



***Введение в проблему
конструирования лекарственных
препаратов***

В.А. Иванисенко, Д.А. Афонников, Н.А. Колчанов



University of Heidelberg

BASF

Drug Research is



the Search for a Needle in a Haystack



Имеется 10^{180} возможных биологически активных веществ, 10^{18} вероятных лекарственных препаратов, 10^7 известных соединений, 10^6 коммерчески доступных соединений, 10^6 соединений в базах данных фирм, 10^4 соединений в базах данных по лекарствам, 10^3 коммерческих лекарств и 10^2 коммерчески выгодных лекарств.



Время	Материалы	Система тестирования
древние времена	растения, яды, минералы	человек
1806	Morphine (морфий)	
1850	Химические	
1890	Синтетические, красители	
1920		Животные, изолированные органы
1970		Ферменты, мембраны
1990		Комбинаторные библиотеки
2000	Сфокусированные библиотеки (focused libraries)	Ultra High Throughput Screening, виртуальный скрининг



Важные результаты в разработке лекарств за период 1806-1981

1806	Morphine	Hypnotic agent
1875	Salicylic acid	Antiinflammatory agent
1884	Cocaine	Stimulant, local anesthetic agent
1888	Phenacetin	Analgesic and antipyretic agent
1899	Acetylsalicylic acid	Analgesic and antipyretic agent
1903	Barbiturates	Sedatives
1909	Arsphenamine	Antisymphilitic agent
1921	Procaine	Local anesthetic agent
1922	Insulin	Antidiabetic agent
1928	Estrone	Female sex hormone
1928	Penicillin	Antibiotic agent
1935	Sulphachrysoidine	Bacteriostatic agent
1944	Streptomycin	Antibiotic agent
1945	Chloroquine	Antimalarial agent
1952	Chlorpromazine	Neuroleptic agent
1956	Tolbutamide	Oral antidiabetic agent
1960	Chlordiazepoxide	Tranquillizer
1962	Verapamil	Calcium channel blocker
1963	Propranolol	Antihypertensive agent (beta-blocker)
1964	Furosemide	Diuretic agent
1971	L-Dopa	Anti-Parkinson agent
1975	Nifedipine	Calcium channel blocker
1976	Cimetidine	Anti-ulcus agent (H2 blocker)
1981	Captopril	Antihypertensive agent (ACE inhibitor)
1981	Ranitidine	Anti-ulcus agent (H2 blocker)



Важные результаты в разработке лекарств за период 1983-2001

1983	Cyclosporin A	Immunosuppressant
1984	Enalapril	Antihypertensive agent (ACE inhibitor)
1985	Mefloquine	Antimalaria agent
1986	Fluoxetine	Antidepressant (5-HT transporter)
1987	Artemisinin	Antimalaria agent
1987	Lovastatin	Cholesterol biosynthesis inhibitor
1988	Omeprazole	Anti-ulcus agent (H/K-ATPase inhibitor)
1990	Ondansetron	Antiemetic agent (5-HT ₃ blocker)
1991	Sumatriptan	Anti-migraine agent (5-HT ₁ blocker)
1993	Risperidon	Antipsychotic agent (D ₂ /5-HT ₂ blocker)
1994	Famciclovir	Anti-herpes (DNA polymerase inhibitor)
1995	Losartan	Antihypertensive agent (A II antagonist)
1995	Dorzolamide	Glaucoma (Carboanhydrase inhibitor)
1996	Meloxicam	Anti-arthritis agent (COX 2 inhibitor)
1996	Nevirapin	HIV reverse transcriptase inhibitor
1996	Indinavir, Ritonavir, Saquinavir	HIV protease inhibitors
1997	Nelfinavir	HIV protease inhibitor
1997	Finasteride	Hair loss
1997	Sibutramine	Adipositas (uptake blocker)
1998	Orlistat	Adipositas (lipase inhibitor)
1998	Sildenafil	Erectile dysfunction (PDE inhibition)
1999	Celecoxib, Rofecoxib	Anti-arthritis agents (COX-2 inhibitors)
1999	Amprenavir	HIV protease inhibitor
1999	Zanamivir, Oseltamivir	Influenza (neuraminidase inhibitors)
2001	Fondaparinux	Thrombosis (synthetic LMWH)
2001	Imatinib	Leukemia (specific abl-TK inhibitor)



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



Основные шаги рационального дизайна лекарств

Биологическая концепция

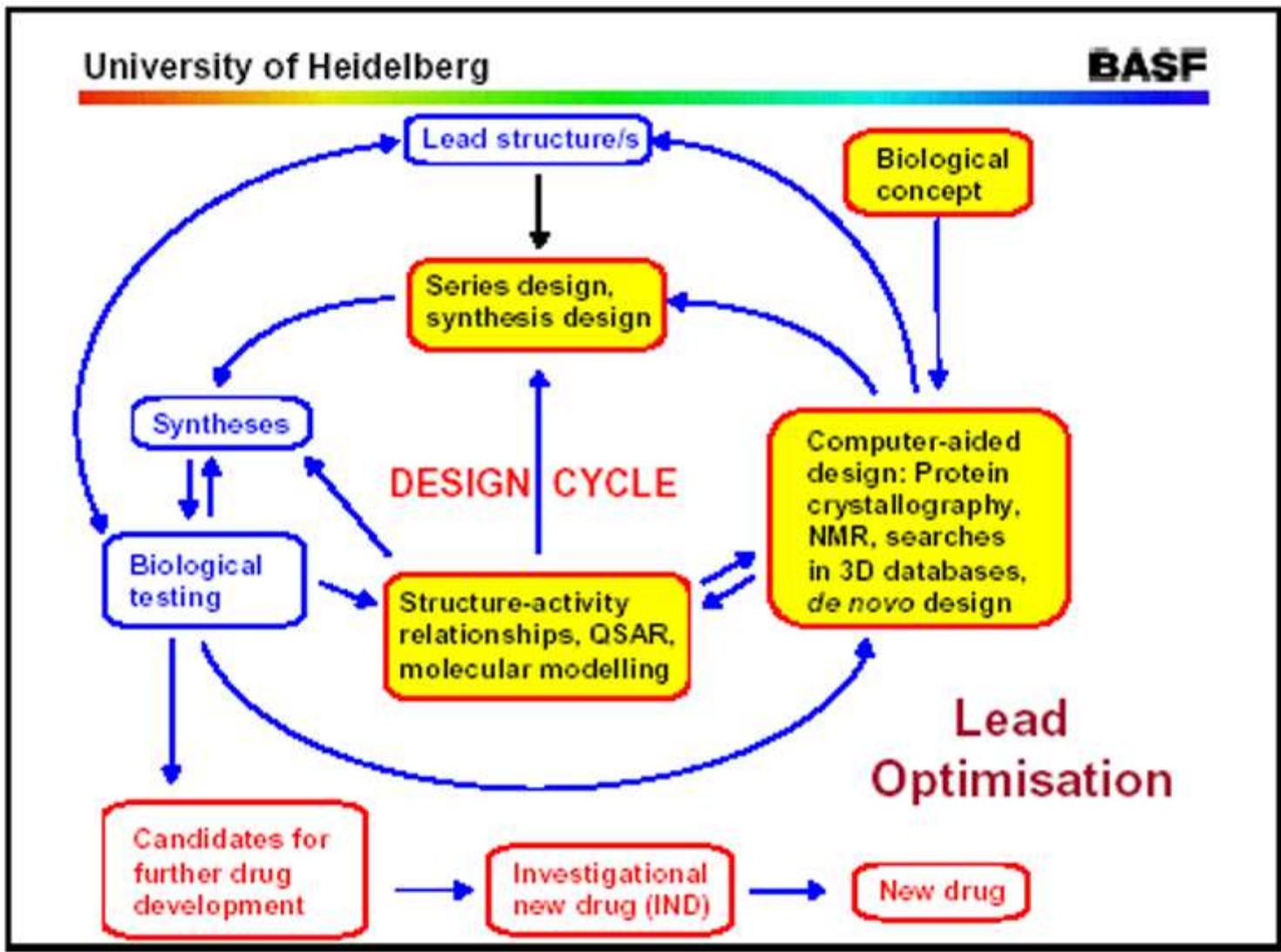
Computer-aided design: белковая кристаллография, ЯМР, поиск по базам третичных структур, de novo design

Анализ взаимосвязи структура-активность, QSAR, молекулярное моделирование



University of Heidelberg

BASF





Биологическая концепция

ГЕНОМНЫЕ ПРОЕКТЫ: ЧЕЛОВЕК

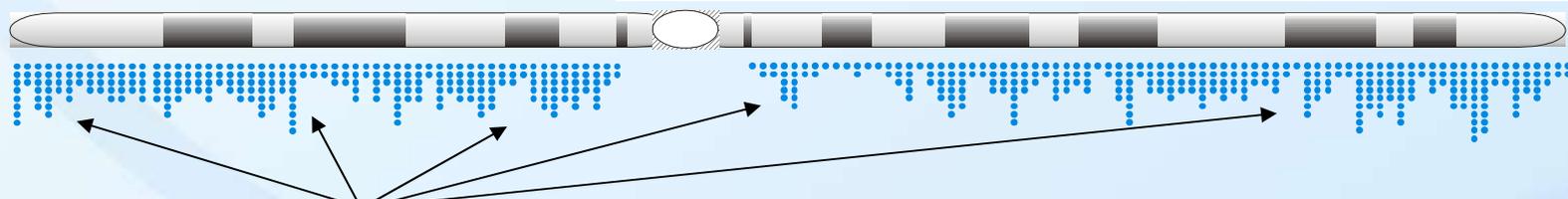
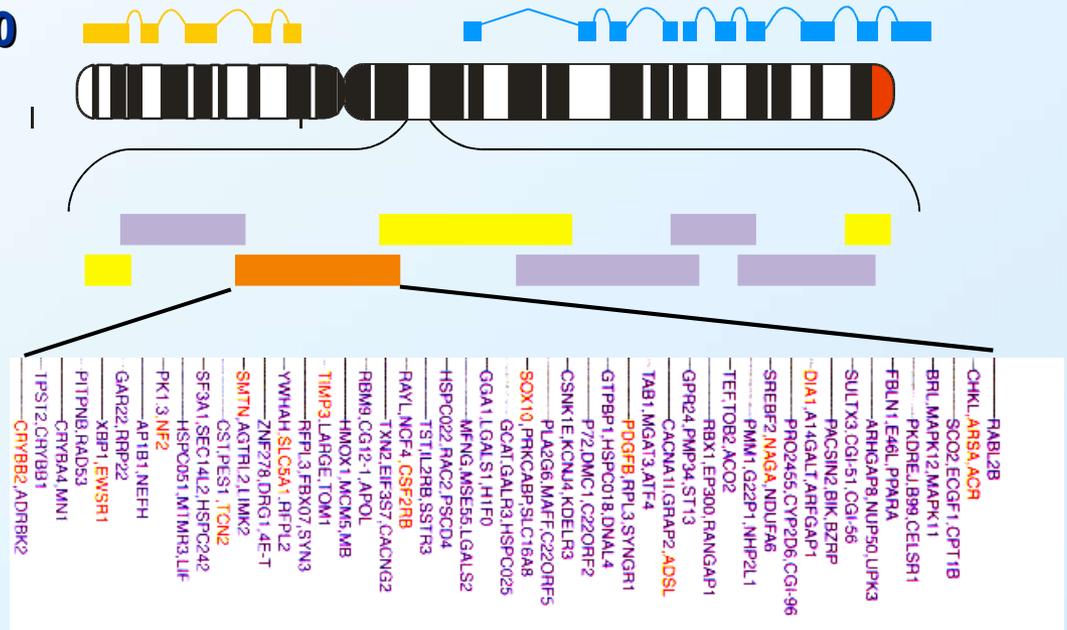
Характеристики генома человека:

размер ~ 3.3 миллиарда п.о. **ATGCCCGGGTTTAATGCGTCAGTGAСТGCACA...**

количество генов > 35000
цитогенетическая карта

физическая карта:

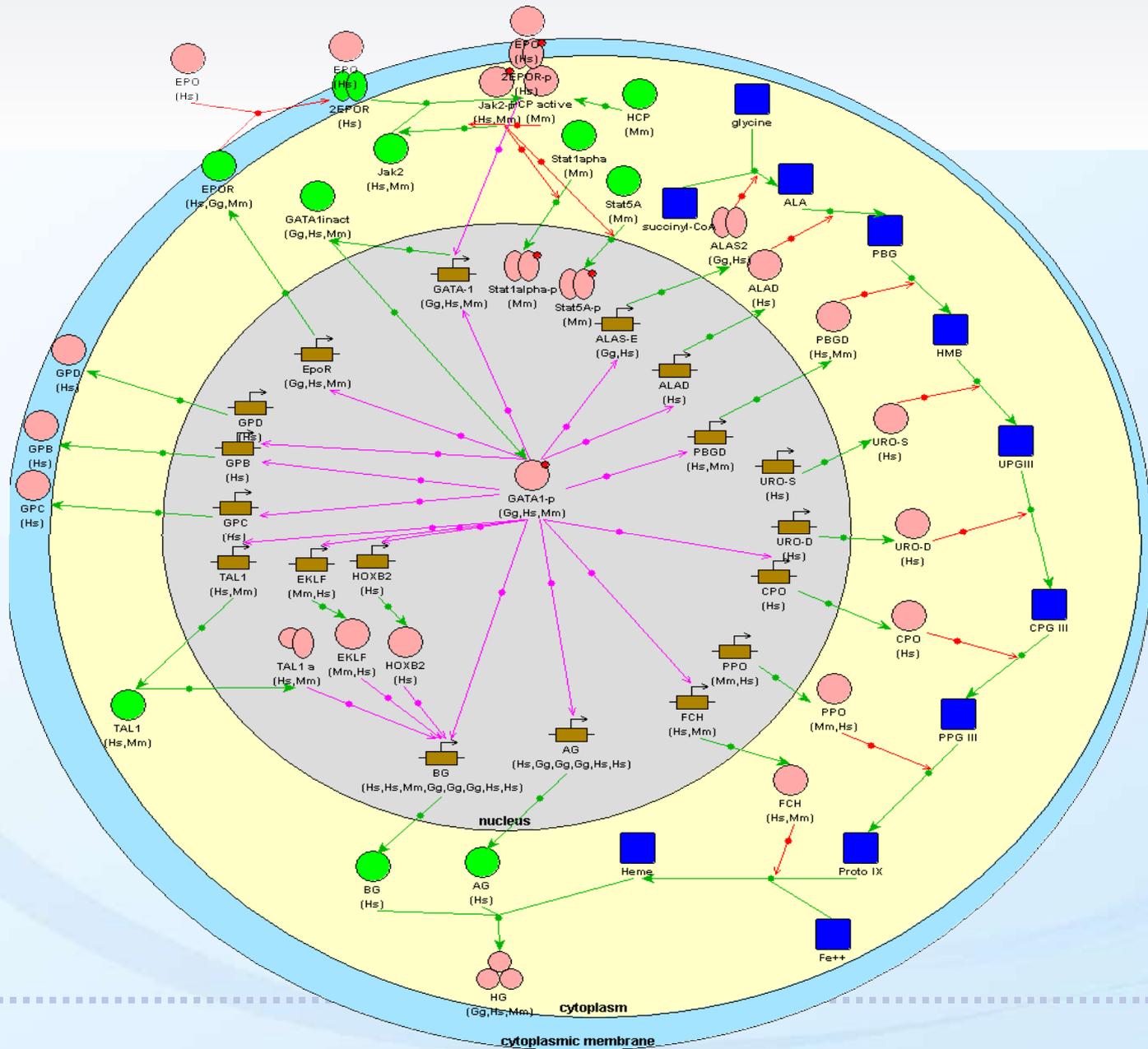
локализация генов



генетические полиморфизмы: более 2.5 миллиона вариантов!!!!

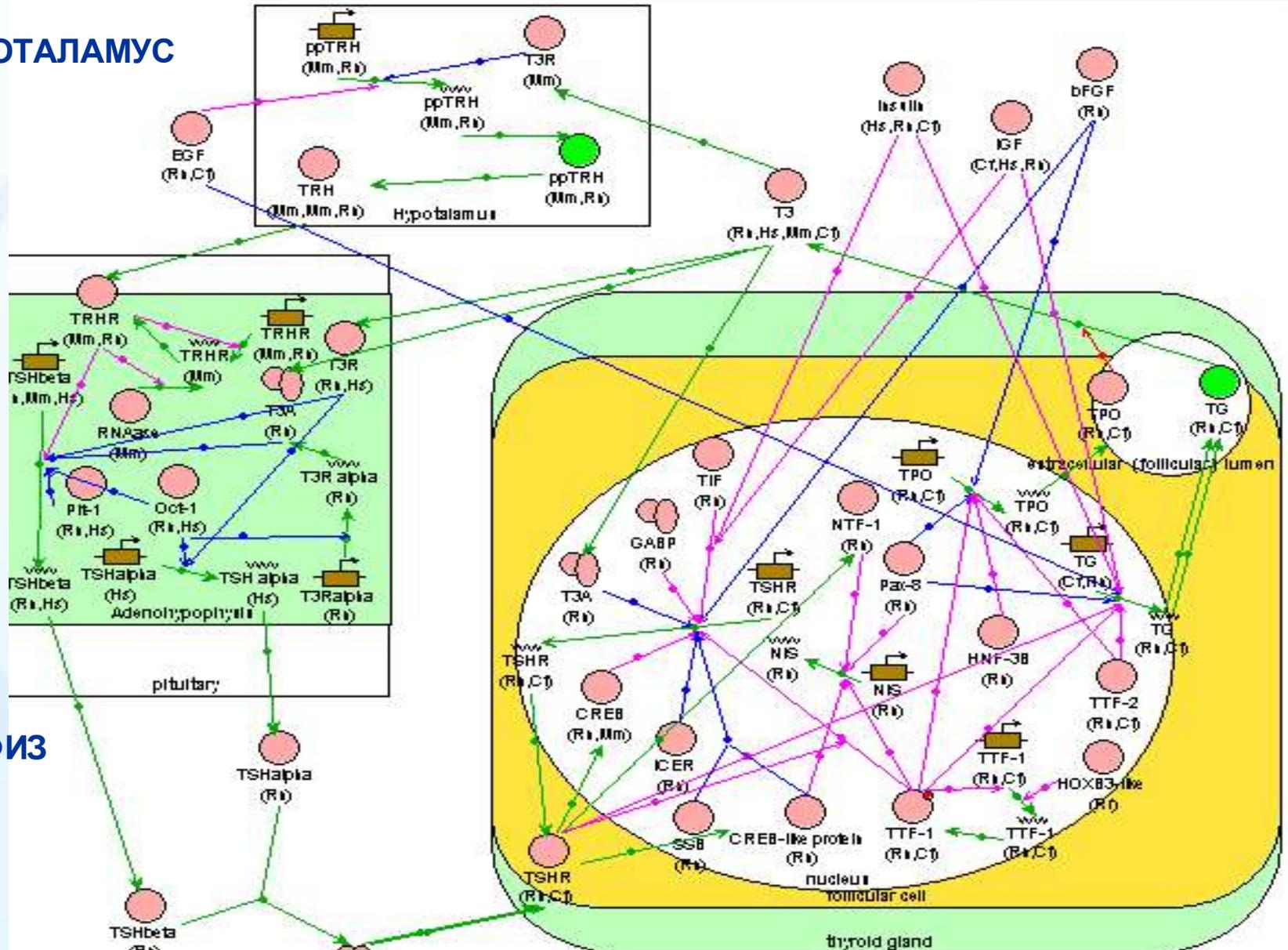


ФРАГМЕНТ ГЕННОЙ СЕТИ СОЗРЕВАНИЯ И ДИФФЕРЕНЦИРОВКИ ЭРИТРОЦИТА



ФРАГМЕНТЫ ГЕННОЙ СЕТИ РЕГУЛЯЦИИ ФУНКЦИИ ЩИТОВИДНОЙ ЖЕЛЕЗЫ

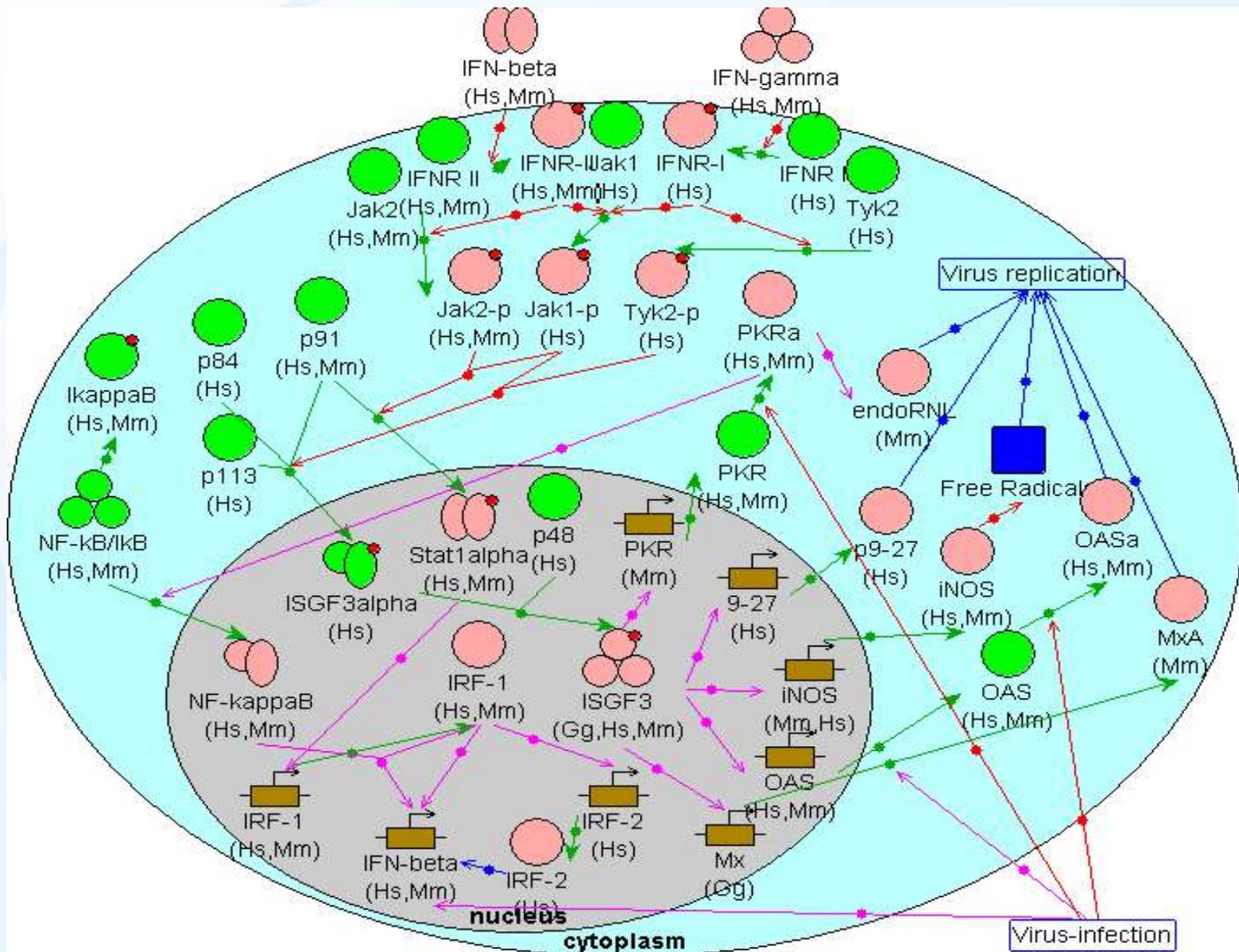
ГИПОТАЛАМУС



ГИПОФИЗ

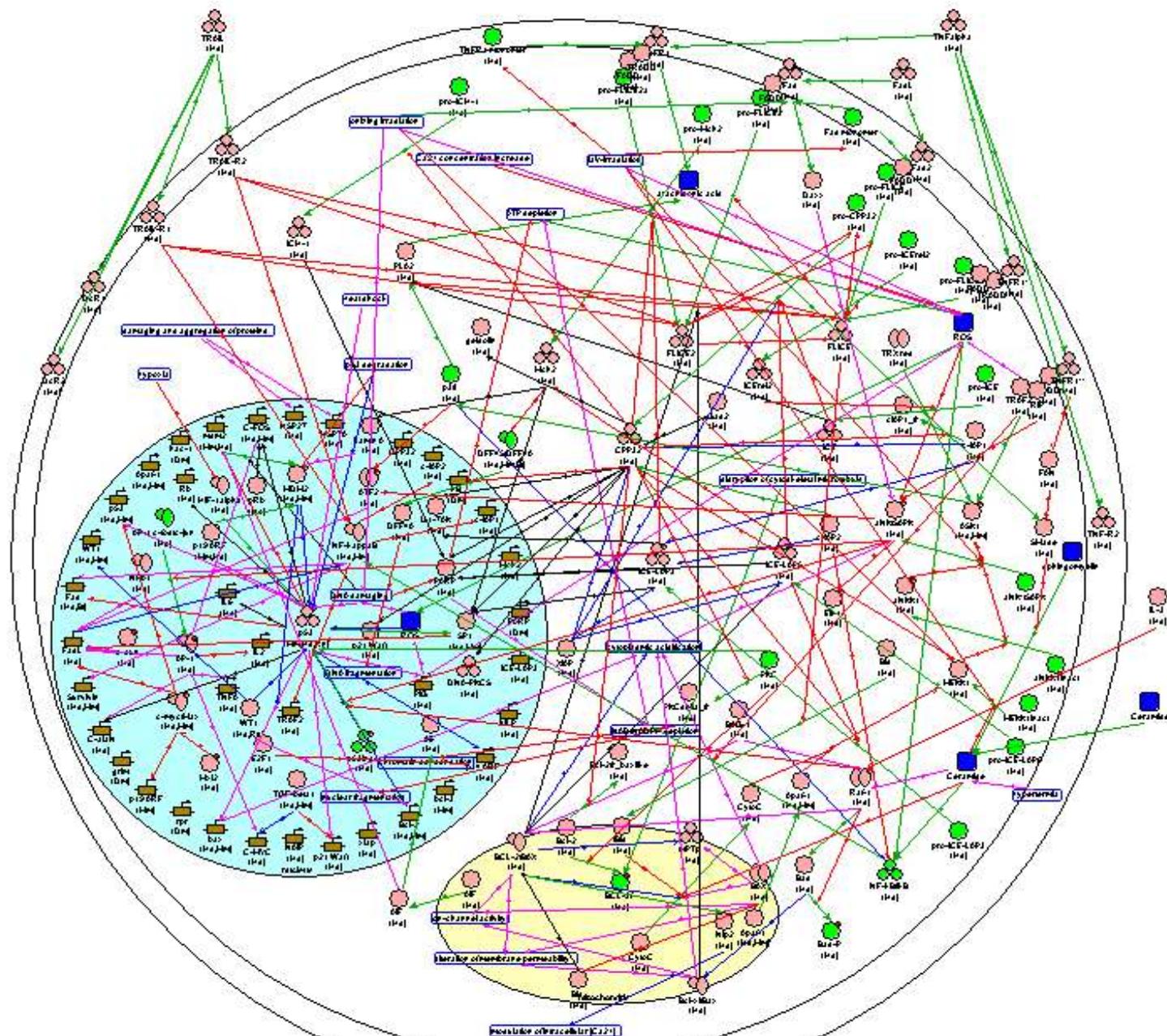
ЩИТОВИДНАЯ ЖЕЛЕЗА

ФРАГМЕНТ ГЕННОЙ СЕТИ АНТИВИРУСНОГО ОТВЕТА НА ОСНОВЕ ИНТЕРФЕРОНОВЫХ ГЕНОВ



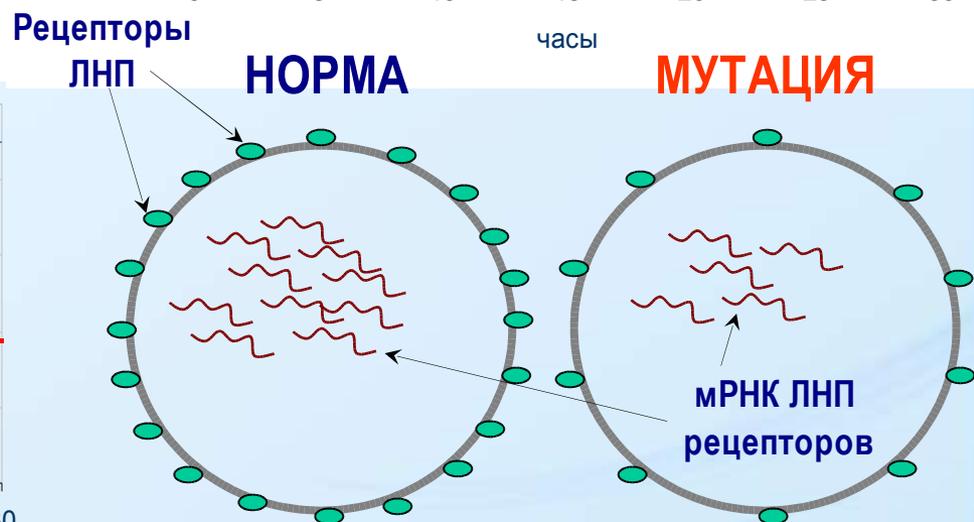
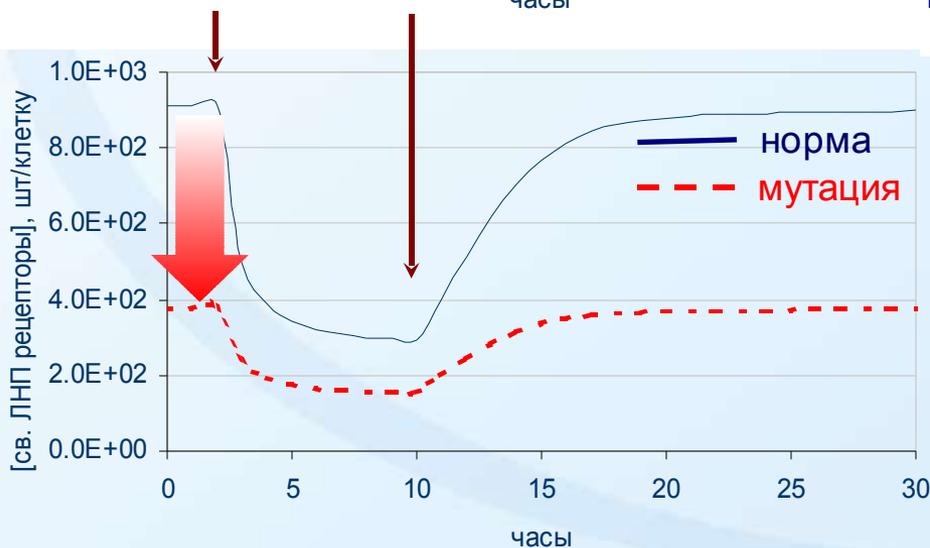
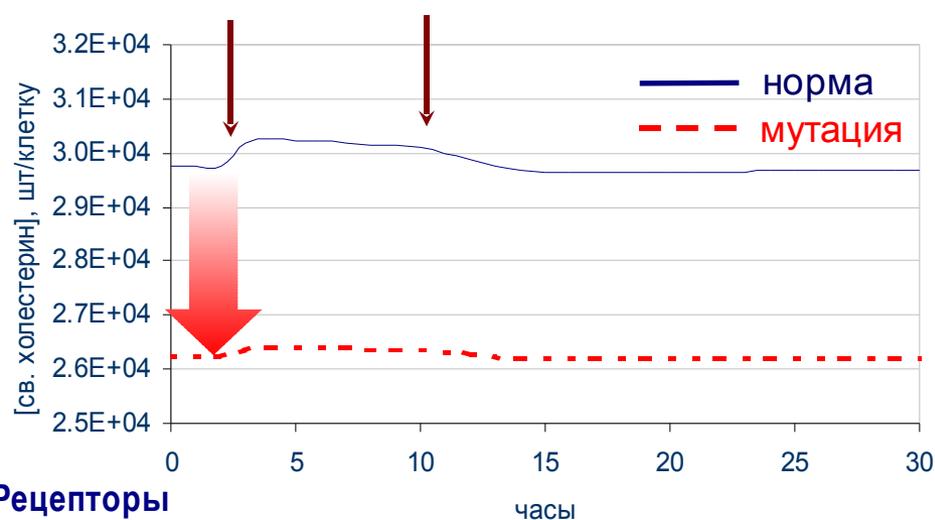
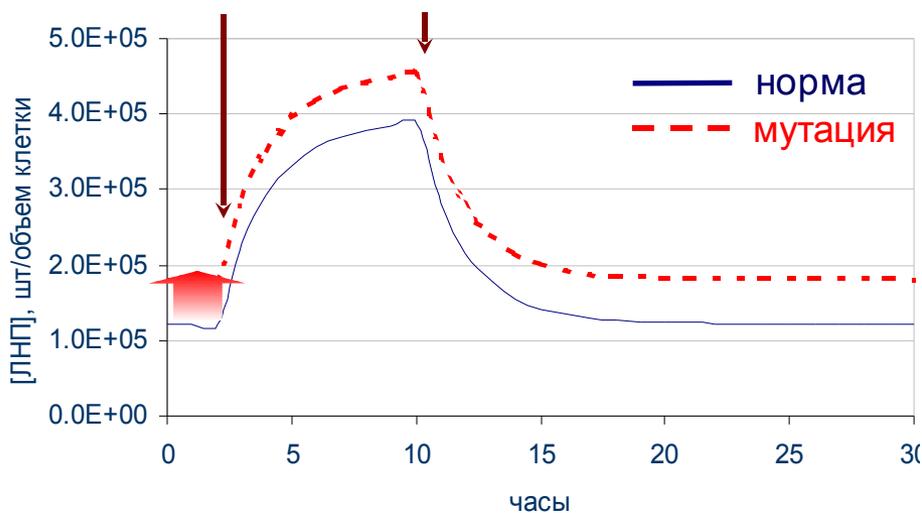


ФРАГМЕНТ ГЕННОЙ СЕТИ, КОНТРОЛИРУЮЩЕЙ ПРОЦЕСС КЛЕТОЧНОЙ СМЕРТИ - АПОПТОЗ



ПОИСК ЛИМИТИРУЮЩИХ ЗВЕНЬЕВ: АНАЛИЗ ВЛИЯНИЯ МУТАЦИЙ НА ДИНАМИКУ ГЕННОЙ СЕТИ РЕГУЛЯЦИИ БИОСИНТЕЗА ХОЛЕСТЕРИНА В КЛЕТКЕ

Отклик генной сети на 2-х кратное увеличенное поступления ЛНП в плазму крови в течении 8-ми часов при мутации, 2-Х КРАТНО уменьшающей скорость экспрессии гена ЛНП-рецепторов

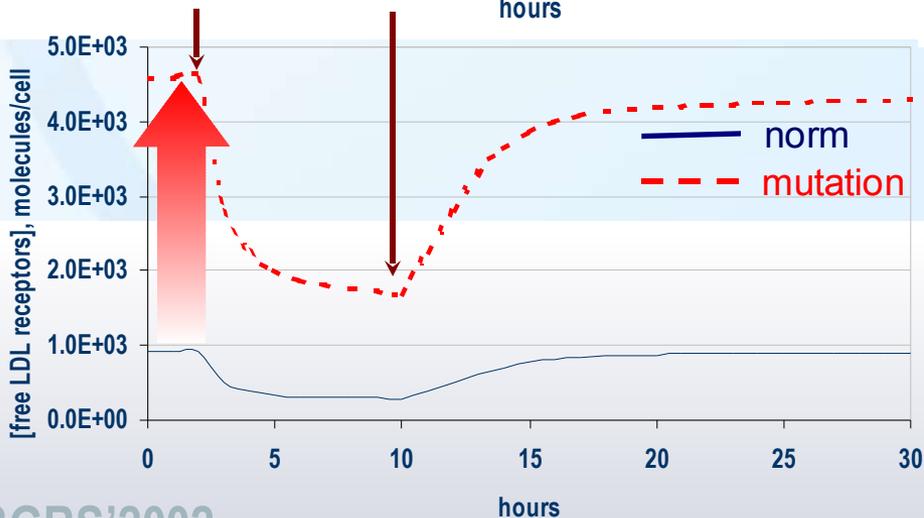
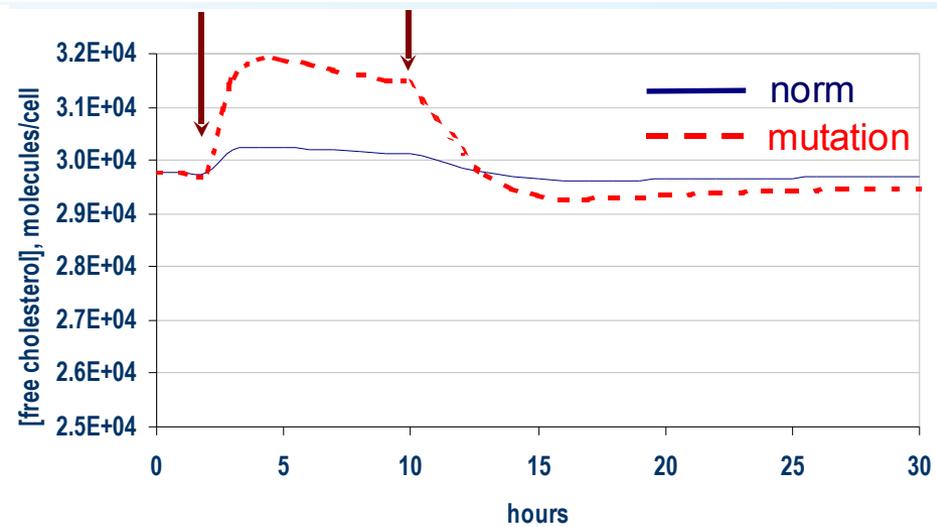
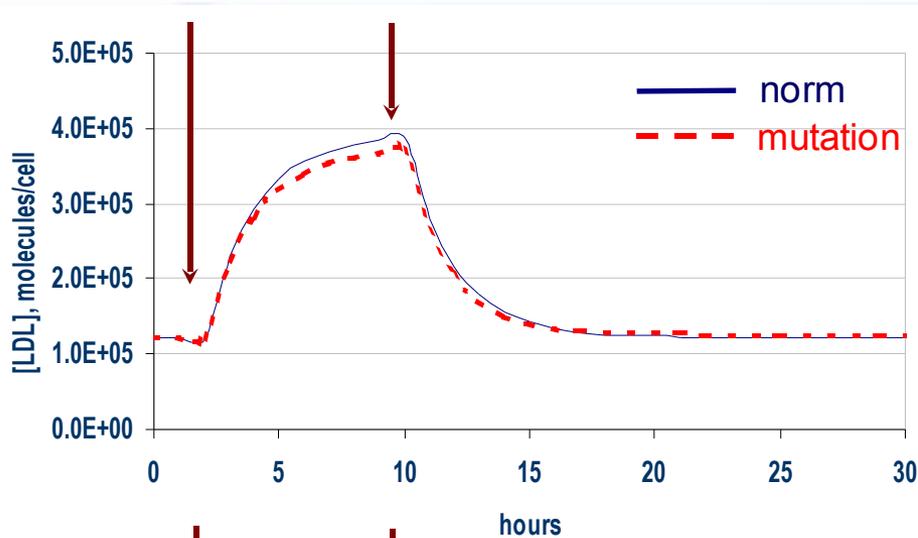


ПОИСК ЛИМИТИРУЮЩИХ ЗВЕНЬЕВ:

АНАЛИЗ ВЛИЯНИЯ МУТАЦИЙ НА ДИНАМИКУ ГЕННОЙ СЕТИ РЕГУЛЯЦИИ

БИОСИНТЕЗА ХОЛЕСТЕРИНА В КЛЕТКЕ

Отклик генной сети на 2-х кратное увеличение поступления ЛНП в плазму крови в течении 8-ми часов при мутации, ПЯТИКРАТНО уменьшающей аффинность ЛНП к своему мембранному рецептору



NORM

MUTATION

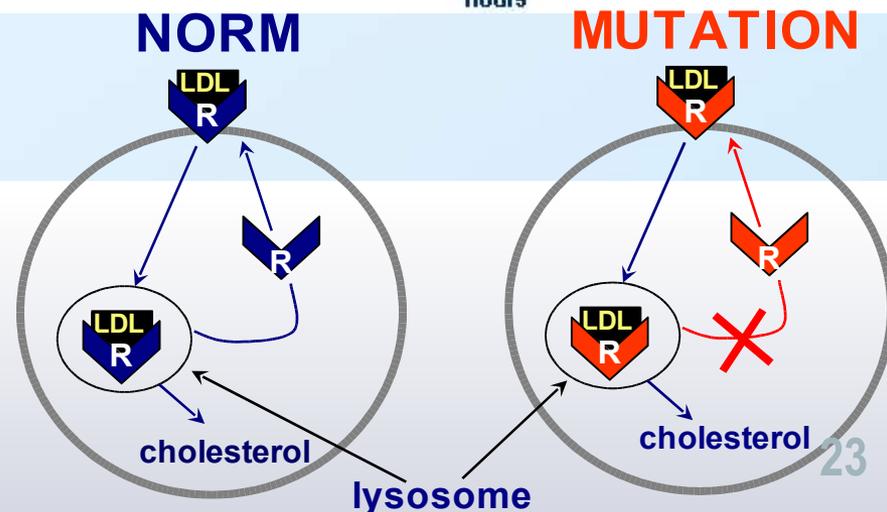
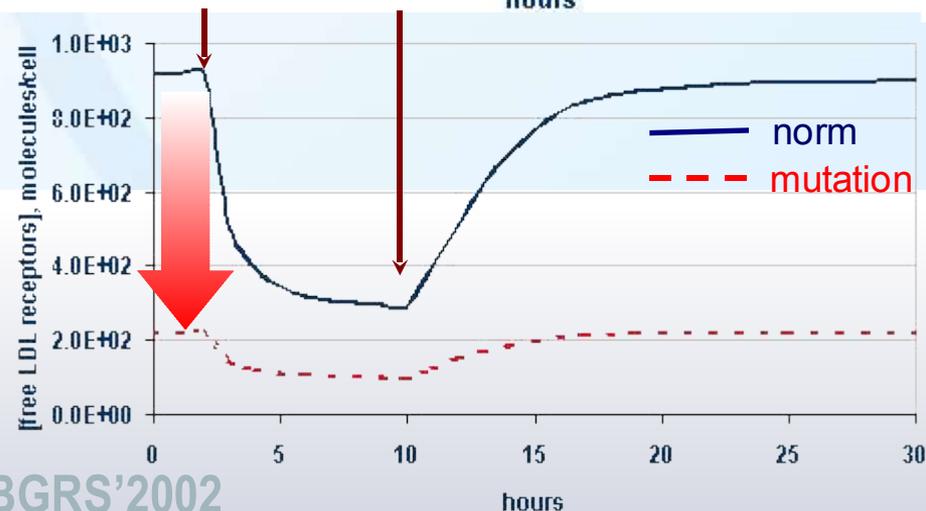
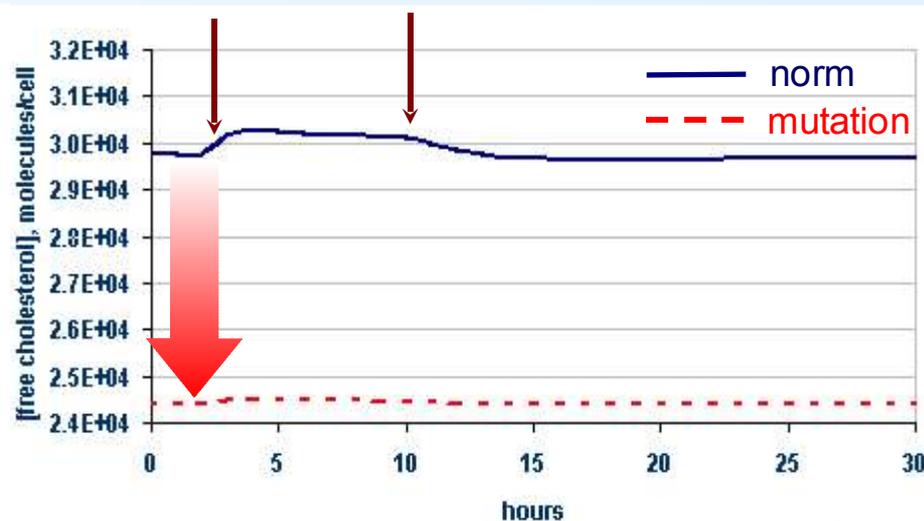
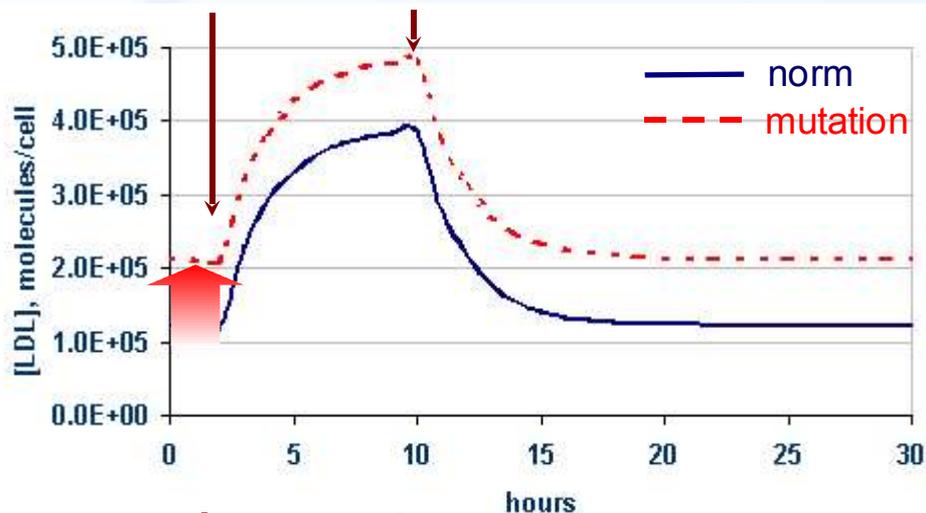


ПОИСК ЛИМИТИРУЮЩИХ ЗВЕНЬЕВ:

АНАЛИЗ ВЛИЯНИЯ МУТАЦИЙ НА ДИНАМИКУ ГЕННОЙ СЕТИ РЕГУЛЯЦИИ

БИОСИНТЕЗА ХОЛЕСТЕРИНА В КЛЕТКЕ

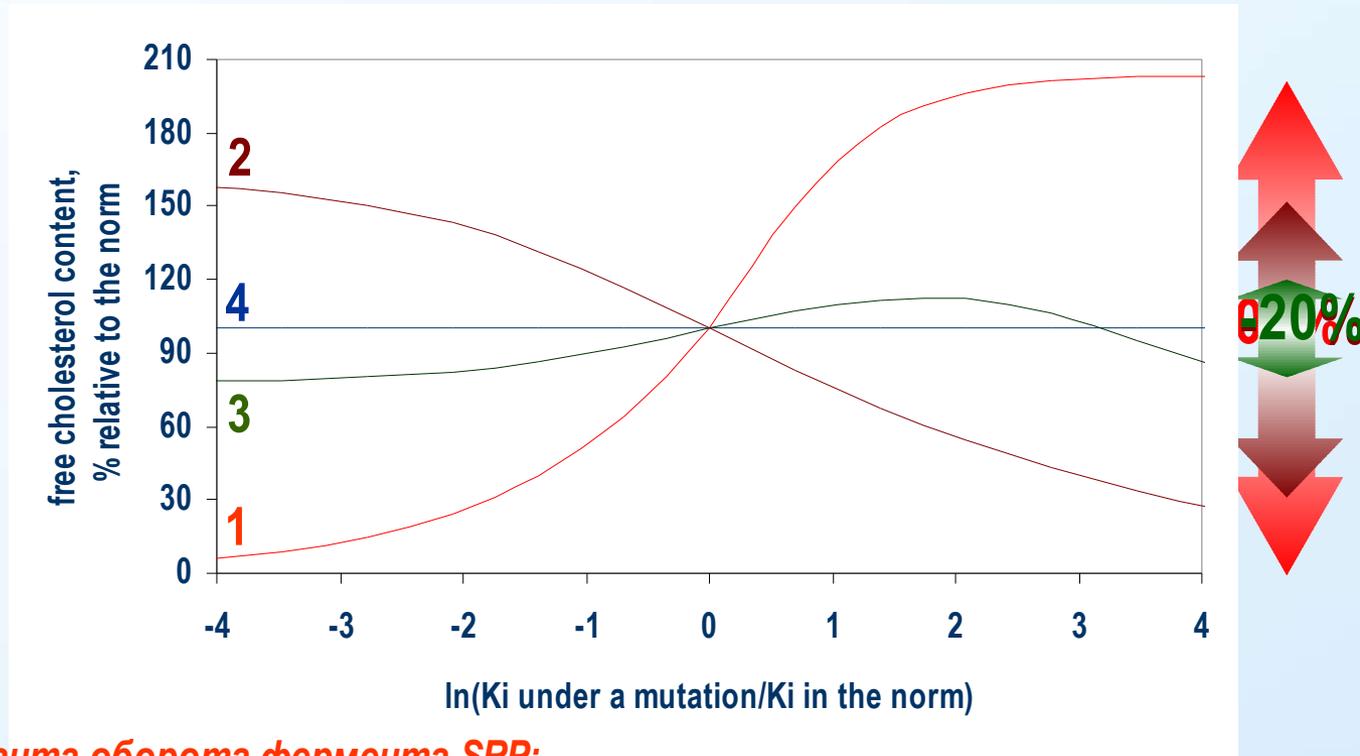
Отклик генной сети на 2-х кратное увеличение поступления ЛНП в плазму крови в течении 8-ми часов при мутации, ДЕСЯТИКРАТНО увеличивающей скорость деградации ЛНП-рецептора в эндосоме



ЛИМИТИРУЮЩИЕ ПРОЦЕССЫ В ГЕННОЙ СЕТИ

регуляции биосинтеза холестерина в клетке

Изменение содержания свободного холестерина в клетке в зависимости от мутационного изменения констант элементарных процессов в генной сети биосинтеза холестерина



(1) Константа оборота фермента SRP;

(2) Константа обратной реакции димеризации SREBP1;

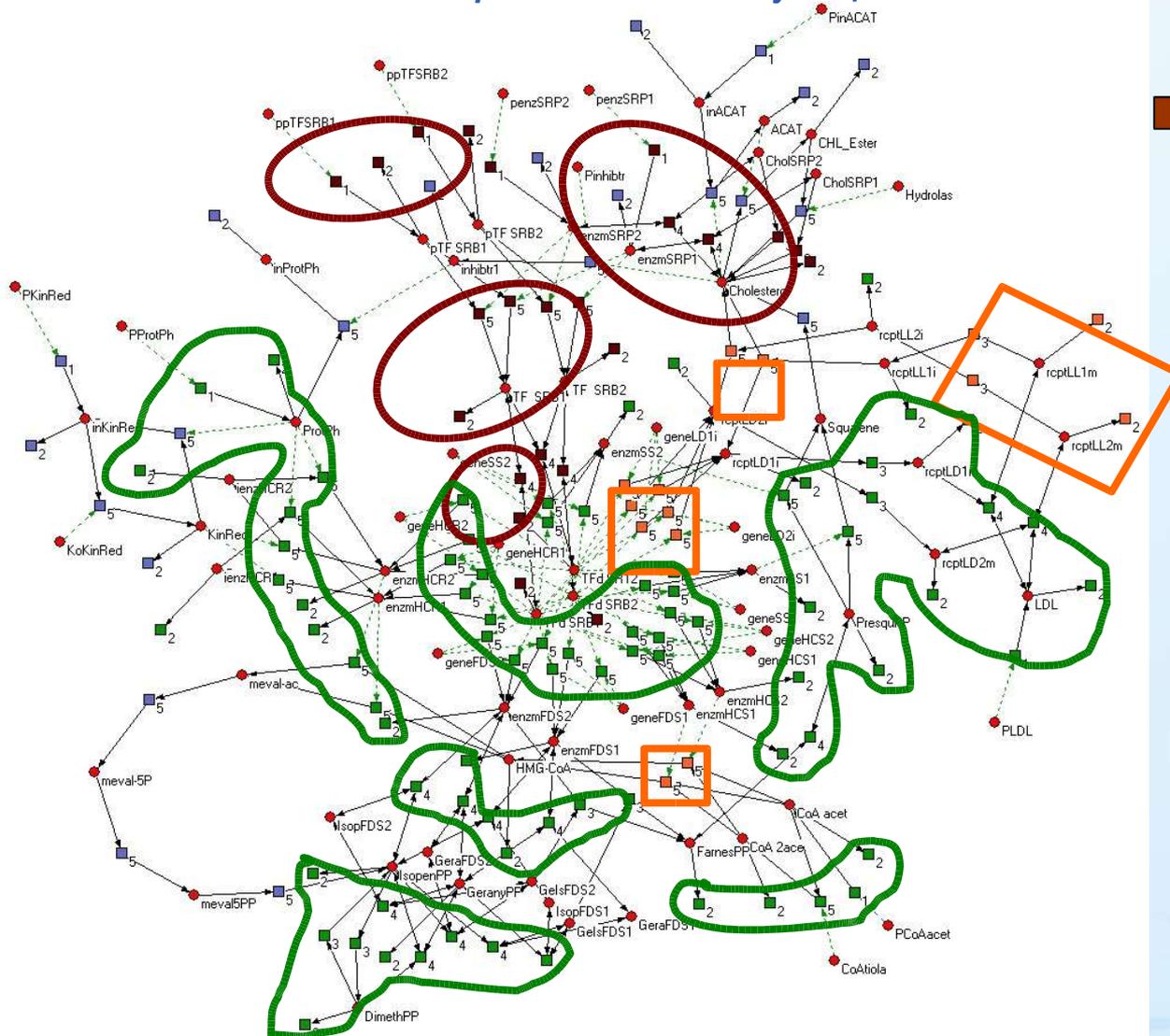
(3) Константа Михаэлиса-Ментен фермента ацетоацетил КоА тиолазы;

(4) Константа оборота фермента АХАТ (ацил-КоА: холестерин ацилтрансфераза)

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.



Задачи дизайна лекарственных препаратов, основанного на анализе структурных данных (Structure-Based Drug design)

Распознавание каталитически-активных центров в третичной структуре белков

Распознавание сайтов связывания аллостерических эффекторов, таких как ионы металлов.

Распознавания сайтов белок-белковых взаимодействий.

Распознавание сайтов белок-ДНК взаимодействий.

Распознавание сайтов белок-РНК взаимодействий.

Распознавание ядер нуклеации третичной структуры для создания ингибиторов белкового фолдинга.



Существующие программы распознавания функциональных сайтов в пространственных структурах белков

WebFEATURE

(Liang, M.P., Banatao, D.R., Klein, T.E., Brutlag, D.L. and Altman, R.B. (2003) WebFEATURE: An interactive web tool for identifying and visualizing functional sites on macromolecular structures. *Nucleic Acids Res.*, 31, 3324-3327.)

DOMPLOT

(Todd, A.E., Orengo, C.A. and Thornton, J.M. (1999) DOMPLOT: a program to generate schematic diagrams of the structural domain organization within proteins, annotated by ligand contacts. *Protein Eng.*, 12, 375-379.)

RELIBASE

(Hendlich, M., Bergner, A., Gunther, J. and Klebe, G. (2003) Relibase: design and development of a database for comprehensive analysis of protein-ligand interactions. *J Mol Biol.*, 326, 607-620.)



Программы распознавания функциональных сайтов в пространственных структурах белков и быстрого предварительного докинга, разработанные в ИЦиГ

Иванисенко В.А. к.б.н., с.н.с.

Пинтус С.С. аспирант

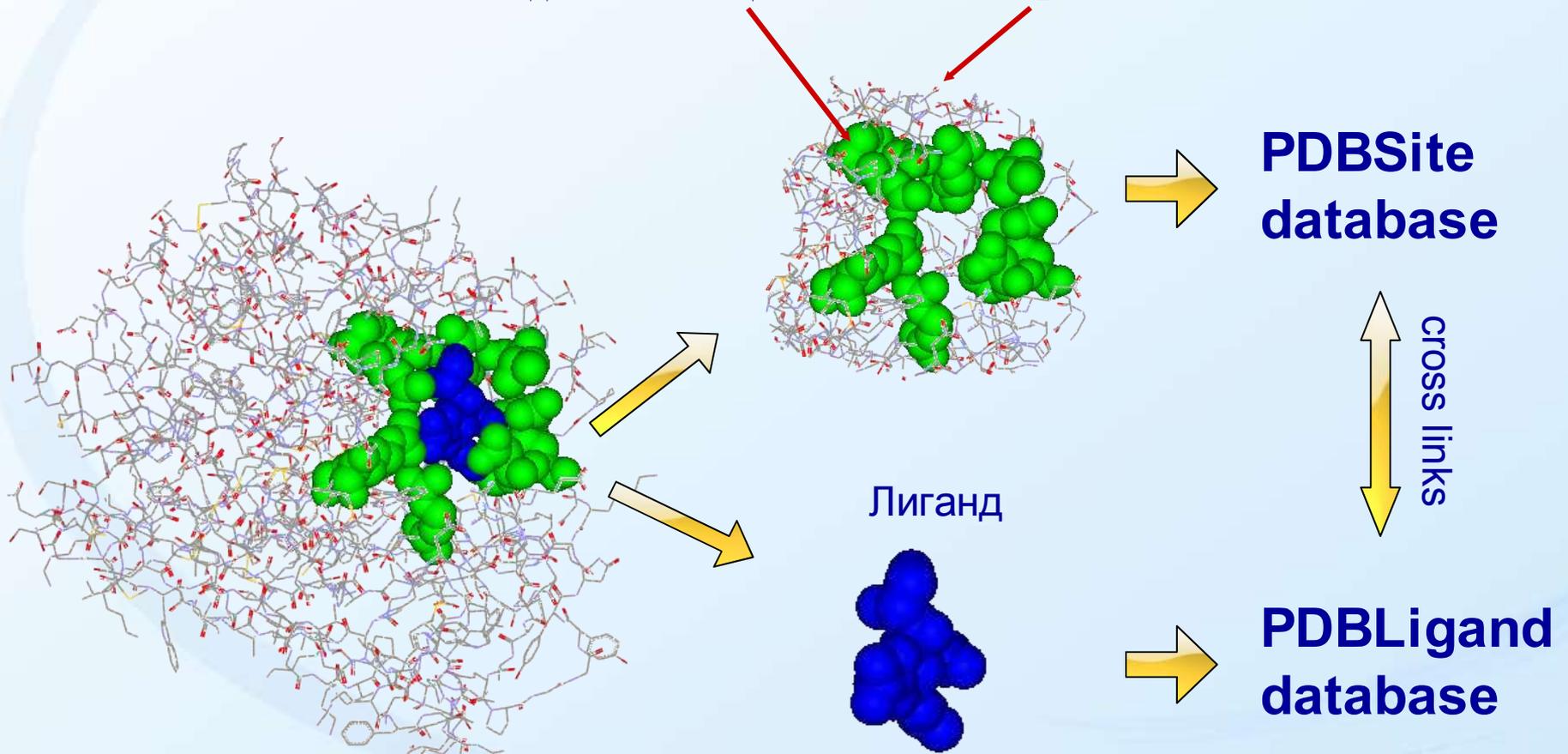
Григорович Д.А. программист



База **PDBSite** накапливает данные по функциональным сайтам белков (<http://www.mgs.bionet.nsc.ru/mgs/systems/fastprot/>)
База **PDBLigand** содержит данные по лигандам для этих сайтов (<http://www.mgs.bionet.nsc.ru/mgs/systems/fastprot/>)

Лиганд-связывающий сайт

Окружение сайта





Критерии выделения сайтов

Использовались два дополняющих друг-друга подхода для получения данных по сайтам:

- Данные извлекались из PDB на основе информации, содержащейся в поле SITE, базы PDB, указывающей на аминокислотные остатки активных центров, сайтов посттрансляционных модификаций белков, сайтов связывания различных лигандов;
- Сайты белок-белковых, белок-РНК и белок-ДНК взаимодействий определялись на основе анализа координат атомов белок-белковых, белок-ДНК и белок-РНК гетерокомплексов. В состав сайта включались аминокислотные остатки, контактирующие с лигандом (белок, РНК или ДНК). Два остатка считались контактными если любые три атома одного остатка находились на расстоянии менее 5 Å от трех атомов другого остатка.

Окружение сайта рассчитывалось по следующему критерию. В окружение сайта включались остатки, хотя бы один атом которых находился на расстоянии ближе 5 Å от любого атома, остатков сайта.



Для каждого сайта в базе PDBSite отводится отдельная запись, содержащая следующие данные, извлеченные из базы PDB

- координаты атомов функционального сайта;
- функциональное описание сайта;
- описание белка, которому принадлежит сайт;
- список, окружающих сайт аминокислотных остатков;
- физико-химические и структурные характеристики сайта и его окружения;
- ссылки с известными молекулярно-генетическими базами данных (SWISS-PROT, EMBL, PDB и др.)



PDBSite содержит более 8000 сайтов

Активные центры

Набор каталитически активных центров различных ферментов

Сайты посттрансляционной модификации белков

Acetylation
Cleavage
Glycosylation
Lipoylation
Myristylation
Phosphorylation

Сайты связывания ионов металлов

Cadmium
Calcium
Cobalt
Copper
Iron
Gallium
Manganese
Mercury
Metal
Molybdenum
Nickel
Potassium
Sodium
Vanadium
Ytterbium
Zinc
Magnesium



Сайты связывания неорганических соединений

Carbohydrate

CO

Phosphate

Sulfur

Xenon

Сайты связывания органических соединений

Ascorbate

ATP

Benzamidine

Benzhydrozamic

Butyramide

Citrate

Cyclosporin

Diaminopimelate

DTPP

FAD

Fluconazole

FMN

Formycin

Ganglioside

Glucose

Glutathione

Glutathione

Glycerol

Hapten

HEC

HEME

Isopropylmalate

Lactose

Lipid

Maltose

Molybdopterin

NAD

NADP

NADPH

Nucleotide

Ornithine

Oxalate

Peptide

Pterin

Purine

Purvalanol

Pyridoxal

Pyrimidine

Pyruvate

Saccharopine

Staurosporine

Succinate

Sugar

Tetracycline

Thiamin

Thymidine

Tris

Trisaccharide

Uracil

Xylan

Zanamivir



Сайты связывания лекарственных препаратов

Cyclosporin

Formycin

Staurosporine

Zanamivir

Сайты белок-белковых взаимодействий

Обширный набор разнообразных сайтов белок-белковых взаимодействий, полученных путем обработки координат атомов белок-белковых комплексов.

Сайты белок-ДНК взаимодействий

Обширный набор разнообразных сайтов белок-ДНК взаимодействий, полученных путем обработки координат атомов белок-ДНК комплексов.

Сайты белок-РНК взаимодействий

Обширный набор разнообразных сайтов белок-РНК взаимодействий, полученных путем обработки координат атомов белок-РНК комплексов.



PDBLigand содержит более 7000 лигандов

База PDBLigand содержит данные по низкомолекулярным лигандам, белкам, ДНК и РНК, которые связываются с сайтами из базы PDBSite.

В базе содержатся координаты атомов лигандов, а также их функциональное описание, извлеченные из базы PDB.

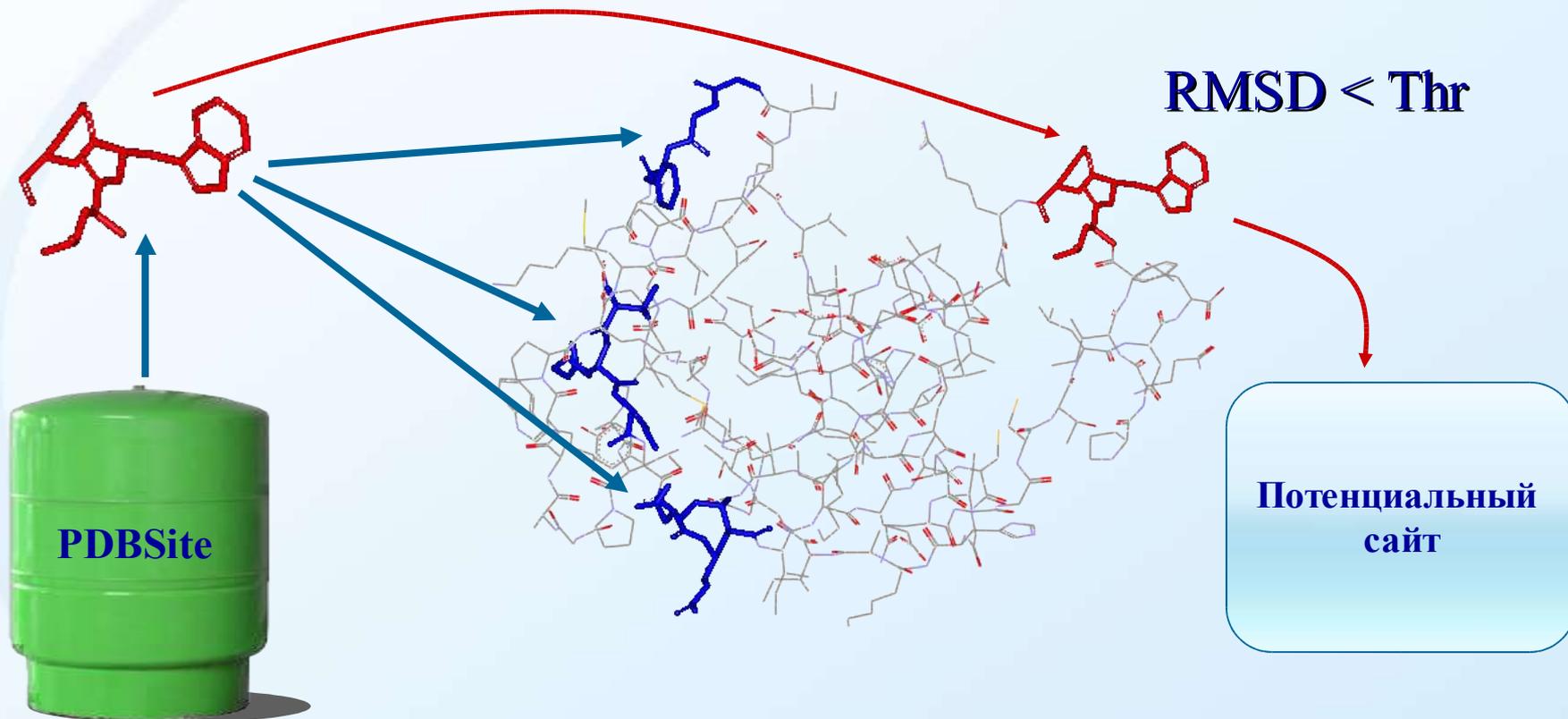
Базы PDBLigand и PDBSite взаимосвязаны между собой

Каждая запись базы данных PDBLigand, содержащая информацию о отдельно взятом лиганде линкована с записью базы PDBSite, содержащей информацию о сайте связывания данного лиганда.



PDBSiteScan - программа распознавания функциональных сайтов в третичных структурах белков

<http://www.mgs.bionet.nsc.ru/mgs/systems/fastprot/pdbsitecan.html>





PDBSiteScan Web интерфейс

PdbSiteScan Stage 0 - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное Медиа

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

PDBSiteScan

Enter file in PDB format: [Help](#)

Select chain to process: Chain:

Limit maximum distance to

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

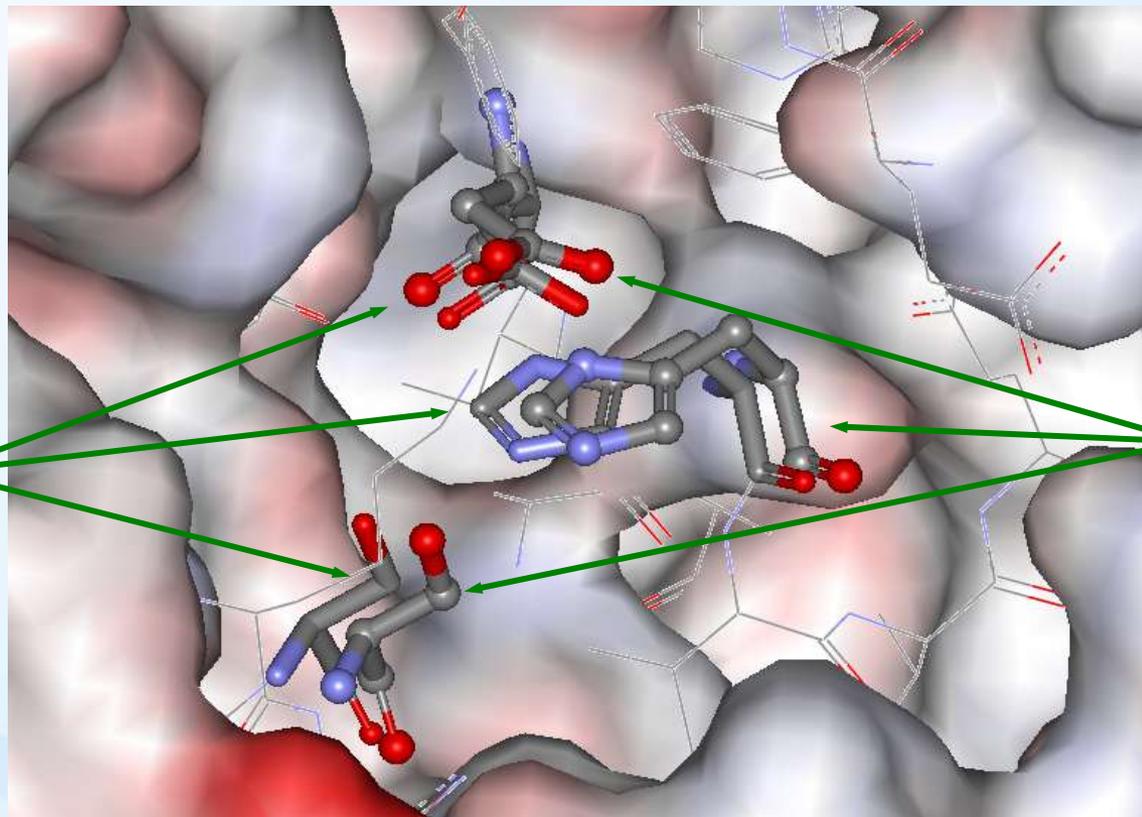
Potassium

Готово Интернет



Пример распознавания каталитического центра в сериновой протеазе человека (1ELV) с помощью структурного выравнивания с одним из каталитических центров сериновых протеаз, содержащихся в базе PDBSite, иллюстрирующий работу программы PDBSiteScan.

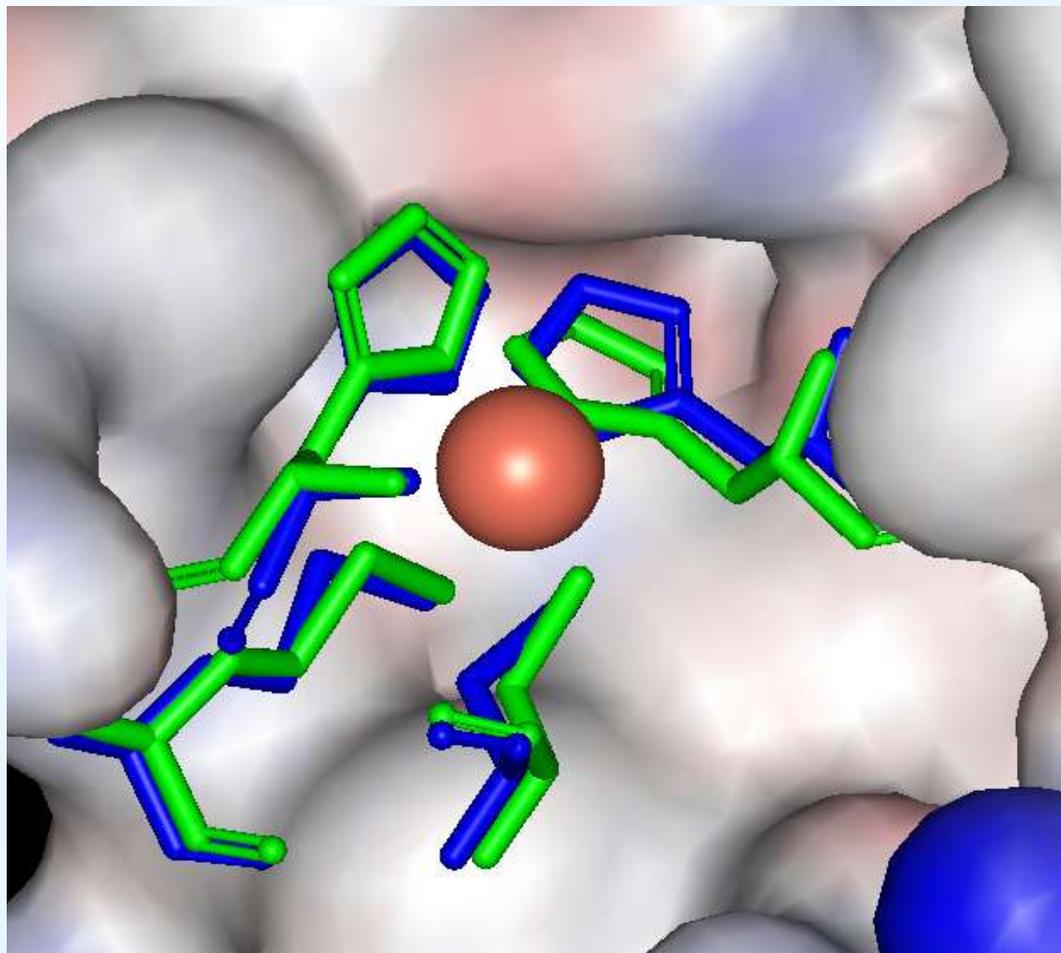
Аминокислотные остатки сериновой протеазы (1ELV), формирующие структуру, аналогичную структуре каталитического сайта из базы PDBSite.



Аминокислотные остатки каталитического центра сериновой протеазы (1BQY) из базы PDBSite.



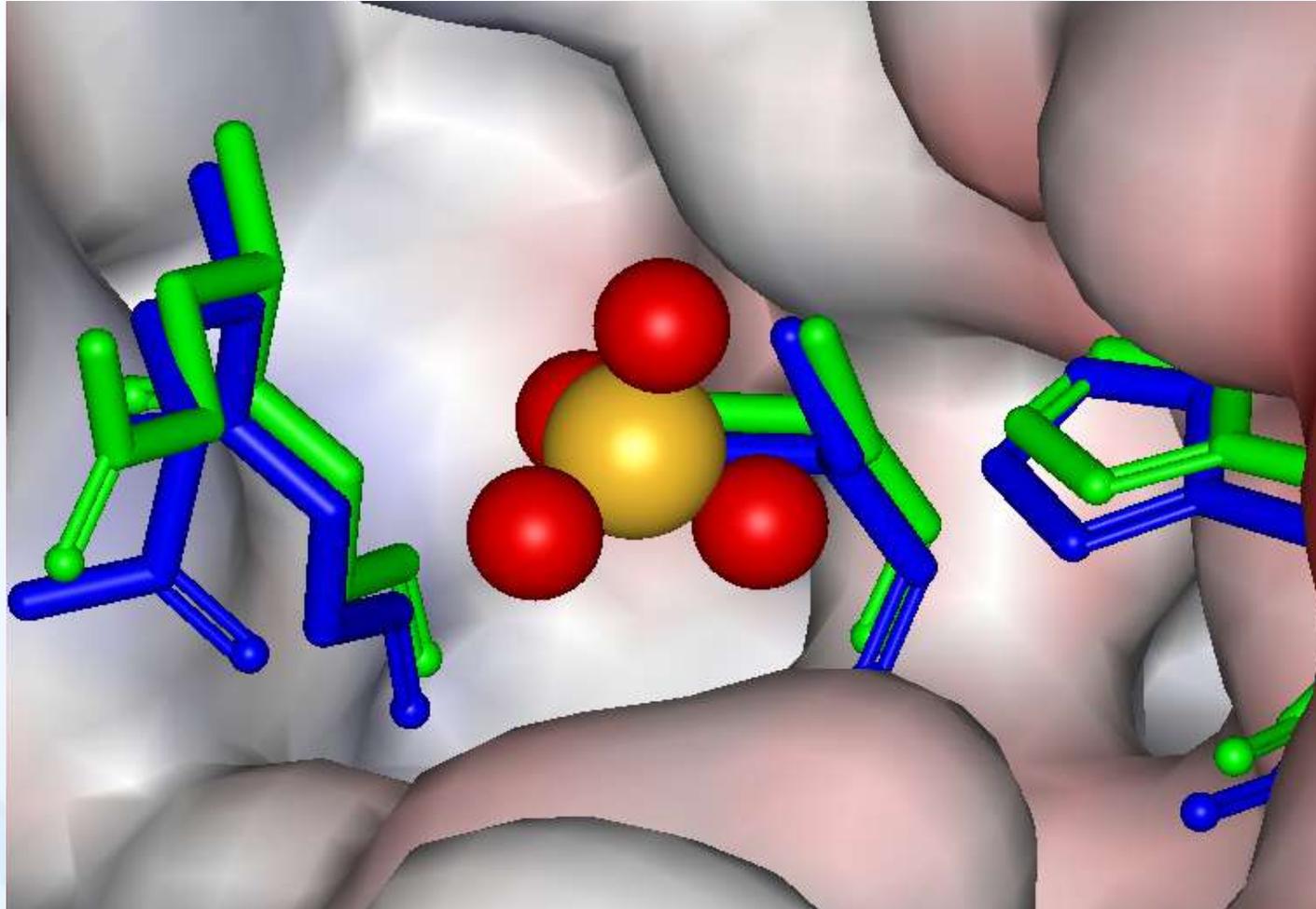
Распознавание сайта связывания иона меди в белке plastocyanin программой PDBSiteScan



Белок plastocyanin (PDB ID 1BXU) представлен в виде фрагмента поверхности. Остатки медь-связывающего сайта показаны как Stick. Зеленым цветом изображены остатки распознанного сайта в белке plastocyanin, Синим цветом изображены остатки сайта-темплата из базы PDBSite (ID 1B3ICU). Ион меди показан оранжевым шариком.



Распознавание сайта связывания иона SO_4 в белке hydrolase программой PDBSiteScan



Hydrolase (PDB ID 1EAX) представлена в виде фрагмента поверхности. Остатки SO_4 -связывающего сайта показаны как Stick. Зеленым цветом изображены остатки распознанного сайта в hydrolase, синим цветом изображены остатки сайта-темплата из базы PDBSite (ID 1QL8AC2). Ион SO_4 показан шариками.



**После обнаружения необходимых сайтов на поверхности белков
встает задача предсказания малых молекул, способных связываться
с выявленными сайтами**

Существует два основных компьютерных подхода:

- виртуальный скрининг
- выращивание молекулы в кармане (de novo structure generation).

Виртуальный скрининг осуществляется с помощью:

Правил (например, набор критериев биоактивности Липински), нейронных сетей, анализа фармакофоров, анализа сходства, поиска структурных классов, а также докинга и скоринга.



Существующие программы:

AutoDock is a suite of automated docking tools. It is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure.

CAESA (Computer Assisted Estimation of Synthetic Accessibility) is a computer program to automatically rank sets of molecules according to their ease of synthesis.

DDASL This program utilizes large existing pharmaceutical databases as input for a new type of structure/activity correlation methodology in order to calculate a large set of new and traditional descriptors to create improved Quantitative Structure-Activity Relationship (QSAR) models that characterize and predict important biological responses.

Flo is a molecular modelling program written by Colin McMartin that helps chemist to visualize molecules and aids the design of lead drug compounds. Flo is very often used in conjunction with the drug design program "Growmol", written by Regine Bohacek. Flo is the front end graphical program that will set up molecules for a drug design run thru Growmol.

GrowMol, generates organic structures that are both spatially and chemically complementary to the target binding site. By "growing" molecules an atom at a time to fill the various nooks and crannies of a binding site, GrowMol can generate structures with exquisite complementary to the host. At each step, the position and type of atom to be added are randomly selected using Boltzmann statistics to bias acceptance toward atoms that can form favorable interactions with the binding site.

и др.

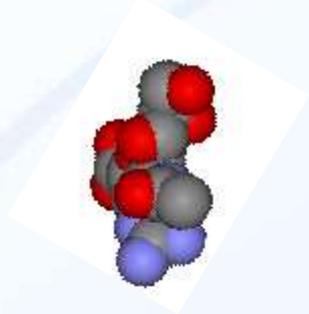


Краткая схема алгоритма быстрого предварительного докинга

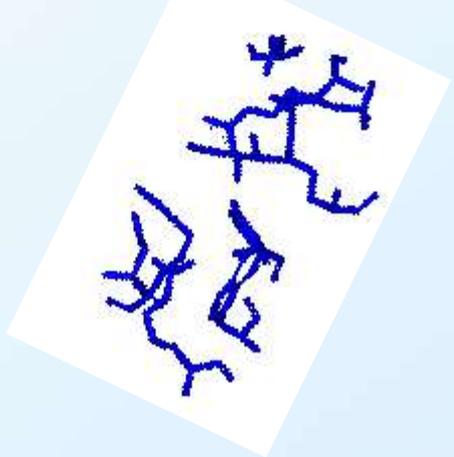
В базе PDBSite содержатся сайты-матрицы, для которых известны координаты атомов их комплексов с лигандами из базы PDBLigand. Быстрый докинг осуществляется путем переноса этих лигандов совместно с сайтом-матрицы при структурном выравнивании сайта-матрицы с белком.



PDBLigand

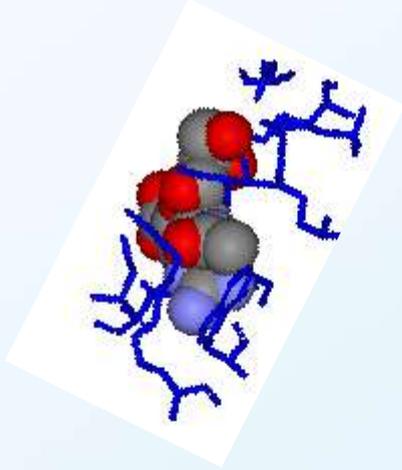


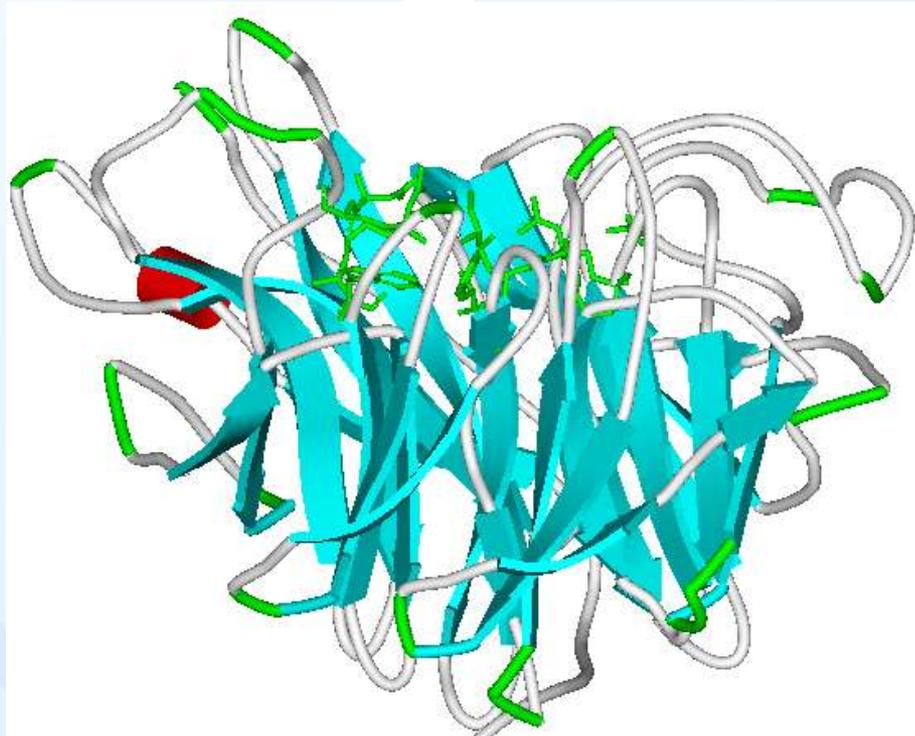
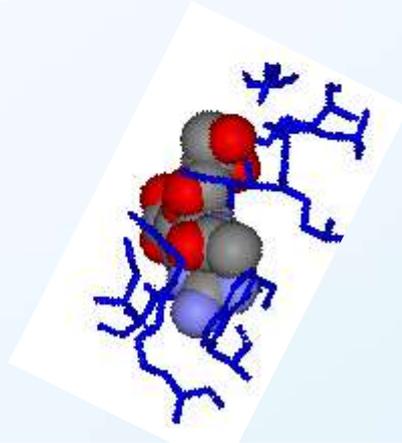
PDBSite

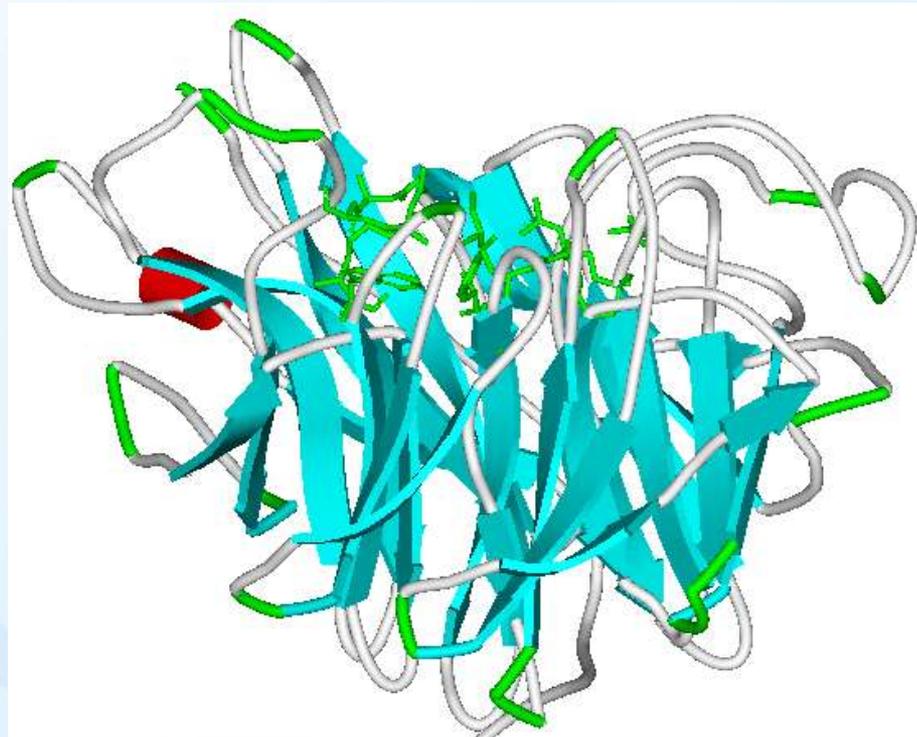
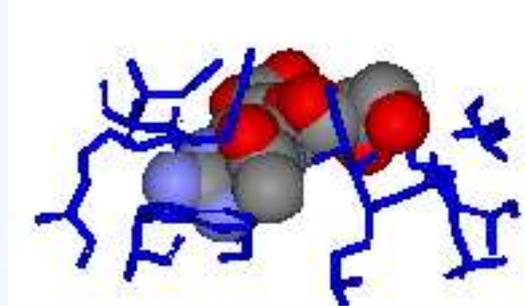


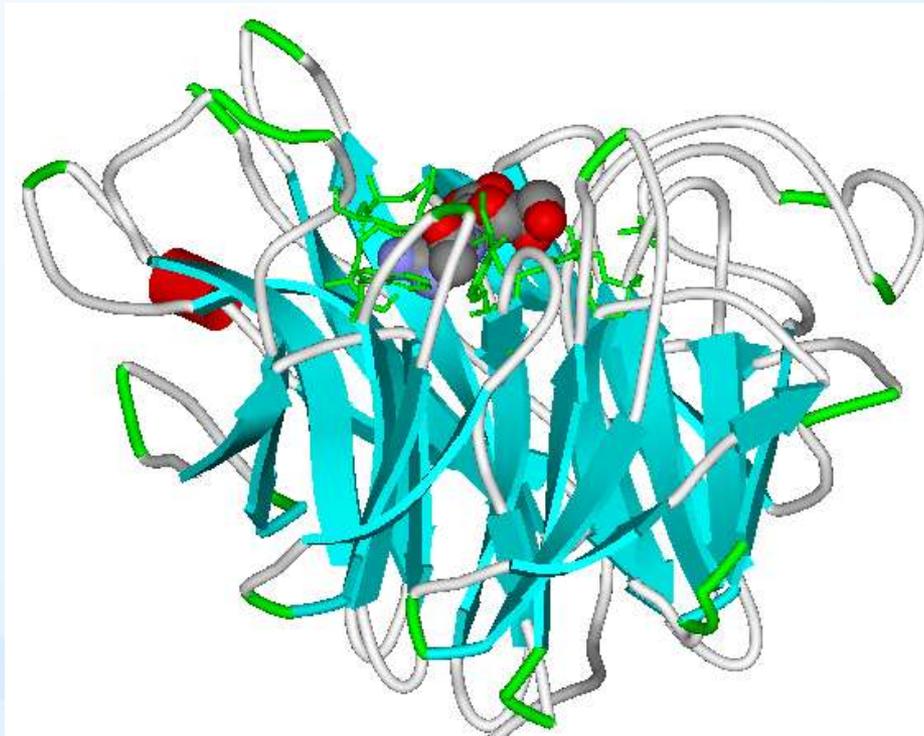


Комплекс сайт-лиганд



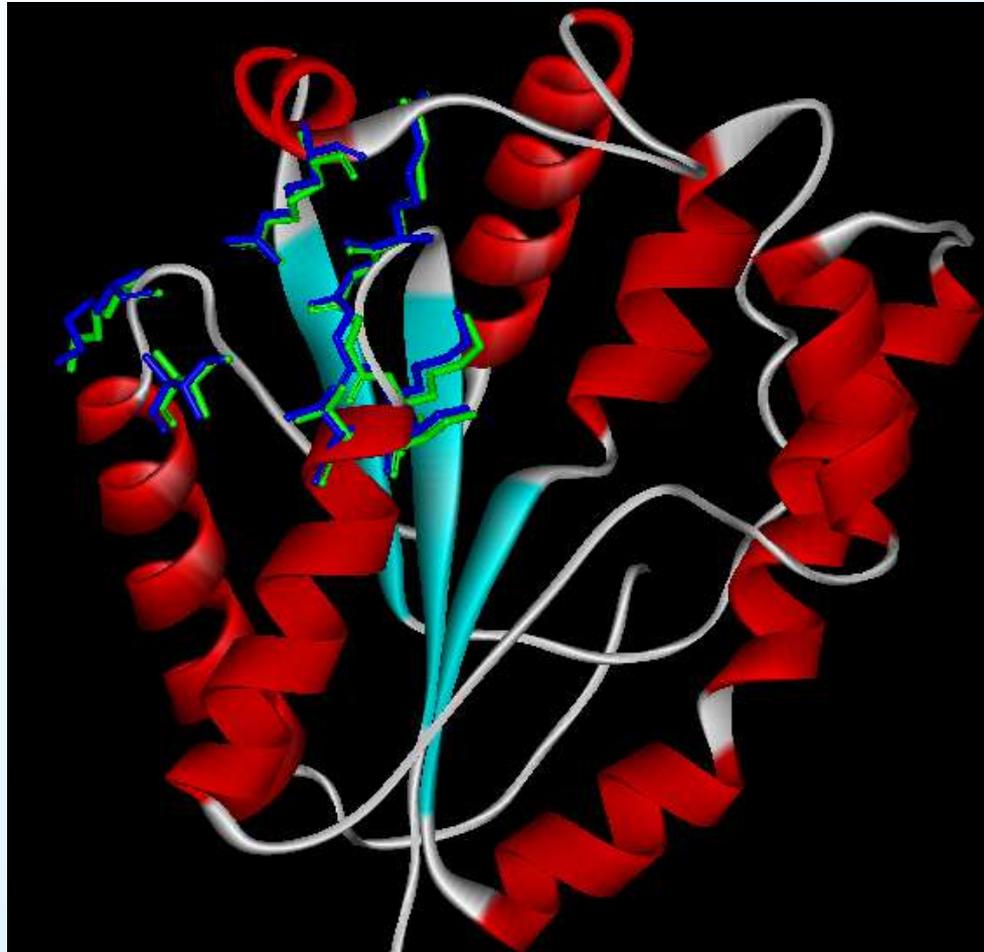








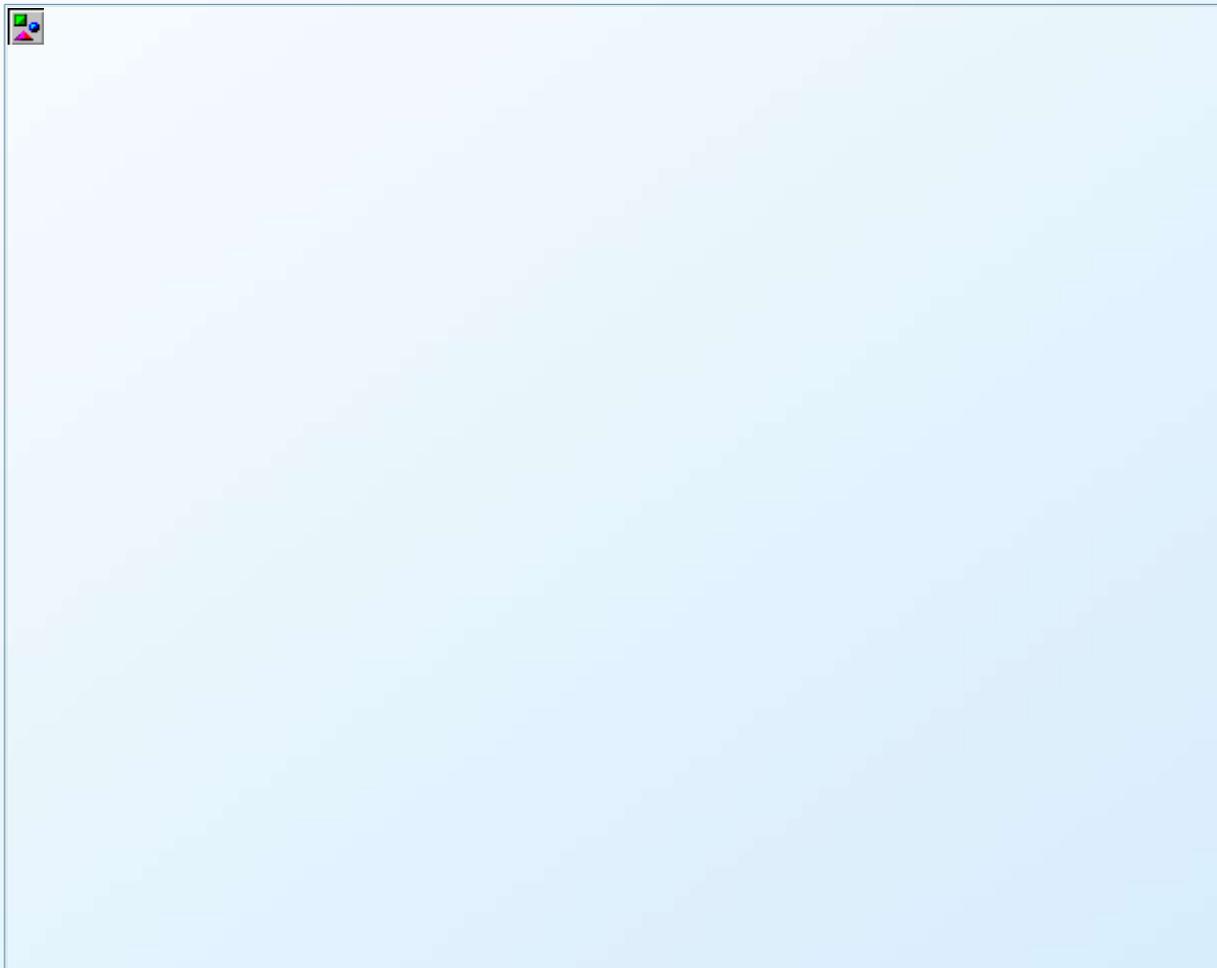
Распознавание АДФ-связывающего сайта в phosphotransferase человека программой PDBSiteScan



Phosphotransferase (PDB ID 1E98) представлена как Solid ribbon. Остатки АДФ-связывающего сайта показаны как Stick. Зеленым цветом изображены остатки распознанного сайта в Phosphotransferase, синим цветом изображены остатки сайта-темплата из базы PDBSite (ID 1E9CADP).



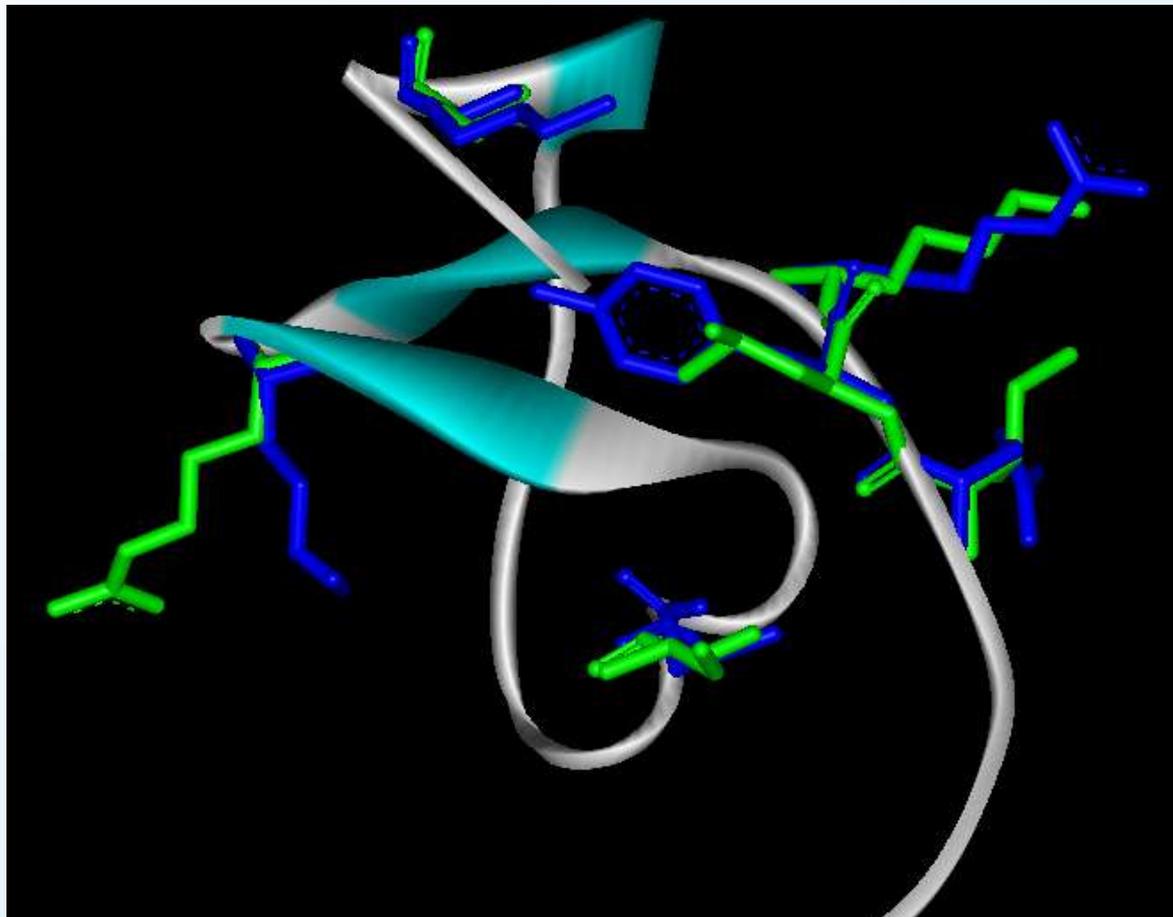
Быстрый докинг, выполненный PDBSiteScan, показывающий потенциальное взаимодействие между ADP и Phosphotransferase



Фрагмент Phosphotransferase (PDB ID 1E98) представлен в виде поверхности, молекула ADP показана шариками. Быстрый докинг выполнен путем переноса (сдвига и поворота) ADP- в пространстве с параметрами переноса, полученными для совмещения распознанного АТФ-связывающего сайта Phosphotransferase с сайтом матрицей из базы PDBSite (ID 1E9CADP).



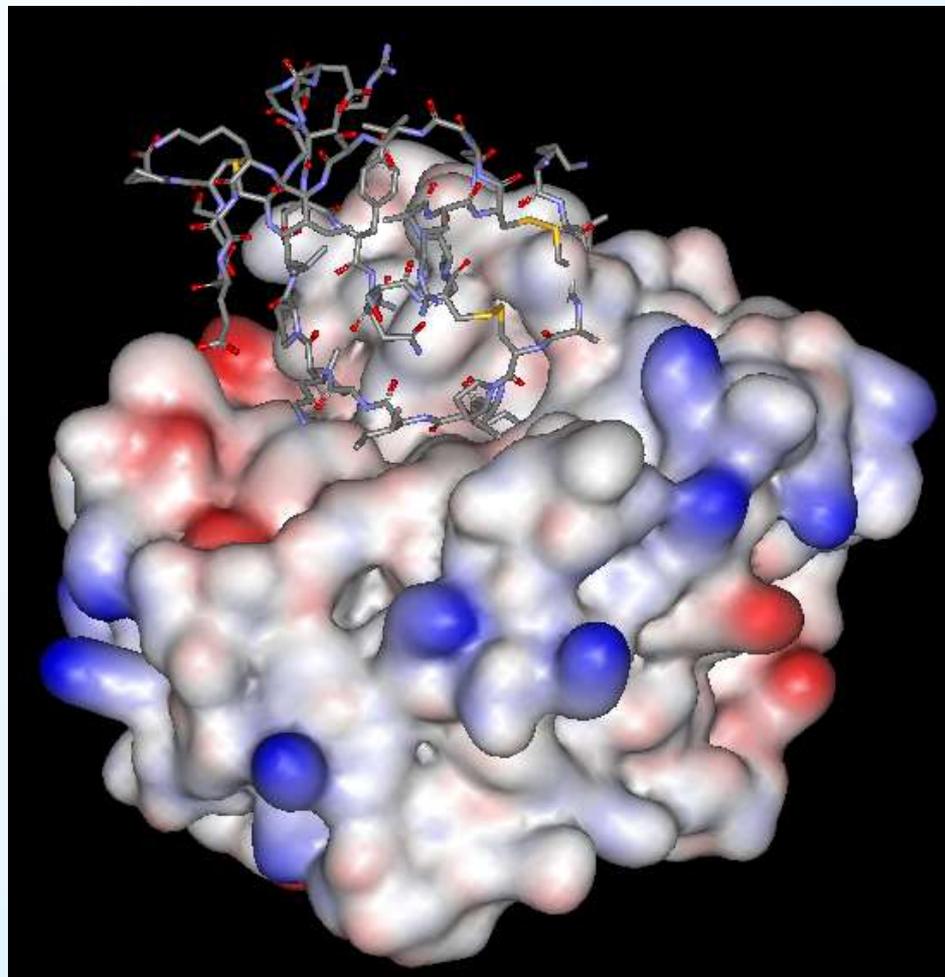
Распознавание белок-связывающего сайта в tryptase inhibitor (PDB ID 1AN1) программой PDBSiteScan



Белок tryptase inhibitor (PDB ID 1AN1) показан как Solid ribbon. Зеленом цветом показаны остатки предсказанного сайта связывания белка в tryptase inhibitor. Синим цветом показаны остатки реального сайта-«МАТРИЦЫ» ovomucoid inhibitor (PDB ID 1HJA), использовавшегося для распознавания сайтов.



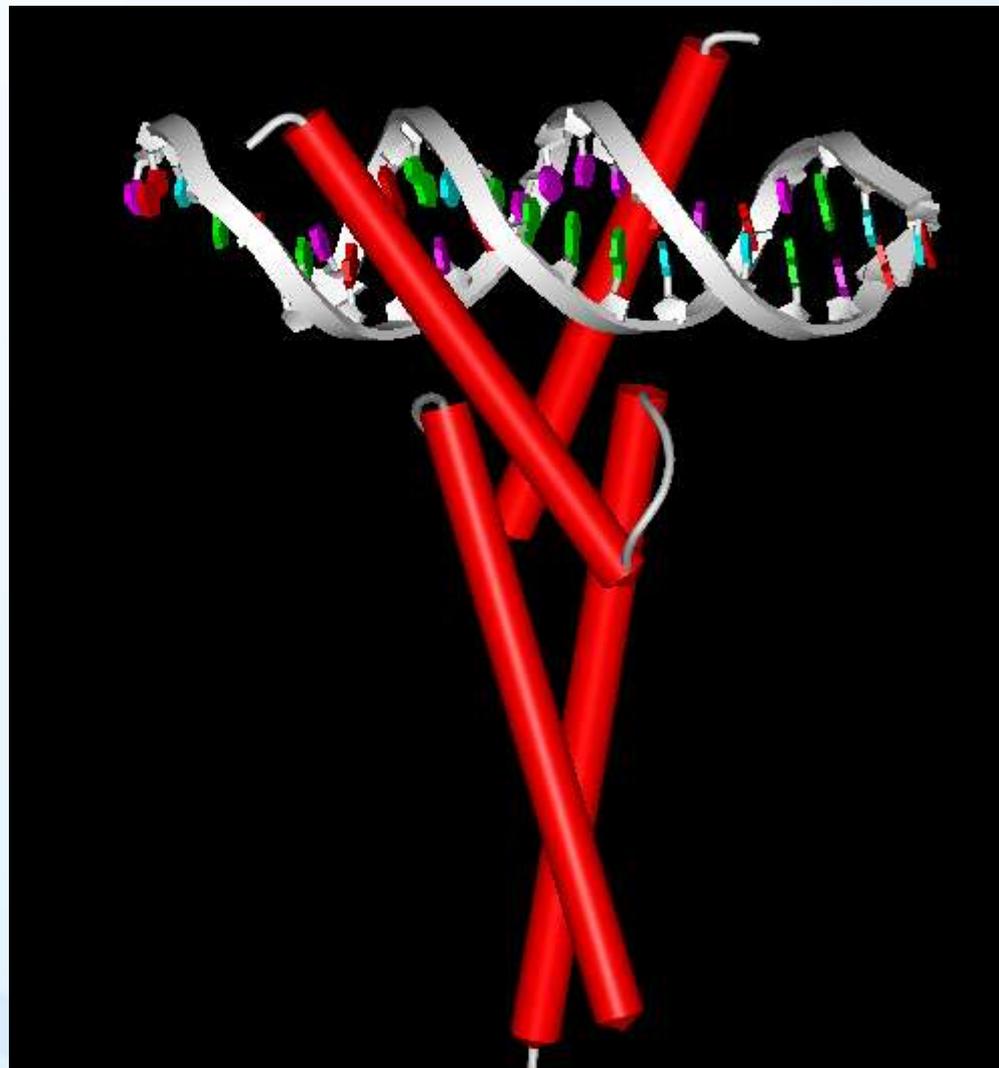
Быстрый белковый докинг, выполненный PDBSiteScan, показывающий потенциальное взаимодействие между tryptase inhibitor и alpha-chymotrypsin



Белок alpha-chymotrypsin (PDB ID 1HJA) представлен в виде поверхности, tryptase inhibitor (PDB ID 1AN1) показан как Stick. Быстрый докинг выполнен путем переноса (сдвига и поворота) alpha-chymotrypsin в пространстве с параметрами переноса, полученными для совмещения белок-связывающего сайта tryptase inhibitor и белок-связывающего сайта ovomucoid inhibitor.



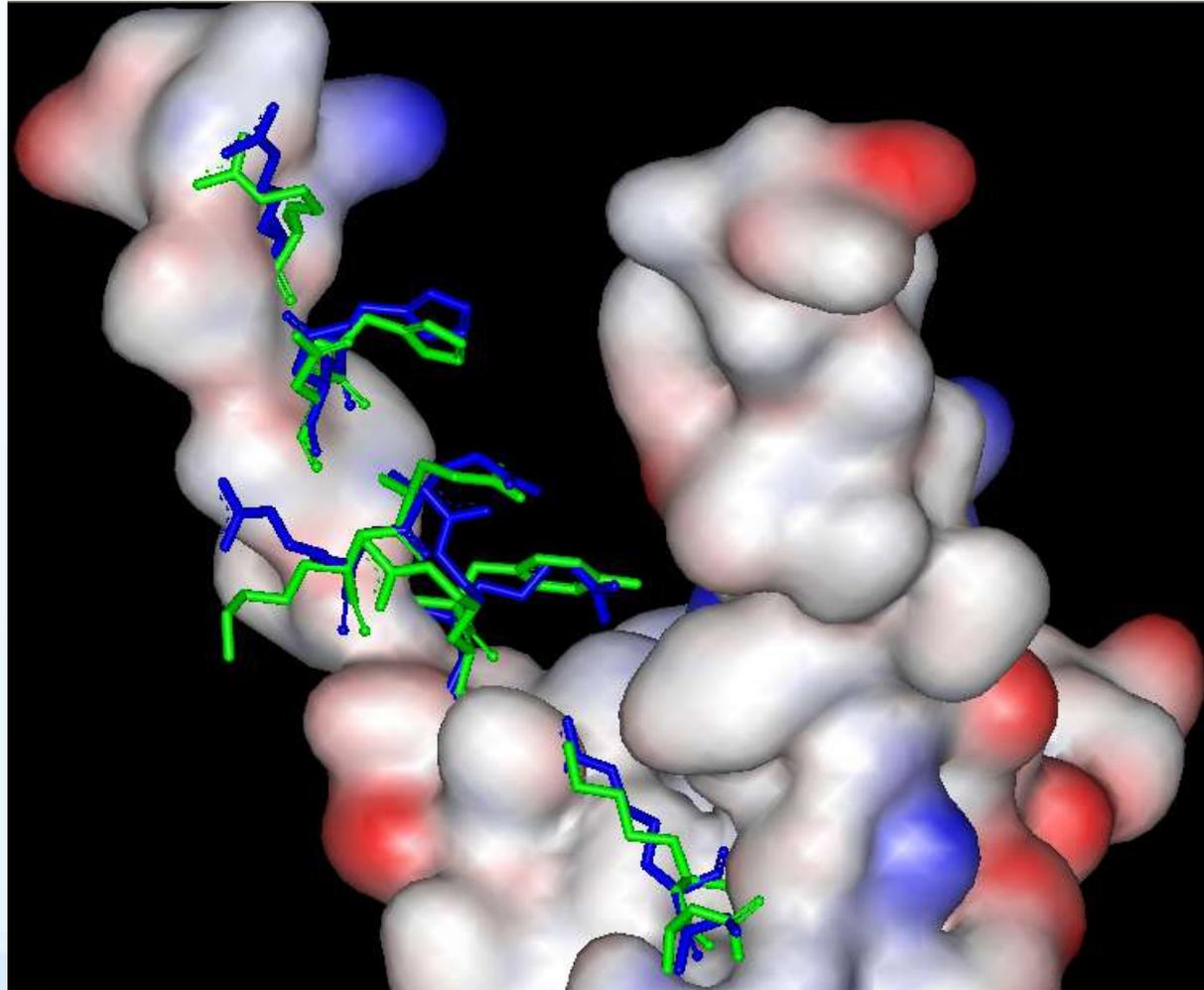
Пространственная структура комплекса человеческого SREBP-1A, связанного с промотором LDL-рецептора (PDB ID 1AM9).



Sterol Regulatory Element Binding Protein (Srebp-1a), Basic-helix-loop-helix-leucine Zipper



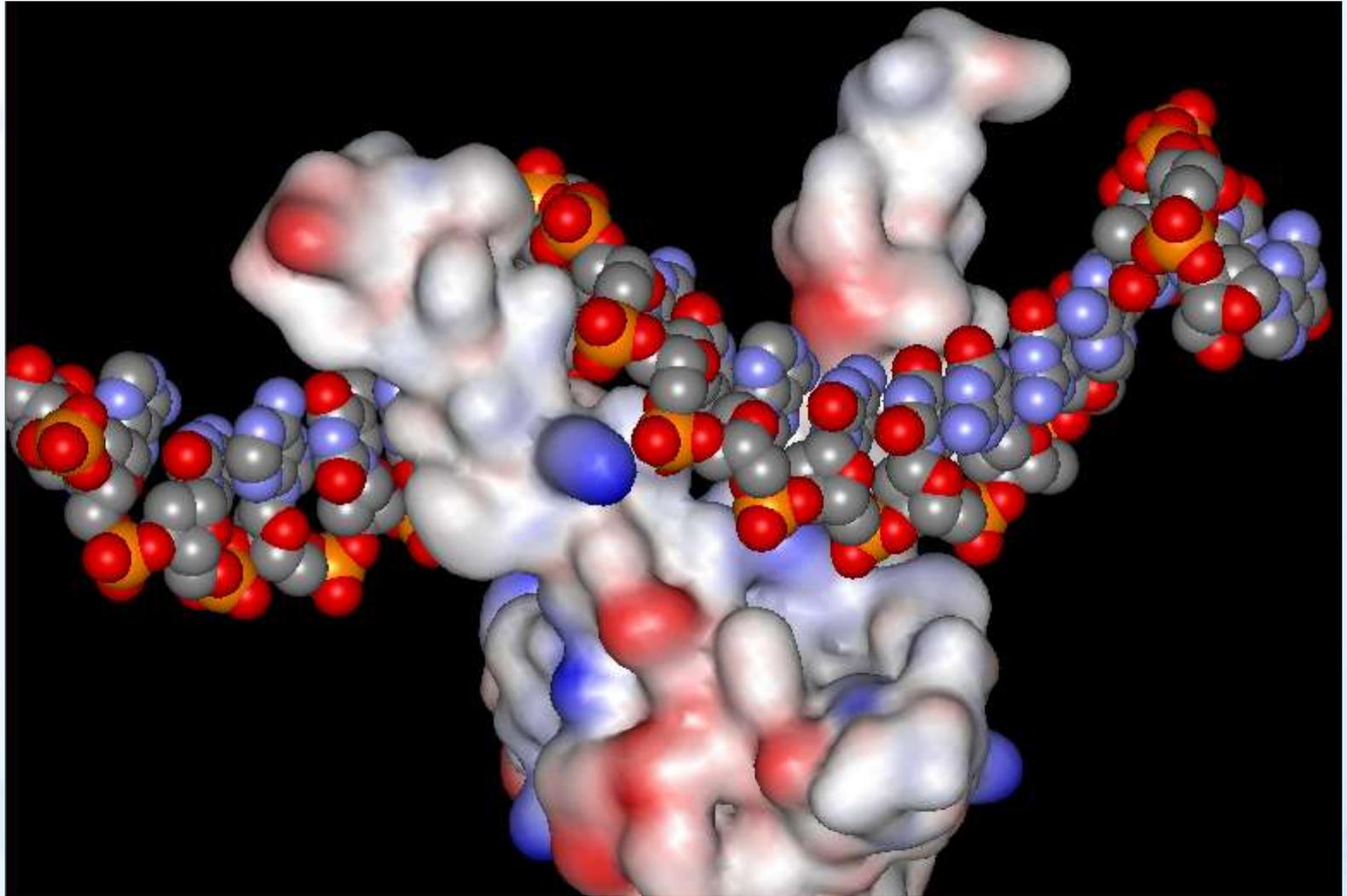
Распознавание ДНК-связывающего сайта в белке SREBP-1A программой PDBSiteScan.



Зеленым цветом отмечены остатки ДНК-связывающего сайта белка SREBP-1A, синим цветом отмечены остатки ДНК-связывающего сайта мышинного белка MAX (PDB ID 1AN2), являющимся сайтом-матрицей для распознавания



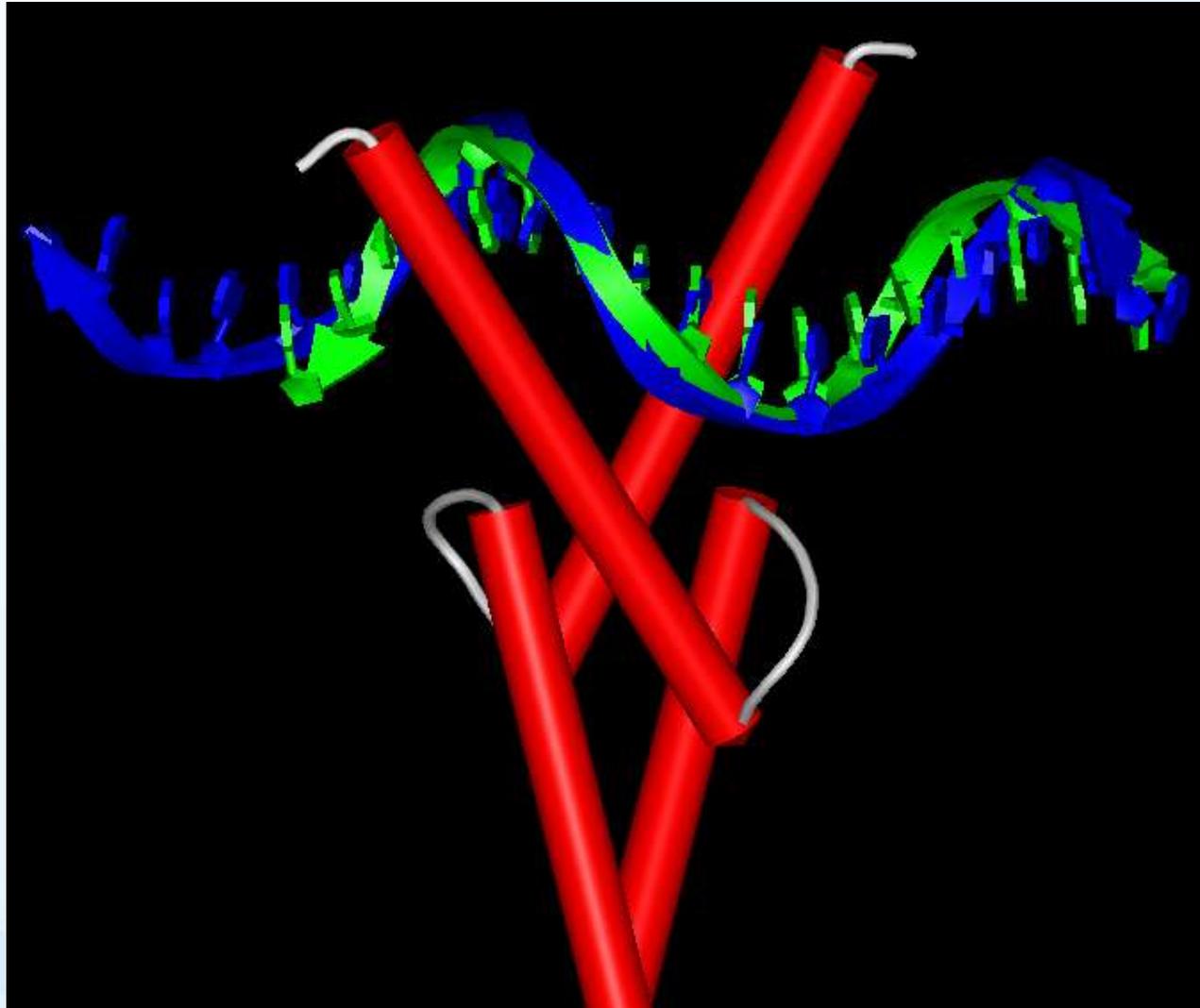
Быстрый докинг белка SREBP-1A с фрагментом ДНК, выполненный PDBSiteScan



Белок SREBP-1A представлен в виде поверхности, цепь ДНК показана шариками. ДНК взята из комплекса белка MAX с ДНК (PDB ID 1AN2). Быстрый докинг выполнен путем переноса (сдвига и поворота) цепи ДНК в пространстве с параметрами переноса, полученными для совмещения ДНК-связывающего сайта белка MAX и белка SREBP-1A.



Наложение структуры цепи ДНК, полученной в результате предварительного докинга, с кристаллографическими данными комплекса белок SREBP-1A /ДНК



Синим цветом отмечена цепь ДНК, взятая из структуры PDB ID 1AN2, зеленым цветом показана цепь ДНК реального комплекса белок-ДНК (PDB ID 1AM9).



Примеры применения PDBSiteScan для поиска сайтов в белках, как потенциальных мишеней для лекарственных препаратов

Цель исследования

Поиск новых сайтов в природных мутантных формах белка p53 человека, возникших в результате мутаций, ассоциированных с развитием опухолей

Краткая схема исследования

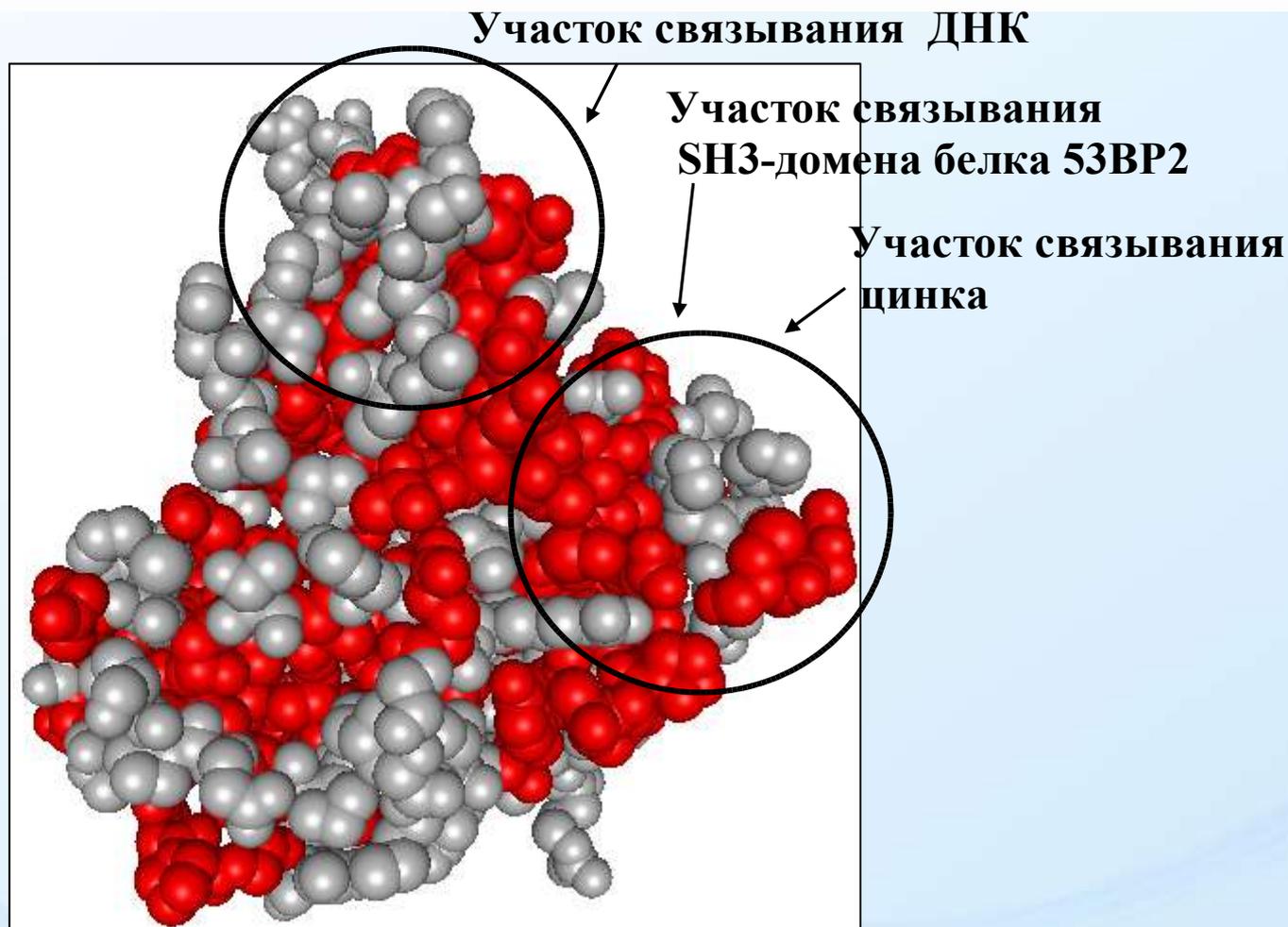
Мутации белка p53 брались из базы данных SWISS-PROT. Проанализировано 181 аминокислотная замена в p53, ассоциированная с различными опухолями человека;

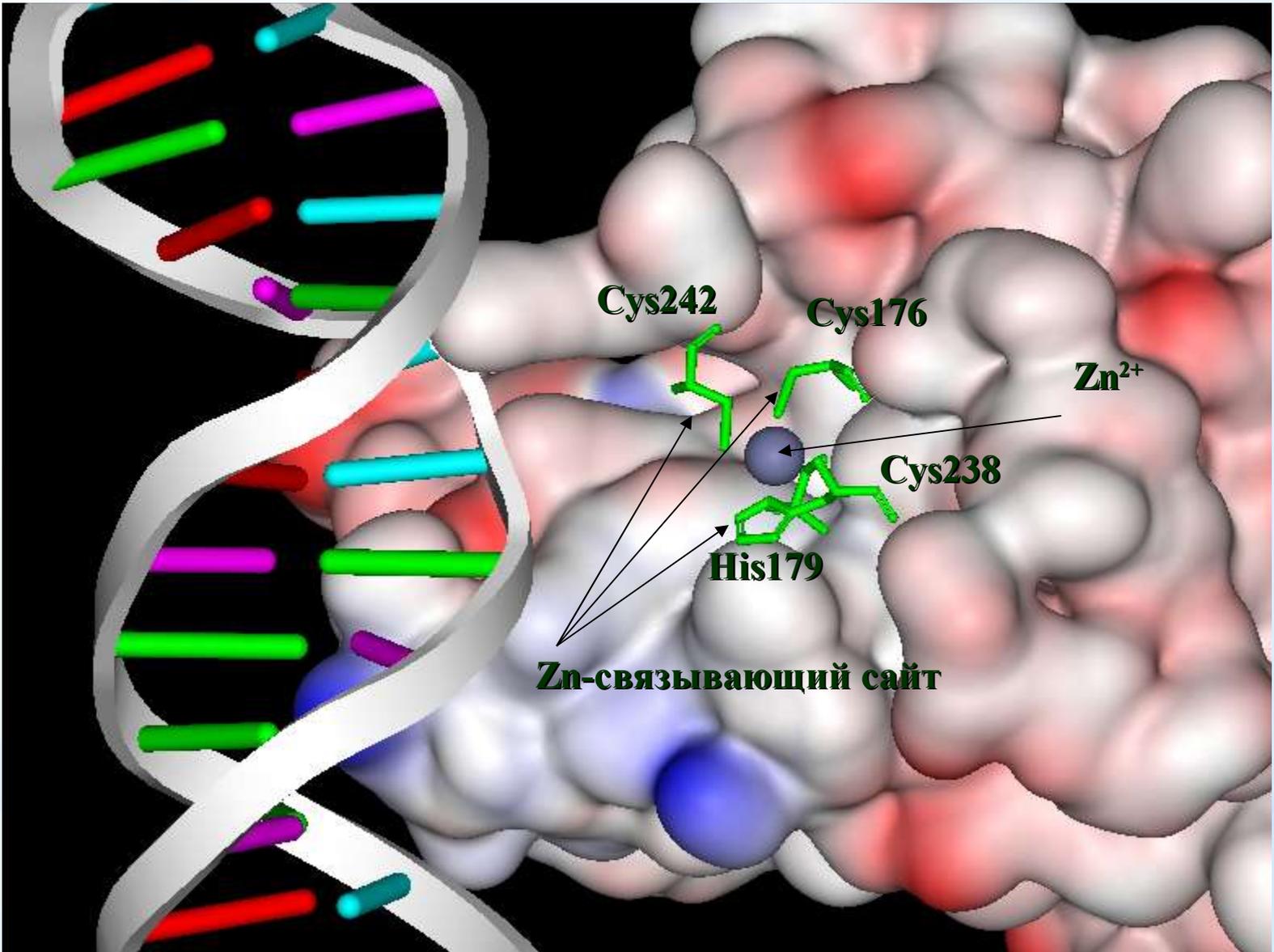
Третичные структуры мутантных форм ДНК-связывающего домена белка p53 человека моделировались с использованием сервера SWISS-MODEL;

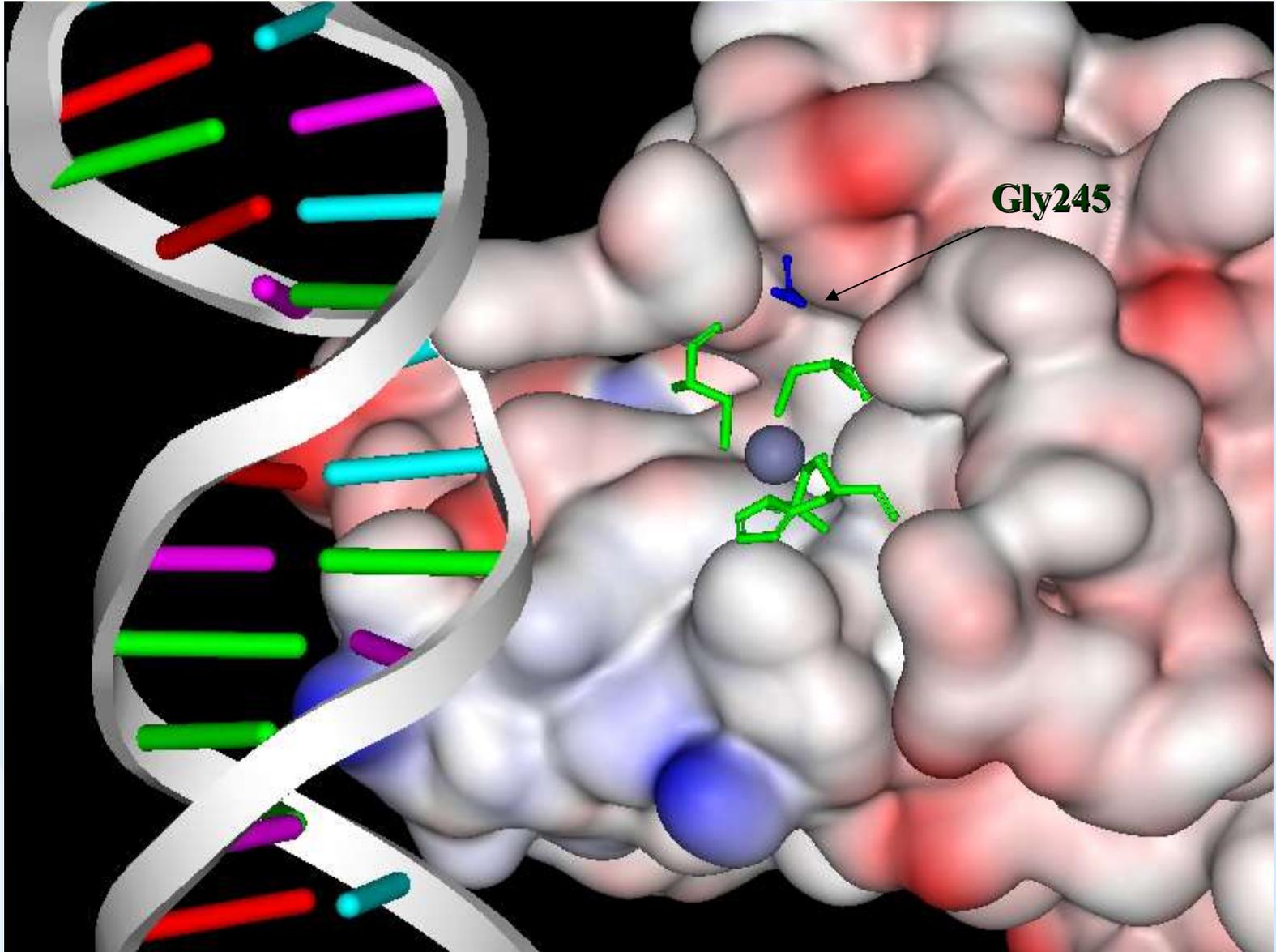
Для каждой мутации получали третичную структуру и с помощью PDBSiteScan проводили поиск новых сайтов.

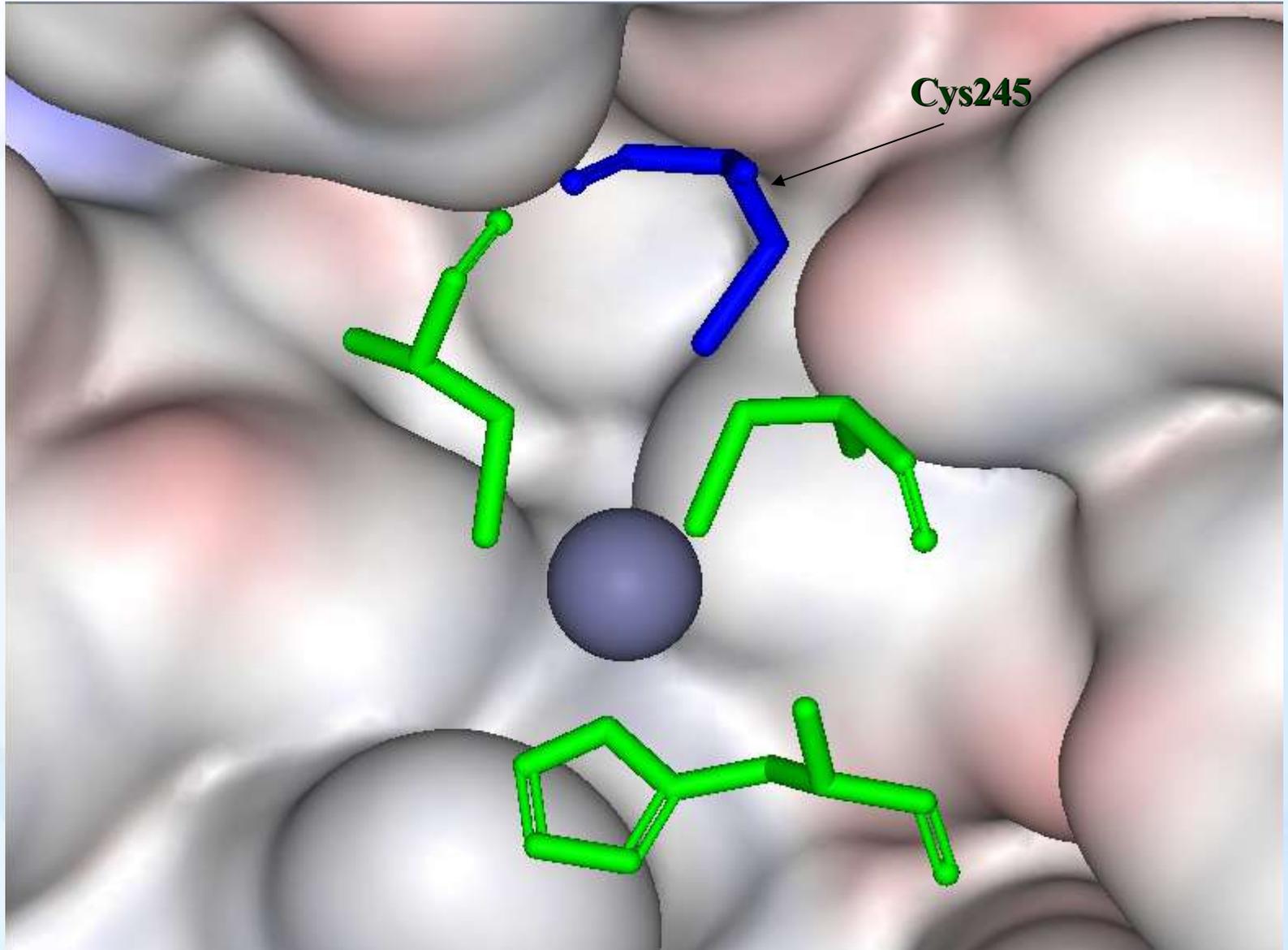


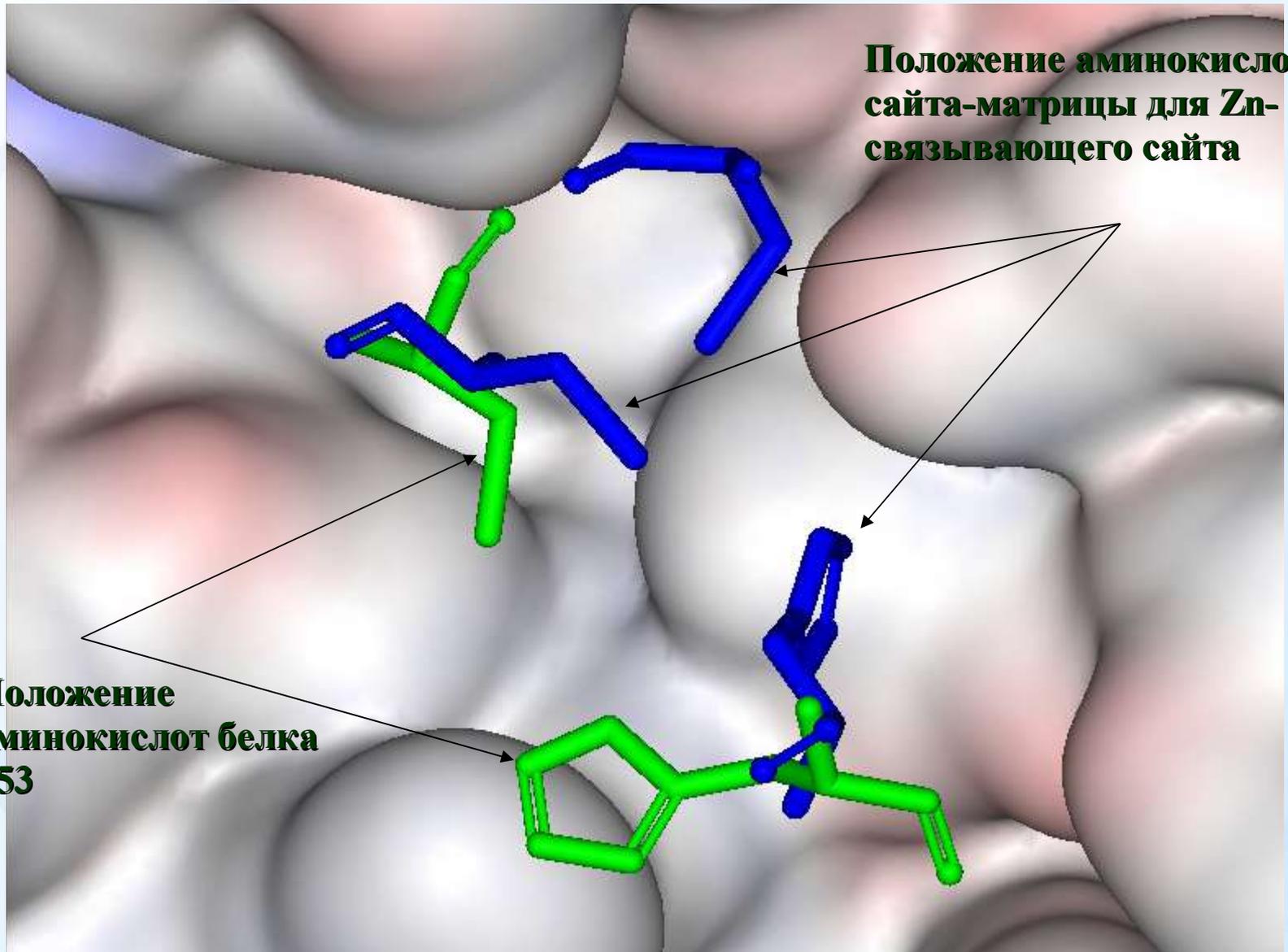
Распределение остатков в пространственной структуре ДНК-связывающего домена белка p53, мутации которых связаны с раковыми заболеваниями

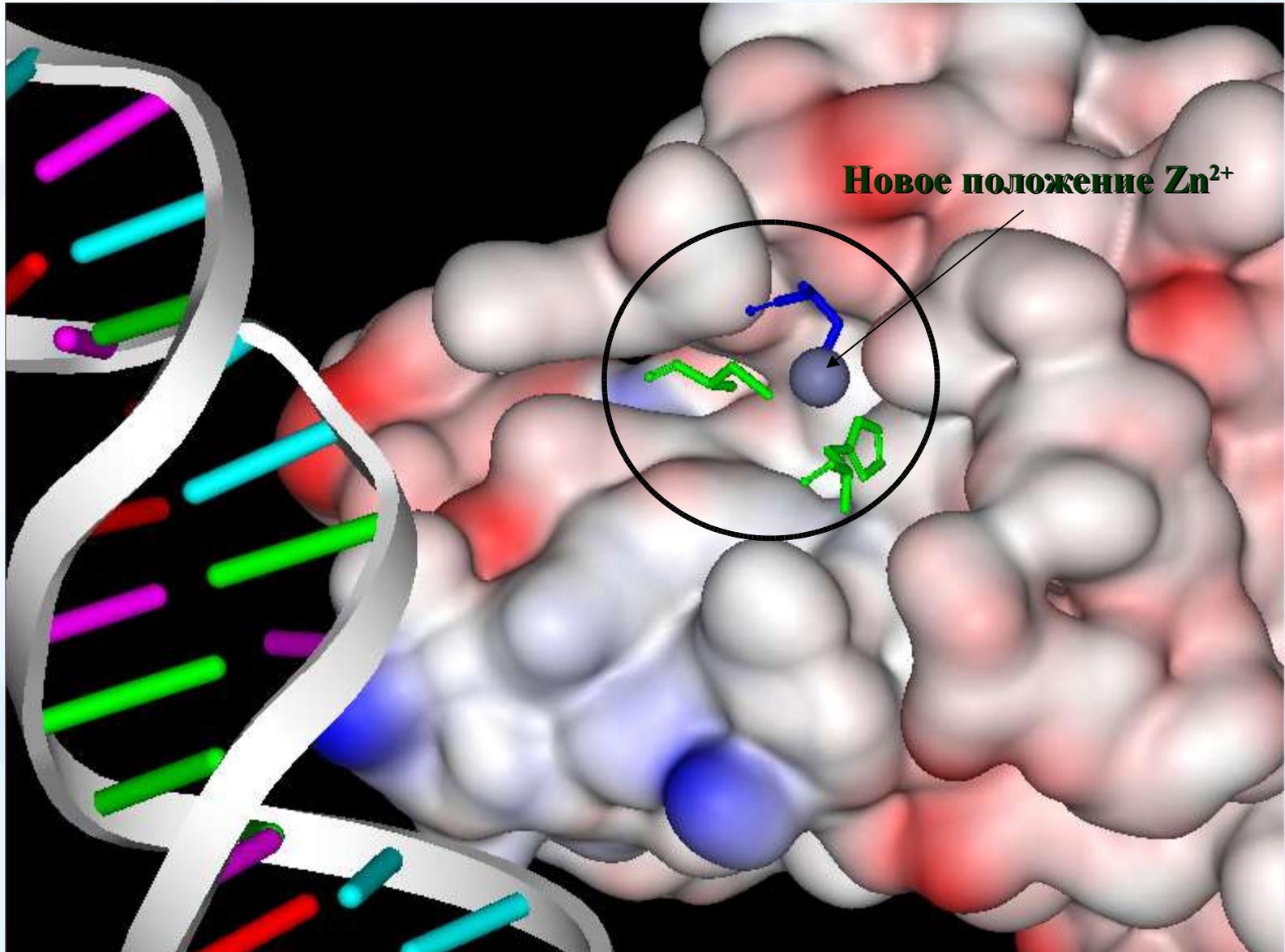


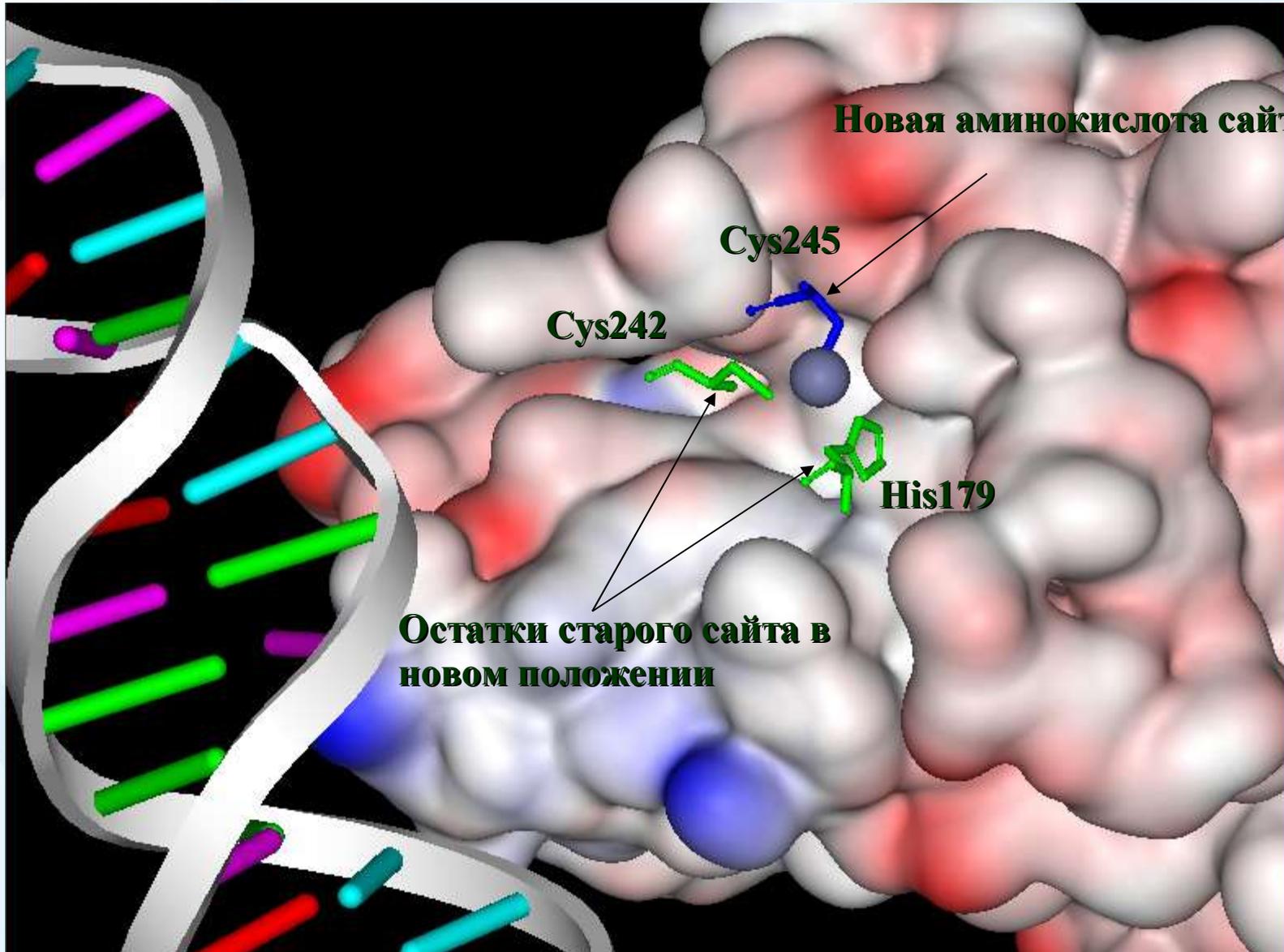


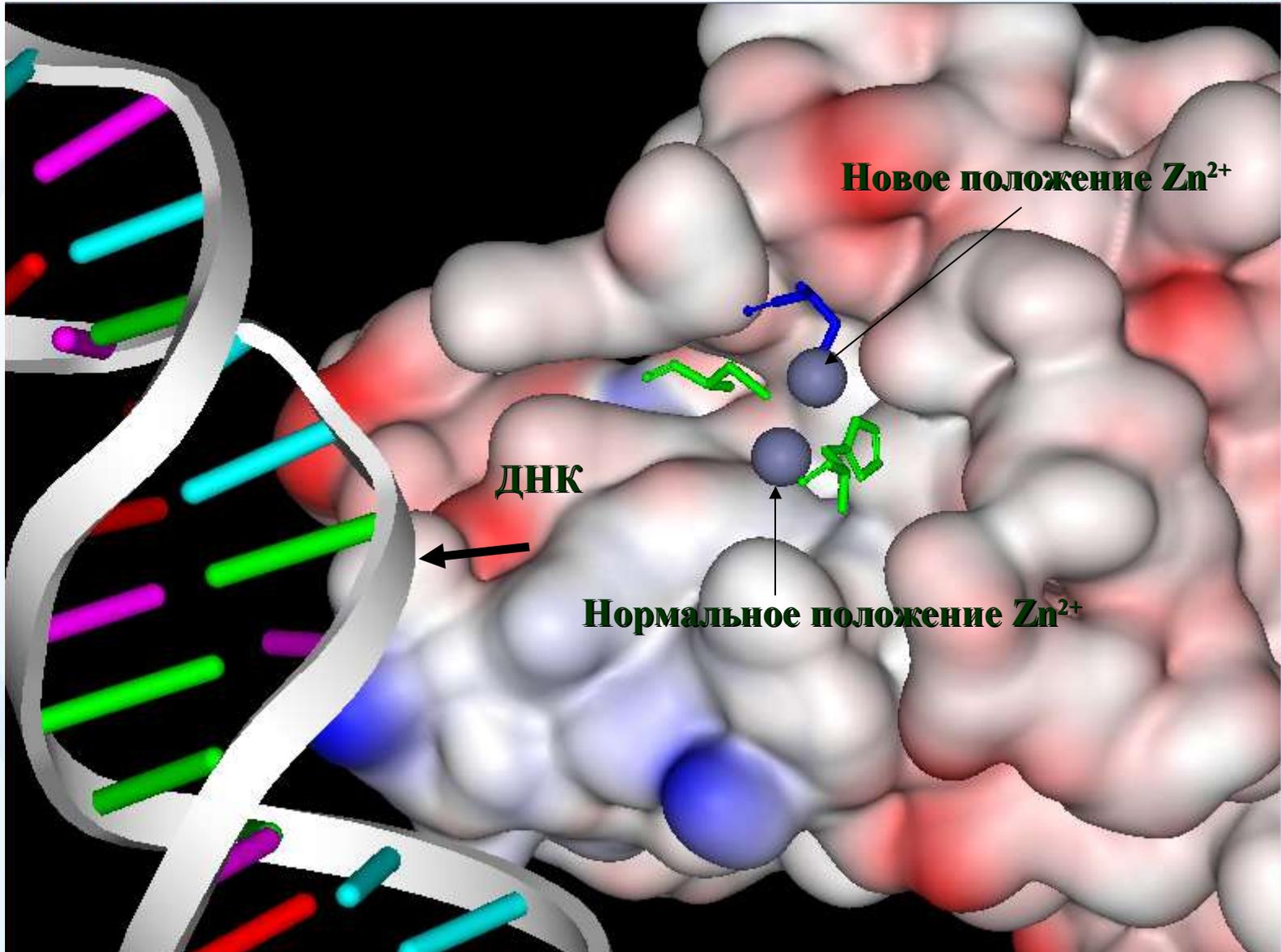


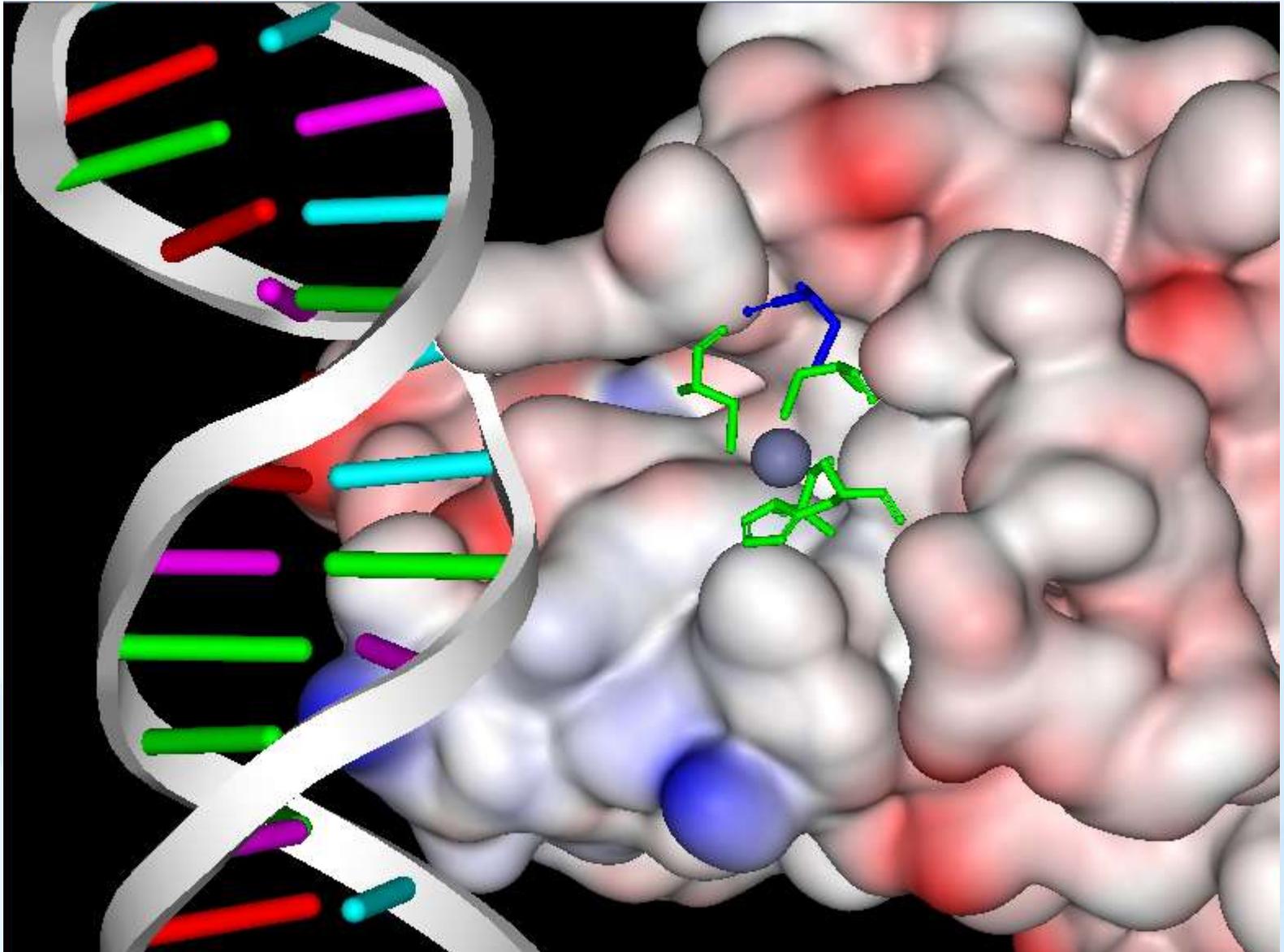


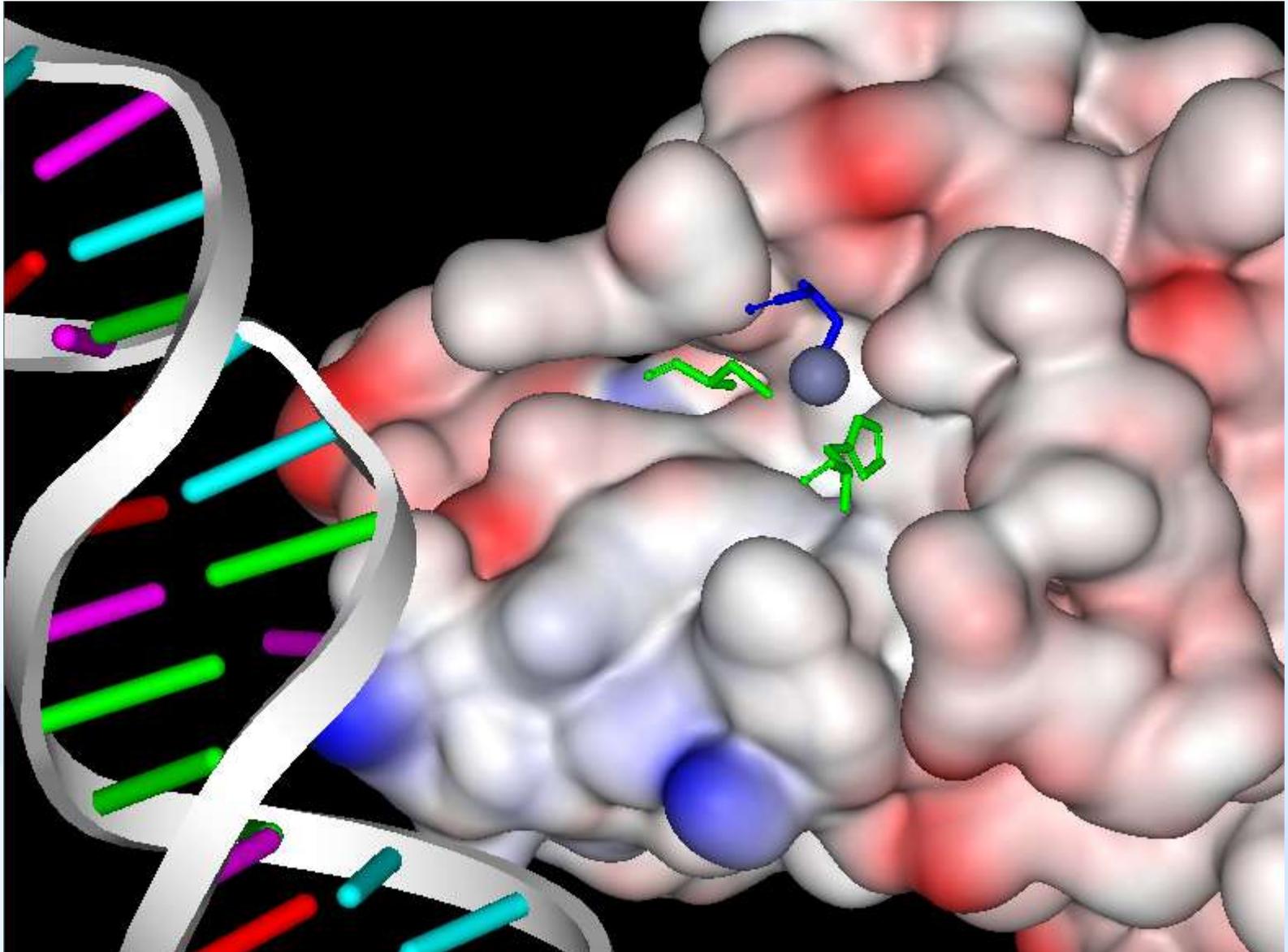


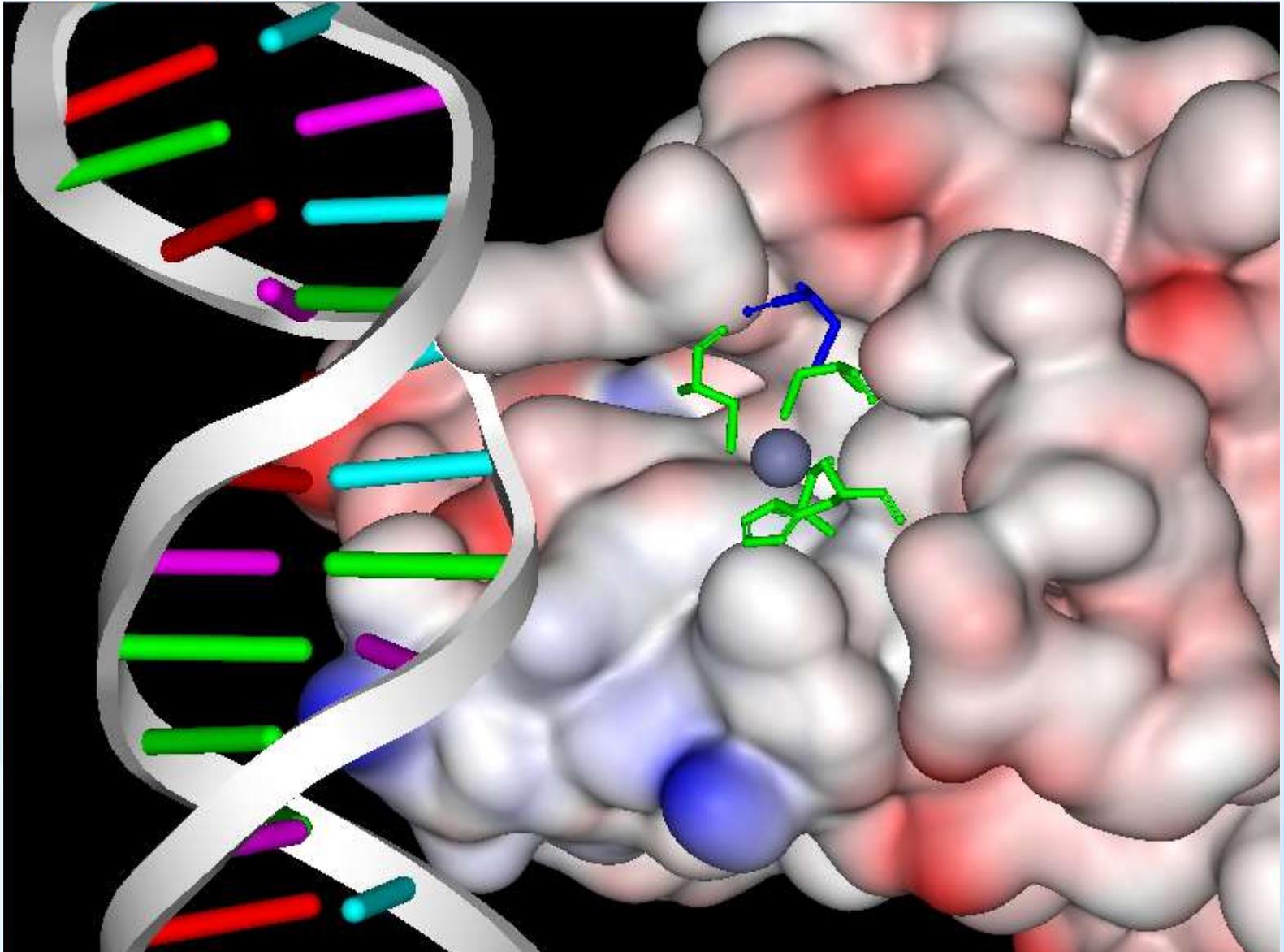


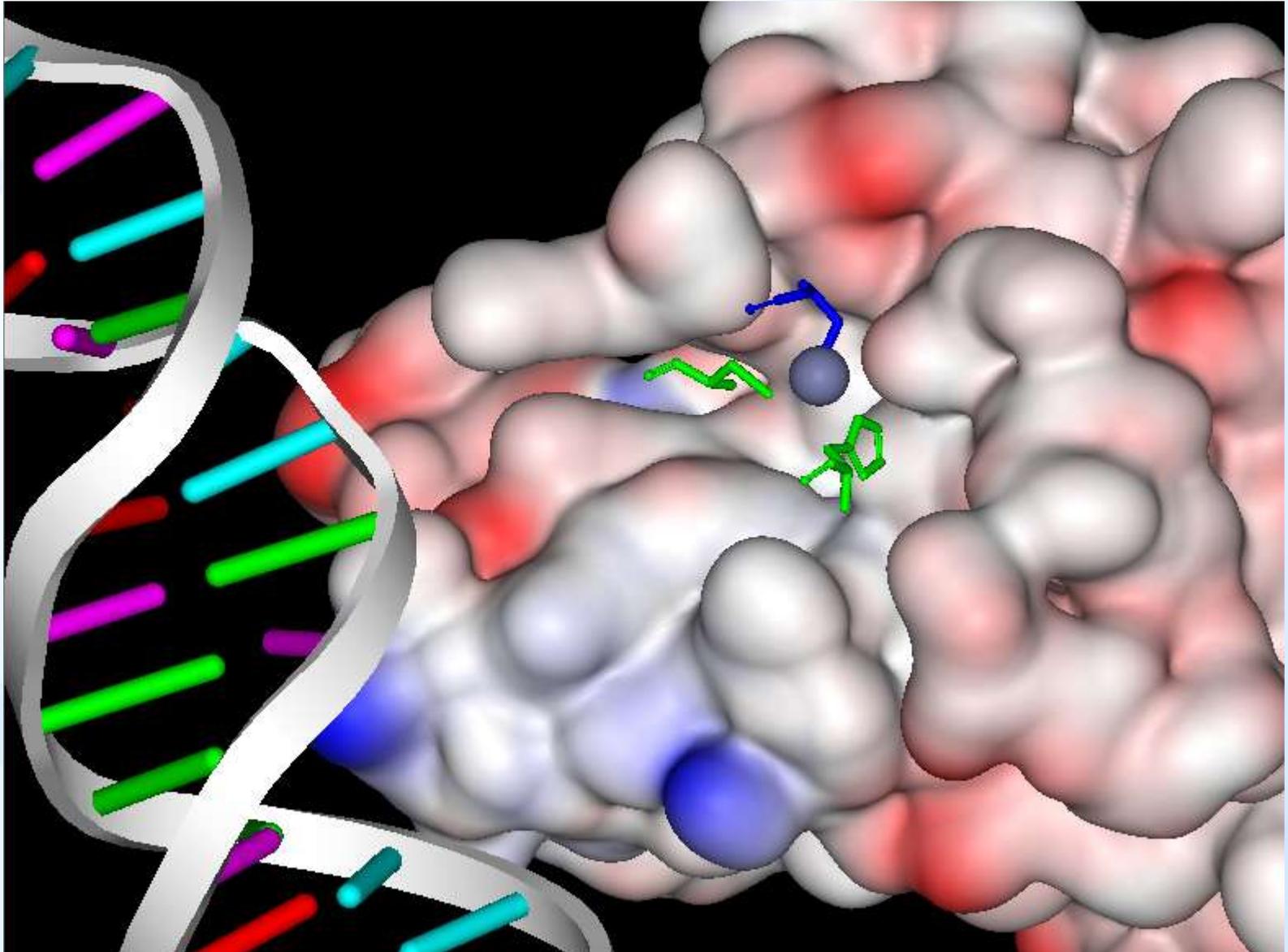


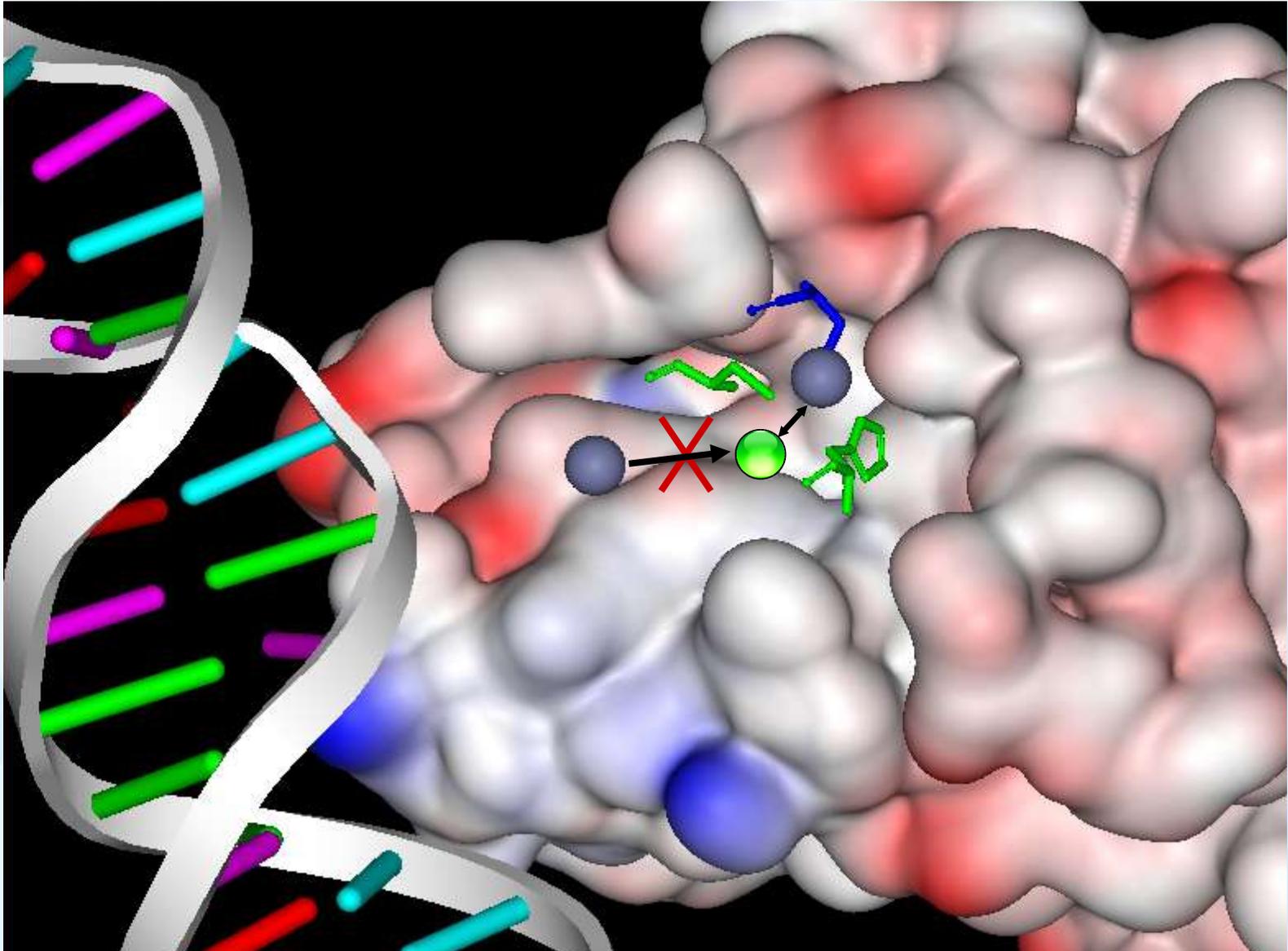






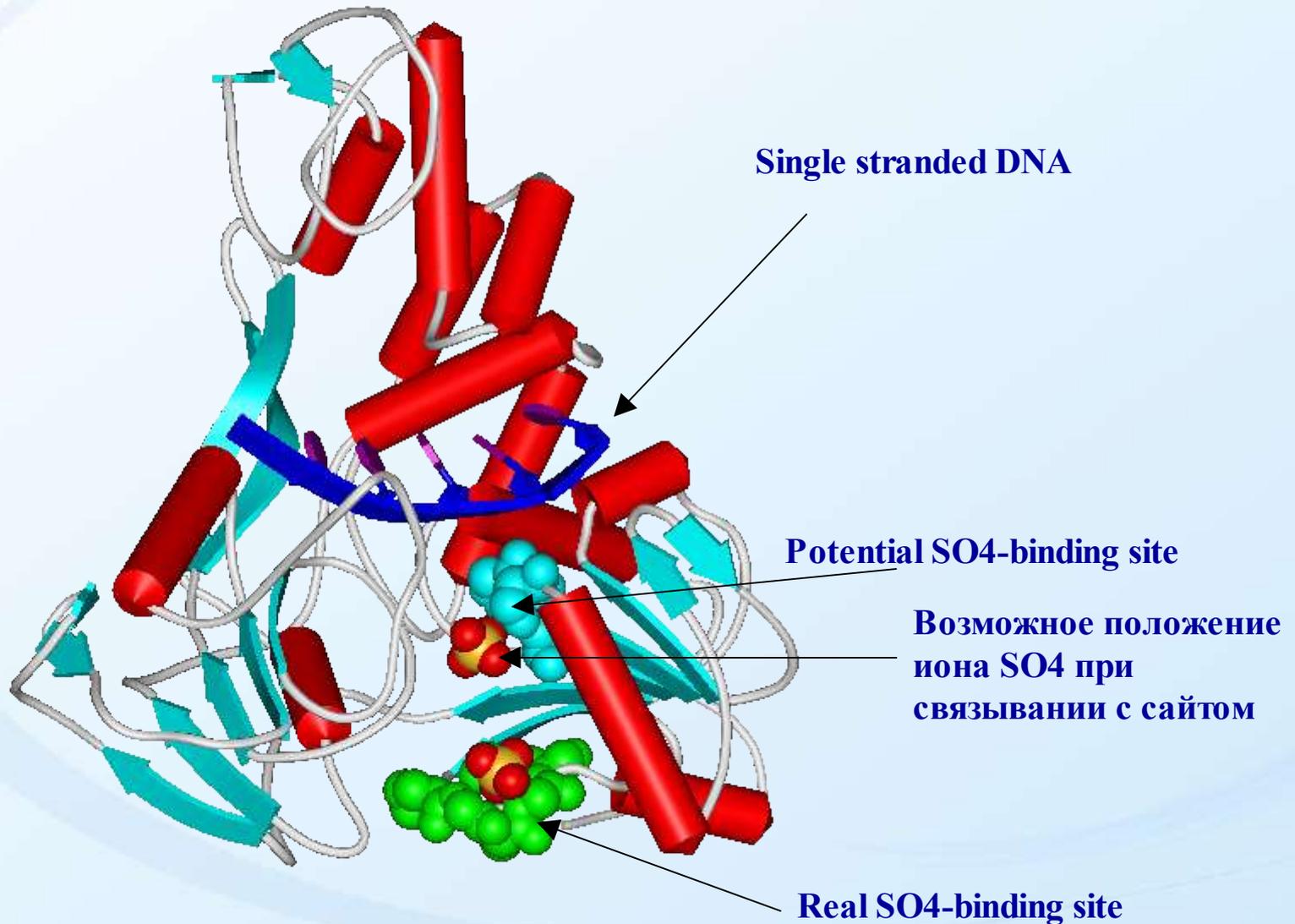






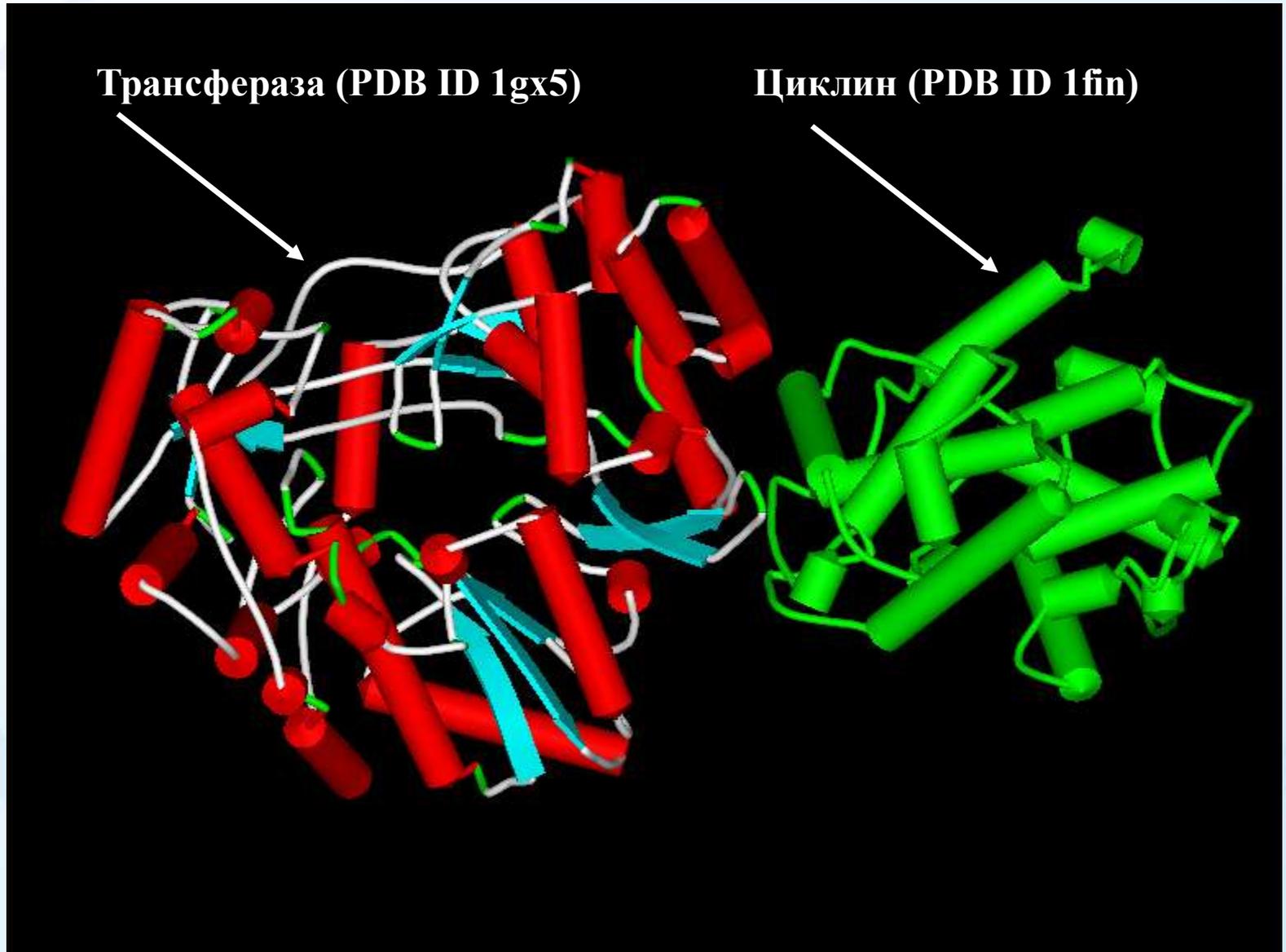


Hepatitis C virus NS3 helicase domain complexed with single stranded DNA



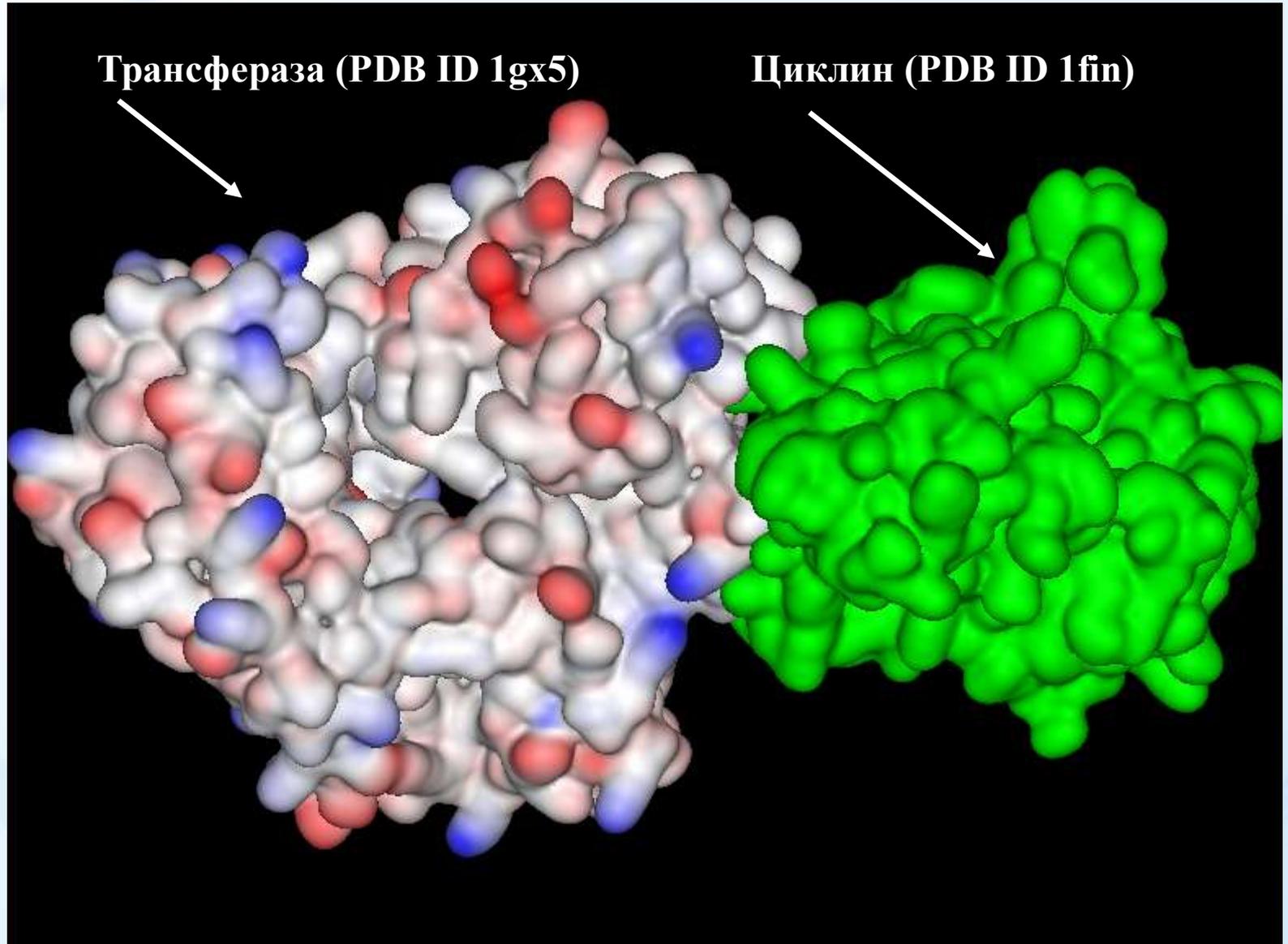


Предсказанное взаимодействие трансферазы вируса гепатита-С с циклином человека





Предсказанное взаимодействие трансферазы вируса гепатита-С с циклином человека





Анализ взаимосвязи структура-активность в наборах гомологичных белков

Иванисенко В.А. к.б.н., с.н.с.





Лекарственный препарат в виде низкомолекулярного ингибитора или модифицированного рекомбинантного белка?

Использование рекомбинантных белков в качестве лекарственных препаратов имеет много привлекательных сторон. В частности, белки обладают многими функциями, поэтому, ингибирование одной функции может приводить к потере белком его остальных функций, а с помощью мутаций можно получить варианты белков с избирательно сниженной активностью для заданной функции.

Кроме того, для борьбы с некоторыми заболеваниями часто стоит задача обеспечения пролонгированного действия того или иного белка. Такой эффект, например, может быть обеспечен путем мутаций, препятствующих быстрому протеолизу. Так мутация Arg31 на Gln в TNF-alpha человека не производит эффекта на цитолитическую активность этого цитокина *in a test in murine fibroblasts L929*, но увеличивает его устойчивость к протеолитическим ферментам (Korepanova A.V. et al, 1994).



Web interface программы WebProanalyst

WebProanalyst - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное Медиа

Адрес: <http://www.mgs.bionet.nsc.ru/mgs/programs/panalyst/> Переход Ссылки

WebProanalyst

Quantitative Structure-Activity Relationship analysis of related protein families

Enter sequence alignment in FASTA format

from Screen (*cut & paste*)...

from File: Обзор...

Protein activity values (for example 1.5 -10 45.66 etc):

Site properties:

- Hydrophobicity (Bogardt)
- Volume (Bogardt)
- Polarity (Bogardt)
- Isoelectric point (Bogardt)
- Charge
- Alpha-helix periodicity of Eisenberg's Hydrophobicity
- Beta-strand periodicity of Eisenberg's Hydrophobicity

Fragment for analysis from to

Готово Интернет

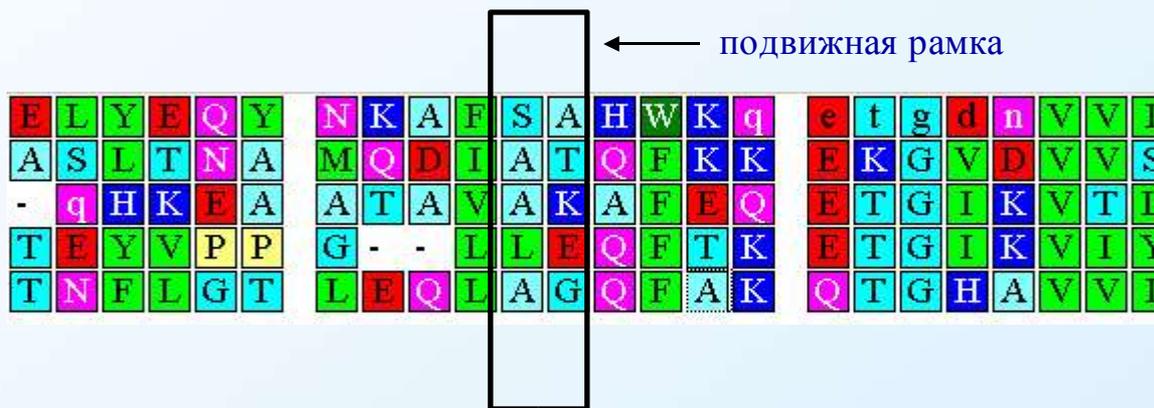


Метод количественного анализа взаимосвязи структура-активность в семействах гомологичных белков, реализованный в WebProanalyst

Множественное выравнивание последовательностей белков

Активности

белок 1
белок 2
белок 3
белок 4
белок 5



y1
y2
y3
y4
y5

сайт

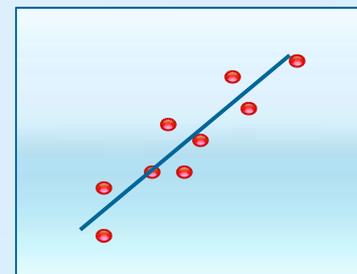
От аминокислот к их физико-химическим свойствам



x1
x2
x3
x4
x5

Поиск статистической зависимости $Y = F(X)$

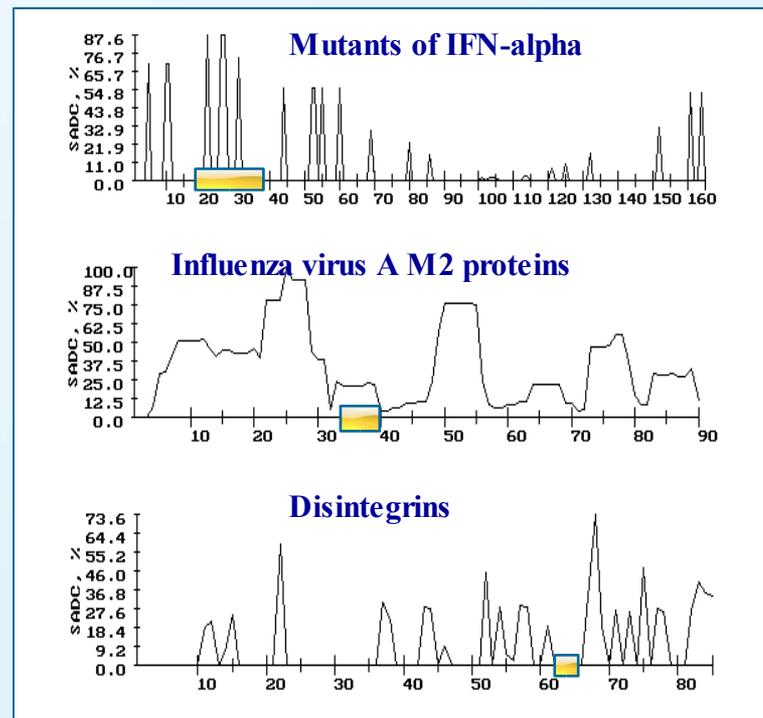
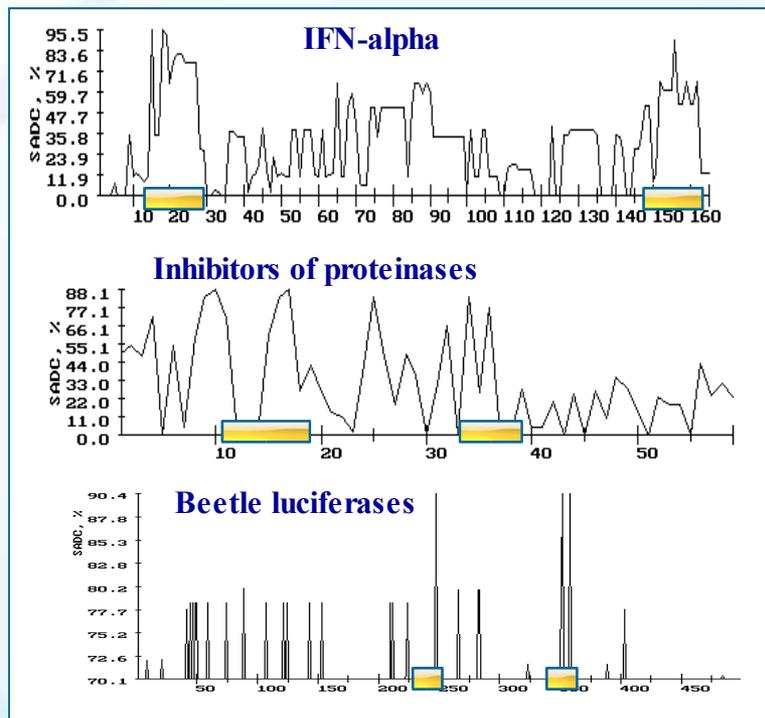
Активности белков



Физико-химическая характеристика сайта



Профили корреляций структура-активность для различных семейств белков



■ - обозначены известные функциональные районы

Для шести белковых семейств показано, что максимально высокие пики на профилях корреляций в основном расположены вблизи функциональных районов, но также существуют и отдаленные от них модулирующие сайты

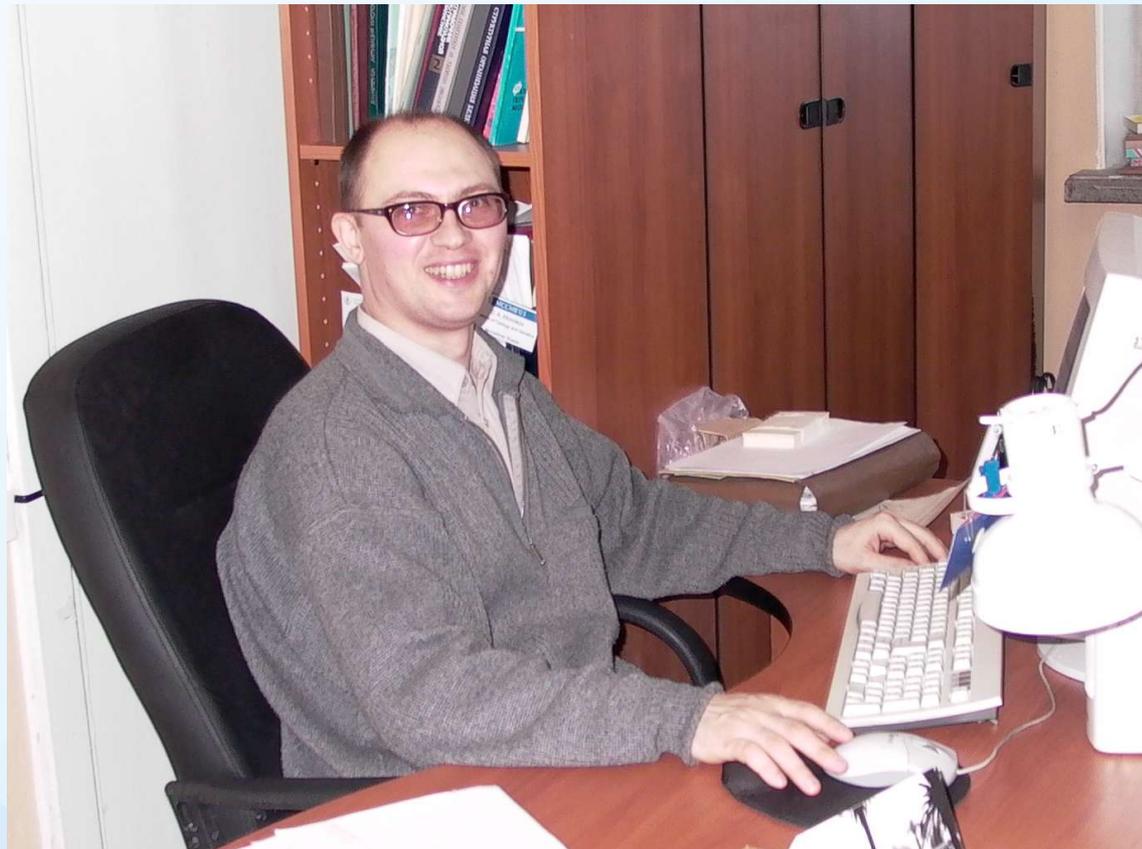


PANCREATIC TRYPsin INHIBITOR (PDB ID 1B0C) из семейства Inhibitors of proteinases





**Анализ координированных замен
аминокислотных остатков.
Афонников Д.А., к.б.н., н.с.**





Для моделирования мутаций в белках, направленно изменяющих их активность, необходимо учитывать сети взаимодействующих остатков, изменения в которых могут сказываться на стабильности белка, а так же влиять на его активность за счет кооперативного эффекта.

Если структура белка известна, то ее анализ позволяет определять такие взаимодействия на уровне структуры.

Если структура неизвестна, то такие взаимодействия приходится определять на уровне последовательностей белка. Однако, анализ на уровне последовательностей может дополнить анализ на уровне структуры.

Одним из подходов является сравнительный анализ наборов последовательностей белков, выполняющих одинаковые функции. Предполагается, что для таких белков критические взаимодействия сохраняются. Поэтому, критические взаимодействия могут быть выявлены на основе анализа консервативных характеристик последовательностей гомологичных белков.



Три режима поддержания критических взаимодействий в белках

LIVKSM DGAL

SIMKCADLIT

GISKNSDQLI

LIMKVVDGYA

Фиксация остатка в позиции белка. Такие остатки являются облигатными, определяют структуру и функцию белка.

LIVKSM DGAL

SVMRCAELIT

GLSHNSDQLI

LIMKVVDGYA

Консервативные замены остатков: сохраняют физико-химическое свойство остатка в одной позиции. Остатки в таких позициях модулируют стабильность или активность белка, определяют его специфичность.

LIVKSM DGAL

STMECARLIT

GTSDNSHQLI

LIMKVVDGYA

Коррелированные замены в группах позиций белка: сохраняют суммарную физико-химическую величину нескольких позиций белка. Остатки в таких позициях модулируют стабильность и активность, стабилизируют структуру активного сайта белка.

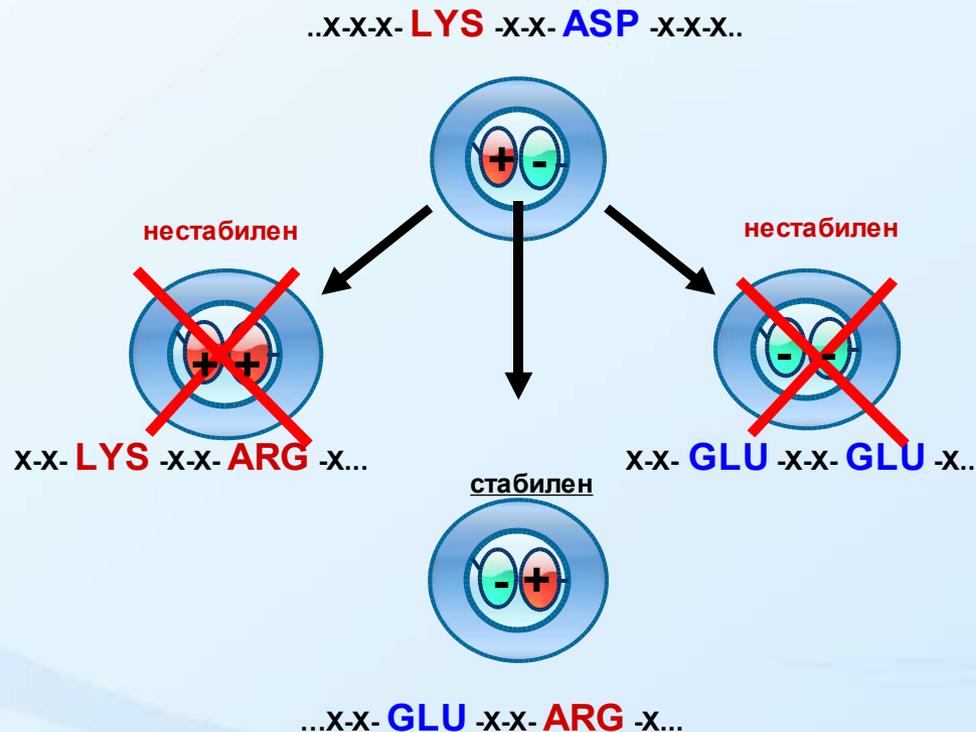


Пример фиксации замен остатков в паре позиций белка, формирующих солевой мостик

LIVKSM^DGAL
STM^ECA^RLIT
GTS^DNS^HQLI
LIM^KV^VDGYA

Анализ множественного
выравнивания
последовательностей
белкового семейства

В результате таких замен
сохраняется суммарный заряд в
паре позиций белка.





Для анализа координированных замен была разработана программа **CRASP**.

Адрес программы <http://www.mgs.bionet.nsc.ru/mgs/programs/crasp/>.

CRASP Start Page - Microsoft Internet Explorer

Файл Правка Вид Переход Избранное Справка

Адрес <http://www.mgs.bionet.nsc.ru/mgs/programs/crasp/> Ссылки

CRASP
Correlation Analysis of Protein Sequences

Correlation analysis of the amino acid substitutions in protein sequences.

[Overview](#)

 [Analysis of pairwise positional correlations](#)

 [Analysis of protein integral physico-chemical characteristics](#)

 [Result examples](#)

 
Molecular Biological Server

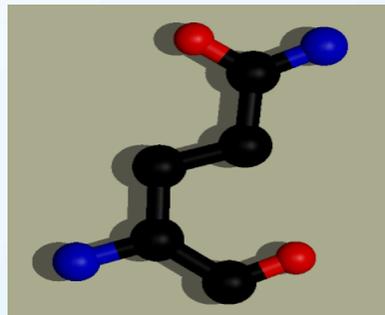
This resource has been developed in Institute of Cytology and Genetics. Novosibirsk, Russia

Пуск Windows Comman... Microsoft Word - R... Microsoft PowerPo... CRASP Start P... Зона Интернета 12:05



Выбор меры зависимости аминокислотных замен в позициях белка

Использовались величины физико-химических свойств аминокислот, поскольку они отражают различные типы взаимодействий остатков



Гидрофобность

Предпочтение ко вторичной структуре

Размер

Заряд

В качестве меры зависимости позиций белка использовались коэффициенты корреляции между величинами свойств в позициях множественного выравнивания



Пакет программ CRASP: основные возможности

Возможности анализа:

- Оценка парных корреляций
- Кластерный анализ
- Анализ интегральных характеристик

Набор выравненных последовательностей

Выбор физико-химической характеристики остатков

Оценка парных коэффициентов корреляции величин физико-химического свойства

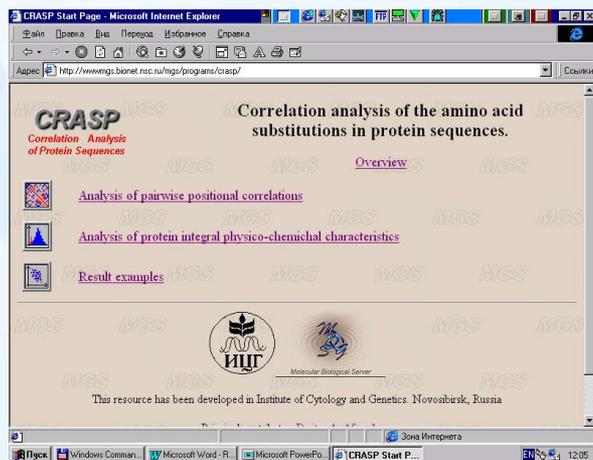
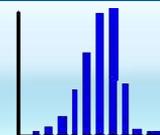
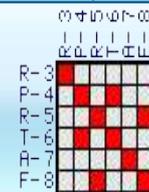
Анализ матрицы парных коэффициентов корреляции. Выбор интегральных характеристик белка

Анализ вклада координи-рованных замен в посто-янство интегральной характеристики

```
ASDRTFNHNKLIYRRMNL-  
GAAQWCCV-RHNGTYKLS  
DFGSSR---WWRQACHMP  
DFGSSR---WWRQACHMP
```

```
+4.0;+6.2;-1.8;  
+5.9;+3.2;+5.6;  
-3.4;-2.1;+4.0;  
-6.9;+2.3;+5.0;
```

$$r_{ij} = \frac{S_{ij}}{\sqrt{S_{ii}S_{jj}}}$$





Для анализа используются данные в виде наборов выровненных последовательностей белков.

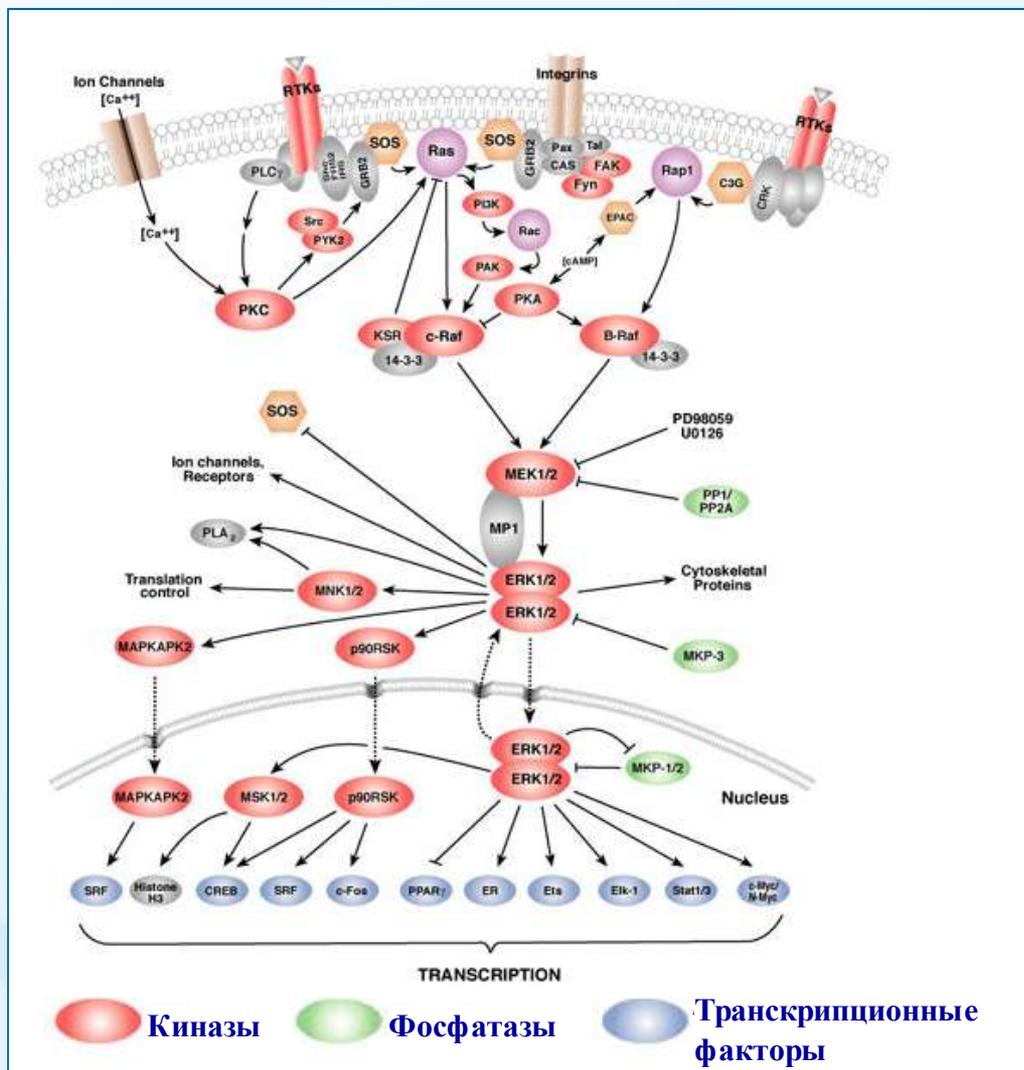
Эволюционные данные -последовательности выровненных гомологичных белков.

Экспериментальные данные – последовательности полученные в результате экспериментов по искусственной селекции.



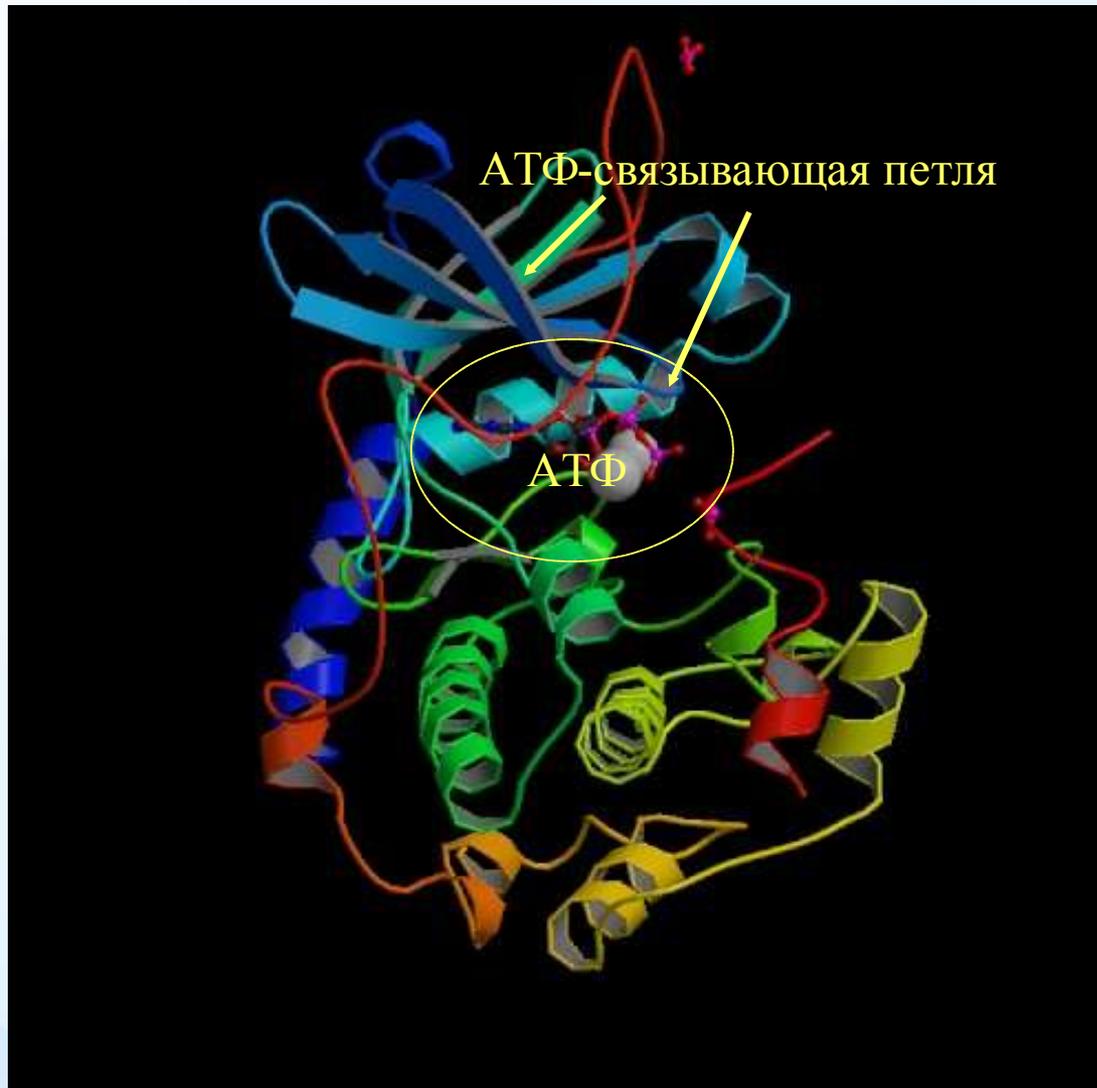
Пример анализа – каталитический домен протеин-киназ.

Протеин-киназы участвуют в передаче сигналов в клетках и являются мишенями для лекарственных препаратов. Ниже приведена схема контроля роста клетки, в котором принимают участие киназы.





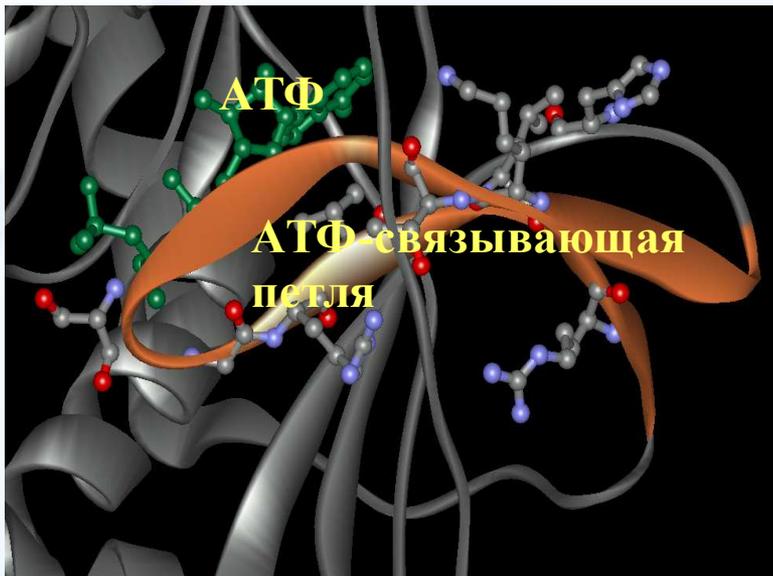
Основной мишенью при разработке ингибиторов протеин-киназ является сайт связывания АТФ.



В связывании с АТФ принимает участие АТФ-связывающая петля



Анализ вклада координированных замен в постоянство интегральных характеристик АТФ-связывающей петли протеин-киназ



Для анализа использовалось множественное выравнивание 388 гомологичных протеин-киназ. Анализировались суммарные значения физико-химических свойств остатков, формирующих петлю. Оценивался вклад координированных замен в консервативность этих свойств.

Было проанализировано 20 физико-химических характеристик аминокислотных остатков, включающих гидрофобность, объем, заряд, вероятность встраивания во вторичную структуру, полярность и др. (база данных AAIndex (Tomii and Kanehiza, 1996)).

Консервативность свойства оценивалась параметром λ (отношение варибельности в приближении независимых замен остатков к реальной варибельности свойства). Если $\lambda > 1$, то физико-химическое свойство является консервативным и определяет ключевые взаимодействия петли.



Анализ вклада координированных замен в постоянство интегральных характеристик АТФ-связывающей петли протеин-киназ

В таблице приведены наиболее консервативные физико-химические свойства остатков АТФ связывающей петли протеин-киназ. Приведены оценки уровней значимости консервативности этих свойств и значения параметра λ

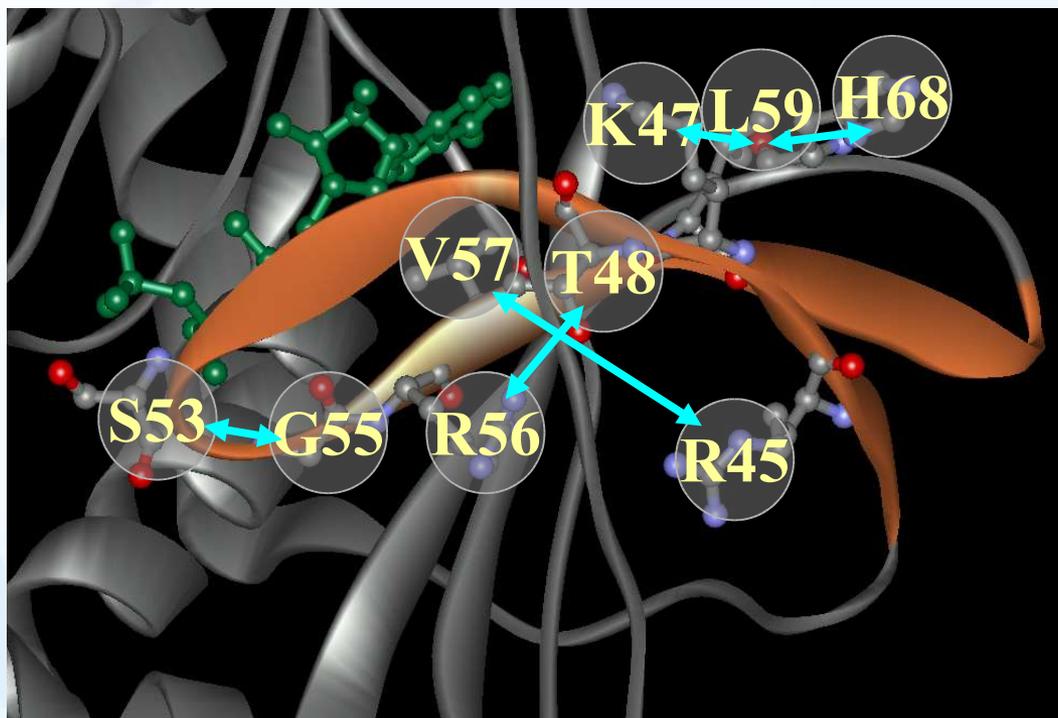
Физико-химическое свойство	λ	$p <$
Hydrophobicity (Eisenberg et al., 1984)	1.45	0.00000
Isoelectric point (Zimmerman et al., 1968)	1.34	0.00003
Normalized frequency of beta-sheet with weights (Levitt, 1978)	1.28	0.00033

Среди наиболее консервативных оказались свойства остатков, характеризующие их гидрофобность, заряд (изоэлектрическая точка) и вероятность встраивания в β -структуру.

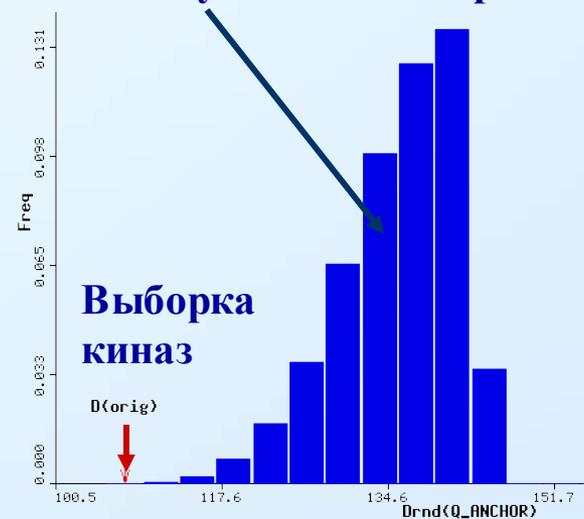
Вывод: для АТФ-связывающей петли наиболее важны гидрофобные, электростатические и β -стабилизирующие взаимодействия.



Пример: анализ вклада координированных замен в постоянство суммарного заряда АТФ-связывающей петли



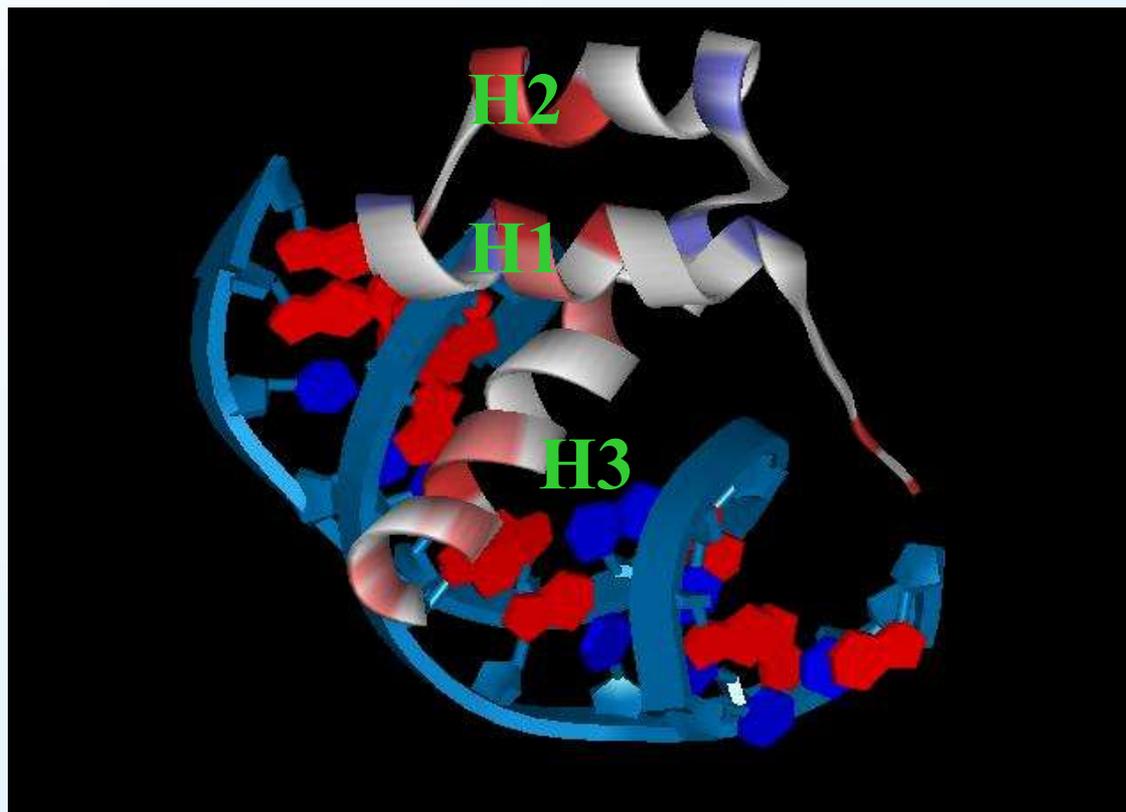
Дисперсия суммарного заряда петли в случайных выборках



На рисунке слева показаны остатки АТФ-связывающей петли киназы сАРКа, замены в парах которых происходят компенсаторным образом по отношению к заряду (значимые отрицательные корреляции заряда показаны голубыми стрелками). Показано, что в реальных белках дисперсия суммарного заряда петли значительно меньше, чем в сгенерированных случайных выборках с отсутствием корреляций (признак постоянства заряда).



Исследование координированных замен в ДНК-связывающих доменах класса «гомеодомен»

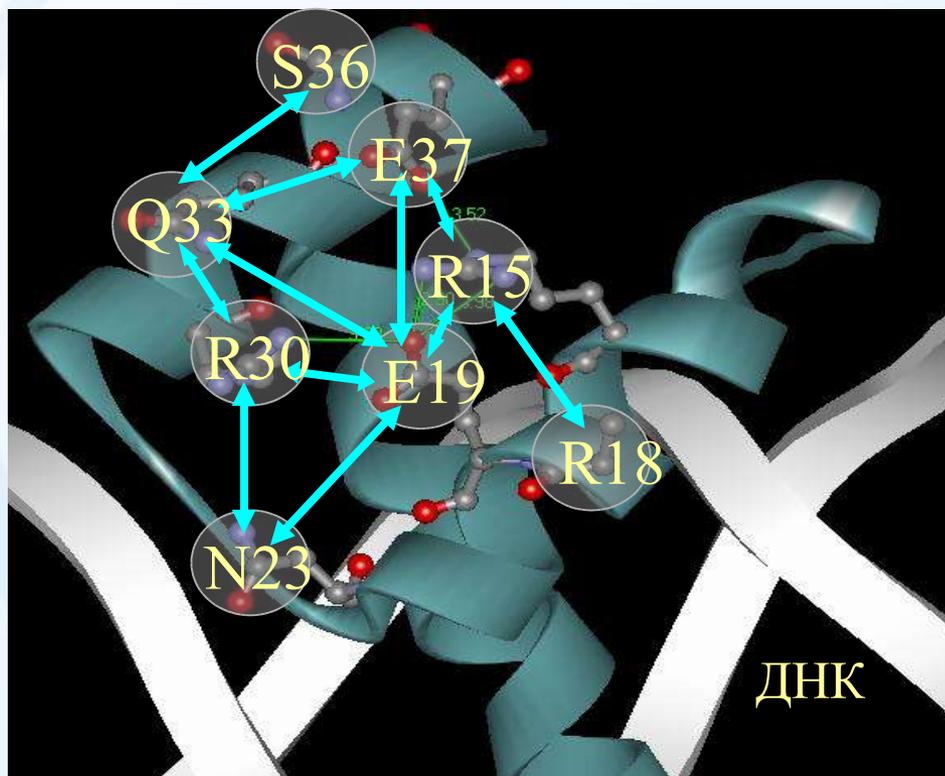


Выборка составила 372 последовательности (из базы Pfam). Филогенетическое дерево построено программой CLUSTAL. Анализировались величины изоэлектрической точки.

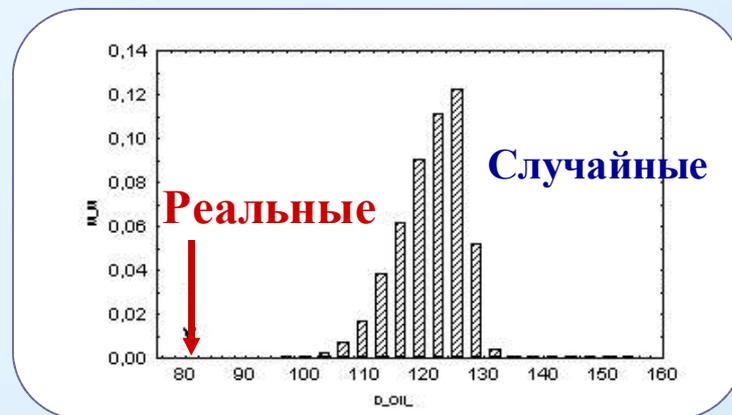


Анализ вклада координированных замен в постоянство суммарного заряда в районе контакта спиралей H1,H2

Для остатков, формирующих контакт между спиральями H1,H2, были выявлены компенсаторные по величине заряда замены (отрицательные парные корреляции показаны голубыми стрелками). Пары R30-E19 и R15-E37 формируют солевые мостики.



Такие замены обеспечивают постоянство суммарного заряда в районе контакта спиралей.



Распределение дисперсии суммарного заряда в случайных выборках и в реальной выборке ($P=1-10^{-5}$)

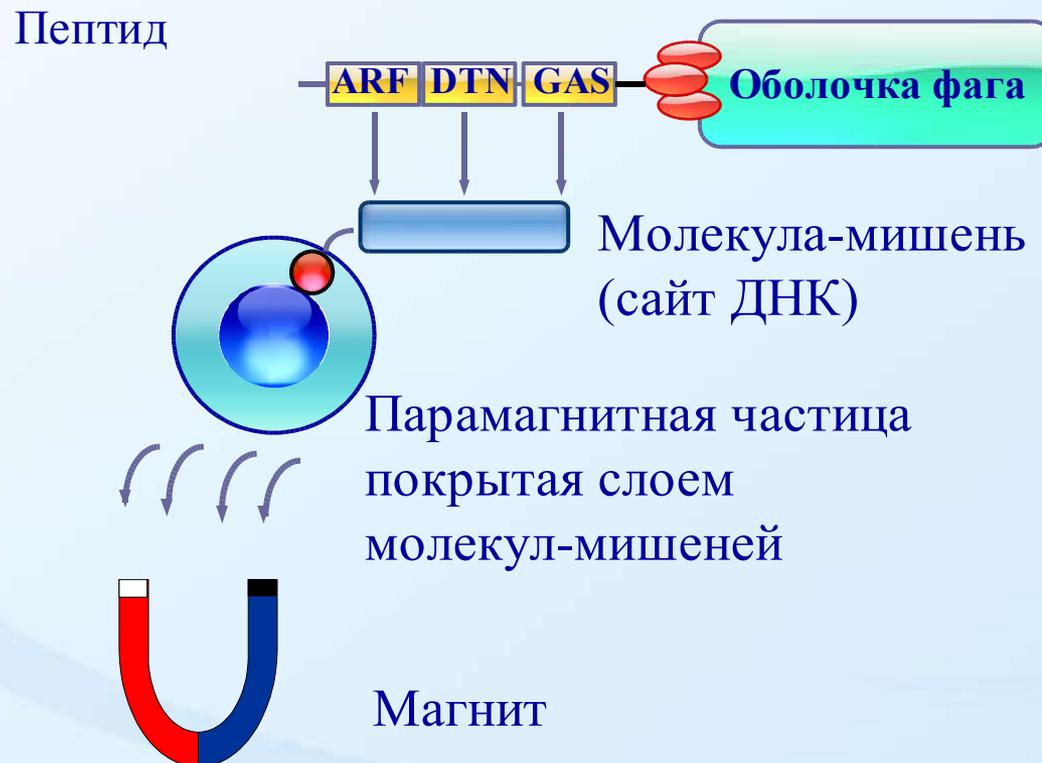
Вывод: координированные замены вносят значимый вклад в постоянство суммарной величины заряда остатков в районе контакта спиралей H1,H2.



Анализ особенностей *in vitro* эволюции пептидов и белков

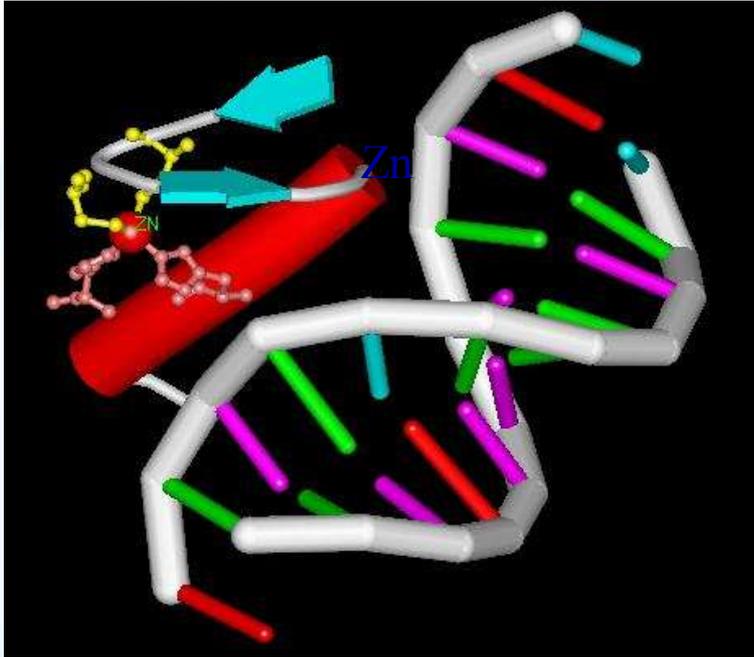
Одной из перспективных методик биотехнологии, в том числе и дизайна лекарств, является метод фагового дисплея, позволяющего проводить искусственную селекцию пептидов и белков.

Схема эксперимента по селекции *in vitro*

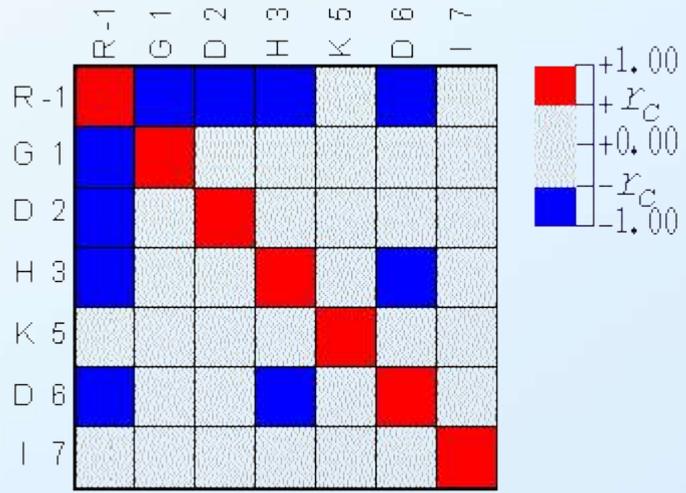




Корреляционный анализ цинковых пальцев C2H2



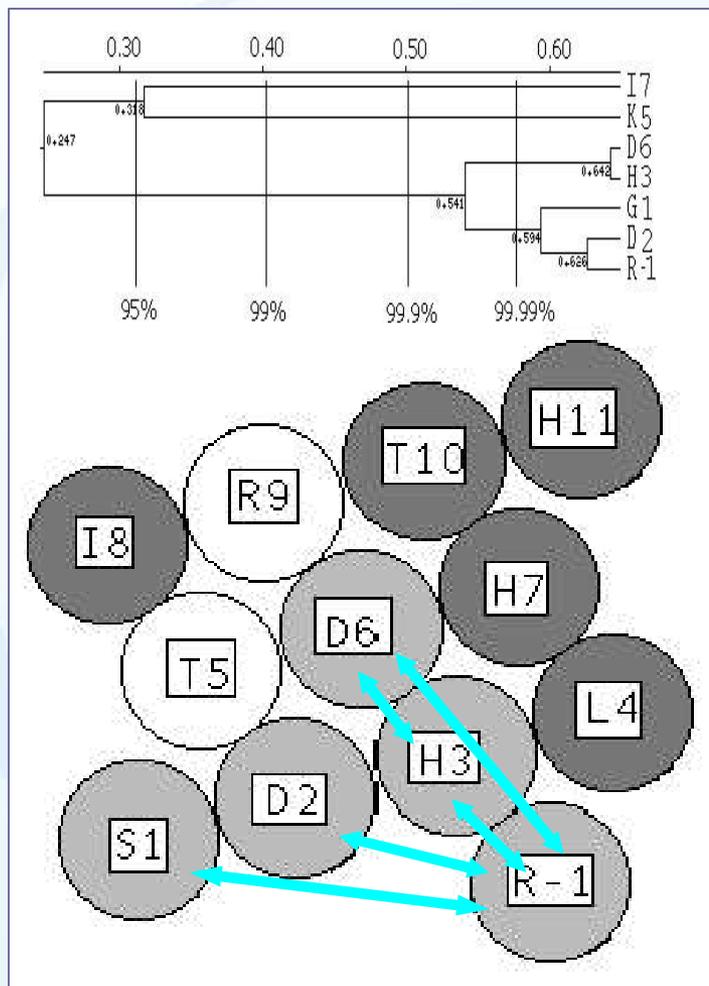
Матрица координированных замен по величине изоэлектрической точки



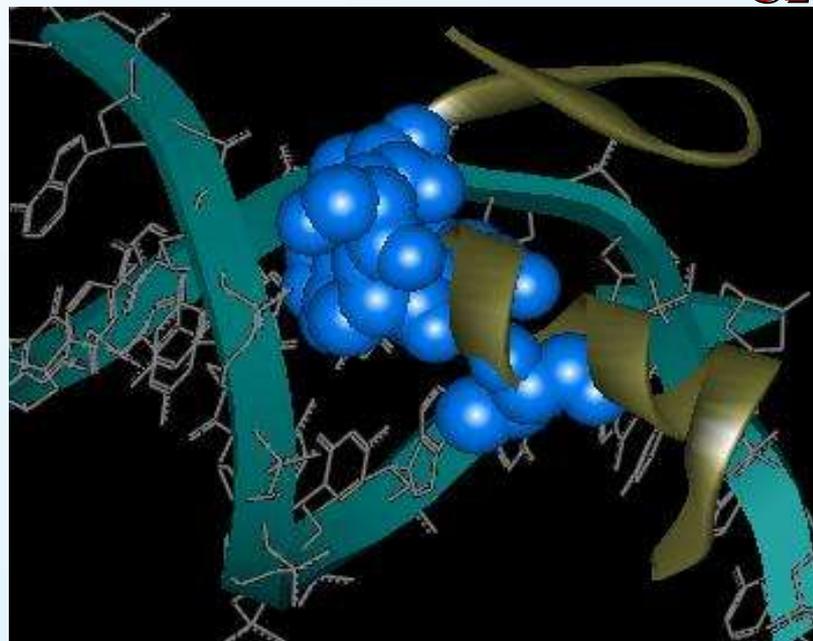
Partial Correlation $r_c = 0.40$
Isoelectric point

45 последовательностей были взяты по результатам эксперимента фаговой селекции Choo & Klug.
Нумерация позиций проведена по отношению к первому остатку альфа-спирали.

Сеть корреляций (по отношению к заряду остатков) в ДНК-связывающем домене класса «цинковый палец C2H2»



Расположение остатков, формирующих сеть на поверхности альфа-спирали (показана цилиндрическая развёртка спирали)

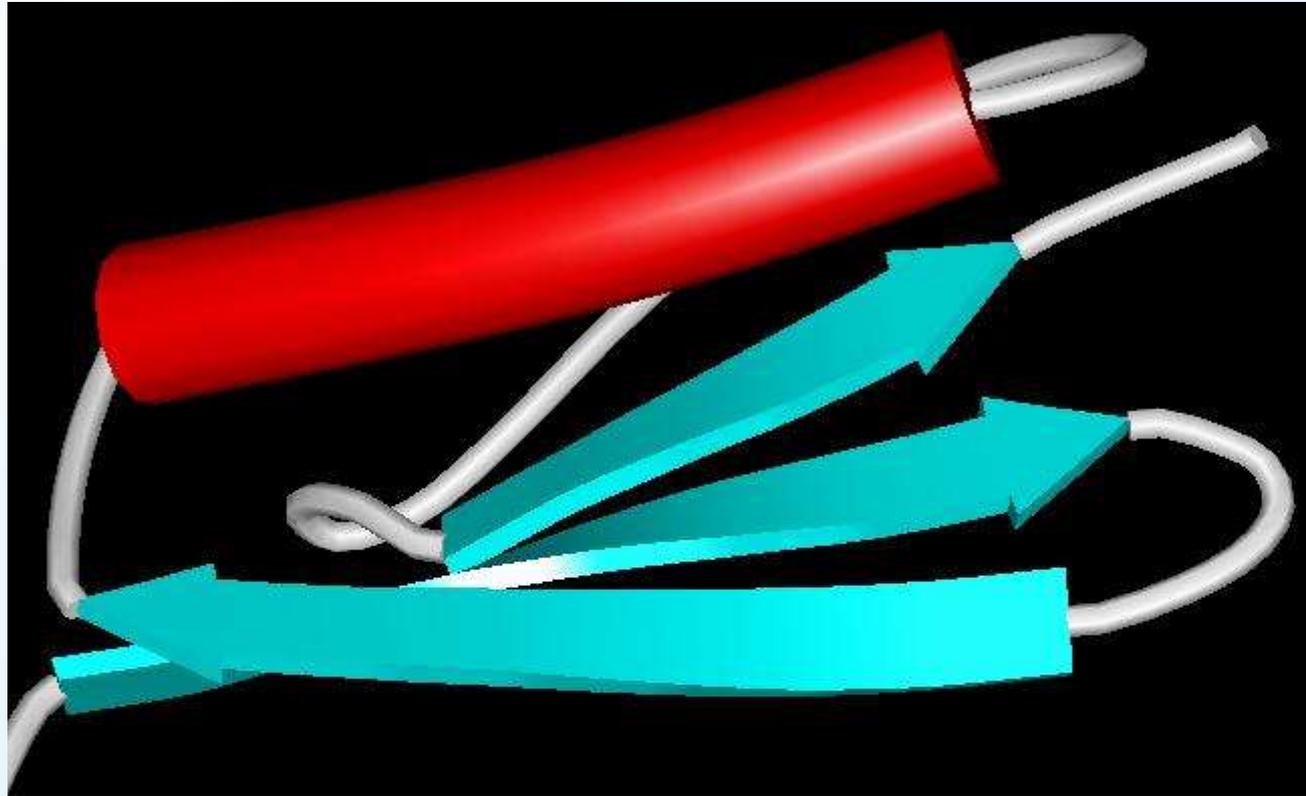


Расположение остатков кластера в пространственной структуре цинкового пальца. Коррелированные замены таковы, что для данных остатков суммарная величина заряда является консервативной.

Возможная роль величины суммарного заряда: стабилизация неспецифических взаимодействий ДНК и белка.



Анализ иммуноглобулин-связывающего домена протеина L



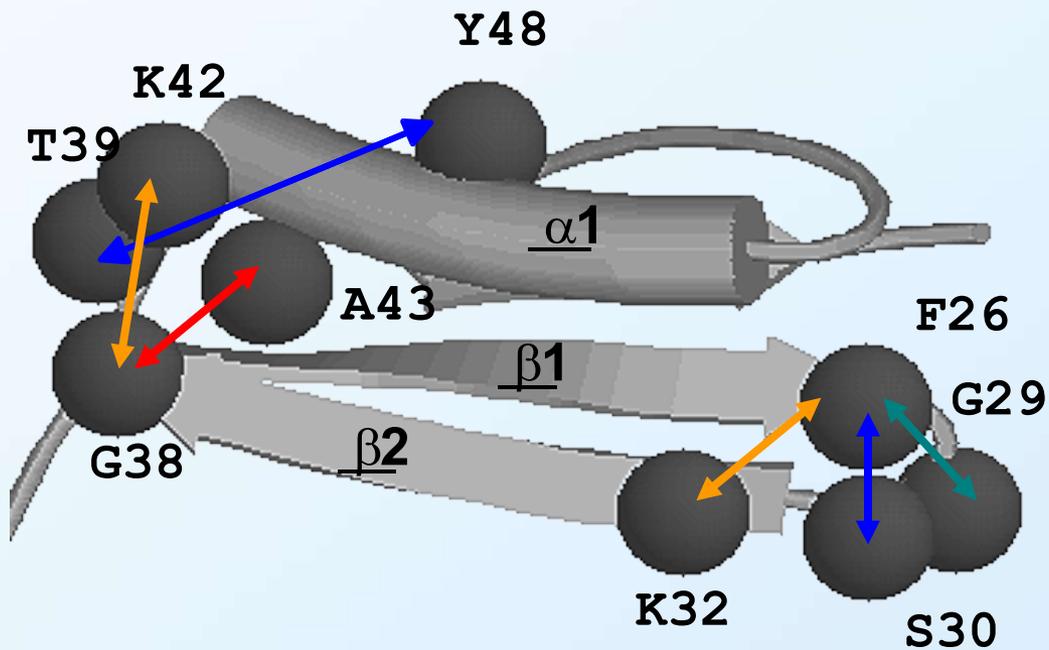
Отбор последовательностей в эксперименте по искусственной селекции производился по критерию быстрой и правильной укладки белка. При анализе корреляций отобранных белков использовались два набора выравниваний из работ:

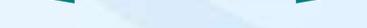
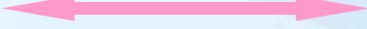
- Gu H., Yi Q., Bray S.T., Riddle D.S., Shiau A.K., Baker,D. Protein Sci. 1995. V. 4, P. 1108-1117.
- Kim D.E., Gu H., Baker D., Proc. Natl. Acad. Sci. USA, 1998. V. 95, P. 4982-4986.

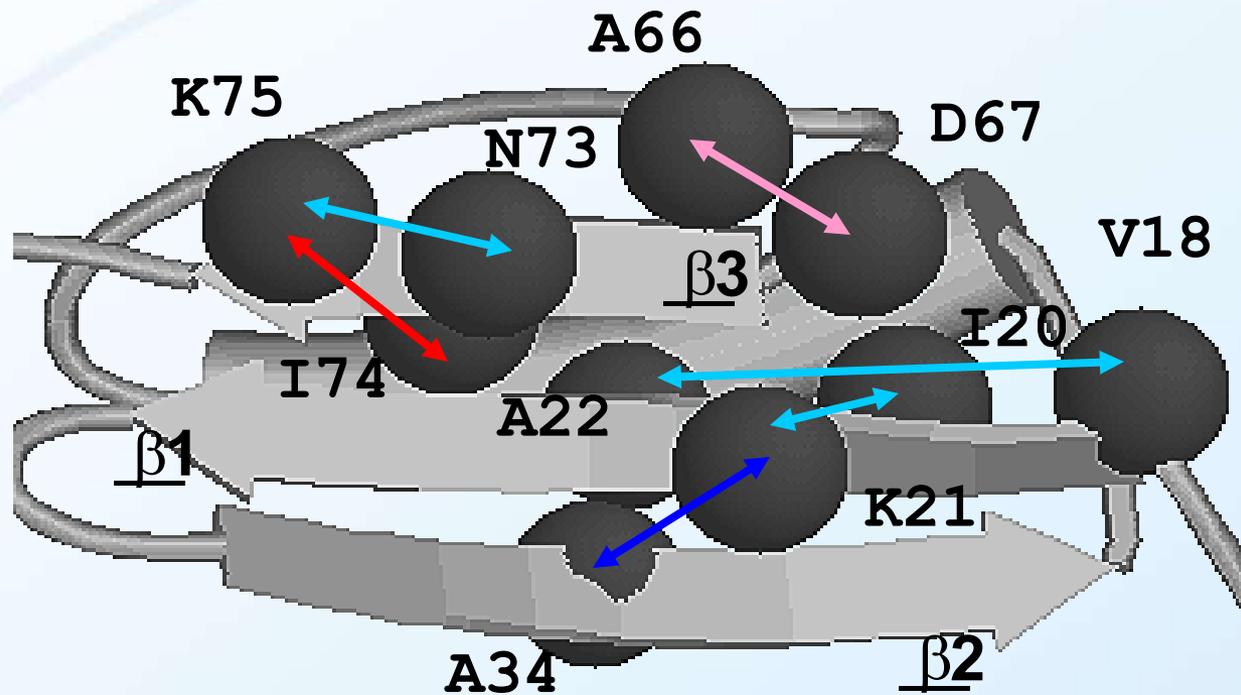


Координированные замены остатков в иммуноглобулин-связывающем домене

Анализовались корреляции величин объема боковых групп, изоэлектрической точки, полярности и гидрофобности.



-  Изоэлектрическая точка, отрицательные
-  Изоэлектрическая точка, положительные
-  объем, отрицательные
-  гидрофобность, отрицательные
-  Гидрофобность, положительные
-  Полярность, положительные



Функциональная роль этих взаимодействий может заключаться в стабилизации структуры белка и определять взаимную укладку элементов вторичной структуры белка.



Массовый анализ последовательностей, полученных в экспериментах по искусственной селекции



ASPD (Artificial Selected Proteins/Peptides Database) is a curated database on selected from randomized pools proteins and peptides. Database access is realised by means of SRS system (Sequence Retrieval System). ASPD is integrated by means of hyperlinks with different databases (SWISS-PROT, PDB, PROSITE ...).

ACCESS to ASPD SRS ACCESS: [ASPD_ALIGN](#) [ASPD_REF](#)
[Blast search ASPD database](#)
[Data submission to ASPD](#)

General information How to cite ASPD? Contact us	General information How to cite ASPD? Contact us	User's guide Brief manual on the database ASPD
User's guide Brief manual on the database ASPD	Current release ASPD is updated on a regular basis. The current release has 103 entries and was indexed 06-Oct-2000.	Additional information Blast search ASPD database Links to other databases and programs ASPD substitution matrix Correlations values of the ASPD substitution matrix with other matrices ASPD amino acid composition
Current release Blast search ASPD database Links to other databases and programs		
Additional information Blast search ASPD database Links to other databases and programs		

Проведен массовый анализ результатов экспериментов по искусственной селекции пептидов и белков.

Выравнивания взяты из базы данных ASPD (Valuev *et al*, 2001). Рассчитывались корреляции по гидрофобности, полярности, объему и изоэлектрической точке. Проанализировано 153 записей из базы данных ASPD. Число проанализированных пар позиций составило 6308 для шкалы изоэлектрической точки и 6334 для трех других характеристик. Результаты анализа в виде электронной коллекции вошли в базу данных ASPD.



Анализ режимов координированных замен в экспериментах по искусственной селекции

Рассматривалось два режима замен: компенсаторный (если коэффициент корреляции $r < 0$) и антикомпенсаторный ($r > 0$). В качестве характеристики режима эволюции рассматривалось отношение доли значимых положительных коэффициентов корреляции к доле значимых отрицательных для интервалов значимости n_{r+}/n_r (при нескольких уровнях P):

P1: $r_c(P=95\%) < r < r_c(P=99\%)$;

P2: $r_c(P=99\%) < r < r_c(P=99.9\%)$;

P3: $r_c(P=99.9\%) < r < r_c(P=99.99\%)$;

P4: $r_c(P=99.99\%) < r$

Вывод: координированные замены являются преимущественно компенсаторными по объему (V) и заряду (pI) и антикомпенсаторными по гидрофобности (H) и полярности (P).



Полученные зависимости могут отражать специфику взаимодействий остатков в белках.