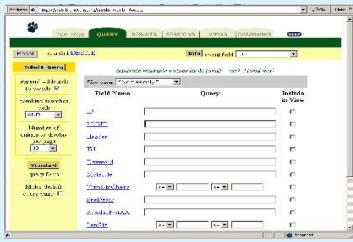


■ -изменение скоростей данных процессов существенно (до 200% относительно нормы) влияет на стационарную концентрацию холестерина;

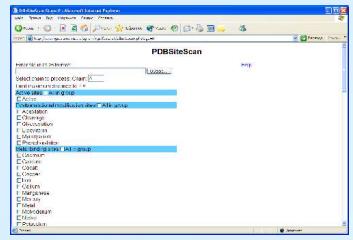
■ - стационарная концентрация холестерина меняется не более чем на 35% от нормы;

■ - стационарная концентрация холестерина меняется не более чем на 25% от нормы.

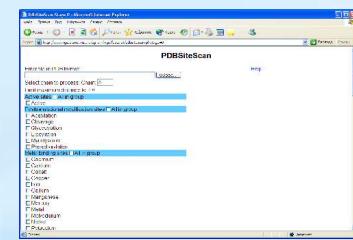
# Инструменты компьютерной структурной биологии



База данных  
функциональных сайтов



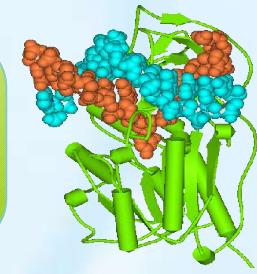
Программа распознавания  
функциональных сайтов



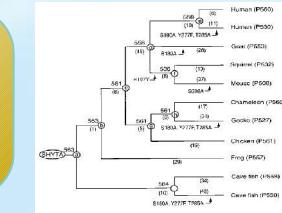
Программа реконструкции  
молекулярных комплексов



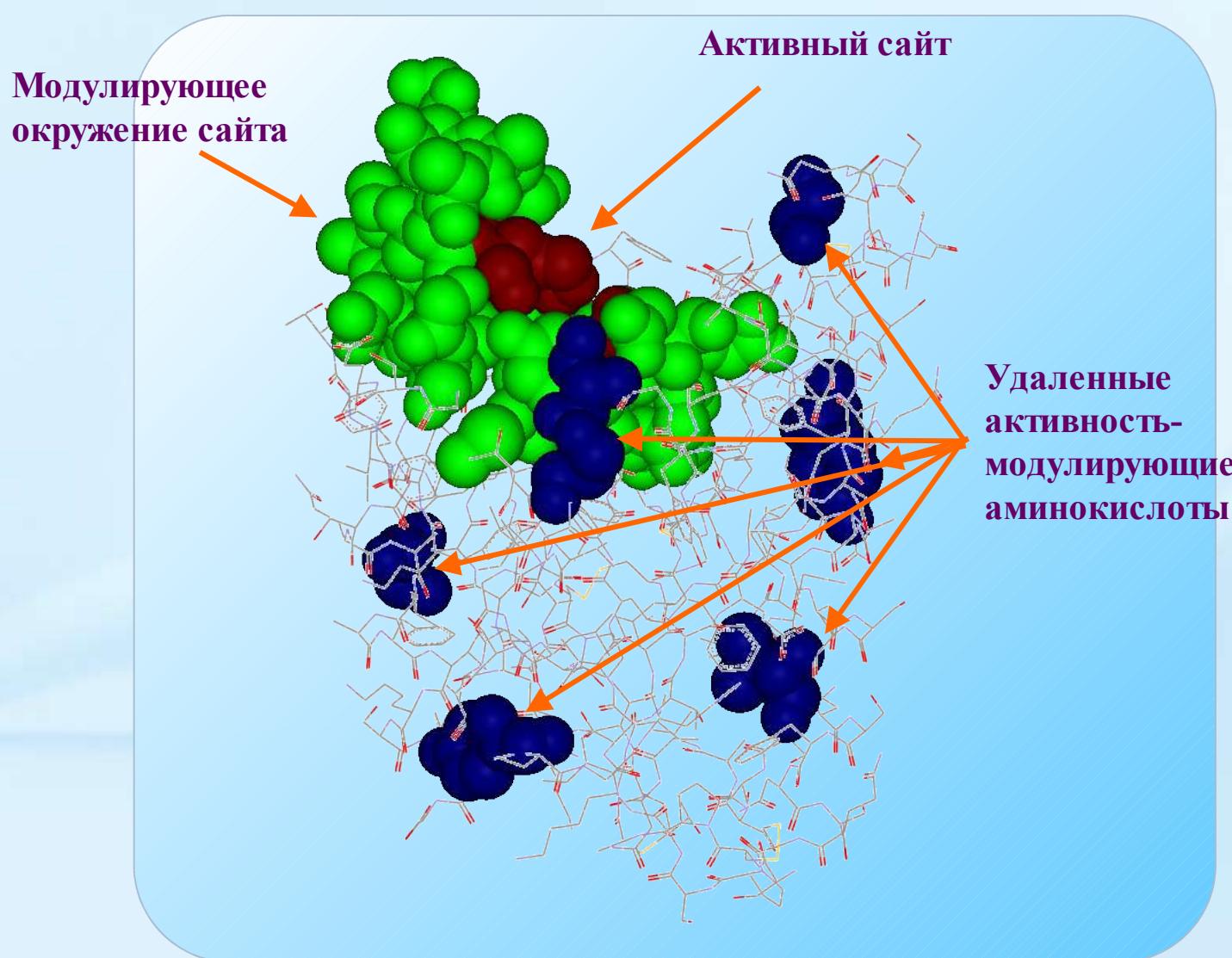
Программа анализа  
взаимосвязи структура-  
активность



Функциональная  
аннотация белков

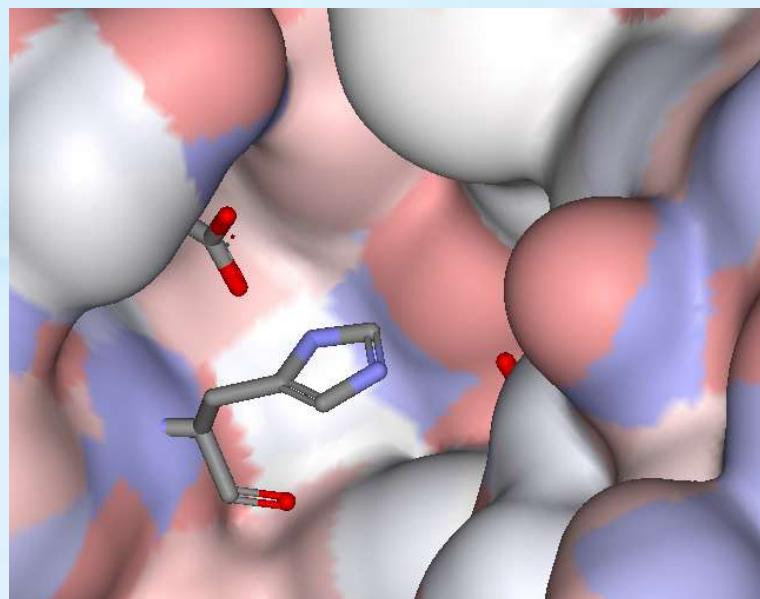
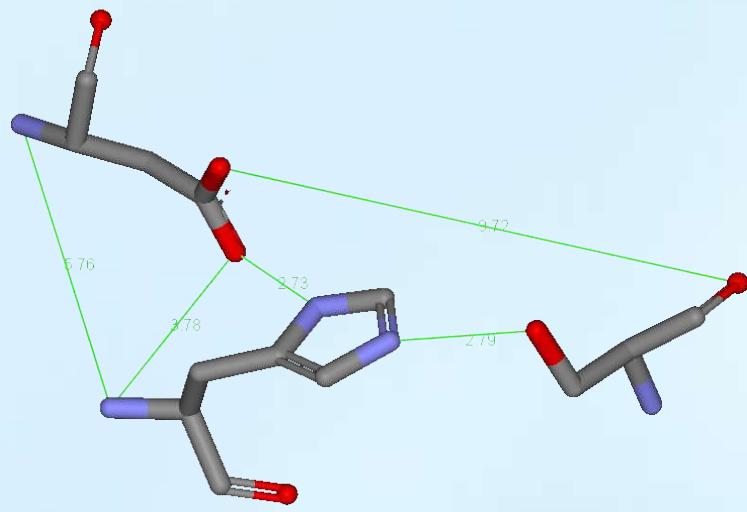


# Концепция структурно-функциональной организации белков



Функциональные центры белков ключевой объект изучения  
в компьютерной протеомике

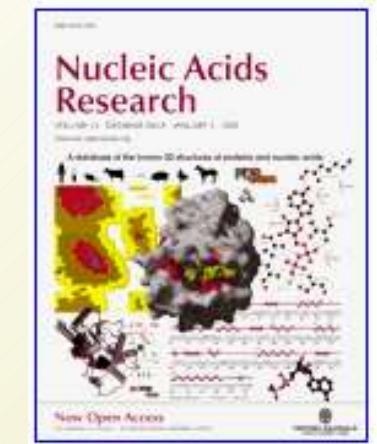
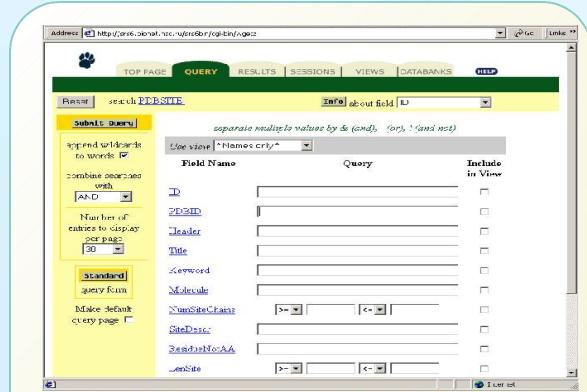
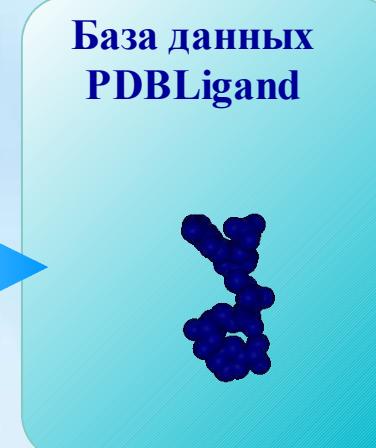
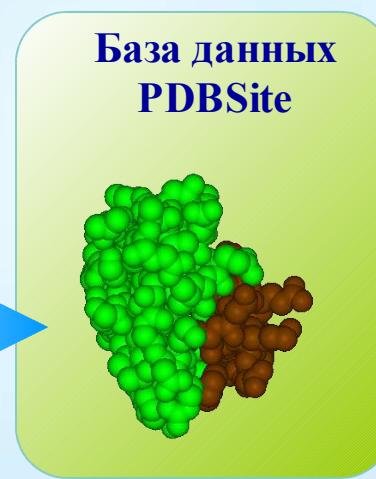
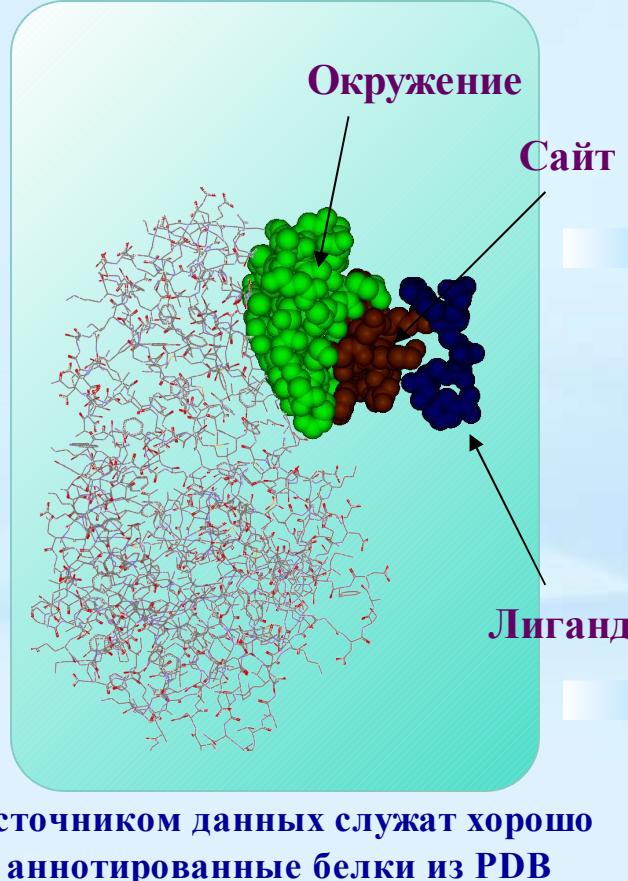
# **Физико-химические и структурные характеристики, определяющие специфичность функциональных сайтов**



- **Физико-химические свойства аминокислот сайта**
- **Особенности расположения аминокислот в первичной структуре**
- **Особенности конформации**
- **Вторичная структура**
- **Доступность для растворителя**
- **Полярность и заряд окружения**

# В базах данных PDBSite и PDBLigand накапливаются пространственные структуры и основные характеристики функциональных сайтов белков и их лигандов

<http://wwwmgs.bionet.nsc.ru/mgs/gnw/pdbsite/>

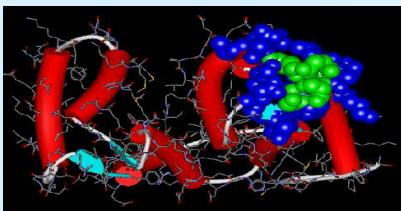


Ivanisenko et al, (2005)  
Nucleic Acids Res.

# Содержание базы данных PDBSite

Фармакологические  
препараты (50)

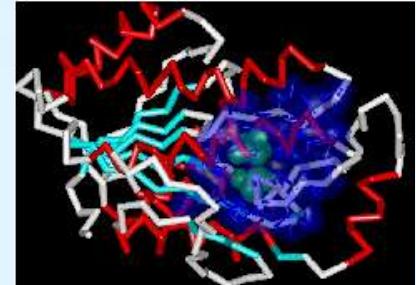
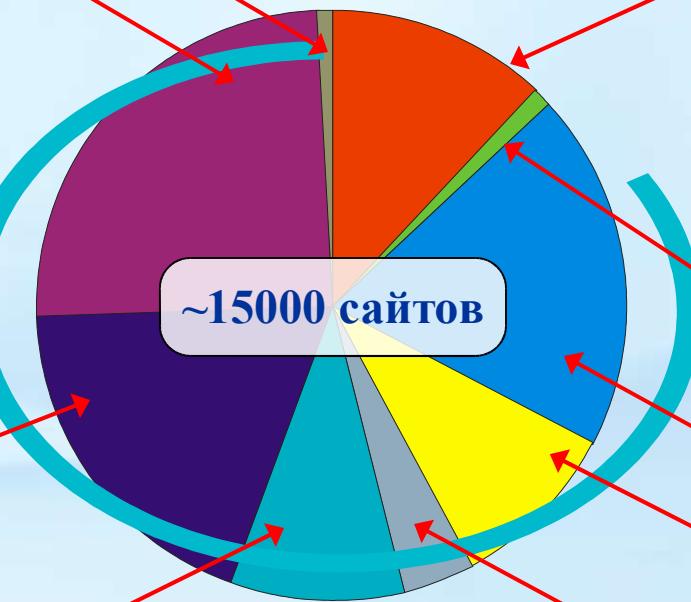
Белок+ДНК (2700)



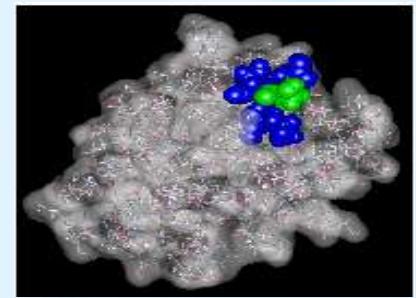
Сайты связывания

Белок+РНК (2000)

Белок+белок  
(1000)



Катализитические  
центры (1300)

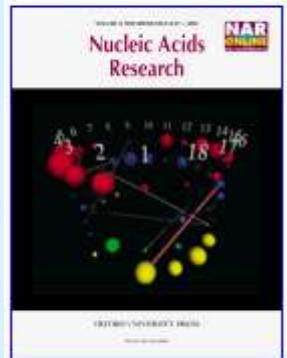


Посттрансляционная  
модификация (100)

Органические  
лиганды (2100)

Металлы (1000)

Неорганические  
лиганды (400)



# PDBSiteScan: программа распознавания функциональных сайтов

<http://wwwmgs.bionet.nsc.ru/mgs/systems/fastprot/pdbsitescan.html>

Ivanisenko et al, (2004)  
Nucleic Acids Res.

PdbSiteScan Stage 0 - Microsoft Internet Explorer

Файл Правка Вид Сборочное Сервис Справка

Назад ⌛ 🔍 Помощь Избранное Медиа

Адрес: http://wwwmgs.bionet.nsc.ru/cgi-bin/mgs/fastprot/pdbsitescan.pl?stage=0

PDBSiteScan

Enter file in PDB format:  Обзор...

Select chain to process: Chain: A

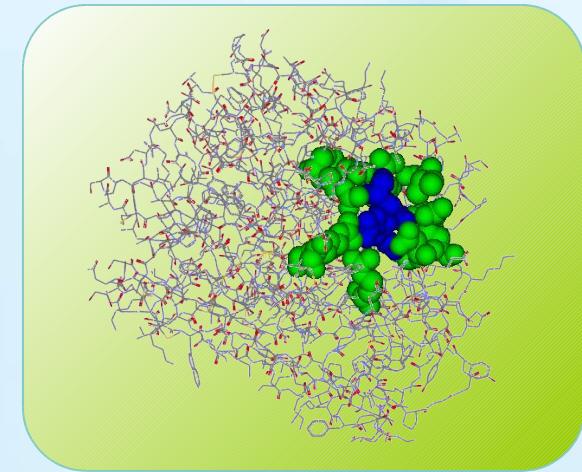
Limit maximum distance to 2.0

Active sites:  All in group  
 Active  
 Posttranslational modification sites  All in group  
 Acetylation  
 Cleavage  
 Glycosylation  
 Lipoylation  
 Myristylation  
 Phosphorylation  
 Metal binding sites  All in group  
 Cadmium  
 Calcium  
 Cobalt  
 Copper  
 Iron  
 Gallium  
 Manganese  
 Mercury  
 Metal  
 Molybdenum  
 Nickel  
 Potassium

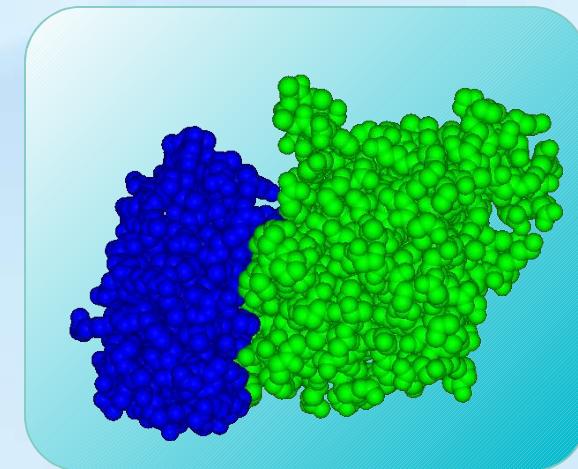
Help

Интернет

Распознавания  
функциональных  
сайтов

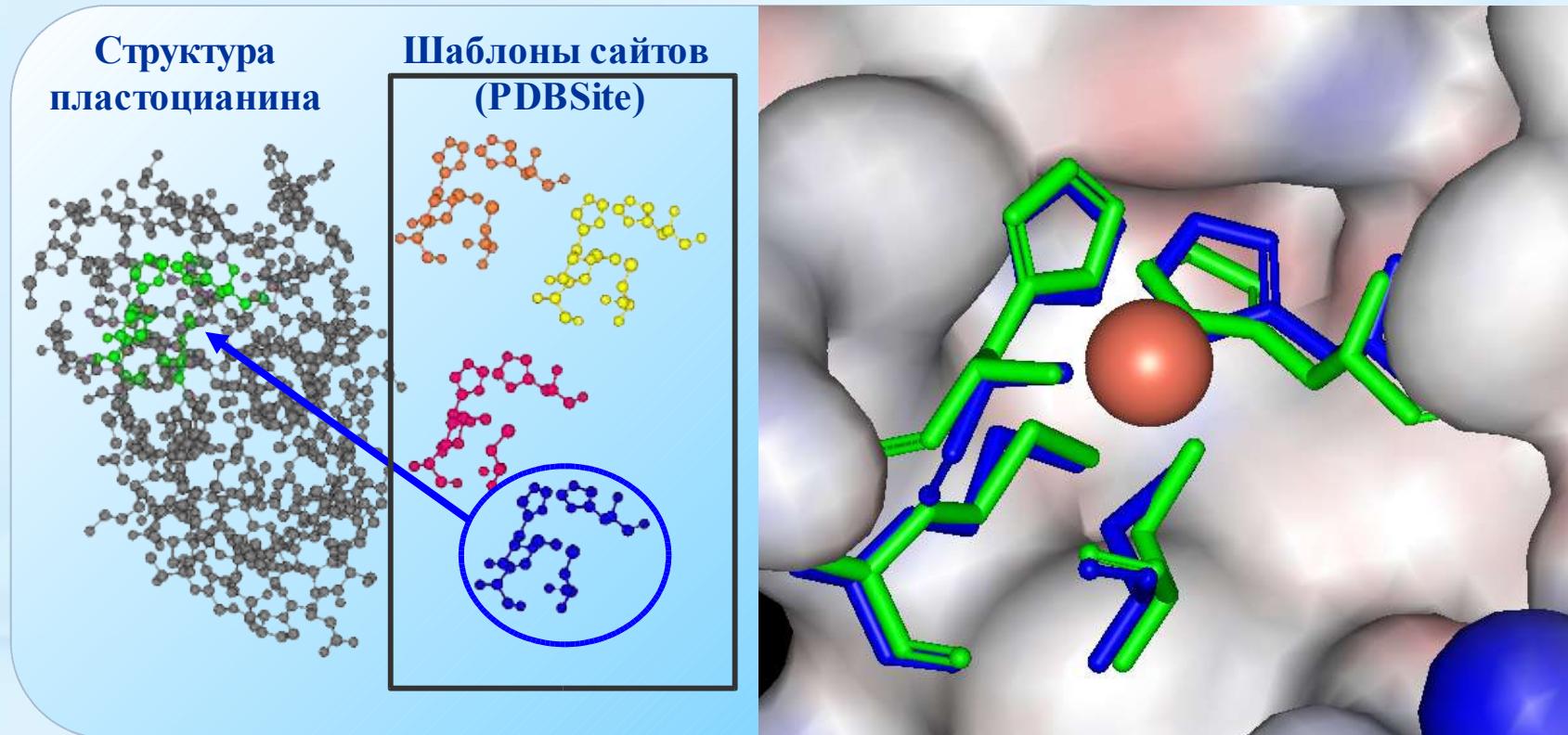


Реконструкция  
молекулярных  
комплексов  
сайт-лиганд



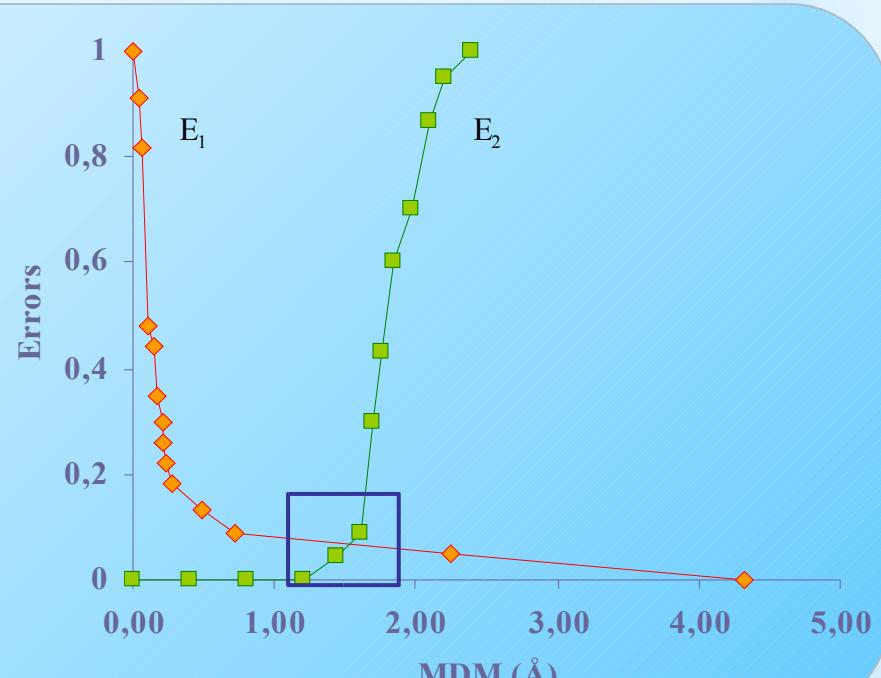
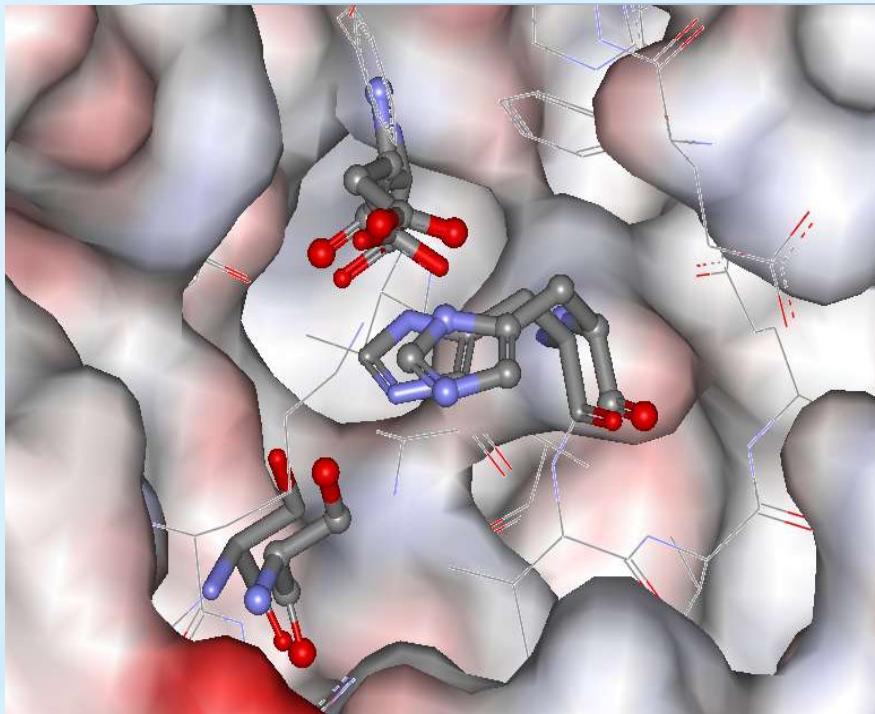
# Программа поиска активных сайтов в структурах белков: PDBSiteScan

Поиск сайта связывания меди в молекуле пластоцианина (PDB ID 1BXU).



Зеленым цветом изображены остатки распознанного сайта в пластоцианине,  
синим цветом изображены остатки сайта-шаблона из базы PDBSite (ID 1B3ICU).  
Ион меди показан оранжевым шариком.

# Анализ точности распознавания катализитических центров в суперсемействе гидролаз



MDM – мера качества совмещения в пространстве шаблона активного сайта из базы PDBSite и фрагмента белка.

E<sub>1</sub> – ошибка I-рода (недопредсказания)

E<sub>2</sub> – ошибка II-рода (перепредсказания)

# **Функциональная аннотация белков**

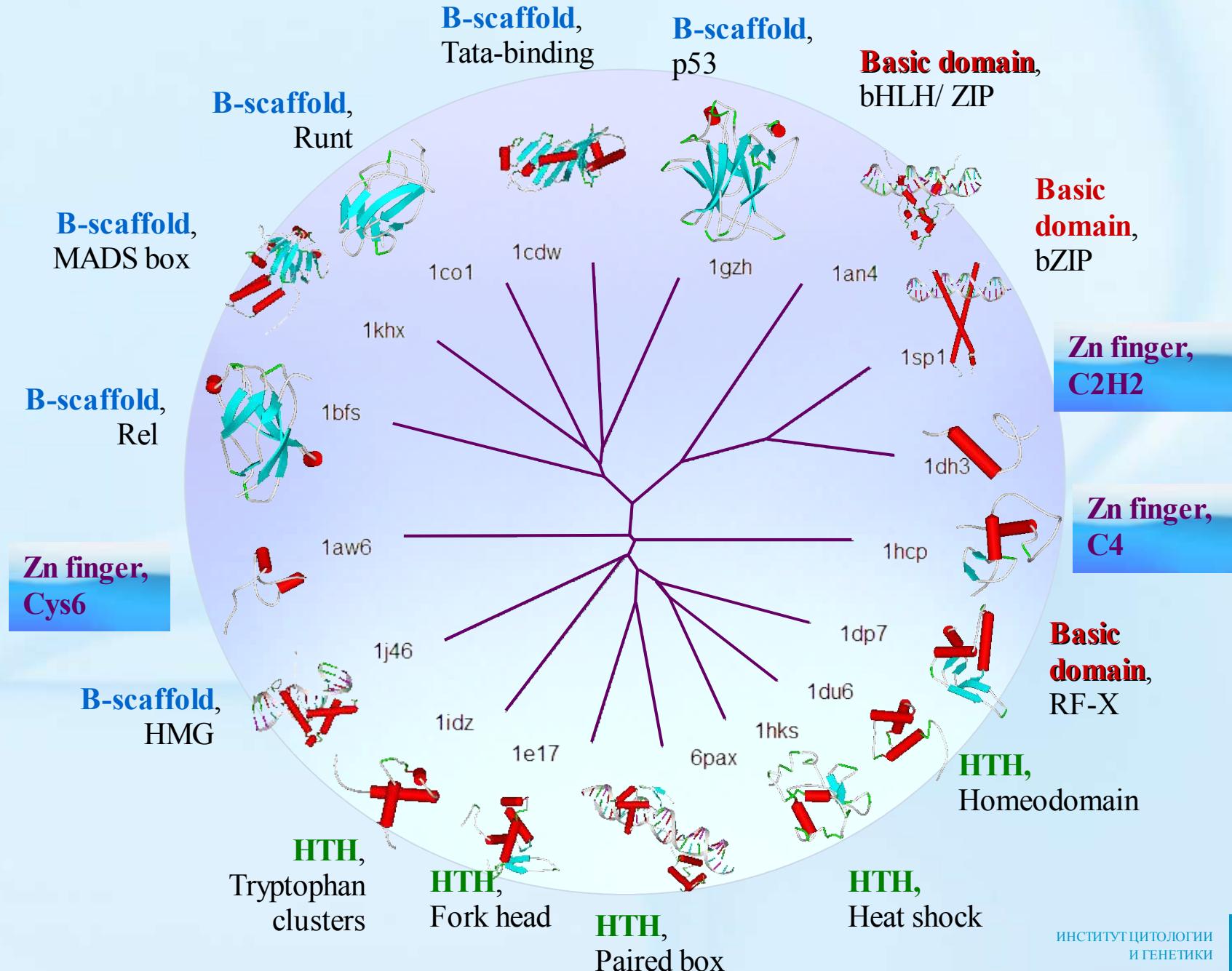


# **Функциональный портрет белков с известной пространственной структурой, реконструированный на основе компьютерного анализа**

- Сайты связывания органических лигандов **82%**
- Сайты связывания ионов металлов **64%**
- Сайты белок-белковых взаимодействий **41%**
- Сайты белок-ДНК/РНК взаимодействий **35%**
- Активные сайты **24%**

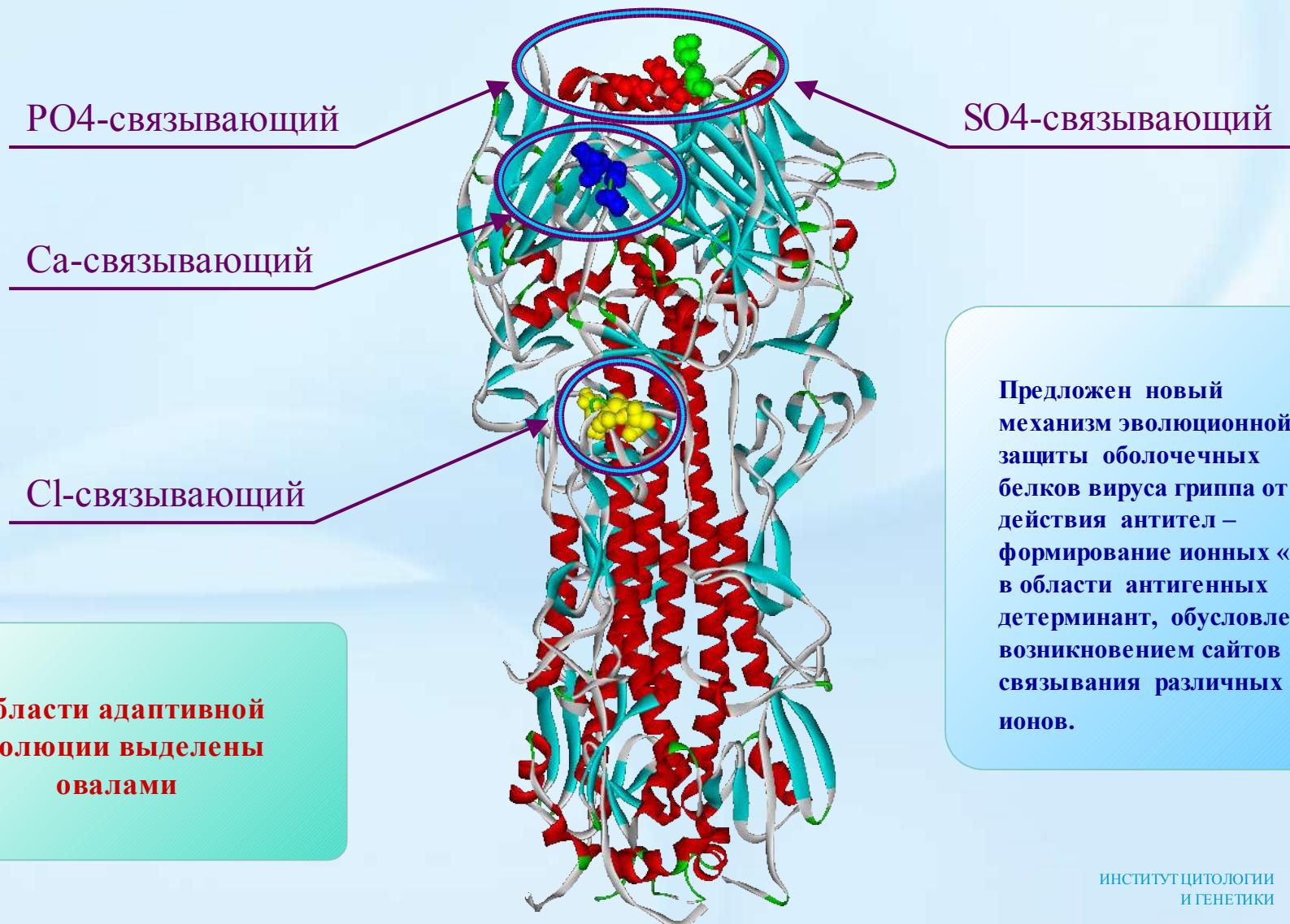


# Классификация ДНК-связывающих доменов транскрипционных факторов

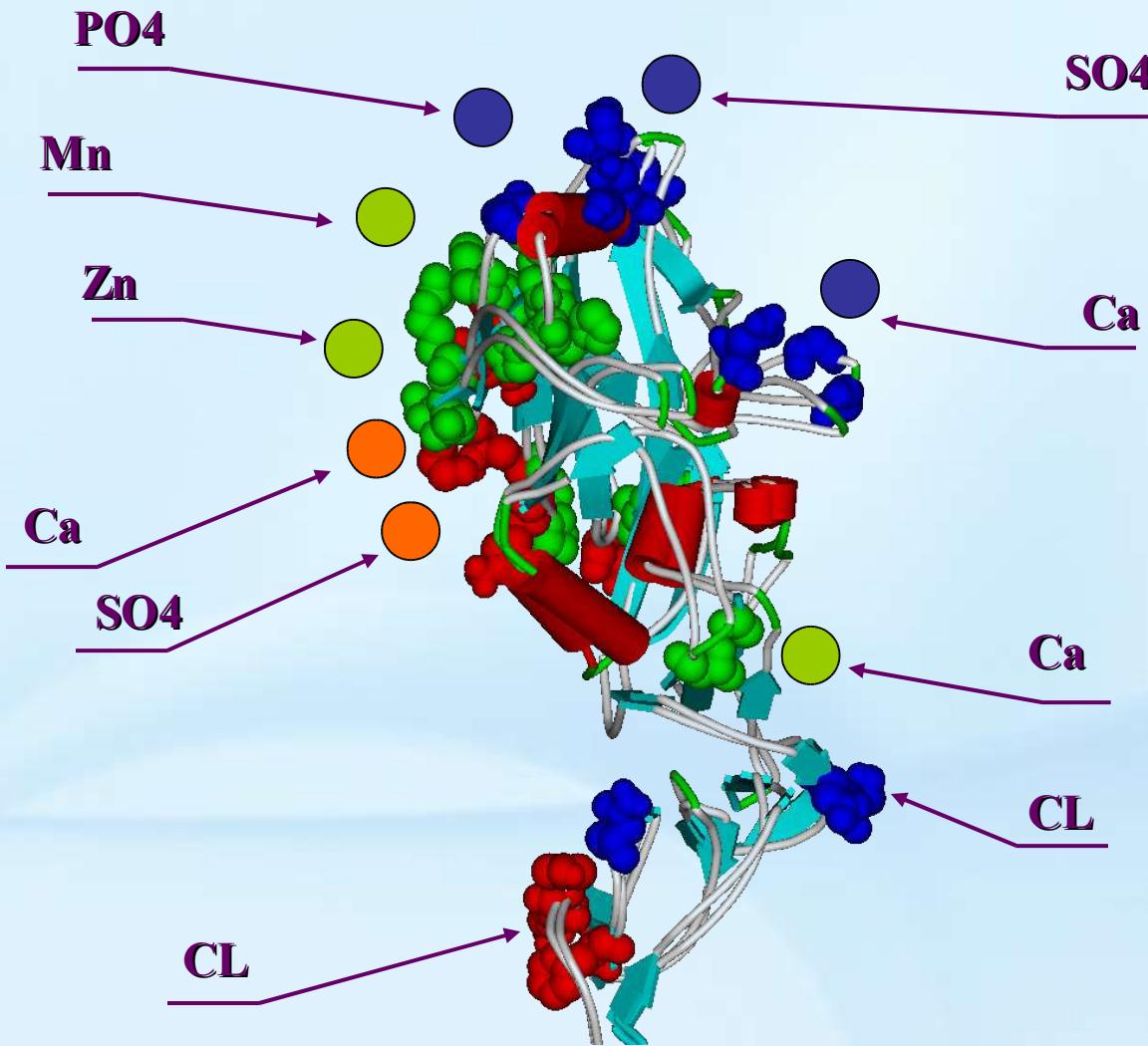


# Адаптивная эволюция : мутационное ускользание вируса от иммунной системы

Сайты связывания ионов в пространственной структуре гемагглютинина вируса гриппа А человека совпадают участками адаптивной эволюции



# **Адаптивная эволюция : мутационное ускользание вируса от иммунной системы. Полиморфизм сайтов связывания ионов в пространственной структуре гемагглютинина вируса гриппа A человека**



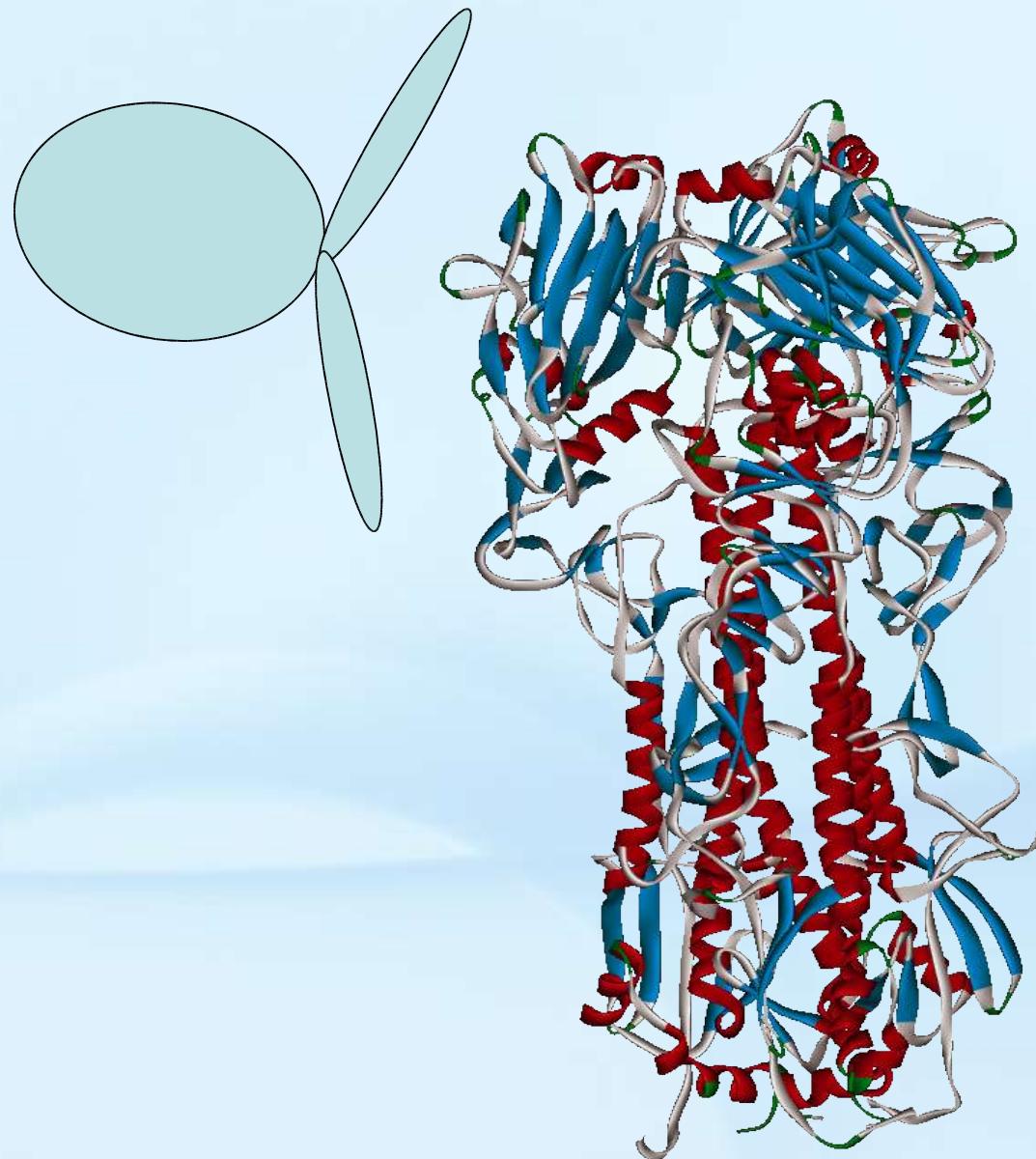
**В пространственной структуре гемагглютининов близкородственных штаммов вируса гриппа А сайты связывания ионов занимают различное положение, что может быть связано с изменением направления атаки имунной системы**

X-31

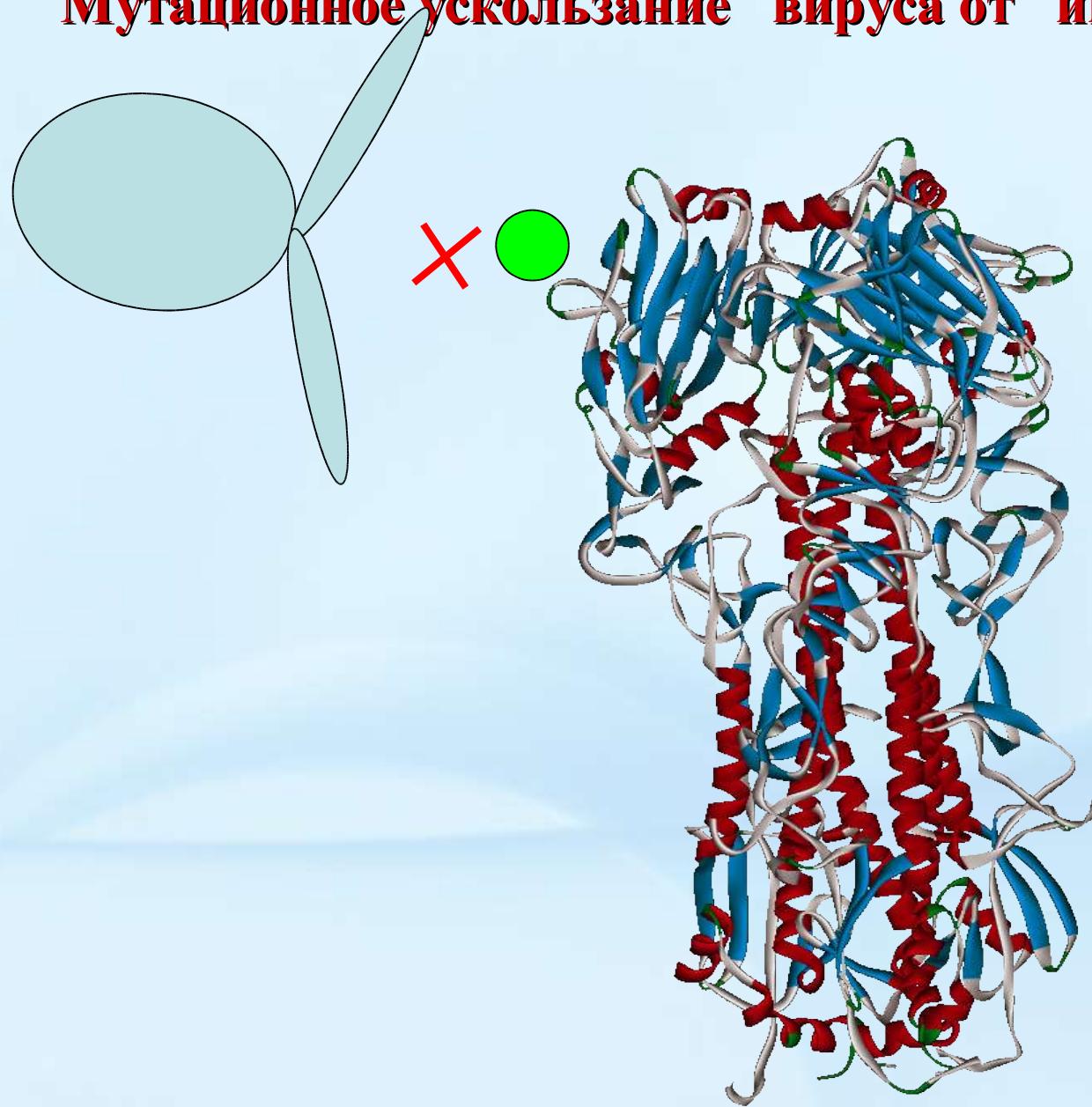
Q67014

Q67100

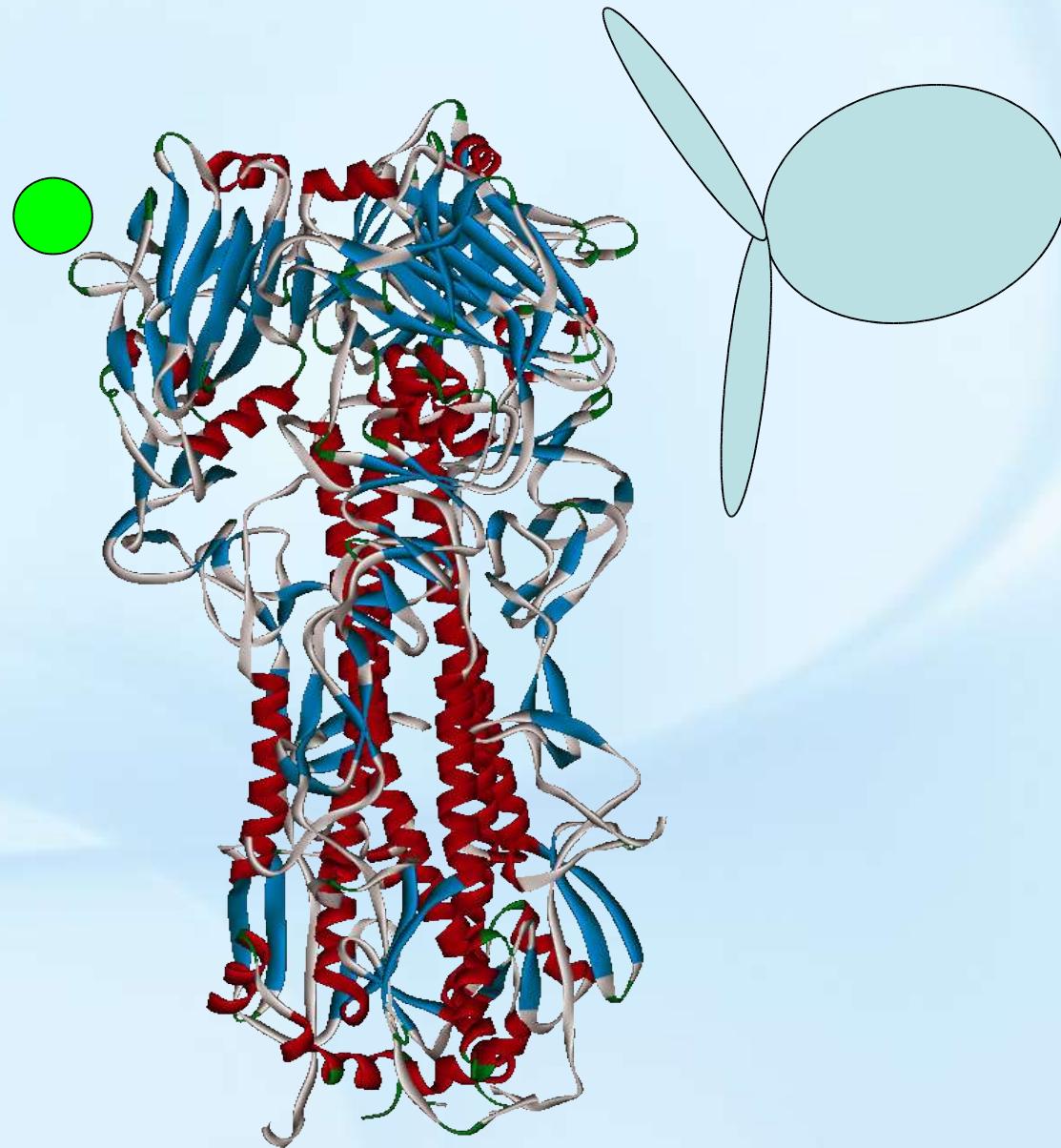
# **Мутационное ускользание вируса от иммунной системы.**



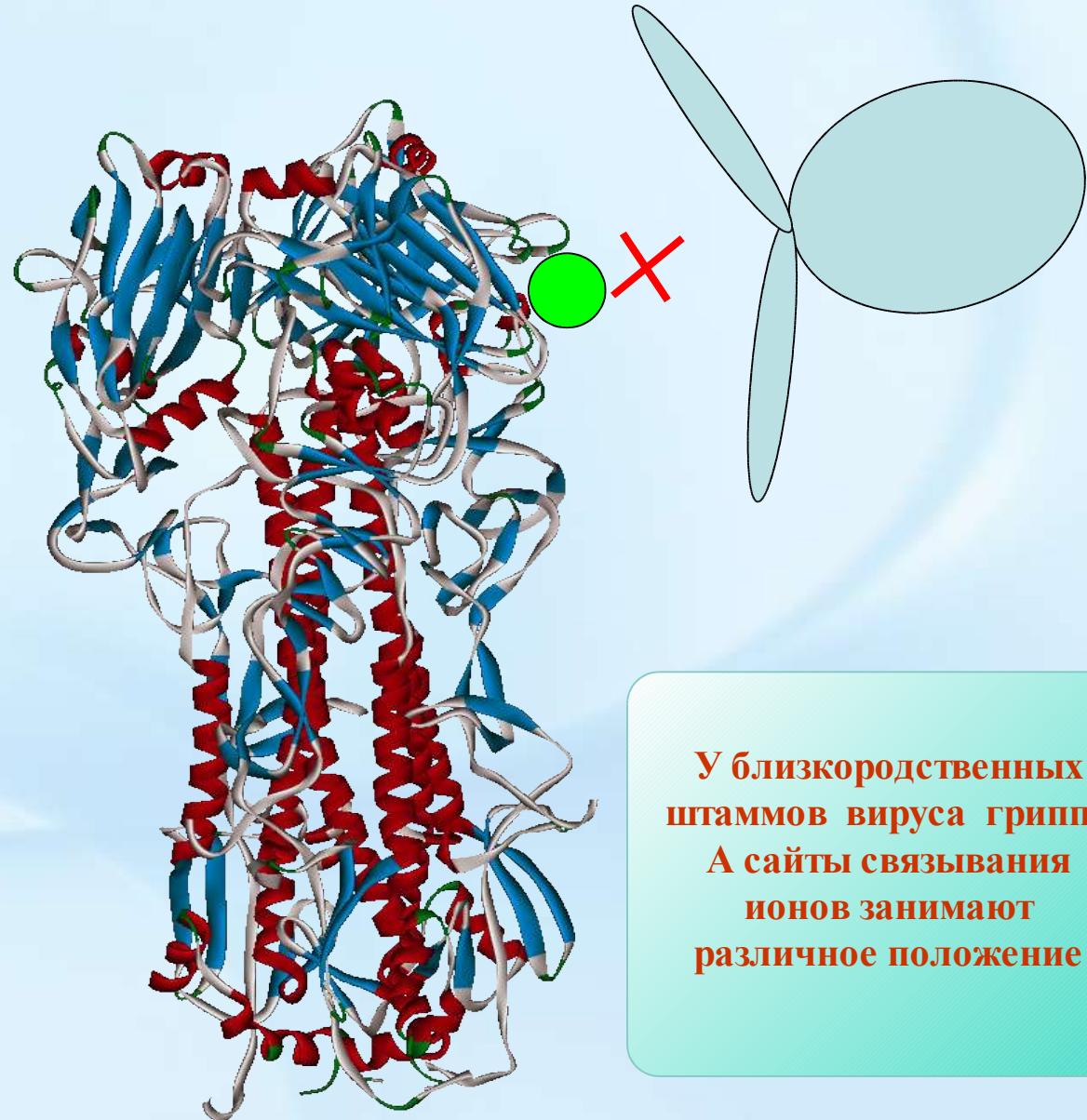
# **Мутационное ускользание вируса от иммунной системы.**



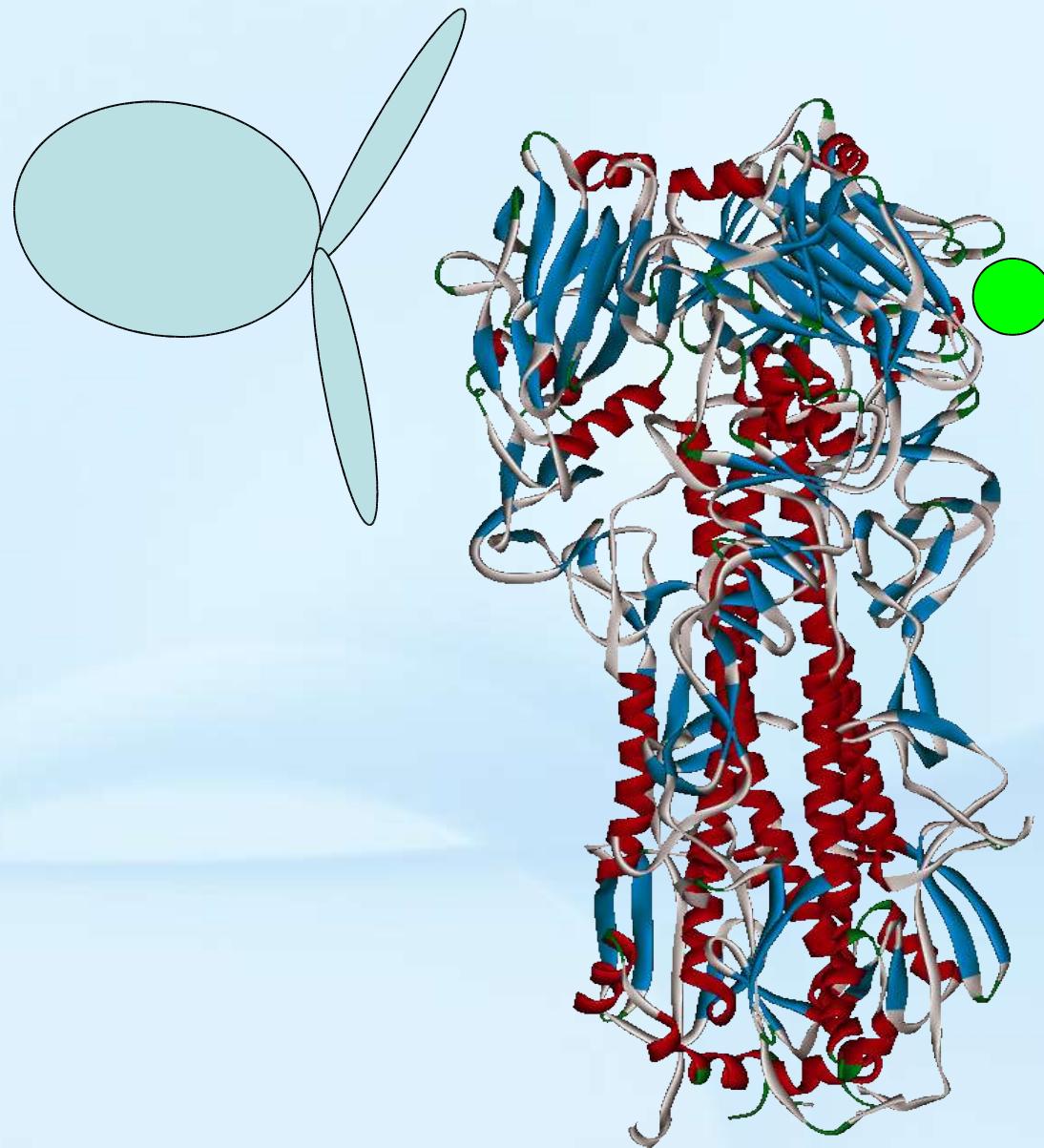
# **Мутационное ускользание вируса от иммунной системы.**



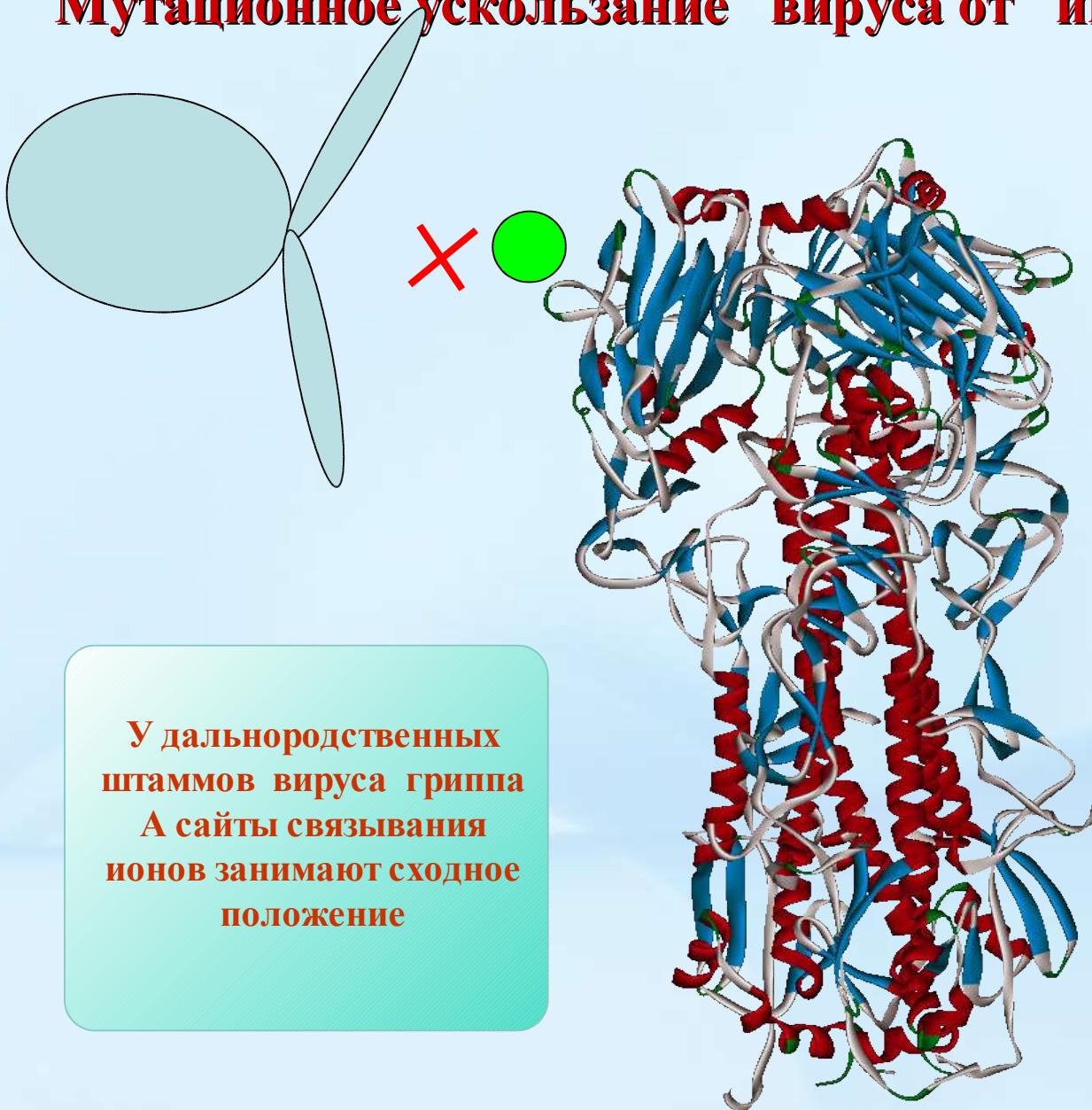
# Мутационное ускользание вируса от иммунной системы.



# **Мутационное ускользание вируса от иммунной системы.**

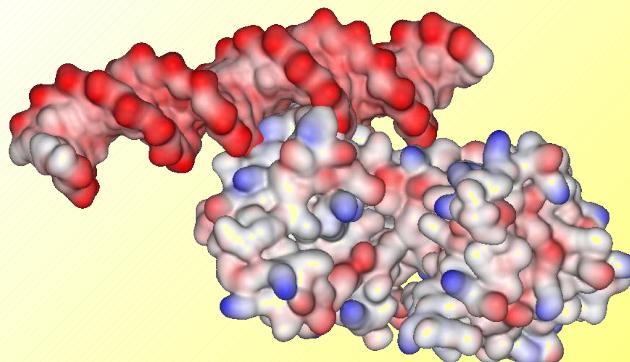


# Мутационное ускользание вируса от иммунной системы.

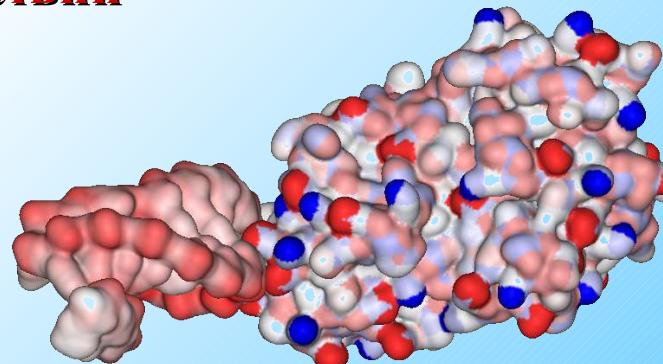


# PDBSiteScan – инструмент для компьютерной интерпретации экспериментальных данных.

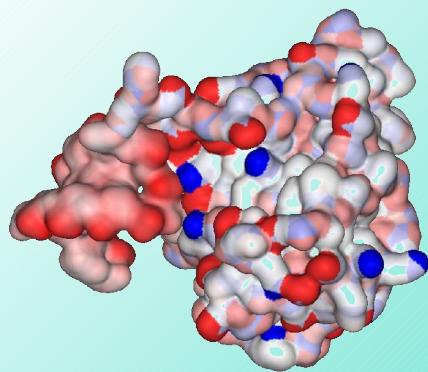
## Реконструкция молекулярных комплексов белок-ДНК/РНК взаимодействий



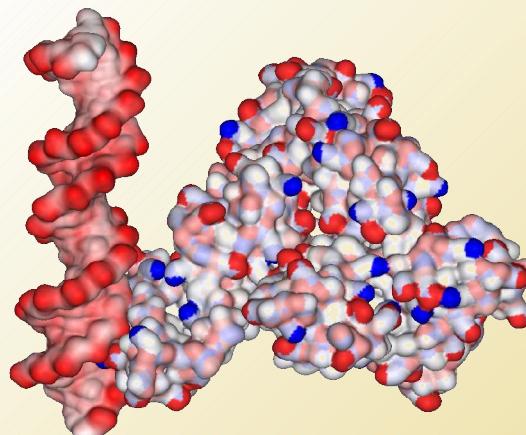
фосфоглицерат киназа 1 - ДНК



глицеральдегид-3-fosfat-дегидрогеназы  
(GPD) - РНК



Н концевой домен белка теплового шока  
HSP90 - РНК



сывороточный альбумин - ДНК



# **Количественный анализ взаимосвязи структура-активность в семействах гомологичных белков**

- Выявление аминокислотных остатков, влияющих на активность белков – активность модулирующих центров.
- Установление зависимости между физико-химическими свойствами активность модулирующих центров и белковой активностью.

## **Задачи:**

Планирование белково-инженерных экспериментов.

Характеризация активностей белков в базах данных белковых последовательностей.

Характеризация активностей искусственных белков.



# WebProAnalyst: программа для количественного анализа взаимосвязи структура-активность в семействах гомологичных белков

<http://wwwmgs.bionet.nsc.ru/mgs/programs/panalyst/>

Множественное выравнивание последовательностей белков

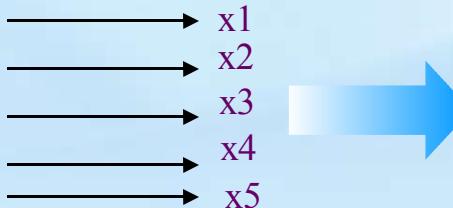
Активности

	подвижная рамка																								
белок 1	E	L	Y	E	Q	Y	N	K	A	F	S	A	H	W	K	q	e	t	g	d	n	V	V	I	
белок 2	A	S	L	T	N	A	M	Q	D	I	A	T	Q	F	K	K	E	K	G	V	D	V	V	S	
белок 3	-	q	H	K	E	A	A	T	A	V	A	K	A	F	E	Q	E	T	G	I	K	V	T	L	
белок 4	T	E	Y	V	P	P	G	-	L	L	E	Q	F	T	K		E	T	G	I	K	V	I	Y	
белок 5	T	N	F	L	G	T	L	E	Q	L	A	G	Q	F	A	K	Q	T	G	H	A	V	V	I	

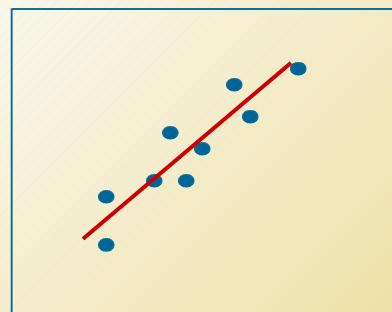
сайт

От аминокислот к их  
физико-химическим свойствам

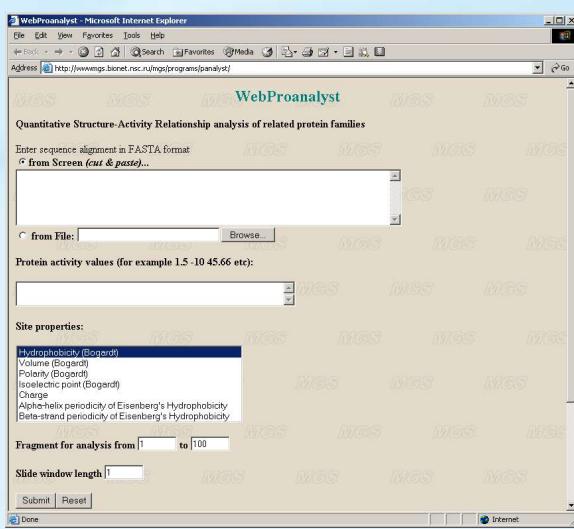
S	A
A	T
A	K
L	E
A	G



Поиск статистической  
зависимости  $Y = F(X)$



Физико-химическая  
характеристика сайтов

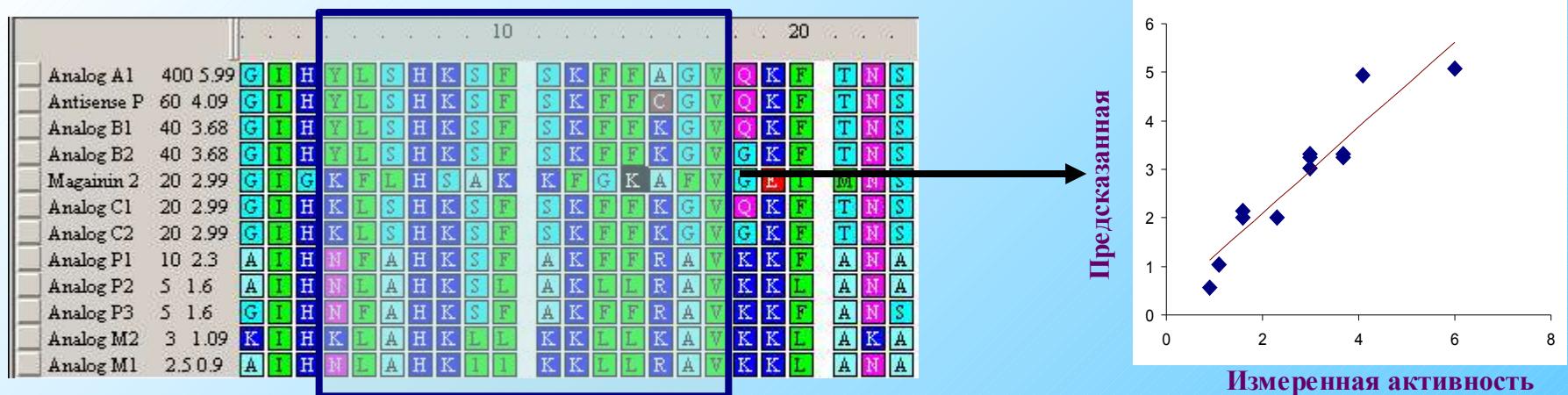


# Количественные зависимости структура-активность в белковых семействах

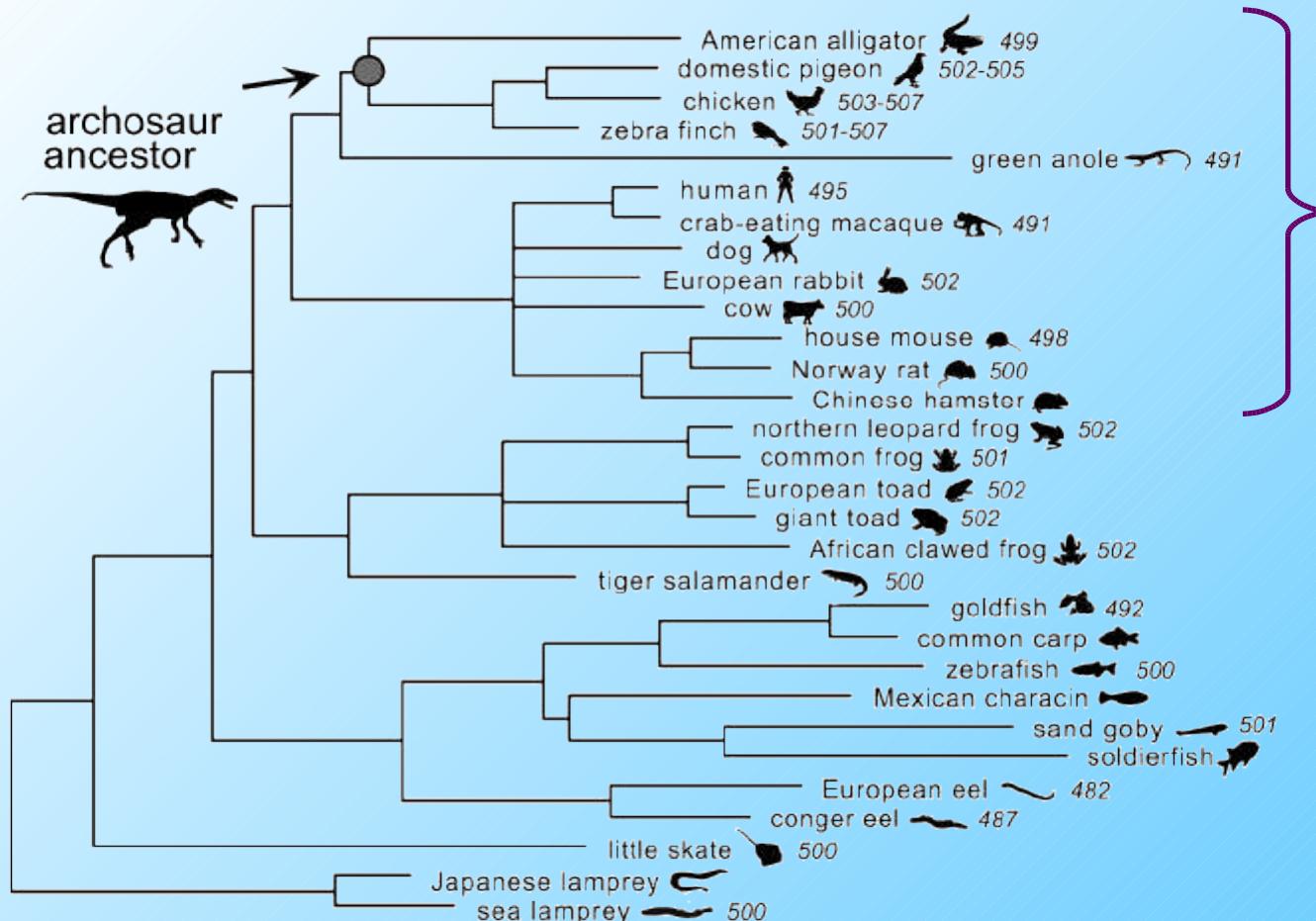
Зависимость между ингибирующей слипание тромбоцитов активностью дезинтегринов и зарядом и альфаспиральным моментом гидрофильности сайта 26-30



Зависимость между антимикробной активностью пептидов и альфаспиральным моментом гидрофобности сайта 3-18

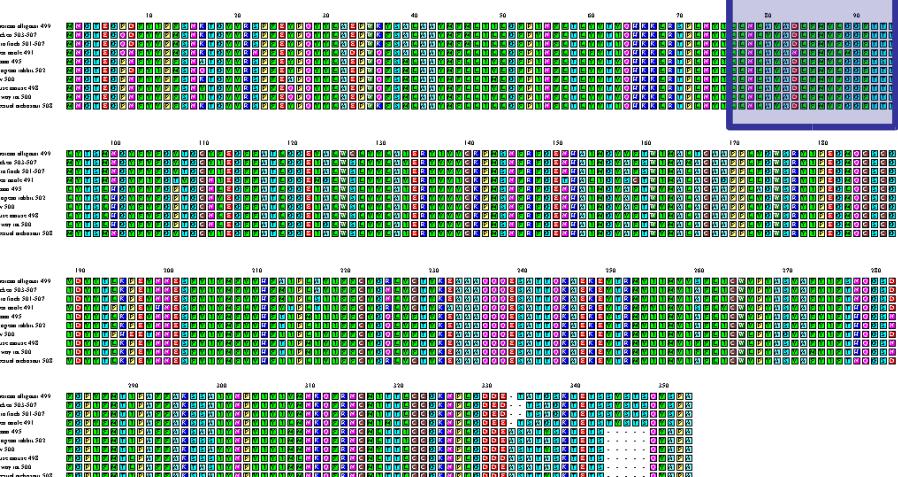
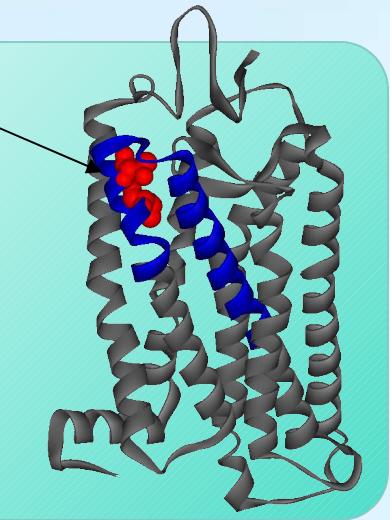


# Предсказание видимого спектра для зрения археозавра (λmax предкового родопсина археозавра)



# Предсказание $\lambda$ тах для родопсина археозавра

Остатки влияющие на сдвиг длины волны, результаты согласуются с данными других авторов *Briscoe, Mol. Biol. Evol. (2001)*.

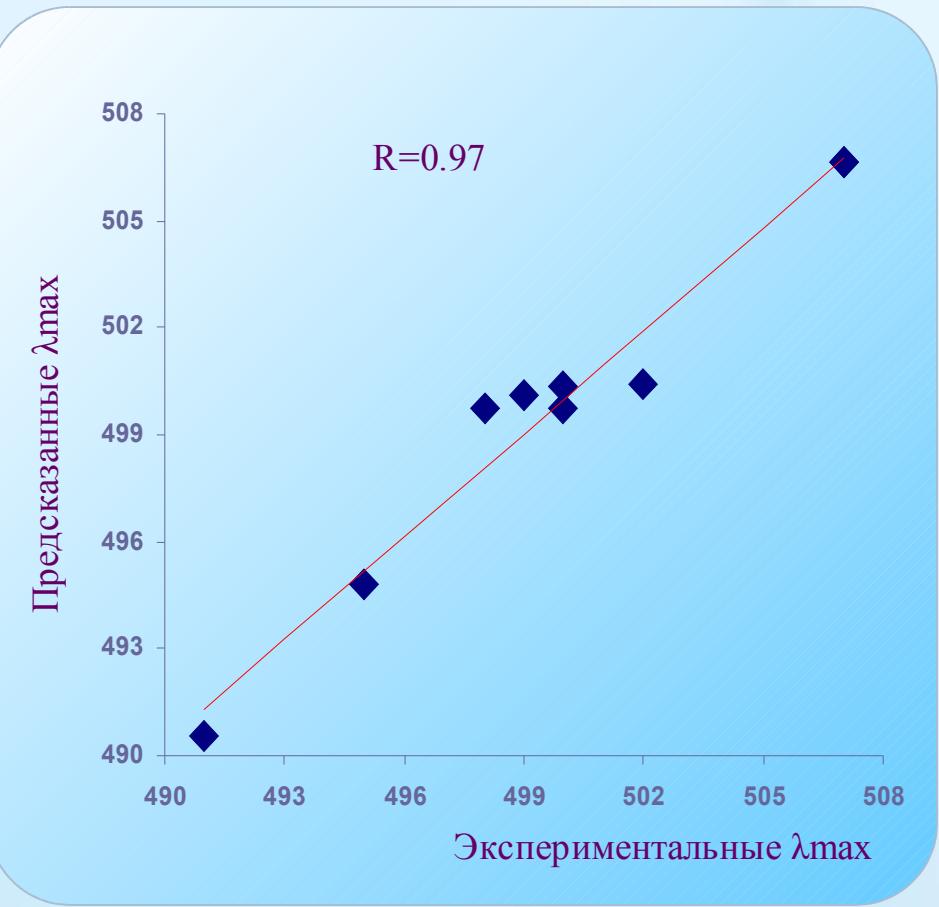


Уравнение регрессии:

$$Y = 15.784 * X_1 - 467.266 * X_2 - 37.661$$

$X_1$  – Среднее значение для изоэлектрической точки (Bogard)

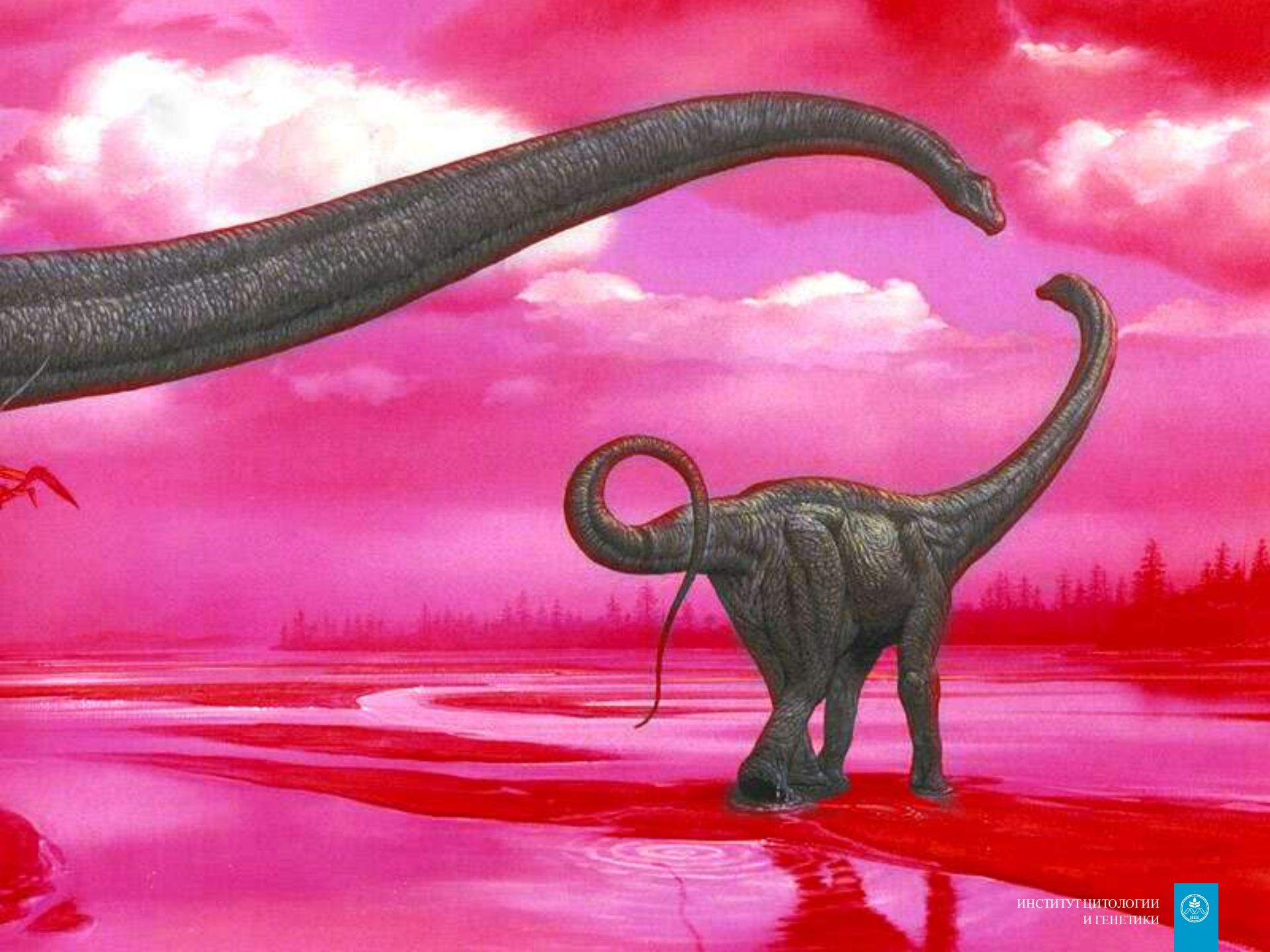
$X_2$  – Момент гидрофобности (Eisenberg)



$\lambda$ тах предсказанное :  $Y = 506.7$

$\lambda$ тах измеренное:  $Y = 508$





# Коэффициент корреляции структура-активность для идеальной регрессионной модели

Анализируемый сайт

Группа 1	ADERKSP	A
	ADERKSP	A
Группа 2	ADERKSP	R
	ADERKSP	R
	ADERKSP	R
Группа 3	ADERKSP	H
	ADERKSP	H

Активности

ADERKSP	$Y_{11}$
ADERKSP	$Y_{12}$
ADERKSP	$Y_{21}$
ADERKSP	$Y_{22}$
ADERKSP	$Y_{23}$
ADERKSP	$Y_{31}$
ADERKSP	$Y_{32}$

$$R^2 = 1 - \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}{\sum_{k=1}^N (Y_k - \bar{Y})^2}$$

$n_i$  - число белков в  $i$ -й группе,

$N$  - общее число белков,

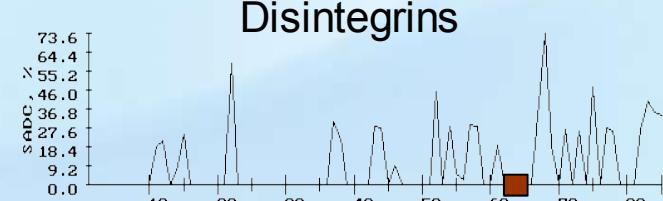
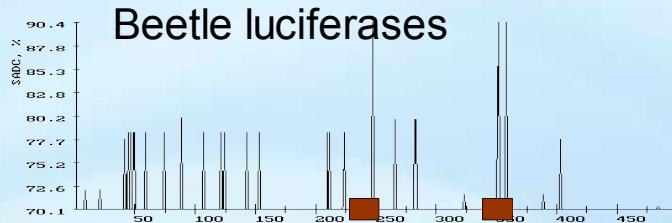
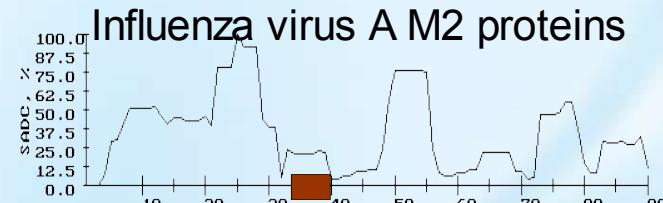
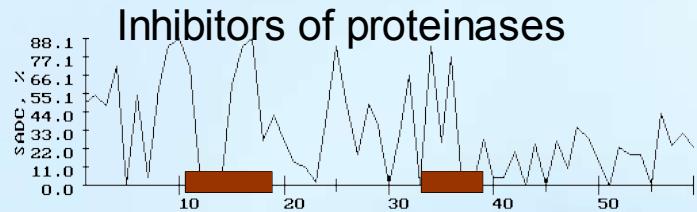
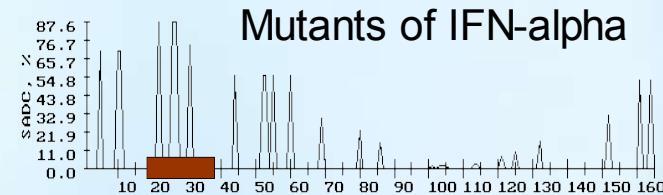
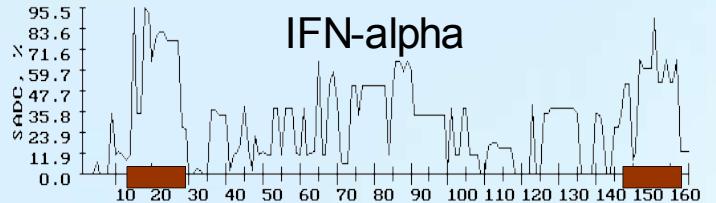
$Y_{ij}$  - активность  $j$ -го белка из  $i$ -й группы,

$\bar{Y}_i$  - среднее значение активности белков в  $i$ -й группе,

$\bar{Y}$  - общее среднее активности по всем белкам.



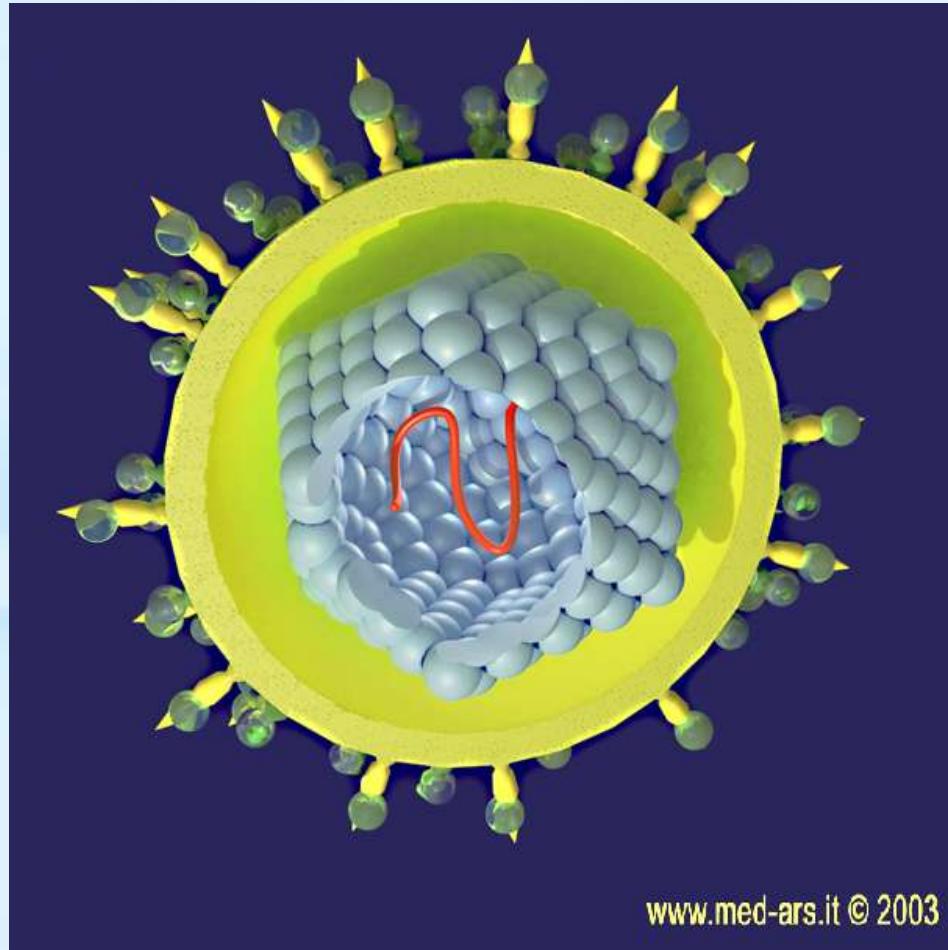
# Профили корреляций структура-активность для различных семейств белков



- обозначены известные функциональные районы

Для шести белковых семейств показано, что пики на профилях корреляций расположены вблизи функциональных районов

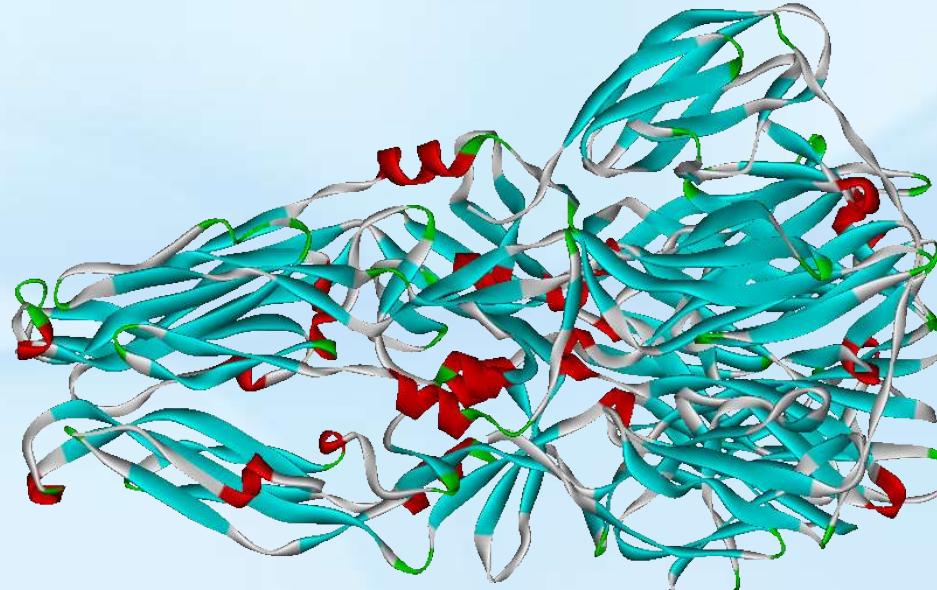
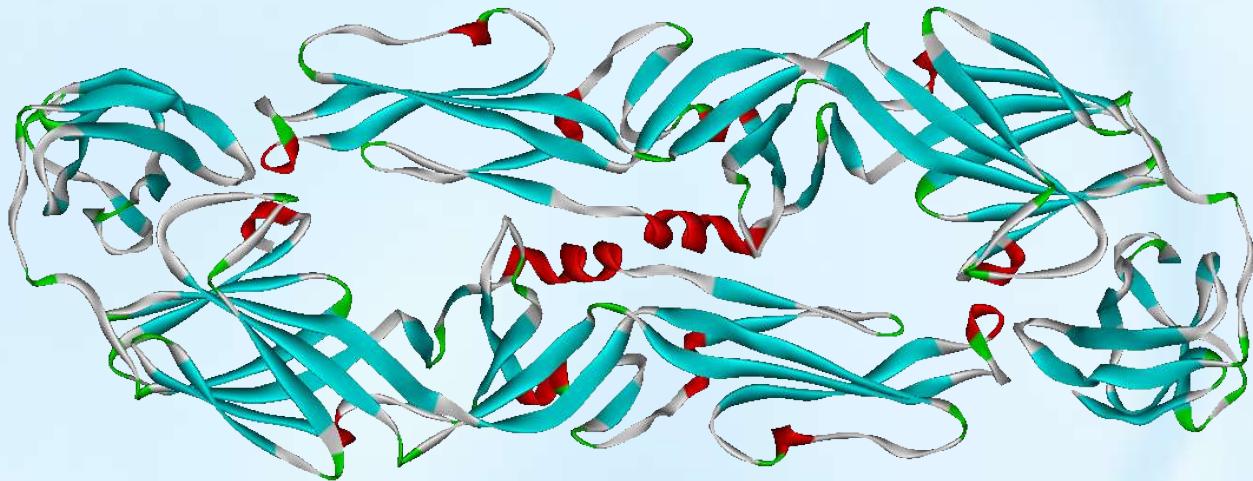
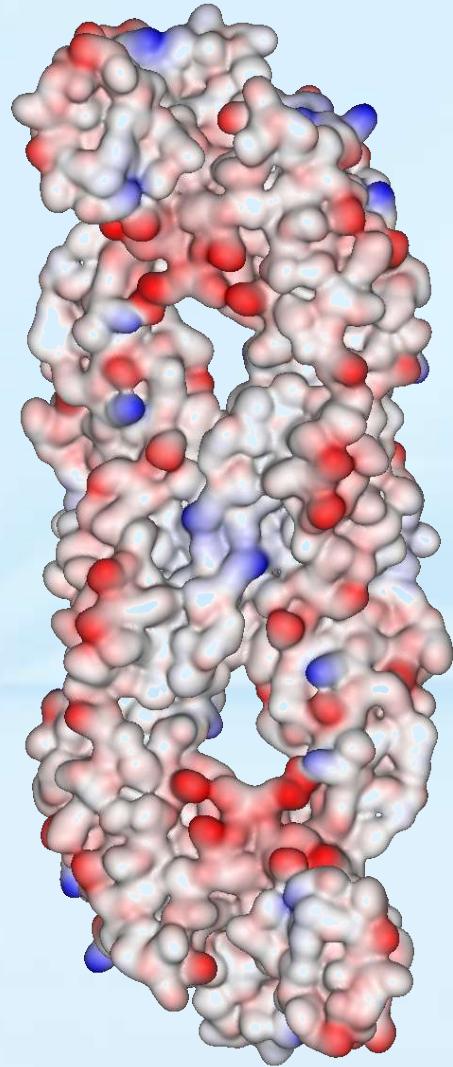
# **Поиск потенциальных мишеней действия противовирусных лекарственных препаратов**



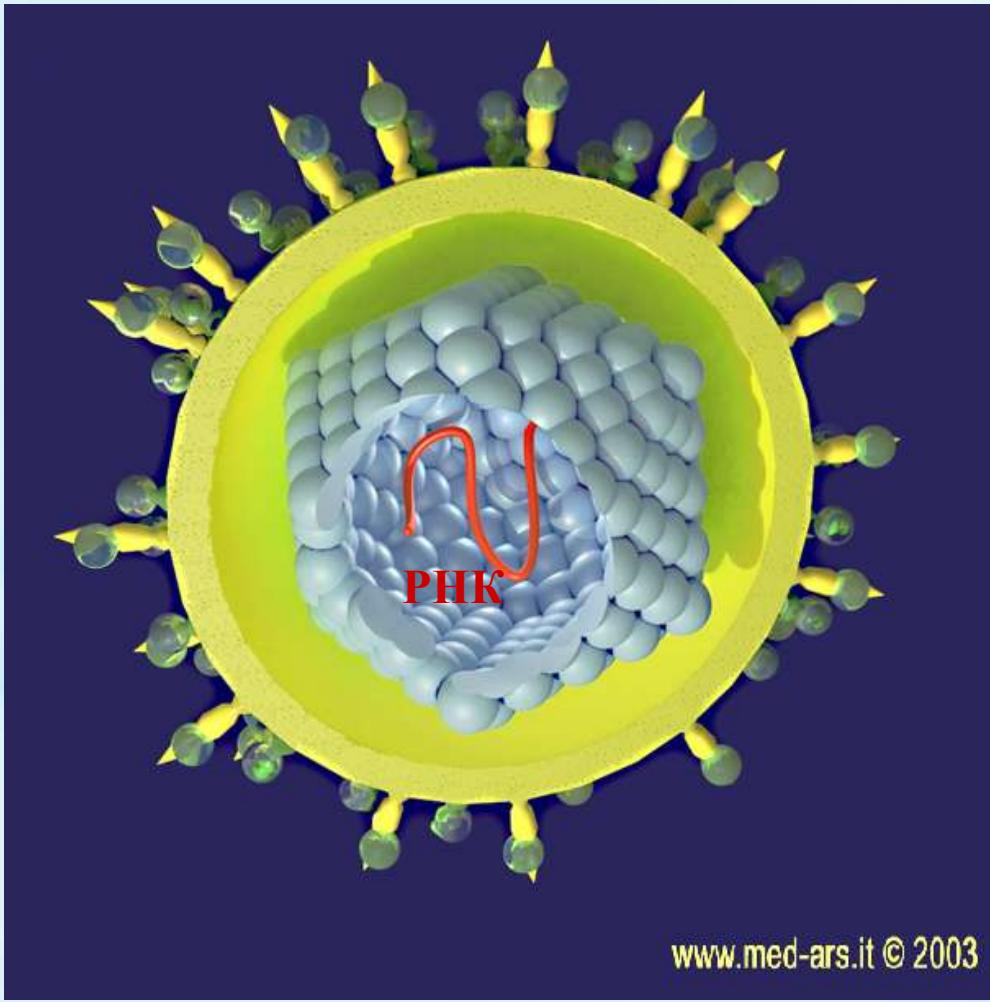
## **Бактериофаг Т4. Проникновение в клетку бактерии.**



# Вирус клещевого энцефалита. Оболочечный белок E.

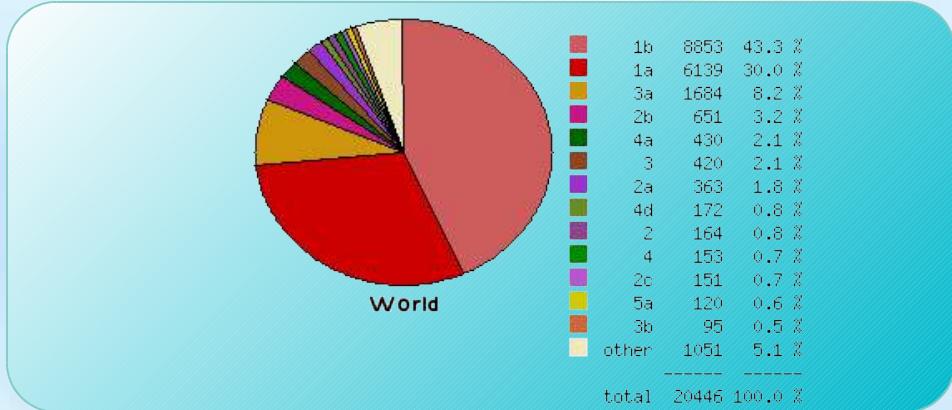
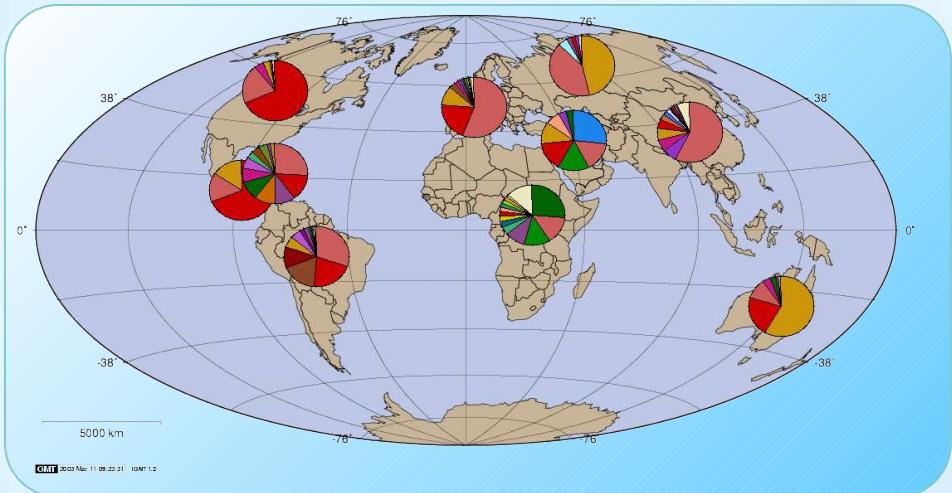
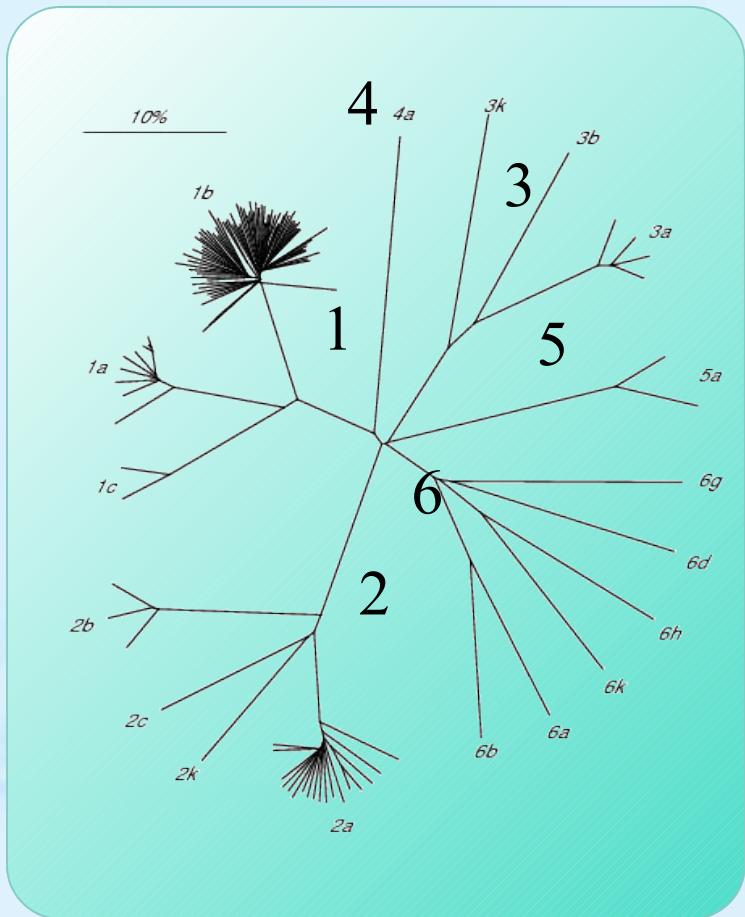


# Вирус гепатита С



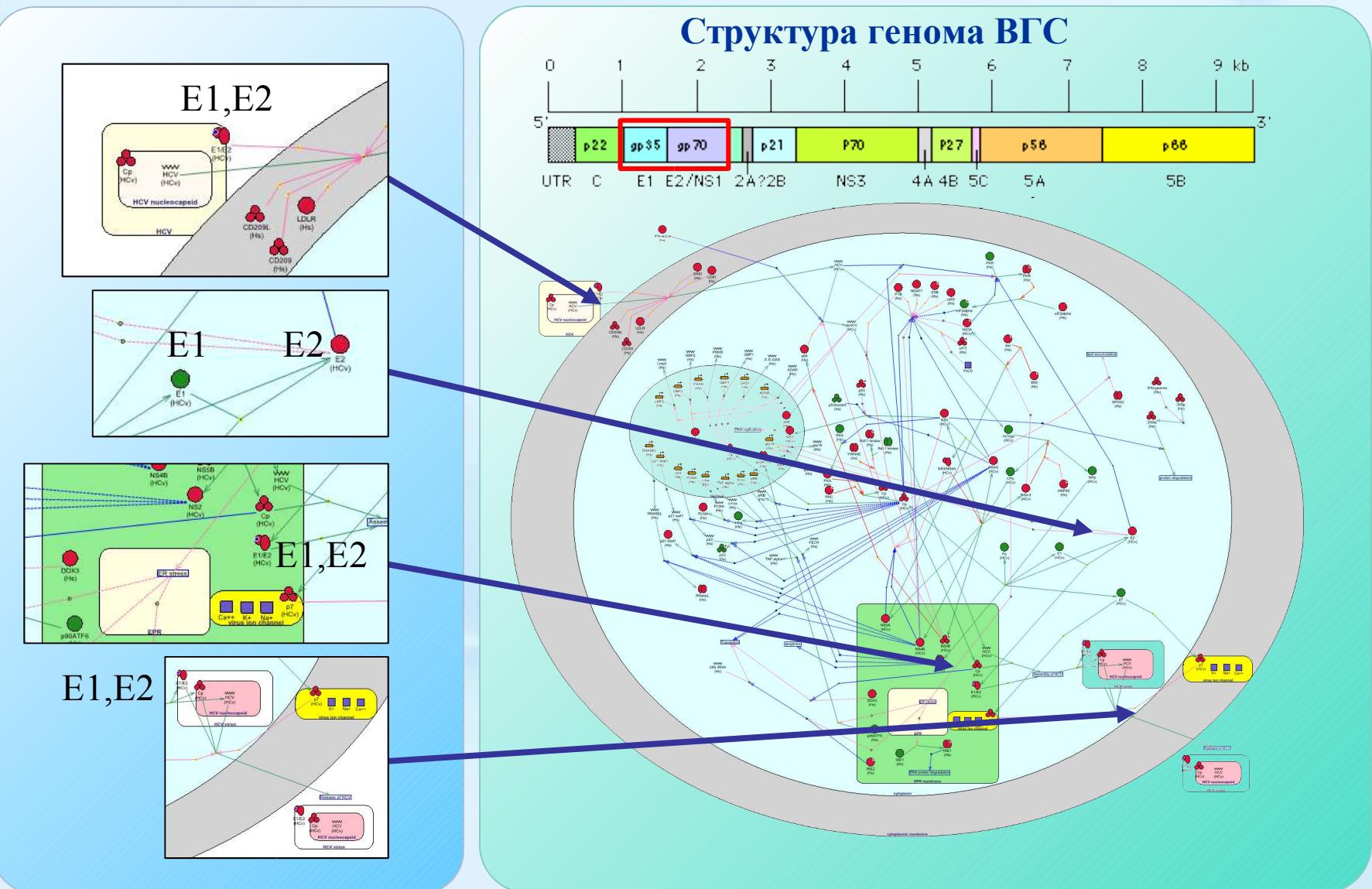
Вирус гепатита С относится к семейству Флавивирусов, род Гепацивирусов. Он попадает в организм путем, сходным с ВИЧ, и поражает печень. Хроническая форма инфекции длится 10-15 лет и вызывает цирроз печени, провоцирует рак, угнетает иммунную систему организма. В России число инфицированных оценивается в 2 млн. человек.

# Основные генотипы вируса гепатита С



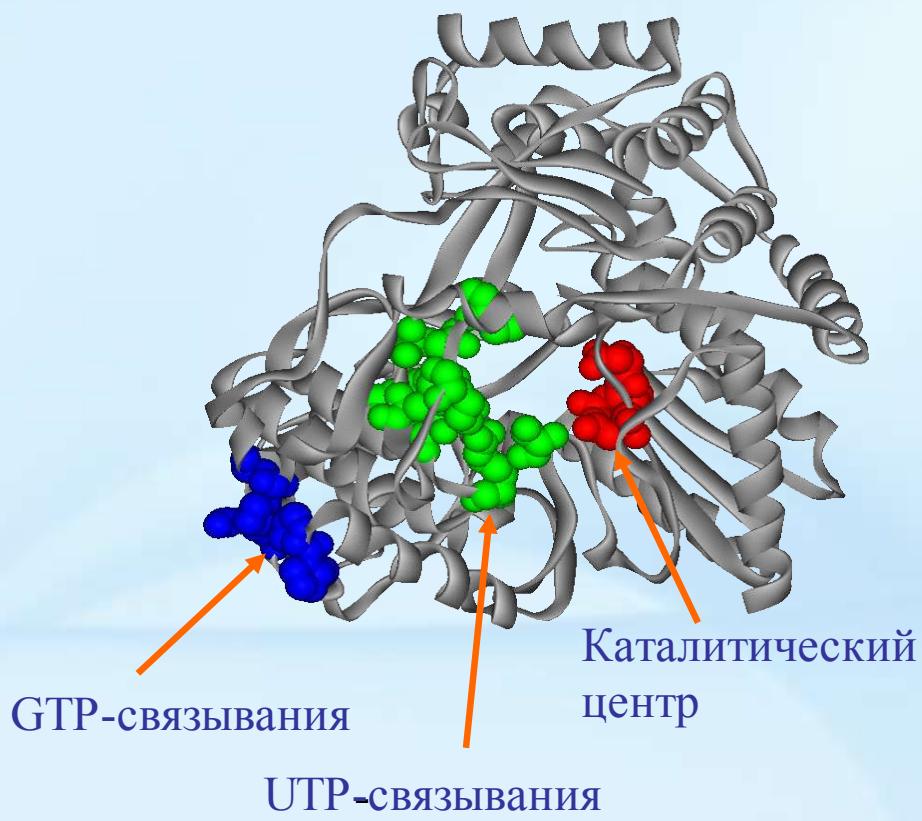
Существует 6 основных генотипов вируса гепатита С и несколько подтипов. Наиболее распространены 1b(40%), 1a (30%), 2a, 2b, 3a.

# Генная сеть жизненного цикла вируса гепатита С

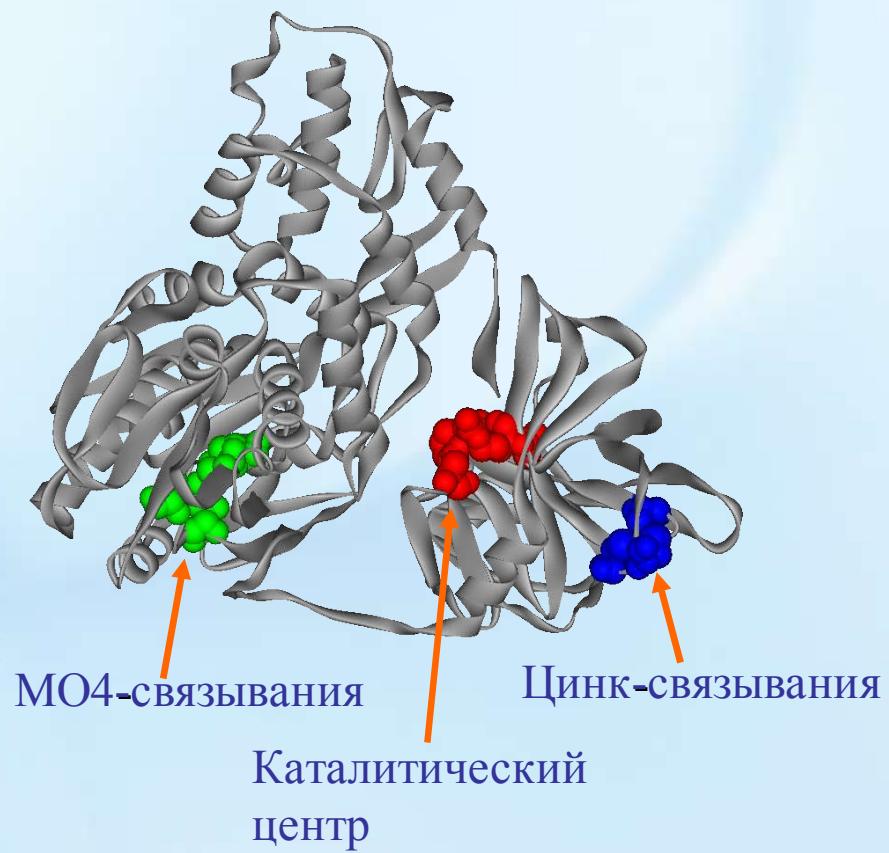


# Потенциальные мишени для действия противовирусных препаратов

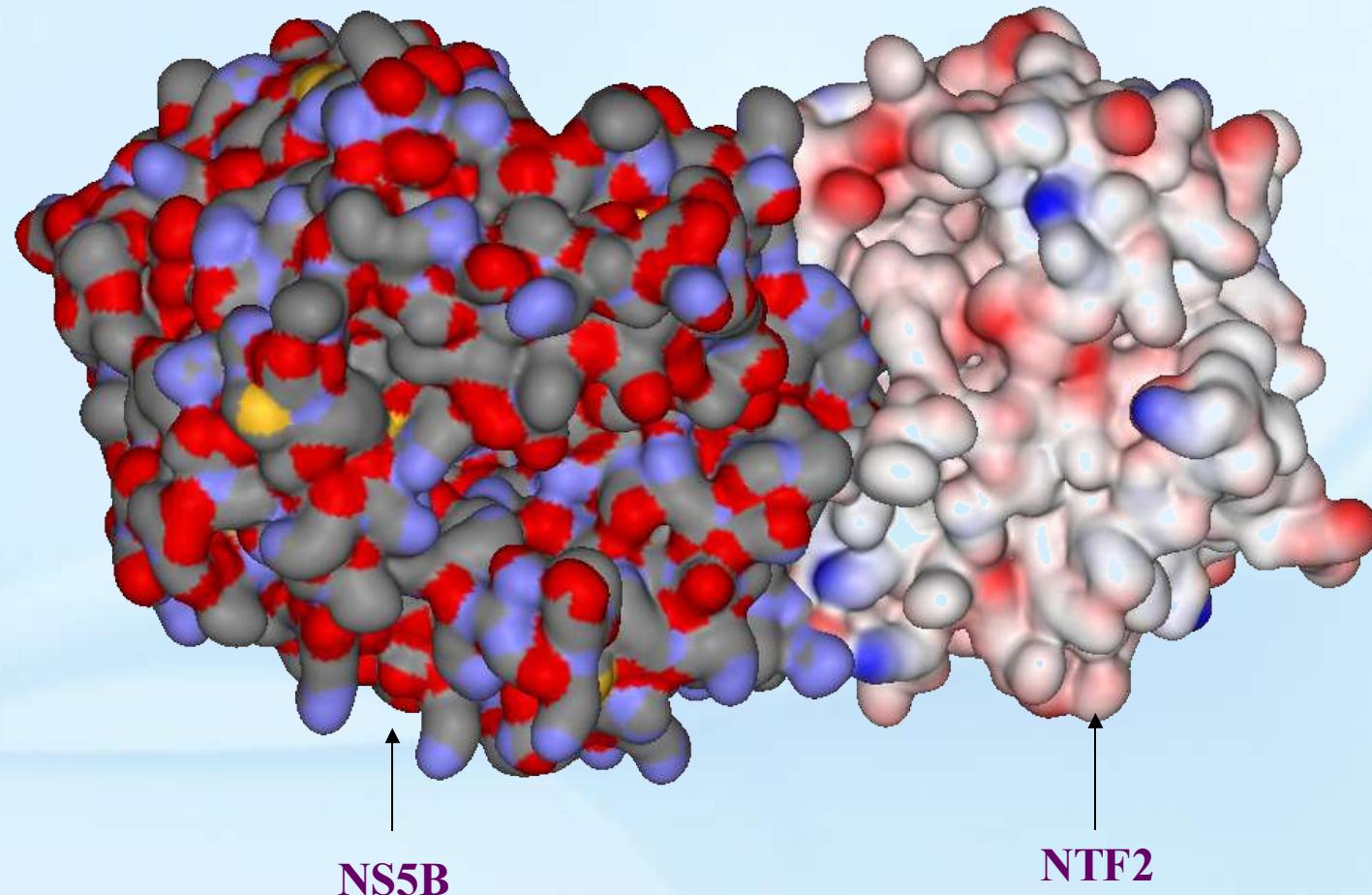
NS5B трансфераза ВГС



NS3 гидролаза ВГС



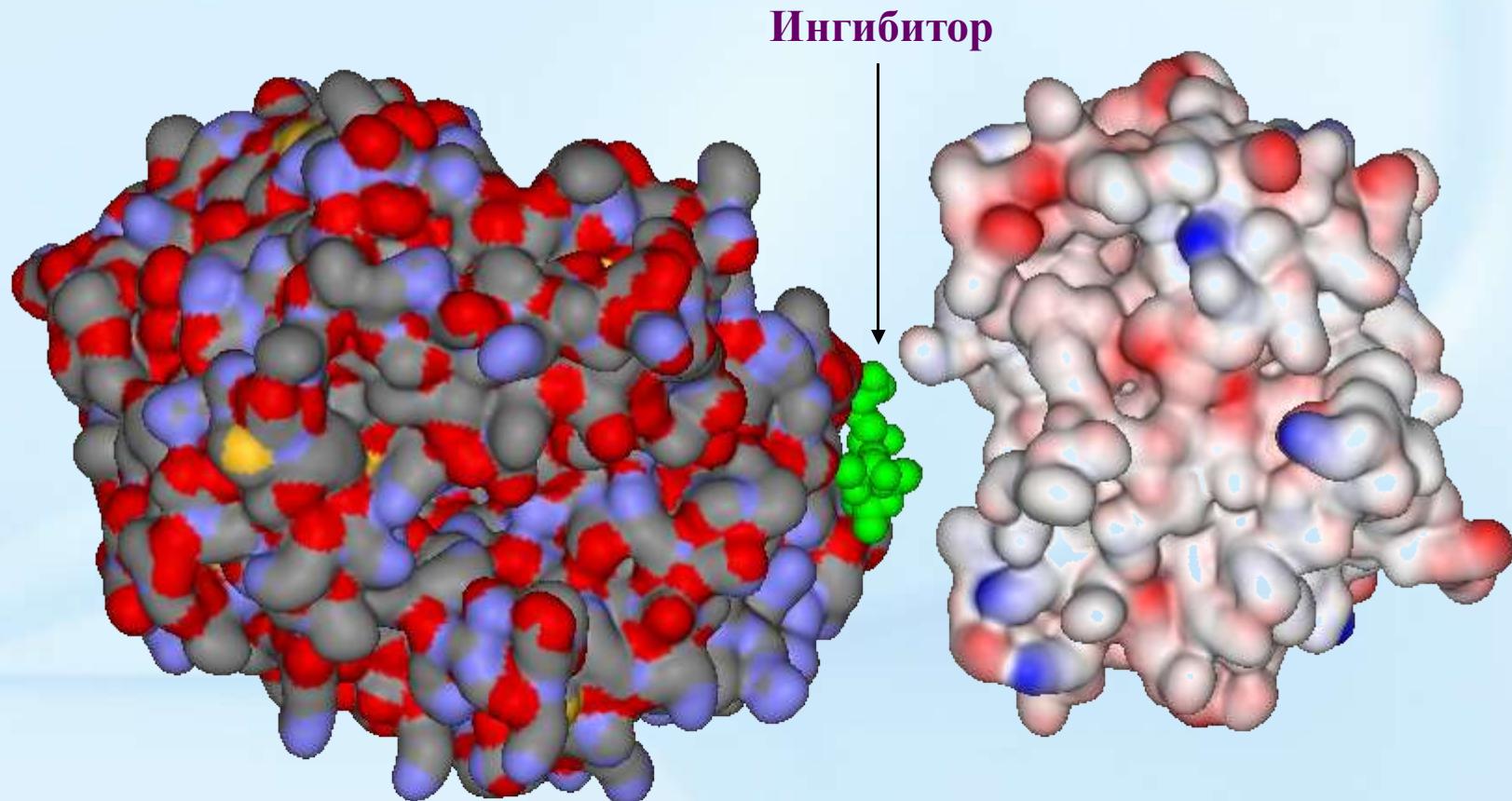
# Потенциальные мишени для действия противовирусных препаратов



Теоретически предсказанный комплекс, образуемый белком NS5B вируса гепатита С с белком человека NTF2, осуществляющим транспорт белков в ядро клетки.



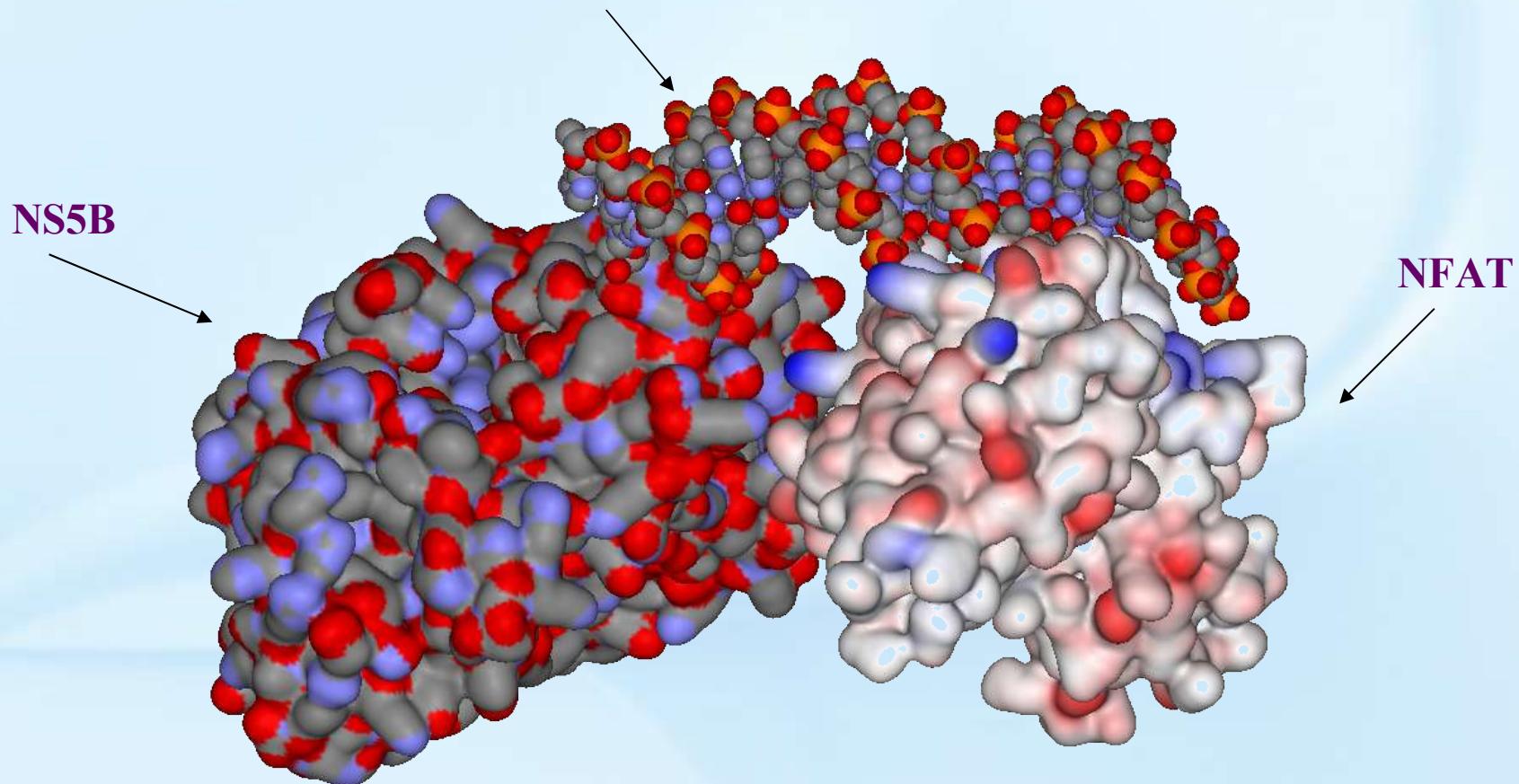
# Потенциальные мишени для действия противовирусных препаратов



Ингибитор образования комплекса NS5B-NTF2 будет препятствовать транспорту NS5B в ядро клетки.

# Потенциальные мишени для действия противовирусных препаратов

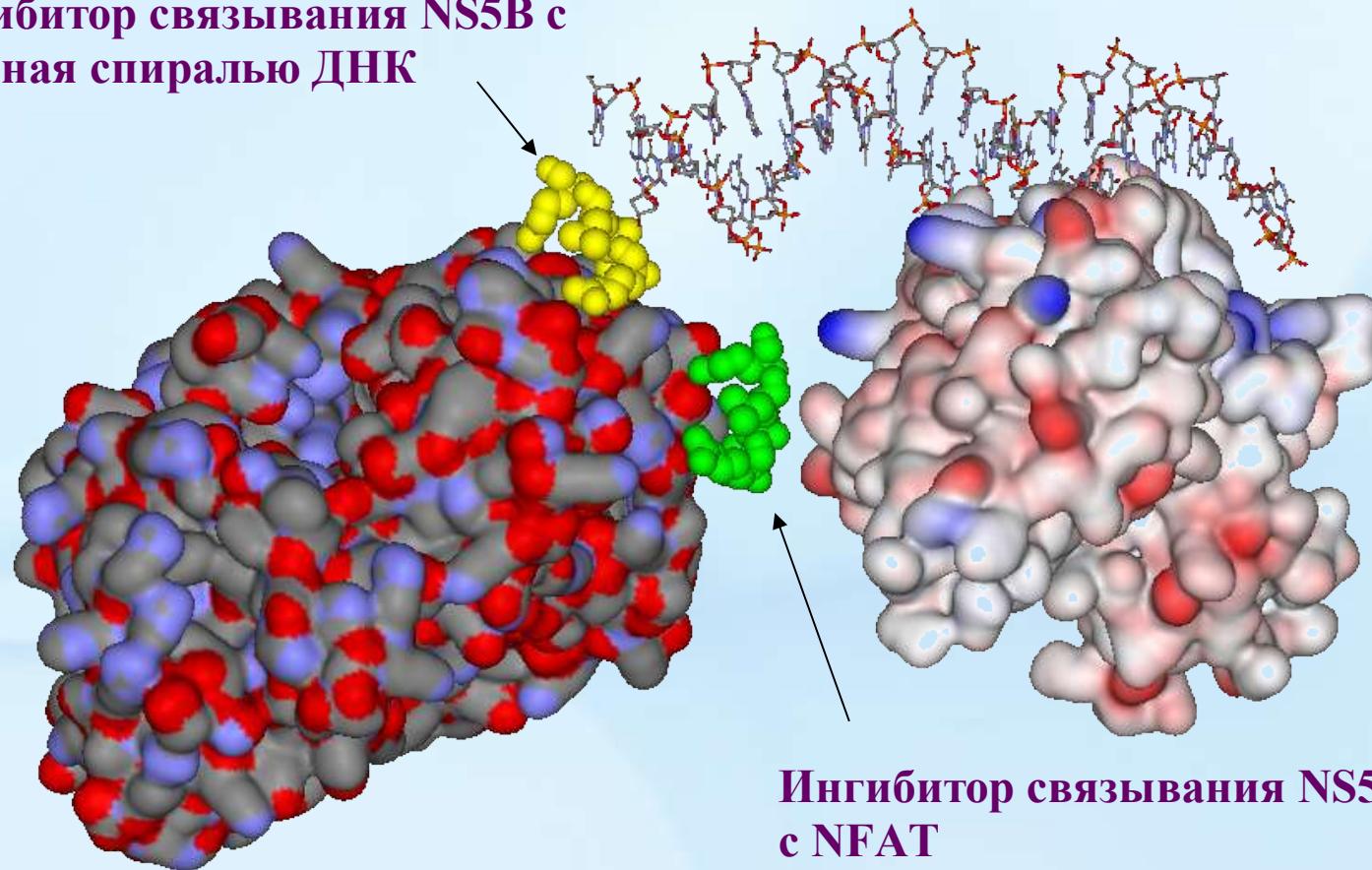
двойная спираль ДНК



Теоретически предсказанный комплекс, образуемый белком NS5B вируса гепатита С, с транскрипционным фактором человека NFAT и двойной спиралью ДНК.

# Потенциальные мишени для действия противовирусных препаратов

Ингибитор связывания NS5B с  
двойной спиралью ДНК

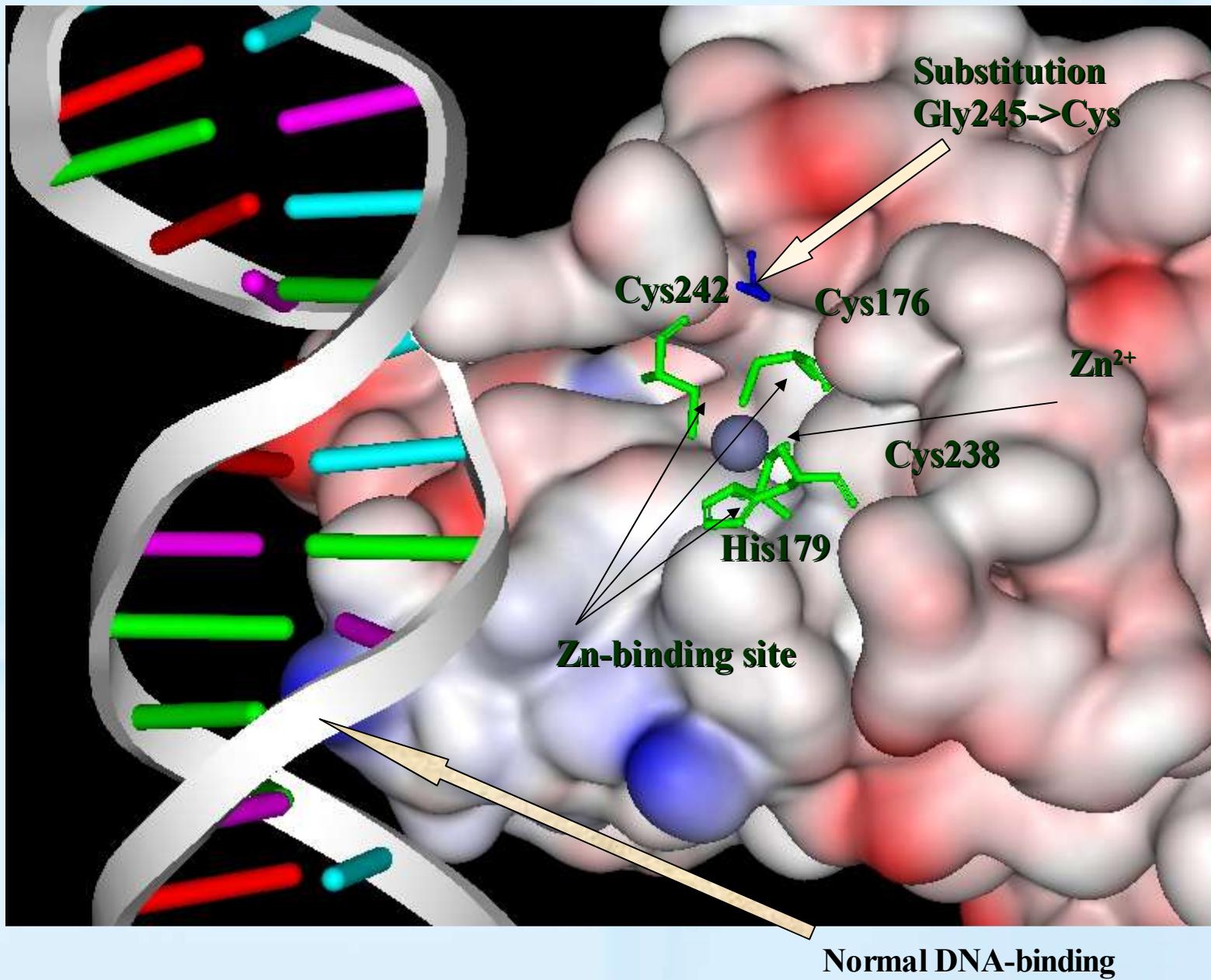


Ингибитор связывания NS5B  
с NFAT

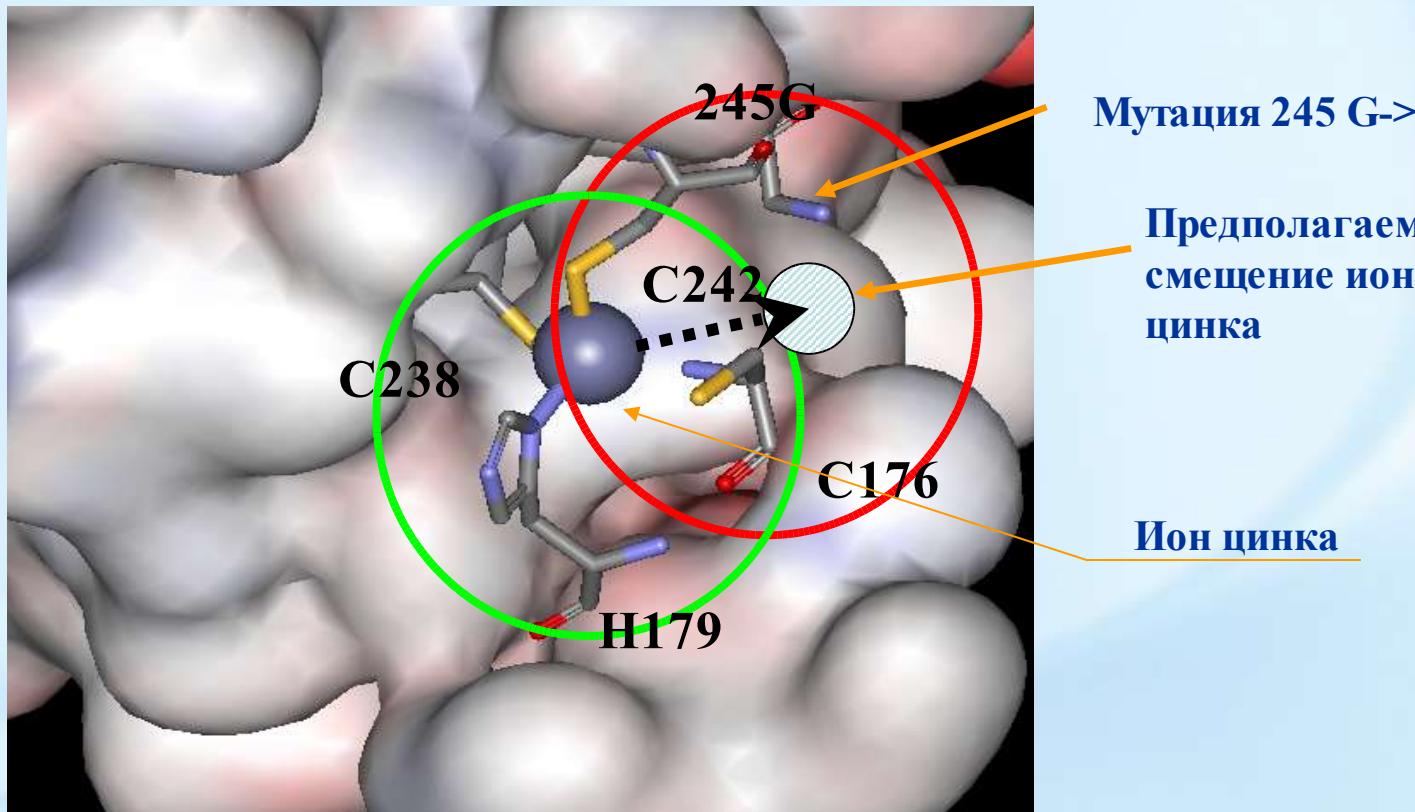
# **Молекулярные механизмы нарушения функции мутантных белков**



# Молекулярные механизмы нарушения функции мутантного белка p53, приводящие к развитию рака



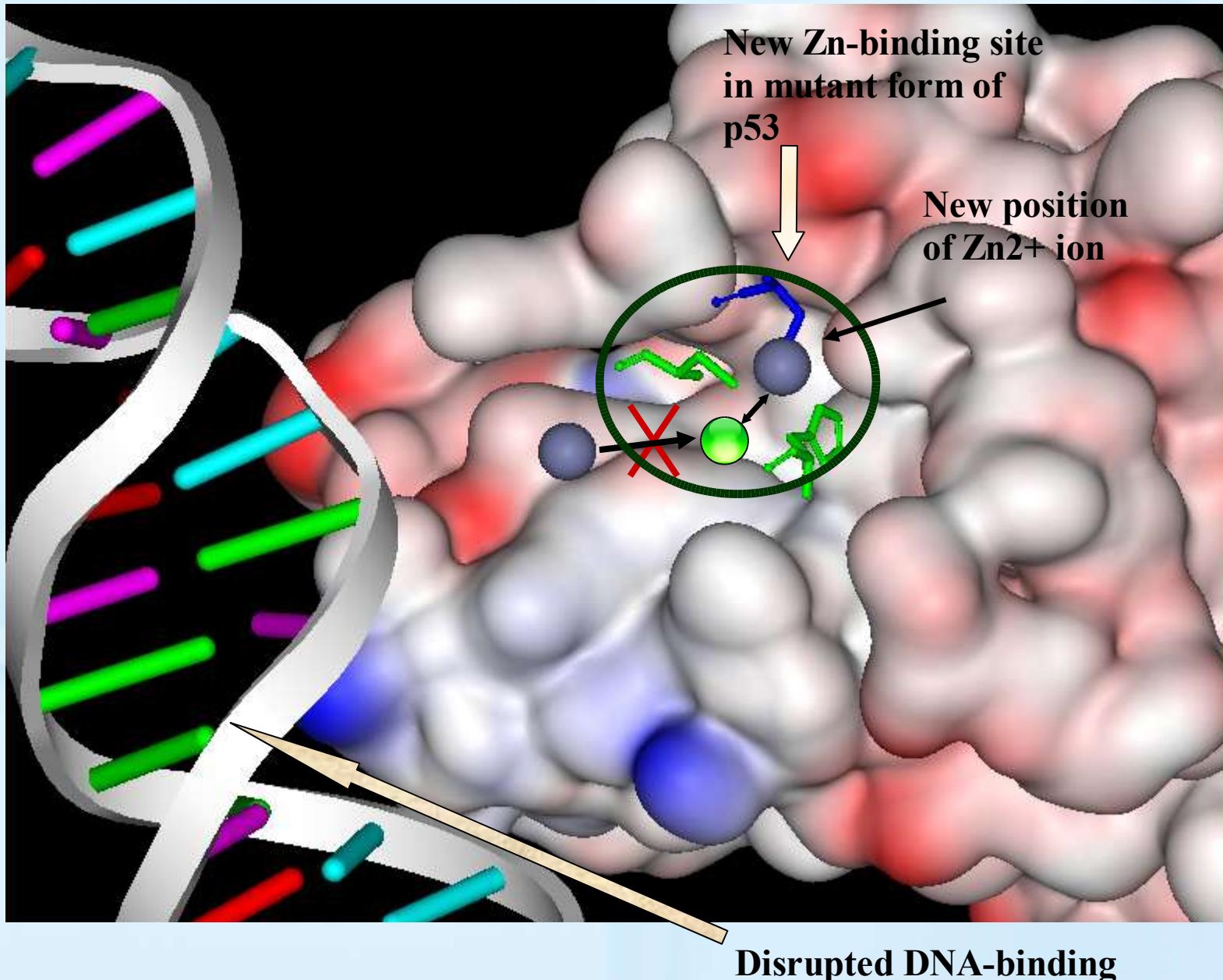
# Молекулярные механизмы нарушения функции мутантного белка p53, приводящие к развитию рака



- Зеленым обведены аминокислоты реального сайта связывания  $Zn^{2+}$
- Красным обведены аминокислоты потенциального сайта связывания  $Zn^{2+}$

Структурный анализ мутации Gly245→Cys в ДНК-связывающем домене белка p53, вызывающую наследственную предрасположенность к раку (синдром Ли-Фраумени).

# Молекулярные механизмы нарушения функции мутантного белка p53, приводящие к развитию рака



# **Молекулярная эволюция функциональных сайтов**



# **Новый подход к молекулярной эволюции функциональных сайтов - «Палеонтологические раскопки» функциональных сайтов в пространственных структурах белков**

**Белки содержат:**

- ✓ Активные функциональные сайты;
- ✓ Пост-сайты - следы ранее активных функциональных сайтов, разрушающиеся в результате фиксации мутаций;
- ✓ Пре-сайты – участки структуры, потенциально способные к образованию в результате фиксации мутаций новых сайтов, несущих функцию *de novo*.

Прототипы функциональных сайтов (пост-сайты и пре-сайты) могут быть обнаружены при анализе пространственной структуры белков;



# «Палеонтологические раскопки» функциональных сайтов в пространственных структурах белков

Активный сайт

D

Y

S

S

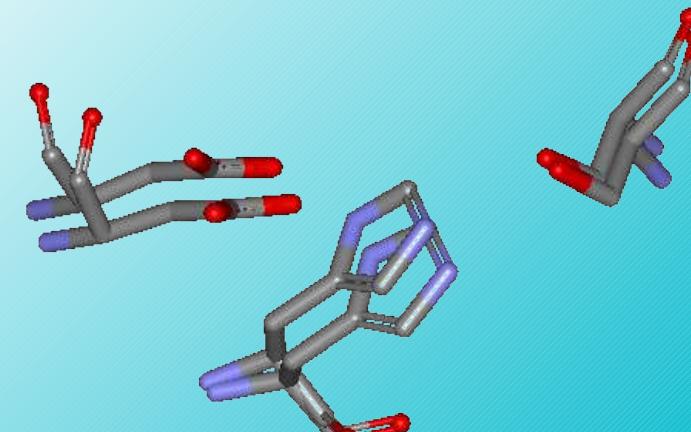
D

A

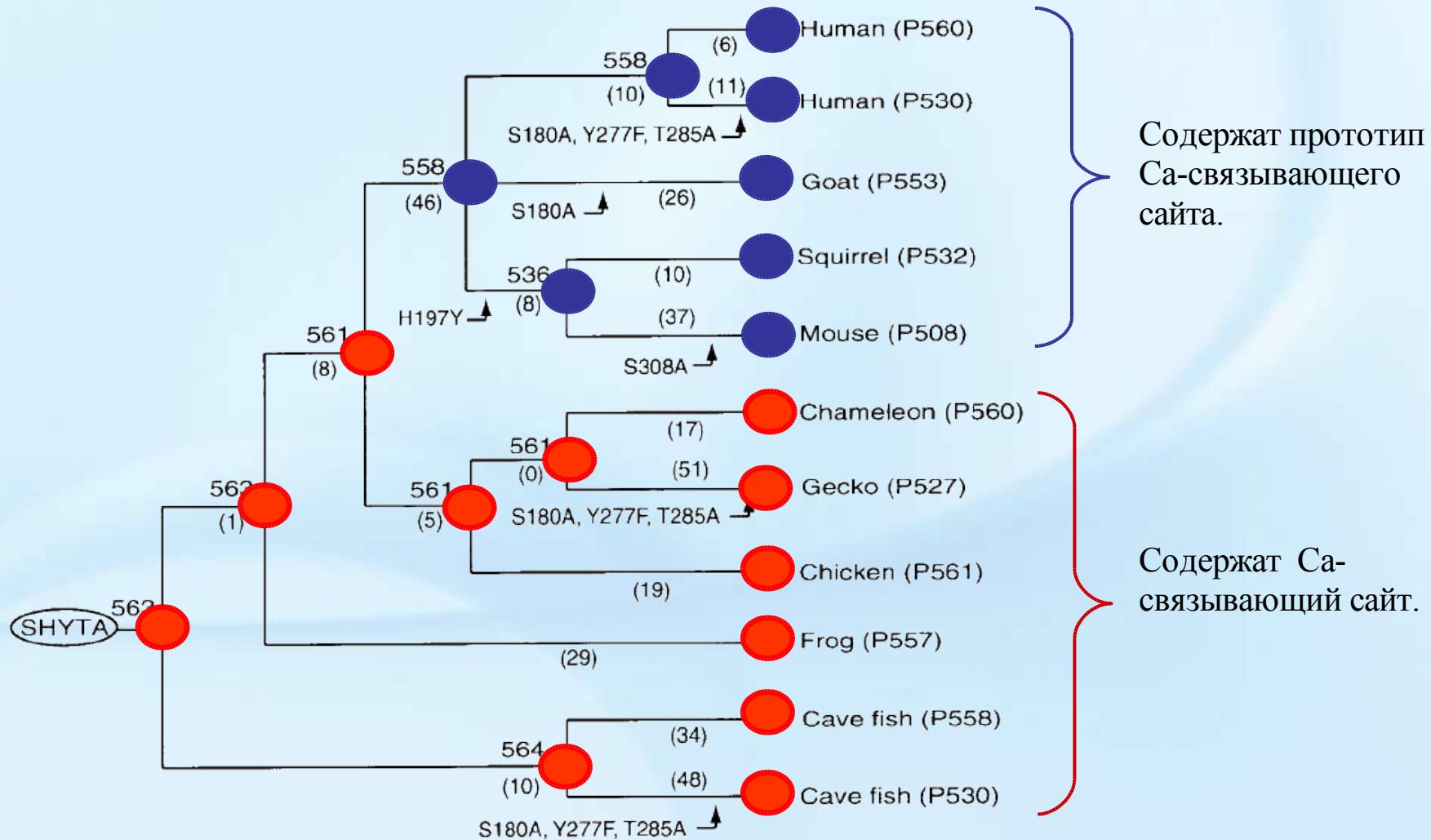
Прототип активного сайта

Мутация A->Y

После мутации A->Y оба сайта обладают одной функцией

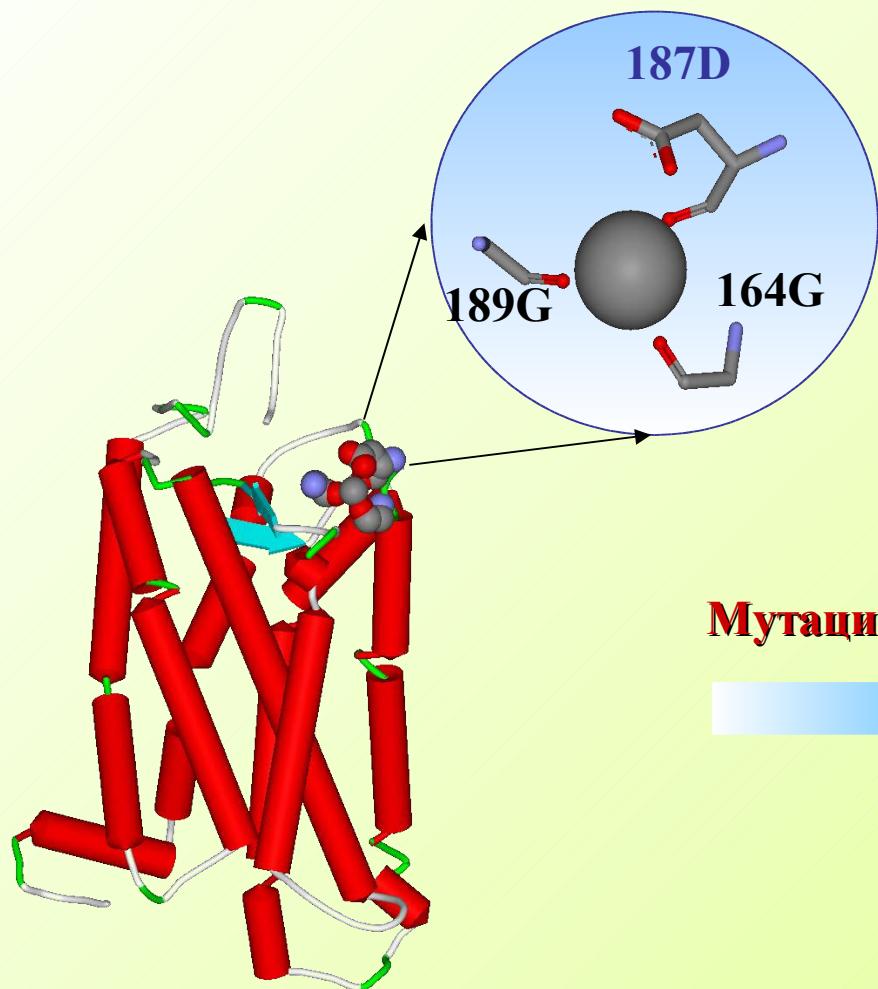


# Топология эволюционного дерева молекул пигмента, чувствительного к средним и длинным волнам для представителей позвоночных организмов



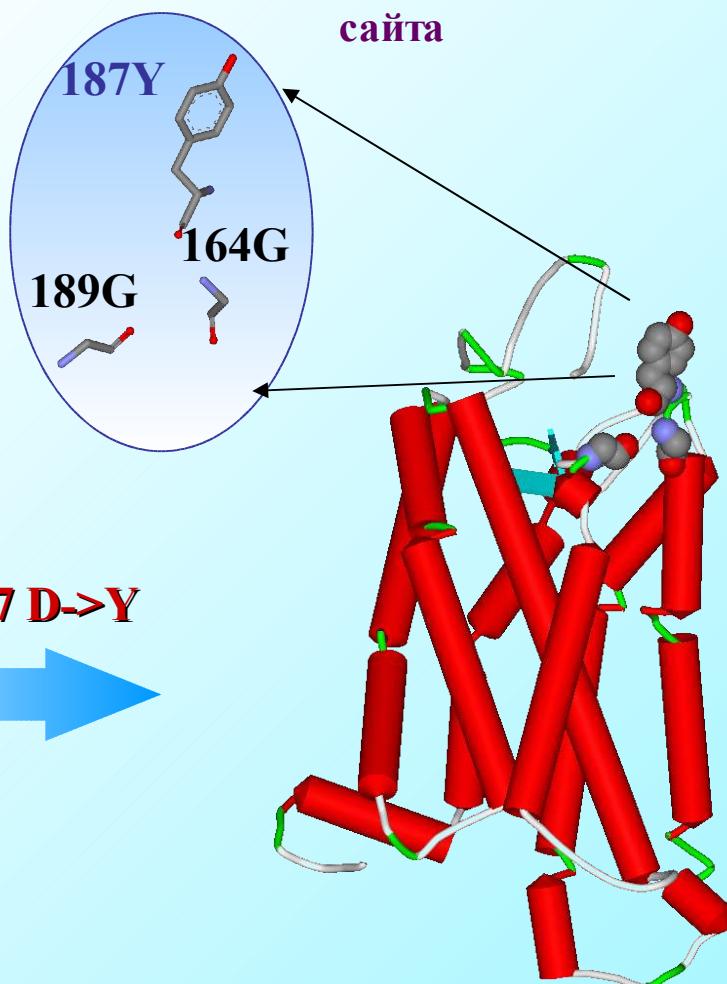
Дерево взято из работы Yokoyama and Radlwimmer, Genetics 158, 2001

## Са-связывающий сайт



Предковая молекула  
пигментов А, В, С, G, H, I

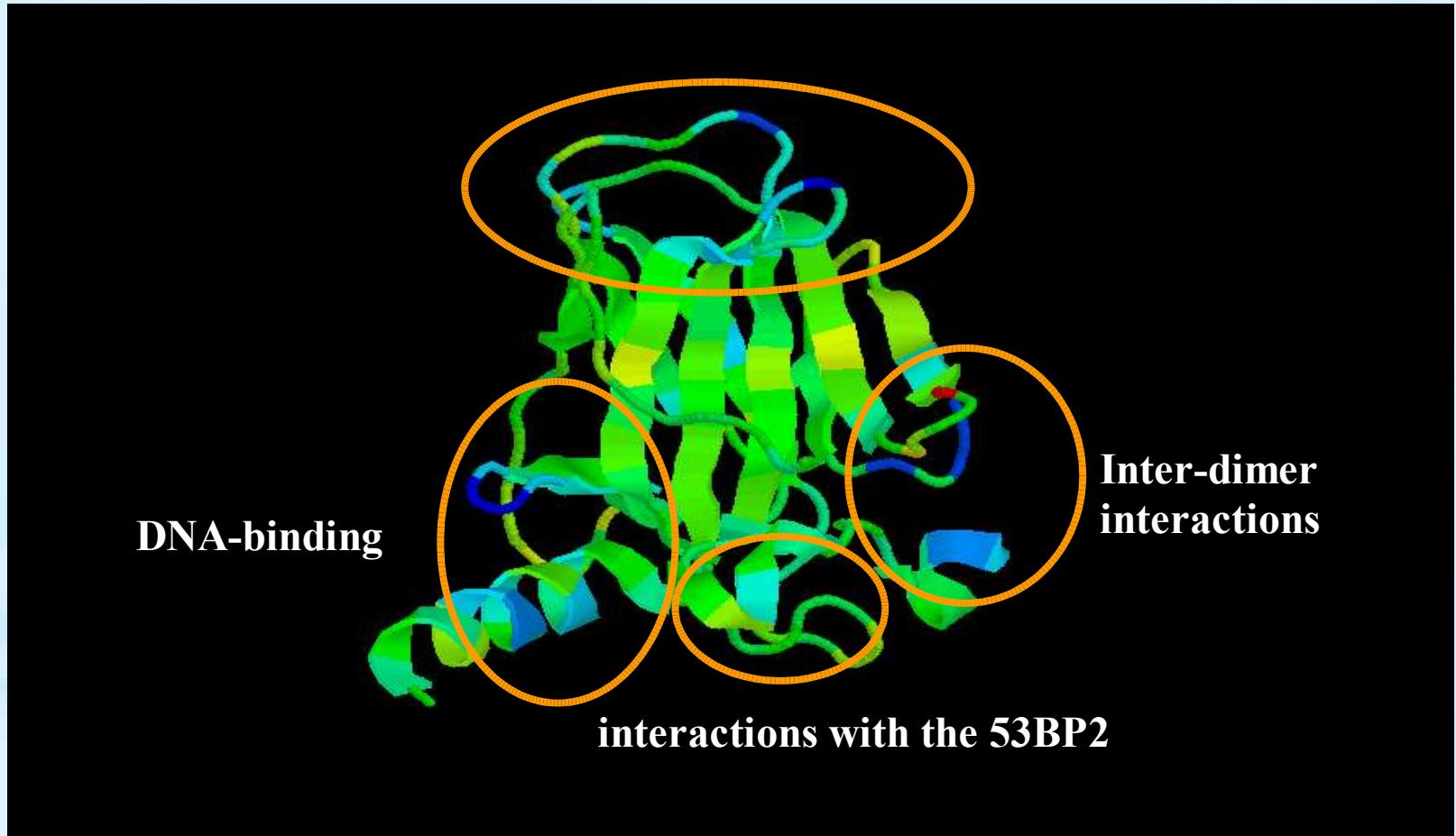
## Прототип Са-связывающего сайта



Пигмент человека (P560)

Мутация 187 D->Y

# Распределение прототипов сайтов в пространственной структуре ДНК-связывающего домена белка p53



Большое эволюционное  
расстояние

Малое эволюционное  
расстояние

## **Заключение**

**Создана компьютерная технология распознавания и анализа функциональных центров белков.**

- **Базы данных пространственных структур и физико-химических особенностей функциональных сайтов;**
- **Программы распознавания функциональных сайтов в пространственных структурах белков;**
- **Программы реконструкции молекулярных комплексов белков с лигандами;**
- **Программы количественного анализа взаимосвязи между структурой и активностью белков.**

**Показана высокая эффективность применения разработанных подходов для решения задач.**

- **Функциональная аннотация белков;**
- **Планирование мутаций направленно меняющих белковую активность;**
- **Поиск фармакологических мишеней;**
- **Изучения молекулярных механизмов нарушения функции белков.**

# **Основные публикации**

1. Ivanisenko VA, Pintus SS, Grigorovich DA, Kolchanov NA. PDBSite: a database of the 3D structure of protein functional sites. *Nucleic Acids Res.*, 2005, V33, D183–D187
2. Ivanisenko V.A., Pintus S.S., Grigorovich D.A., Kolchanov N.A. (2004) PDBSiteScan: a program for searching for active, binding and posttranslational modification sites in the 3D structures of proteins. *Nucleic Acids Res.*, 32, W549-W554
3. Ivanisenko, V.A., Pintus, S.S., Grigorovich, D.A., Ivanisenko, L.N., Debelov, V.A., Matsokin, A.M. PDBSITESCAN: a program searching for functional sites in protein 3D structures. In: Bioinformatics of genome regulation and structure. Ed. By N. Kolchanov and R. Hofestaedt, *Kluwer Academic Publishers*, Boston/Dordrecht/London, 2004, pp. 185-192.
4. Afonnikov DA, Kolchanov NA. (2004) CRASP: a program for analysis of coordinated substitutions in multiple alignments of protein sequences, *Nucleic Acids Res*, V.32 , W64-W68
5. Ivanisenko V.A., Pintus S.S., Krestyanova M.A., Demenkov P.S., Znobisheva E.K., Ivanov E.E., Grigorovich D.A. PDBSITE, PDBLIGAND and PDBSITESCAN: a computational workbench for the recognition of the structural and functional determinants in protein tertiary structures combined with protein draft docking. Proc. of the Fourth International Conference on Bioinformatics of Genome Regulation and Structure. 2004. V.1. P. 269-273.
6. Pintus S.S., Ivanisenko V.A. A molecular mechanism for the structure-functional alterations in mutant forms of human p53 protein. Proc. of the Fourth International Conference on Bioinformatics of Genome Regulation and Structure. 2004. V.1. P. 338-342.
7. Chelobanov B.P., Ivanisenko V.A., Kharkova M.V., Laktionov P.P., Rykova E.Yu., Vlassov V.V. Identification and analysis of cell surface nucleic acids-binding proteins. Proc. of the Fourth International Conference on Bioinformatics of Genome Regulation and Structure. 2004. V.1. P. 248-251.
8. Ivanisenko V.A., Grigorovich D.A. and Kolchanov N.A. PDBSITE: A Database on Protein Active Sites and Their Environment. Third International Conference on Bioinformatics of Genome Regulation and Structure. 2002. V.3. P. 145-148.
9. Ivanisenko V.A., Debelov V.A., Pintus S.S., Matsokin A.M., Nikolaev S.V., Grigorovich D.A., Kolchanov N.A. PDBSITEScan: A Tool For Search For The Best-Matching Superposition In The Database PDBSITE. Third International Conference on Bioinformatics of Genome Regulation and Structure. 2002. V3. P.149-152.
10. Grigorovich D.A., Ivanisenko V.A. ENPDB: A Retrieval System For The Pdb Databank. Third International Conference on Bioinformatics of Genome Regulation and Structure. 2002. V3. P.141-144.



# **Благодарности**

## **Сотрудничество с Институтами СО РАН и НГУ**

- Асп. Челобанов Б.П. (ИХБиФМ СО РАН)
- Проф. Чекмарев С.Ф. (ИТФ СО РАН)
- Проф. Дебелов В.А. (ИВМиМГ СО РАН)
- Научно-Образовательный Центр НГУ «Молекулярный дизайн и экологически безопасные технологии»

## **Работа поддержана грантами Сибирского Отделения:**

- Молодежный Лаврентьевский грант
- Интеграционный проект СО РАН № 142
- Программой фундаментальных исследований Президиума РАН "Происхождение и эволюции биосферы"

## **Работа поддержана грантами Российской Федерации:**

**РФФИ (03-07-96833, 02-07-90359)**

**Минпромнауки (№ 43.073.1.1.1501)**

## **Работа поддержана зарубежными грантами**

**CRDF –НОЦ (Y1-B-08-20 )**

**Национального Института Здоровья США № .2 R 01-HG -01539-04A 2**

