

ВВЕДЕНИЕ В БИОИНФОРМАТИКУ

Лекция №13

Структурная биоинформатика

Новоселецкий Валерий Николаевич
к.ф.-м.н., доц. каф. биоинженерии
valery.novoseletsky@yandex.ru

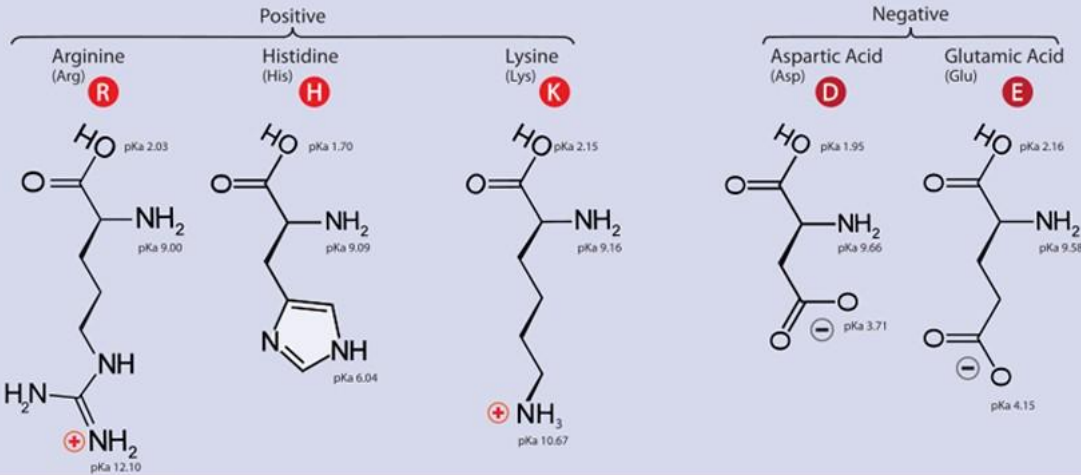
Сайт курса <http://intbio.org/bioinf2018-2019>

Структура аминокислот

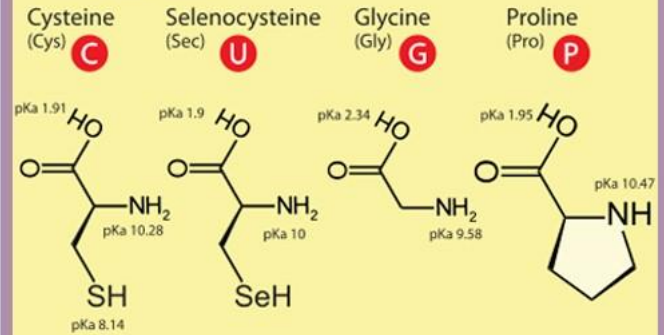
Twenty-One Amino Acids

⊕ Positive ⊖ Negative
 • Side chain charge at physiological pH 7.4

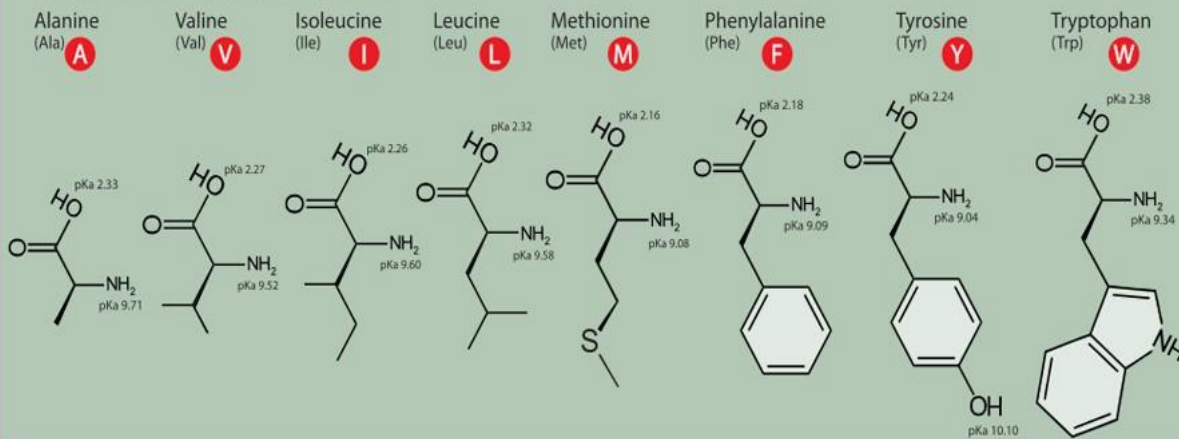
A. Amino Acids with Electrically Charged Side Chains



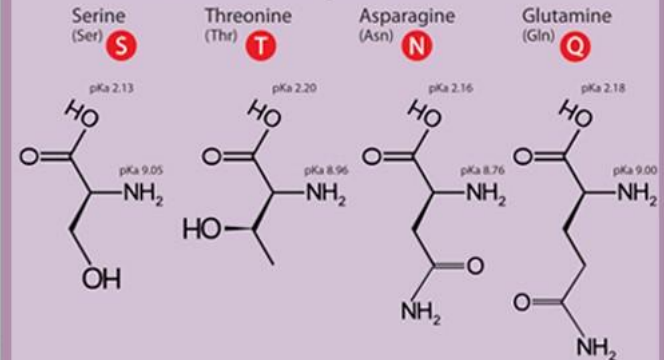
C. Special Cases



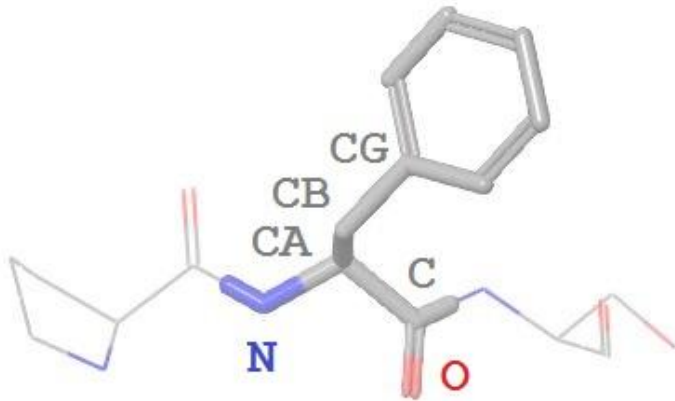
D. Amino Acids with Hydrophobic Side Chain



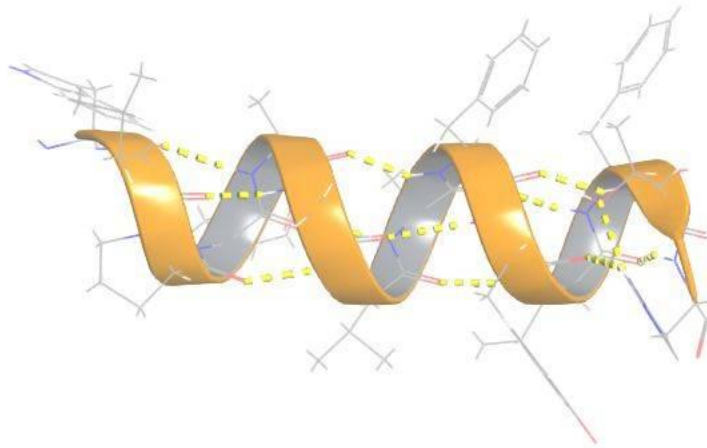
B. Amino Acids with Polar Uncharged Side Chains



Структурная организация белков



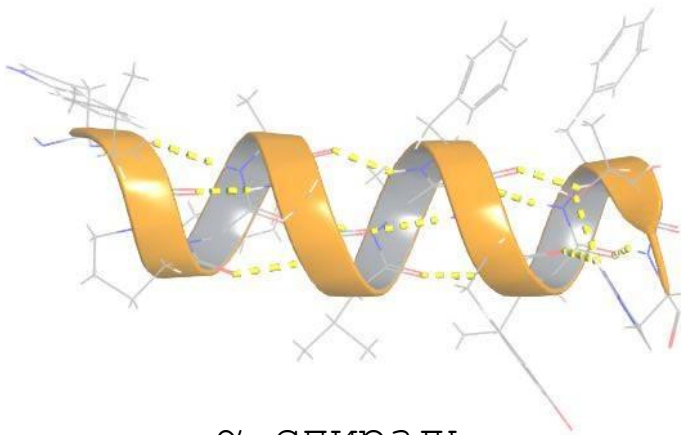
...FYV**PFS**NKT...



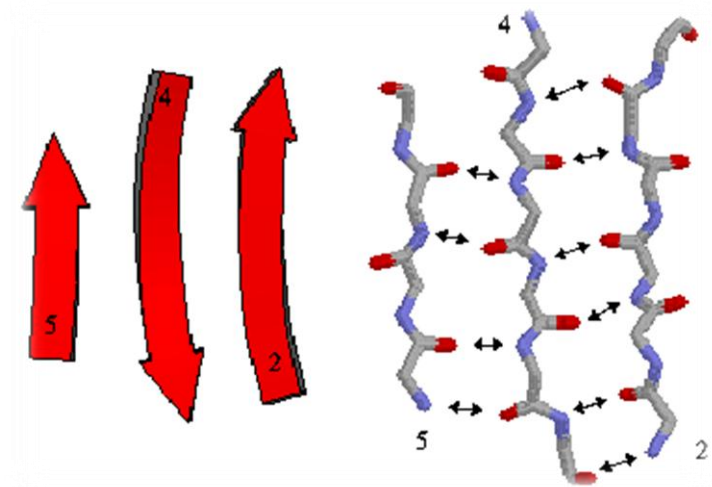
...WLPY**AGVAFYIF**TH...



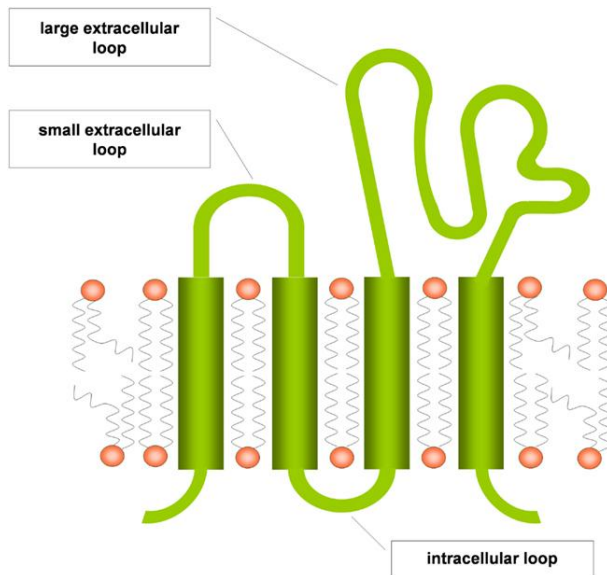
Типы вторичной структуры белков



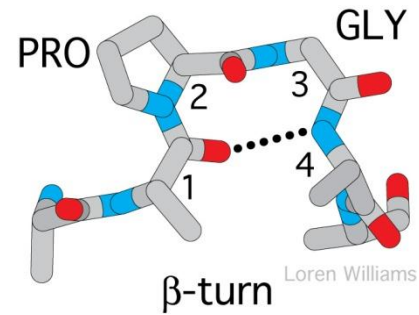
α -спираль



β -лист, состоящий из β -тяжей



ПЕТЛЯ

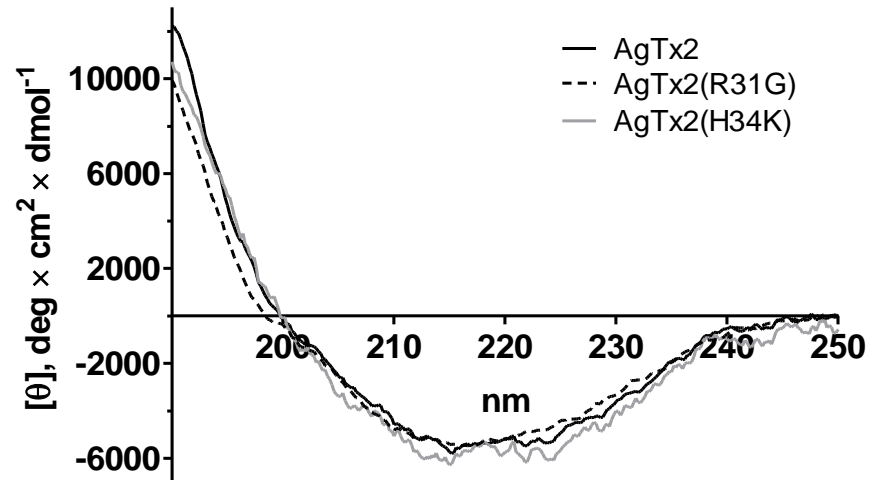
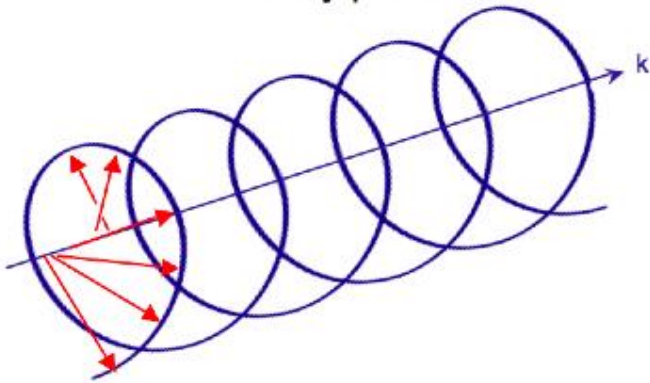


поворот

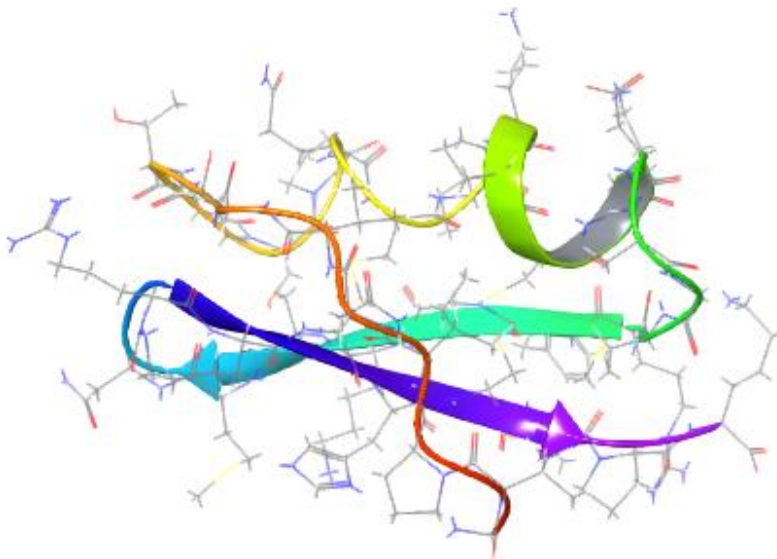
Спектроскопия кругового дихроизма

Экспериментальный метод изучения вторичной структуры белков

Circularly polarized

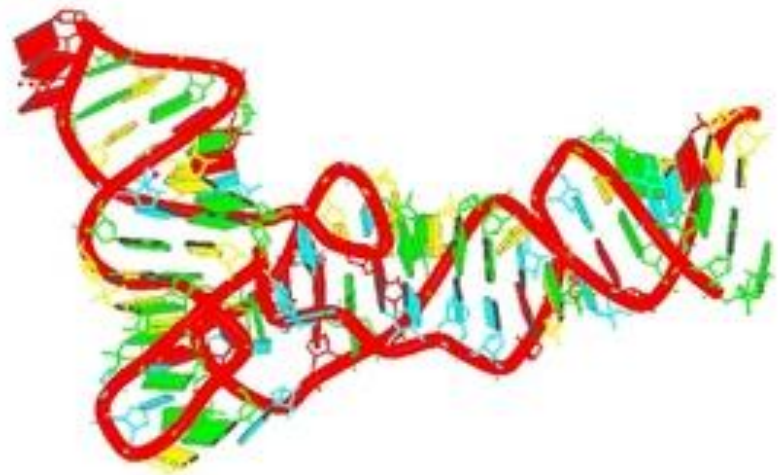
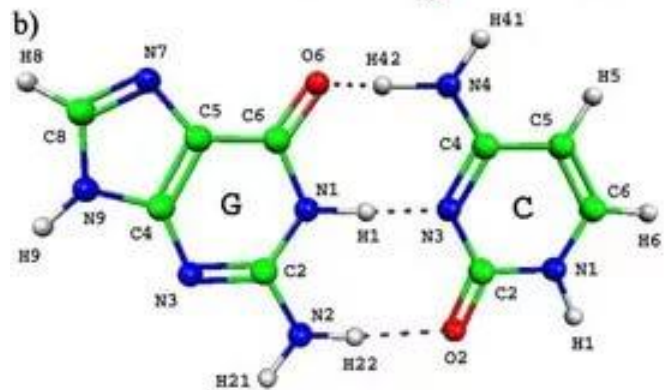
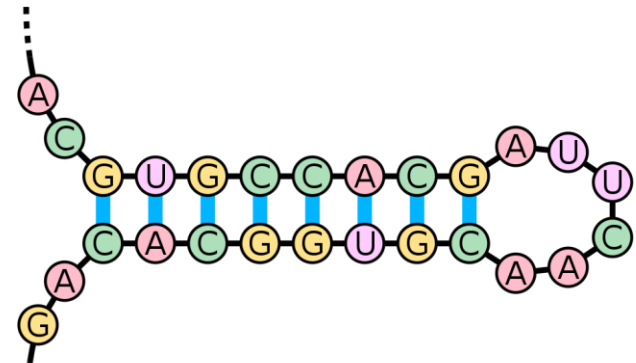
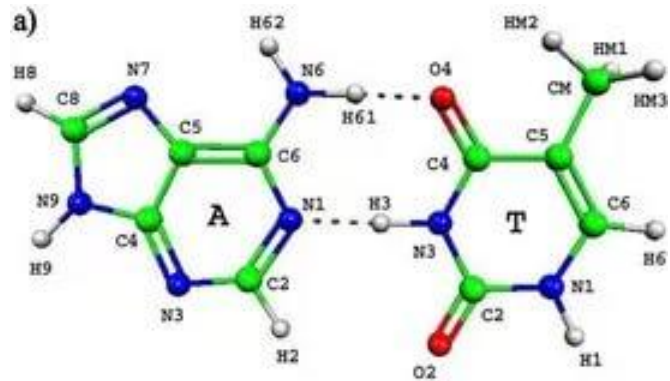


Спектр кругового дихроизма природного аджитоксина и двух его мутантных форм

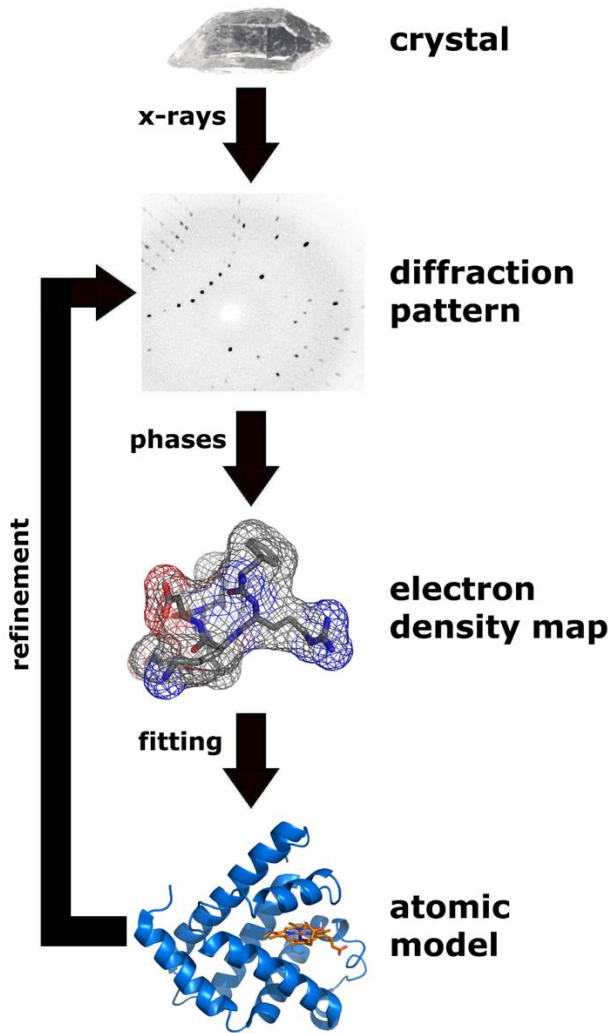


Пространственная структура природного аджитоксина

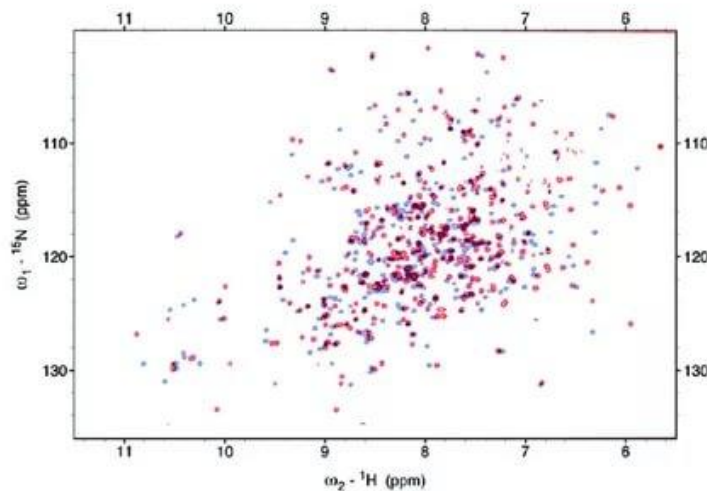
Структурная организация нуклеиновых кислот



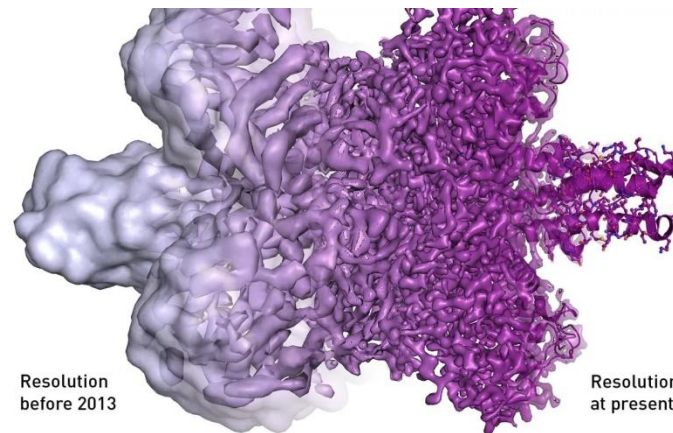
Методы определения структуры макромолекул



Этапы определения структуры с помощью рентгеновской кристаллографии



Спектроскопия ядерного магнитного резонанса (ЯМР)



(Крио) электронная микроскопия

Структурные базы данных

MyPDB

Hide

Login to your Account
Register a New Account

Home

Hide

News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
Sitemap
New Website Features

Deposition

Hide

All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamline
Related Tools

Search

Hide

A Resource for Studying Biological Macromolecules

The PDB archive contains information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies. As a member of the **wwPDB**, the RCSB PDB curates and annotates PDB data according to agreed upon standards.

The RCSB PDB also provides a variety of tools and resources. Users can perform simple and advanced searches based on annotations relating to sequence, structure and function. These molecules are visualized, downloaded, and analyzed by users who range from students to specialized scientists.

[Hide Welcome Message](#)

Featured Molecules

([MoM Category View](#) / [Previous](#))

Hide

Features: [MoM](#) | [PSI](#)

Structural
View of
Biology



Protein Synthesis



Molecule of the Month:

Riboswitches

Why use two or more molecules when one



Customize This Page

New Features

Hide

Improved Domain Display on Sequence Tab

Latest features released:

Website Release Archive:

RCSB PDB News

Hide

[Weekly](#) | [Quarterly](#) | [Yearly](#)

2010-10-12

Categorized View of Molecule of the Month



Explore new Molecule of the Month feature categories.

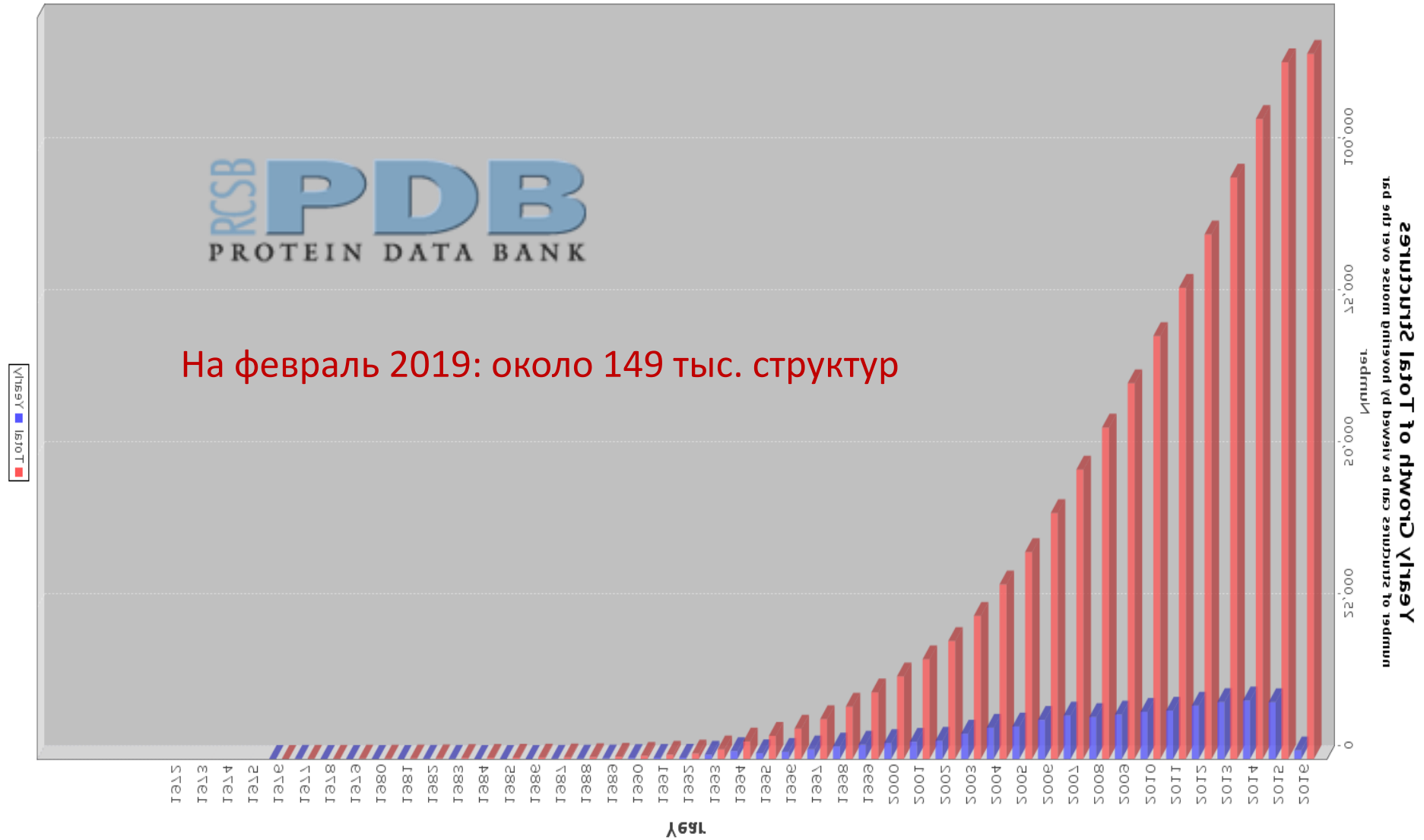
- Latest Website Release
- Poster Prize Awarded at ECM
- Analyze small molecule interactions in the PDB with

PDB (с 1971 по н.в.) – содержит информацию об экспериментально определенных структурах белков, нуклеиновых кислот и различных комплексов.

Структурные базы данных



На февраль 2019: около 149 тыс. структур



Структурные базы данных

POTASSIUM CHANNEL KCSA IN COMPLEX WITH TETRABUTYLAMMONIUM IN HIGH K

MyPDB Hide

Login to your Account
Register a New Account

Home Hide

News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
Sitemap
New Website Features

Deposition Hide

All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamline
Related Tools

Search Hide

Advanced Search
Latest Release
New Structure Papers
Sequence Search
Chemical Components
Unreleased Entries
Browse Database
Help

Summary Sequence Derived Data Seq. Similarity 3D Similarity Literature Biol. & Chem. Methods Geometry Links

POTASSIUM CHANNEL KCSA IN COMPLEX WITH TETRABUTYLAMMONIUM IN HIGH K

2JK5

Display Files ▾
Download Files ▾
Print this Page
Share this Page ▾

DOI: 10.2210/pdb2jk5/pdb

Primary Citation

Structure of a Potassium Channel in Complex with Symmetrical Quaternary Ammonium Compounds Reveals a Novel Hydrophobic Binding Site

Lenaus, M.J., Focia, P.J., Wagner, T., Gross, A.

Journal: To be Published

Not in PubMed

Molecular Description Hide

Classification: Immune System/metal Transport
Structure Weight: 61280.07

Molecule:	ANTIBODY FAB FRAGMENT LIGHT CHAIN	Length:	219
Polymer:	1 Type: polypeptide(L)		
Chains:	A		
Molecule:	ANTIBODY FAB FRAGMENT HEAVY CHAIN	Length:	212
Polymer:	2 Type: polypeptide(L)		
Chains:	B		
Molecule:	VOLTAGE-GATED POTASSIUM CHANNEL	Length:	124
Polymer:	3 Type: polypeptide(L)		
Chains:	C		

Biological Assembly ?



Biological Assembly Image for 2JK5

More Images...

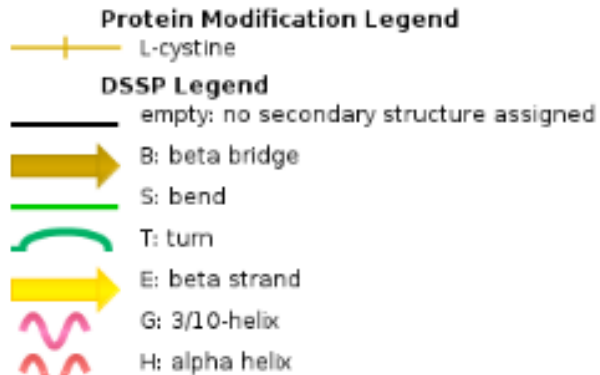
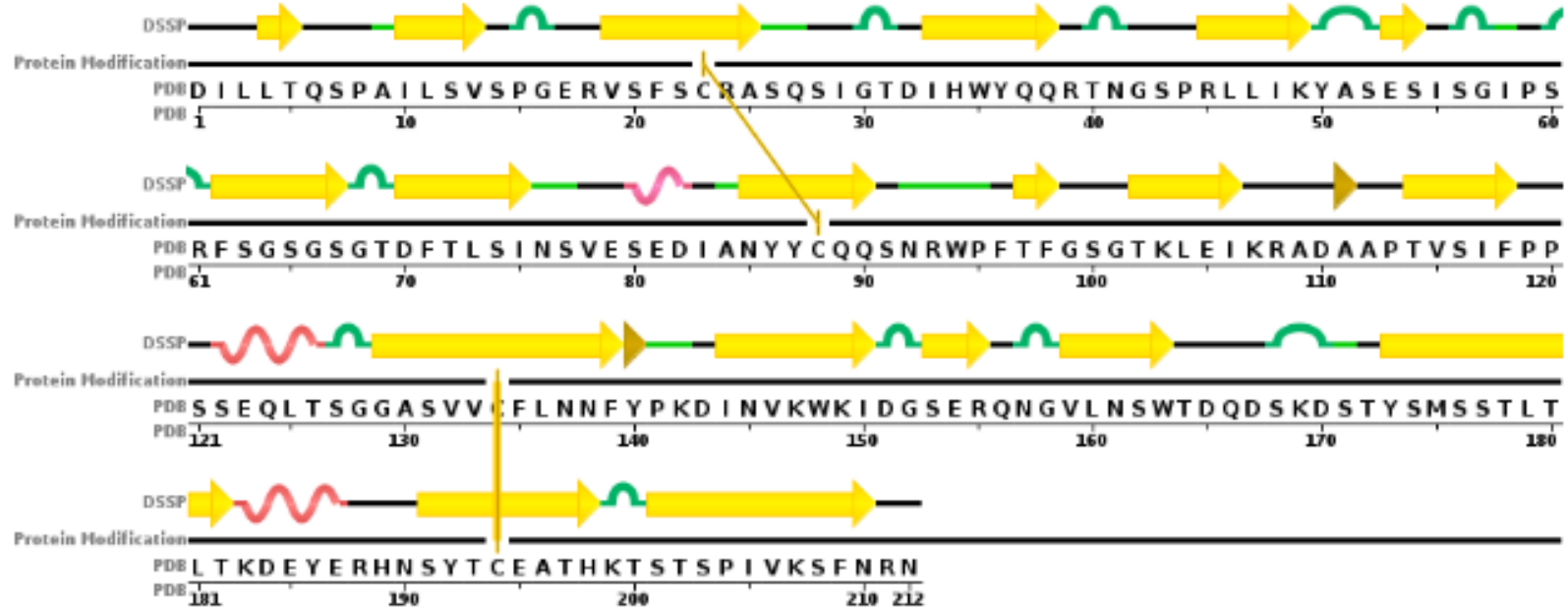
View in Jmol SimpleViewer
Other Viewers ▾ Protein Workshop

Biological assembly generated by PISA (software)

Структурные базы данных

POTASSIUM CHANNEL KCSA IN COMPLEX WITH TETRABUTYLAMMONIUM IN HIGH K

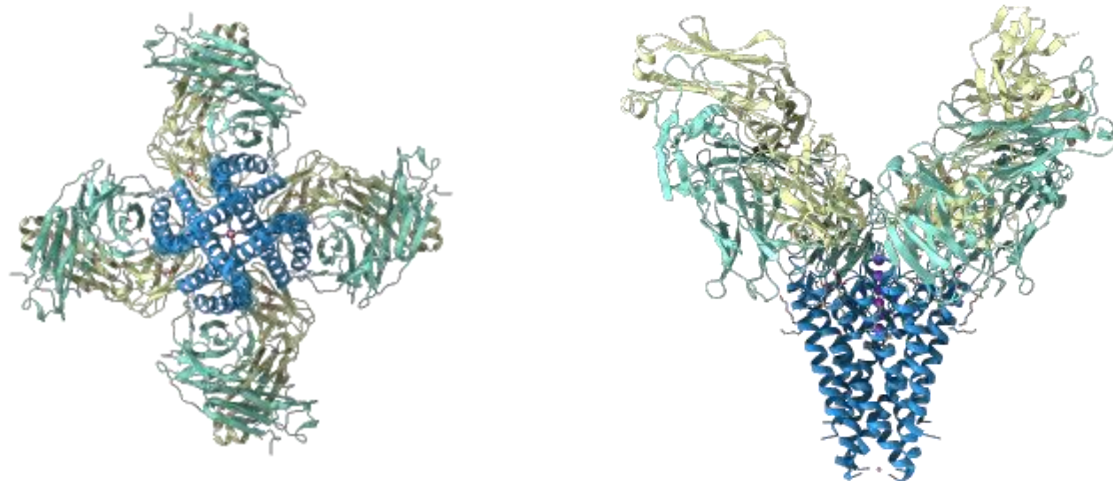
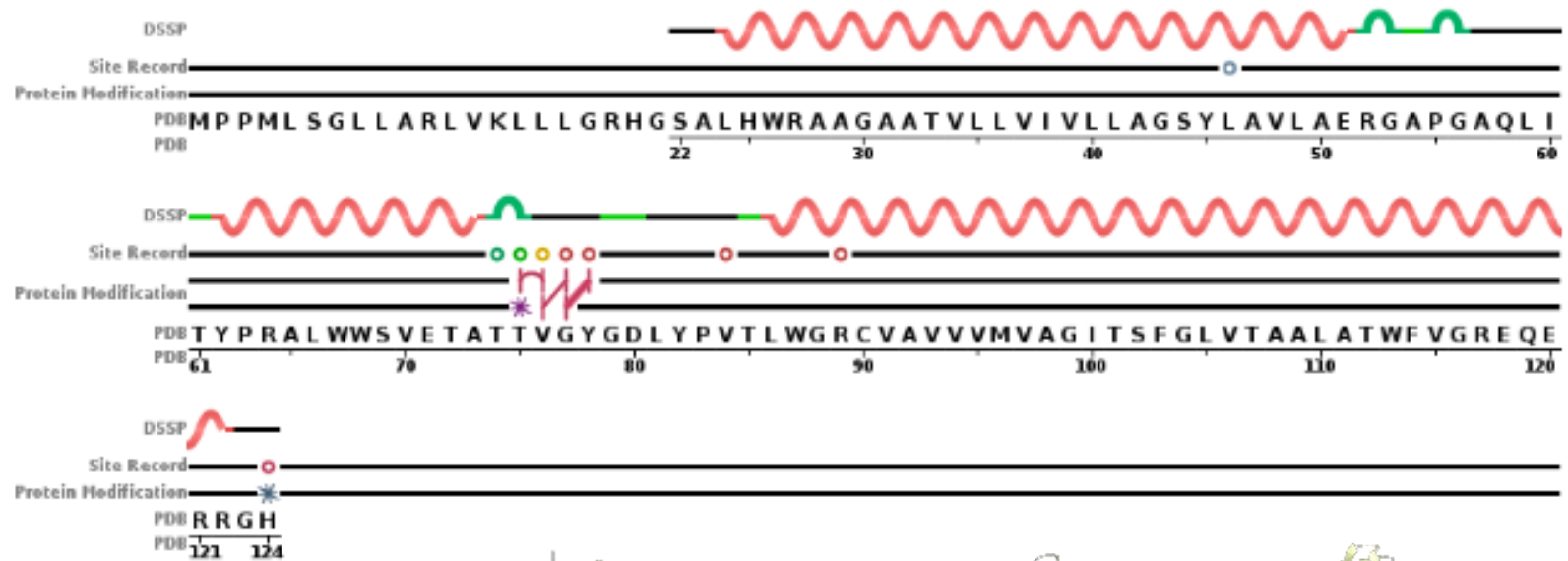
Sequence Chain View



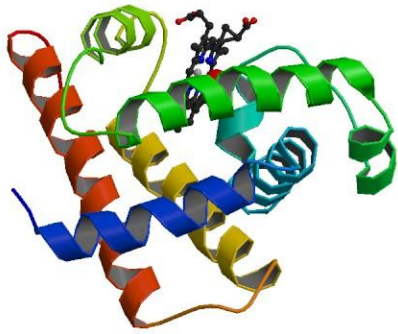
Структурные базы данных

POTASSIUM CHANNEL KCSA IN COMPLEX WITH TETRABUTYLAMMONIUM IN HIGH K

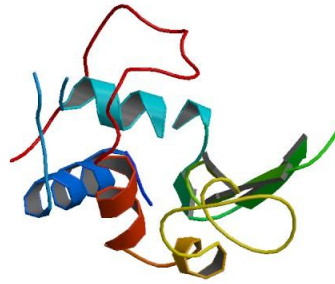
Sequence Chain View



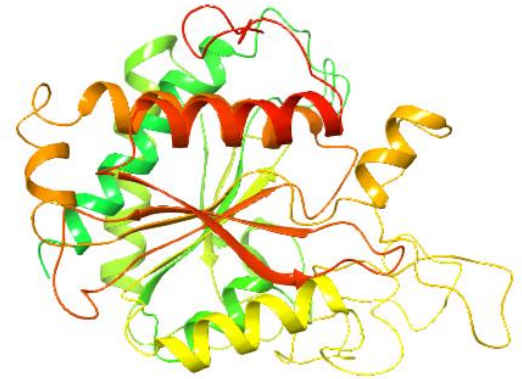
Разнообразиие белковых структур



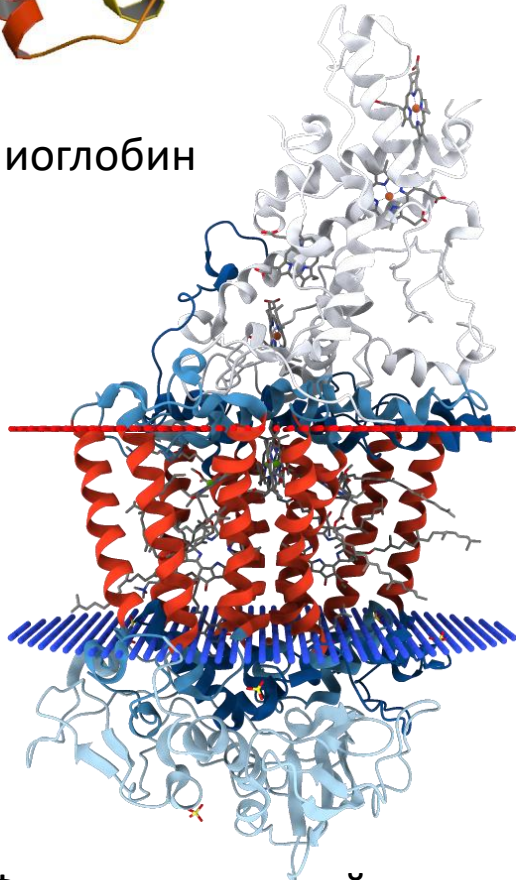
Миоглобин



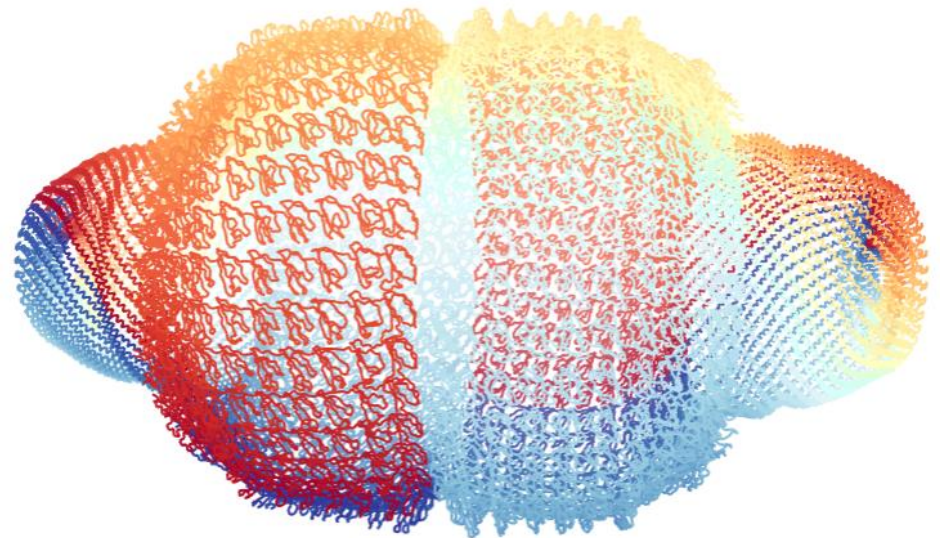
Лизоцим



Карбоксипептидаза



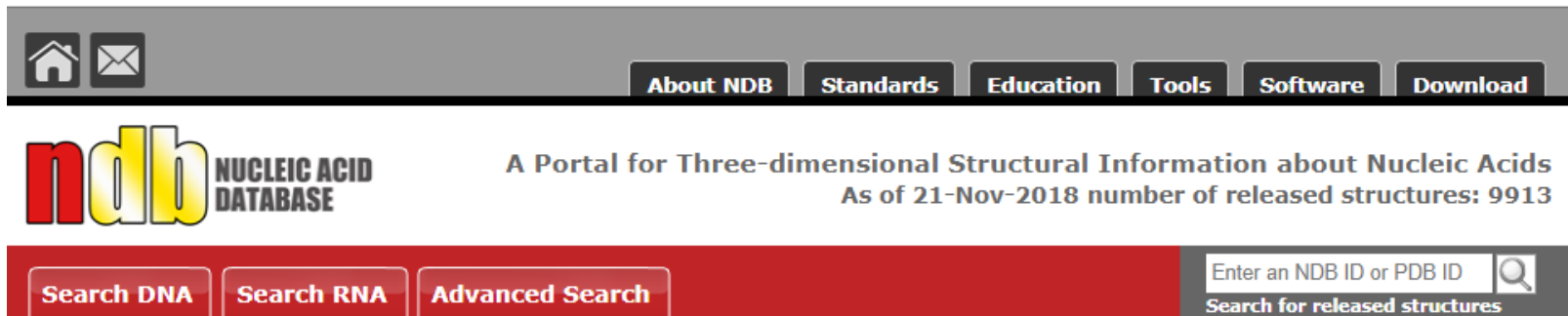
Фотореакционный центр



Структура органеллы vault из печени крысы

Структурные базы данных

NDB (с 1992 по н.в.) – сбор и распространение информации о структуре нуклеиновых кислот. Формат хранения данных идентичен PDB.



The screenshot shows the top navigation bar of the NDB website. It includes a home icon and an email icon on the left. The main navigation menu contains links for 'About NDB', 'Standards', 'Education', 'Tools', 'Software', and 'Download'. Below the navigation bar is the NDB logo, which consists of the letters 'ndb' in a stylized font with a yellow-to-red gradient, followed by the text 'NUCLEIC ACID DATABASE'. To the right of the logo is the text 'A Portal for Three-dimensional Structural Information about Nucleic Acids' and 'As of 21-Nov-2018 number of released structures: 9913'. Below this is a search bar with a magnifying glass icon and the text 'Enter an NDB ID or PDB ID' and 'Search for released structures'. Below the search bar are three buttons: 'Search DNA', 'Search RNA', and 'Advanced Search'.

Welcome to the NDB

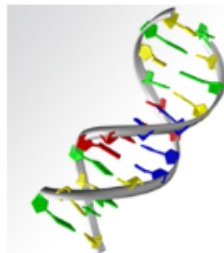
The NDB contains information about experimentally-determined nucleic acids and complex assemblies. Use the NDB to perform searches based on annotations relating to sequence, structure and function, and to download, analyze, and learn about nucleic acids.

Search Structures

Search DNA
Search DNA and its complexes

Search RNA
Search for RNA structures in the NDB archive or in the Non-Redundant list

Advanced Search
Search for structures based on structural features, chemical features, binding modes, citation and experimental information



Featured Tools

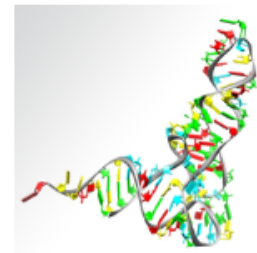
RNA 3D Motif Atlas, a representative collection of RNA 3D internal and hairpin loop motifs

Non-redundant Lists of RNA-containing 3D structures

RNA Base Triple Atlas, a collection of motifs consisting of two RNA basepairs

WebFR3D, a webserver for symbolic and geometric searching of RNA 3D structures

R3D Align, an application for detailed nucleotide to nucleotide alignments of RNA 3D structures



Структурные базы данных

COMPLEX BETWEEN NUCLEOSOME CORE PARTICLE (H3,H4,H2A,H2B) AND 146 BP LONG DNA FRAGMENT

NDB ID: PD0001 **PDB ID:** [1AOI](#) 

Title:

COMPLEX BETWEEN NUCLEOSOME CORE PARTICLE (H3,H4,H2A,H2B) AND 146 BP LONG DNA FRAGMENT

Molecular Description:

COMPLEX BETWEEN NUCLEOSOME CORE PARTICLE (H3
H4
H2A
H2B) AND 146 BP LONG DNA FRAGMENT

Structural Keywords:

B DOUBLE HELIX


Nucleic Acid Sequence:

[Click to show/hide 1 nucleic acid sequences](#)

Protein Sequence:

[Click to show/hide 4 protein sequences](#)

Primary Citation:

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J.
[Crystal structure of the nucleosome core particle at 2.8 Å resolution.](#) 
Nature, **389**, pp. 251 - 260, 1997.

Experimental Information:

X-RAY DIFFRACTION

Space Group:

P 21 21 21

Cell Constants:

a = 106.04 b = 181.78 c = 110.12 (Ångstroms)

$\alpha = 90.0$ $\beta = 90.0$ $\gamma = 90.0$ (degrees)

Refinement:

The structure was refined using the X-PLOR program. The R value is 0.0 for 51237 reflections in the resolution range 25.0 to 2.8 Ångstroms with Fobs > 0.0 sigma(Fobs) and with I > 0.0 sigma(I)

Structural Features

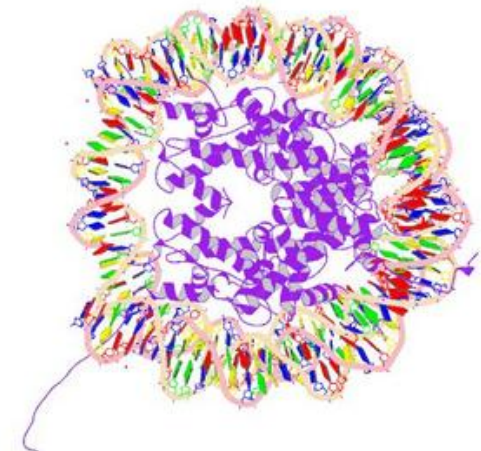
[Base Pair Hydrogen Bonding Classification](#)

[Nucleic Acid Backbone Torsions](#)

[Base Pair Morphology Parameters](#)

[Base Pair Morphology Step Parameters](#)



Biological Assembly 1




[More Images...](#)


Структурные базы данных

EMBL-EBI [Terms of Use](#) | [Privacy](#) | [Cookies](#)

[Databases](#) | [Tools](#) | [Research](#) | [Training](#) | [Industry](#) | [About Us](#) | [Help](#) | [Site Index](#)  

[Home](#) **EMD-1367** [Mirrors](#) [Contact us](#)



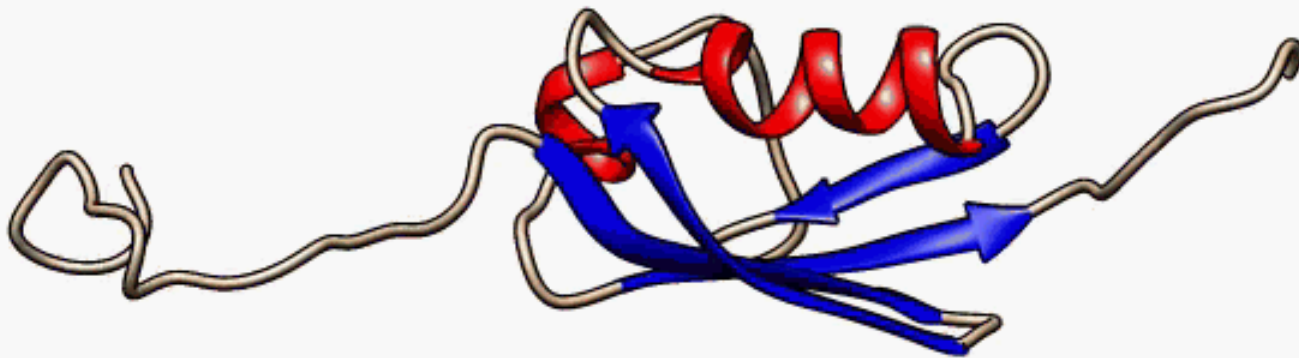
Title: Three-dimensional structure of a voltage-gated potassium channel at 2.5 nm resolution.
Authors: Olga Sokolova, Ludmila Kolmakova-Partensky and Nikolaus Grigorieff
Sample: Shaker B channel
Aggregation state: Single particle (25 angstroms resolution)
 **Latest update:** 2011-05-26

- [Summary](#)
- [Experimental details](#)
- [Visualization](#)
- [Map information](#)
- [Downloads](#)

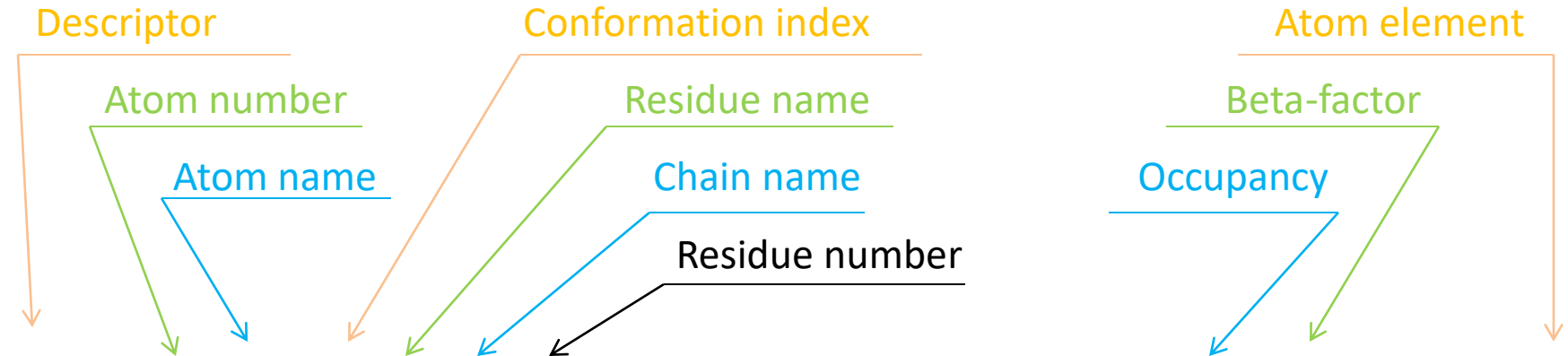
Summary													
Status:	Released												
Deposition date:	2007-05-24												
Header release date:	2007-05-30												
Map release date:	2007-05-30												
Primary citation:	Sokolova O, Kolmakova-Partensky L, Grigorieff N: Three-dimensional structure of a voltage-gated potassium channel at 2.5 nm resolution. <i>STRUCTURE</i> (2001) 9 , pp. 215-220 [PubMed 11286888]												
Sample:	Shaker B channel												
Resolution:	25 Å (determined by FSC at 0.5 cut-off)												
Fitted PDB:	<table border="1"><thead><tr><th>PDB</th><th>Authors</th><th>PubMed</th><th>Status</th></tr></thead><tbody><tr><td>1bl8</td><td>Doyle, D.A., Cabral, J.M., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., Mackinnon, R.</td><td>9525859</td><td>Released</td></tr><tr><td>1a68</td><td>Kreusch, A., Pfaffinger, P.J., Stevens, C.F., Choe, S.</td><td>9582078</td><td>Released</td></tr></tbody></table>	PDB	Authors	PubMed	Status	1bl8	Doyle, D.A., Cabral, J.M., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., Mackinnon, R.	9525859	Released	1a68	Kreusch, A., Pfaffinger, P.J., Stevens, C.F., Choe, S.	9582078	Released
PDB	Authors	PubMed	Status										
1bl8	Doyle, D.A., Cabral, J.M., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., Mackinnon, R.	9525859	Released										
1a68	Kreusch, A., Pfaffinger, P.J., Stevens, C.F., Choe, S.	9582078	Released										

Неупорядоченные белки

Intrinsically disordered proteins - «нарушители» догмы «структура определяет функцию»



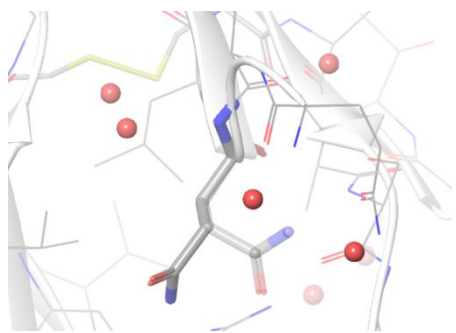
Структурные файлы. *.pdb



```

...
ATOM      22  CG  AGLN  A   3   -30.506  25.807  16.125  0.50  43.30  C
ATOM      23  CD  AGLN  A   3   -29.018  25.546  15.959  0.50  43.01  C
ATOM      24  OE1AGLN A   3   -28.187  26.374  16.332  0.50  43.76  O
ATOM      25  NE2AGLN A   3   -28.675  24.391  15.403  0.50  41.73  N
ATOM      26  N   BGLN  A   3   -33.005  25.819  17.852  0.50  45.16  N
ATOM      27  CA  BGLN  A   3   -32.788  24.834  16.796  0.50  44.25  C
ATOM      28  C   BGLN  A   3   -33.529  23.510  17.024  0.50  43.75  C
...
HETATM   4140 C1  F09  C1131  -21.017  -3.092  -1.563  1.00  60.56  C
HETATM   4141 C2  F09  C1131  -21.015  -1.597  -1.357  1.00  60.40  C
...

```



↑ ↑ ↑
 X Y Z coordinates in Å

Задание поворота в пространстве

Матрица поворота:

$$M_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}$$

$$M_y(\alpha) = \begin{pmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{pmatrix}$$

$$M_z(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Любое вращение в трехмерном пространстве может быть представлено как композиция поворотов вокруг трех ортогональных осей.

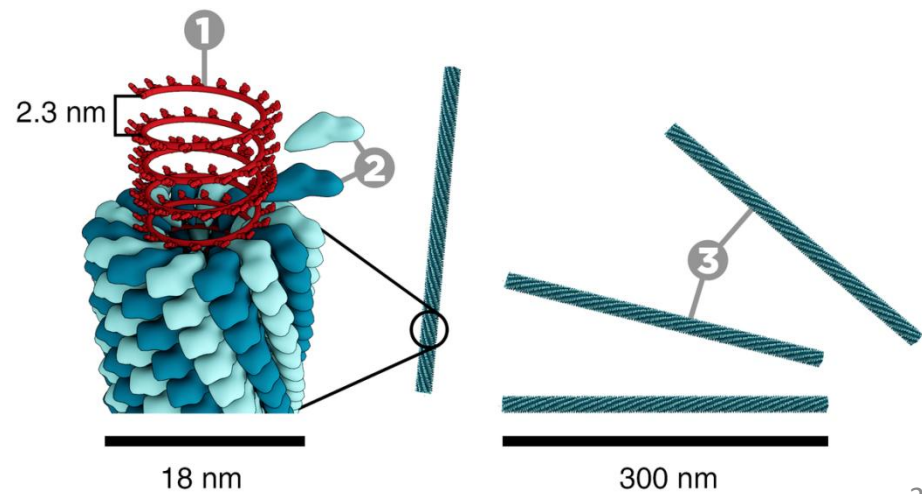
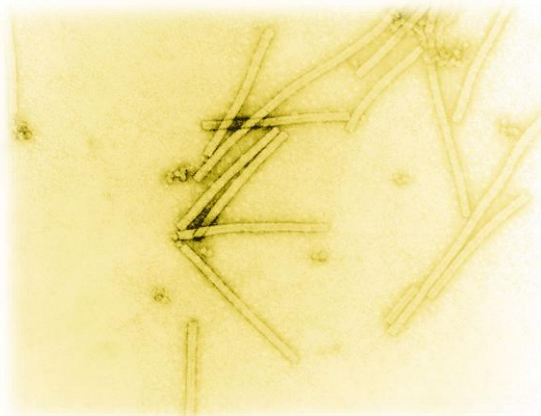
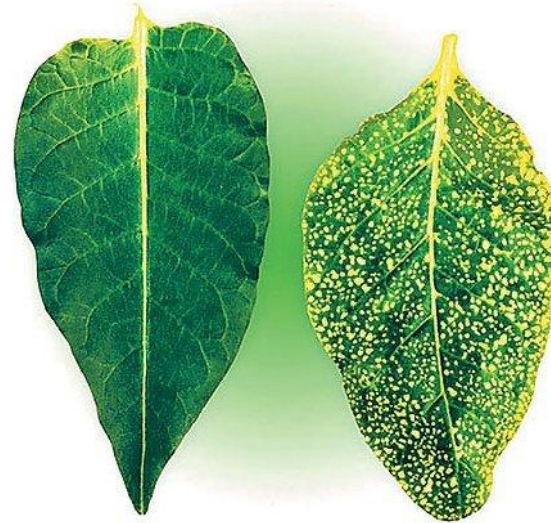
Недостатки: 1) громоздкость вычислений

2) неустойчивость вычислений (ошибки округления)

Вирус табачной мозаики



Д.И. Ивановский
(1864 — 1920)



Вирус табачной мозаики

REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
 REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
 REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
 REMARK 350 GIVEN BELOW. BOTH NON-CRYSTALLOGRAPHIC AND
 REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.

матрица поворота

REMARK 350

REMARK 350 BIOMOLECULE: 1

REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, R

вектор переноса

REMARK 350 BIOMT1 1 1.000000 0.000000 0.000000

0.00000

REMARK 350 BIOMT2 1 0.000000 1.000000 -0.000000

0.00000

REMARK 350 BIOMT3 1 0.000000 -0.000000 1.000000

0.00000

REMARK 350 BIOMT1 2 0.923478 -0.383651 0.000000

-0.00000

REMARK 350 BIOMT2 2 0.383651 0.923478 -0.000000

0.00000

REMARK 350 BIOMT3 2 0.000000 -0.000000 1.000000

-67.58400

...

REMARK 350 BIOMT1 49 0.926988 0.375092 0.000000

-0.00000

REMARK 350 BIOMT2 49 -0.375092 0.926988 -0.000000

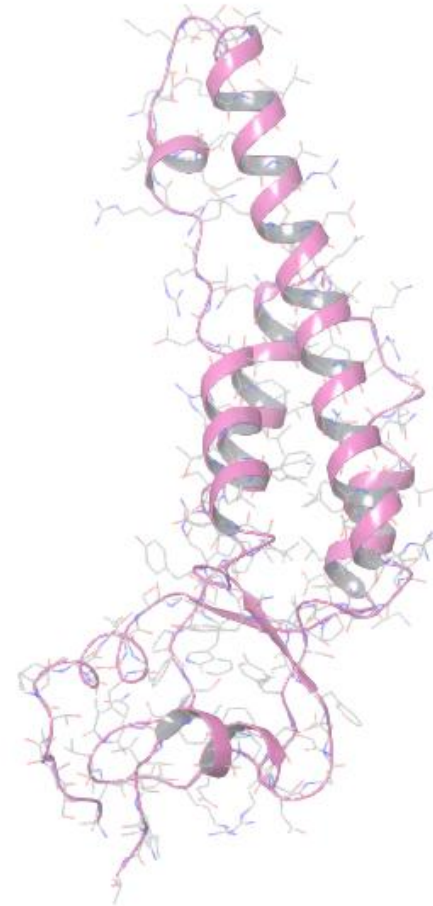
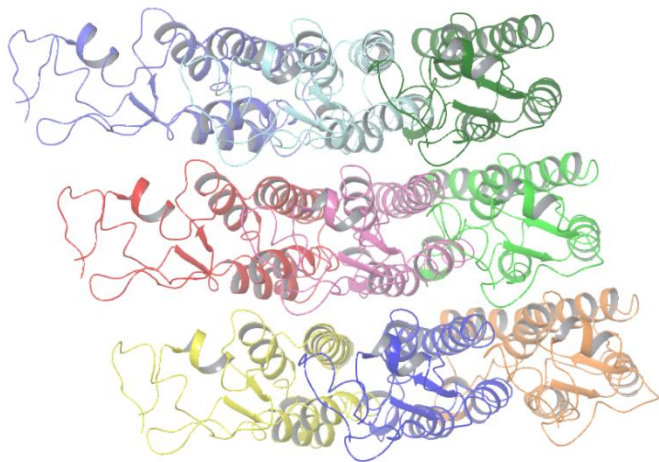
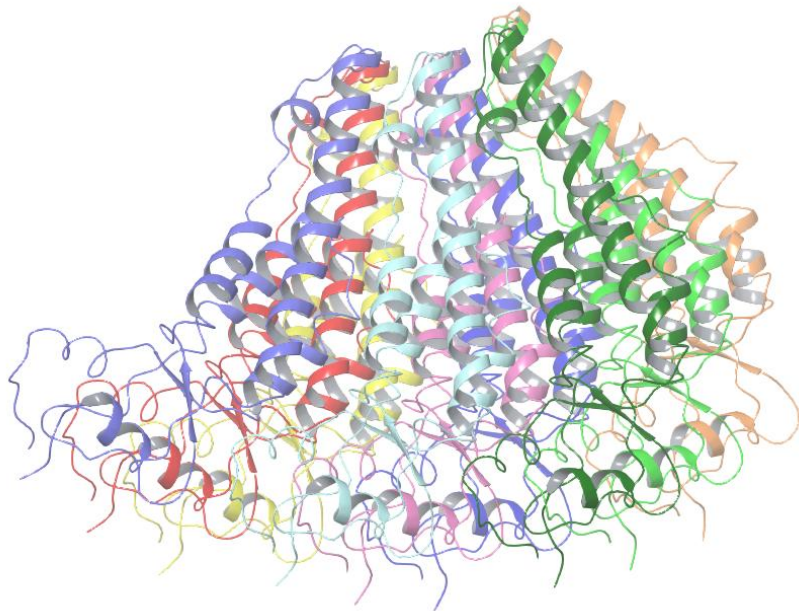
0.00000

REMARK 350 BIOMT3 49 0.000000 -0.000000 1.000000

-1.40800

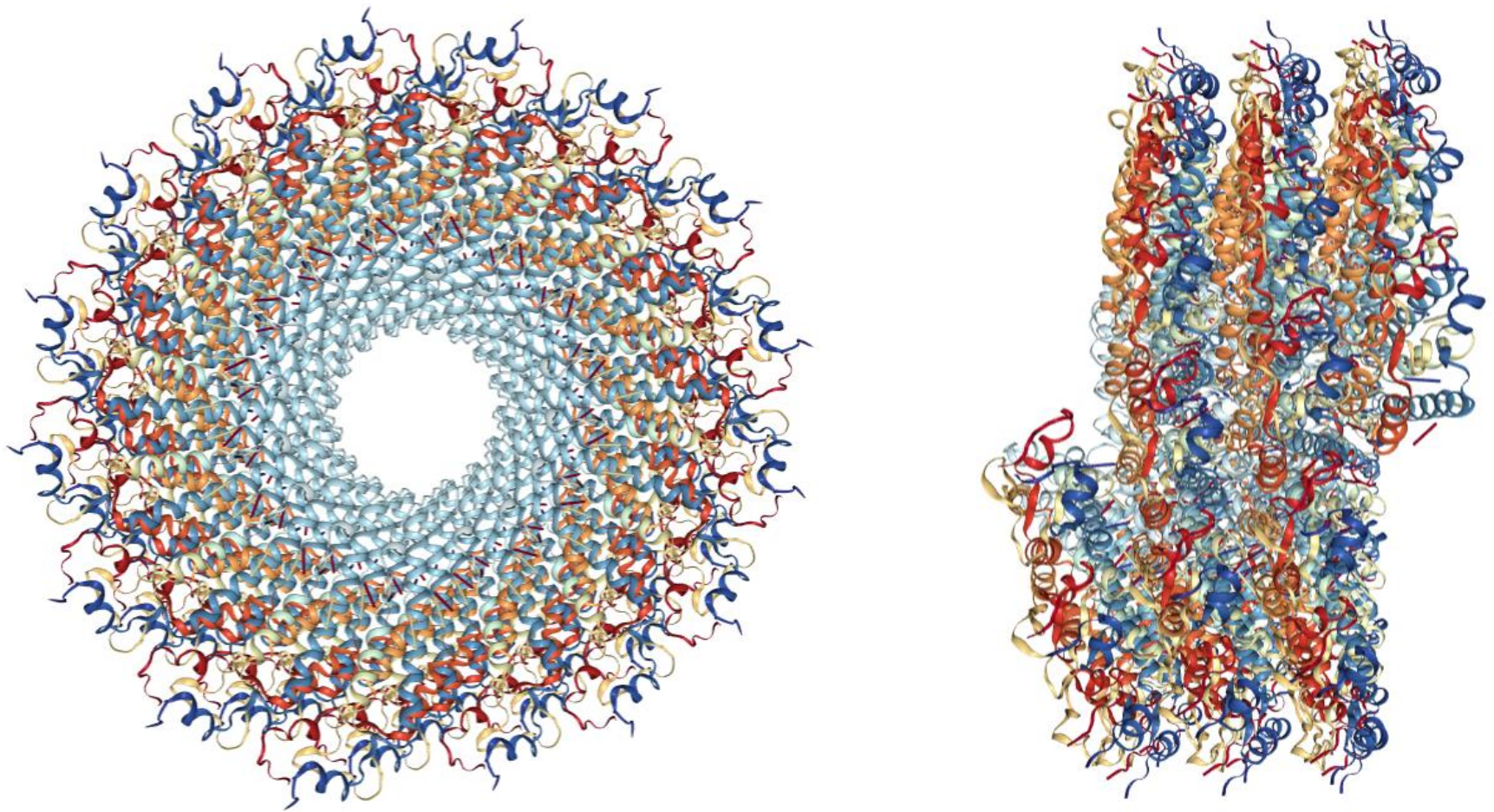
$$\begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} BIOMT1_{2x} & BIOMT1_{2y} & BIOMT1_{2z} \\ BIOMT2_{2x} & BIOMT2_{2y} & BIOMT2_{2z} \\ BIOMT3_{2x} & BIOMT3_{2y} & BIOMT3_{2z} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} BIOMT1_{2s} \\ BIOMT2_{2s} \\ BIOMT3_{2s} \end{pmatrix}$$

Вирус табачной мозаики



(pdb-код 4udv)

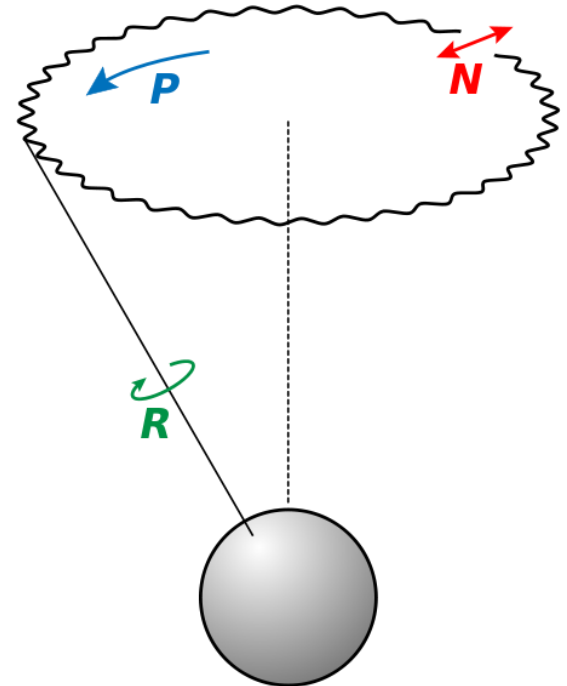
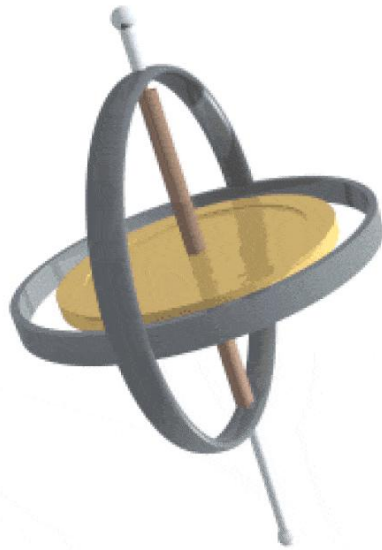
Вирус табачной мозаики



(pdb-код 4udv)

Задание поворота в пространстве

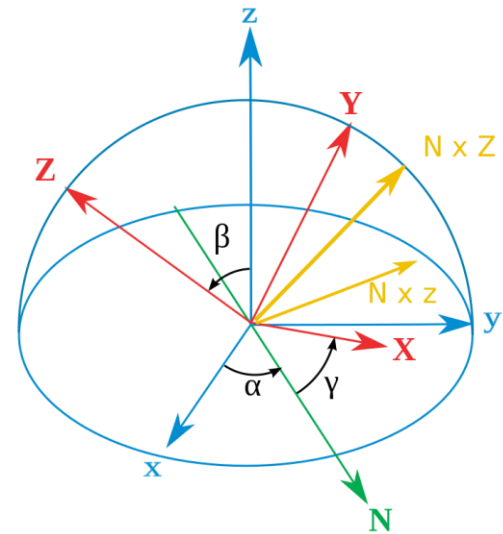
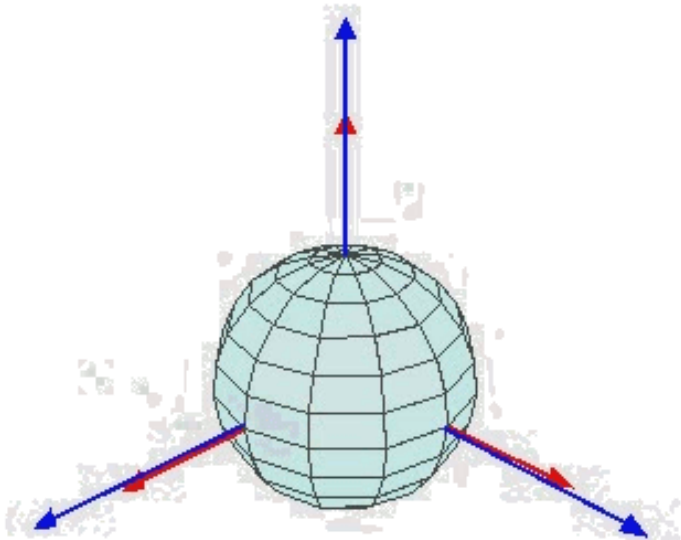
Гироскоп: прецессия, нутация и собственное вращение



<https://www.youtube.com/watch?v=kgcM7MmBj40>

Задание поворота в пространстве

Прецессия гироскопа:



x, y, z – исходные оси

X, Y, Z – конечные оси

N - линия узлов

Задание поворота в пространстве

Углы Эйлера (1776):

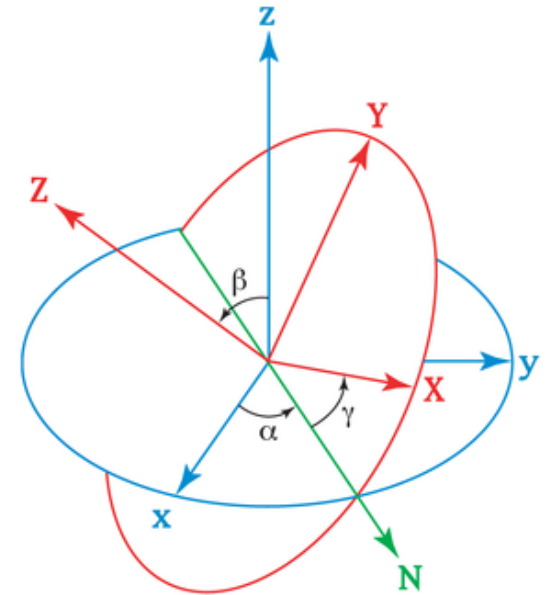
x, y, z – исходные оси

X, Y, Z – конечные оси

Угол α между осью x и линией узлов — угол прецессии (поворот вокруг оси z).

Угол β между осями z и z' — угол нутации (поворот вокруг оси $x' =$ линии узлов).

Угол γ между осью x' и линией узлов — угол собственного вращения (поворот вокруг оси $z'' = Z$).



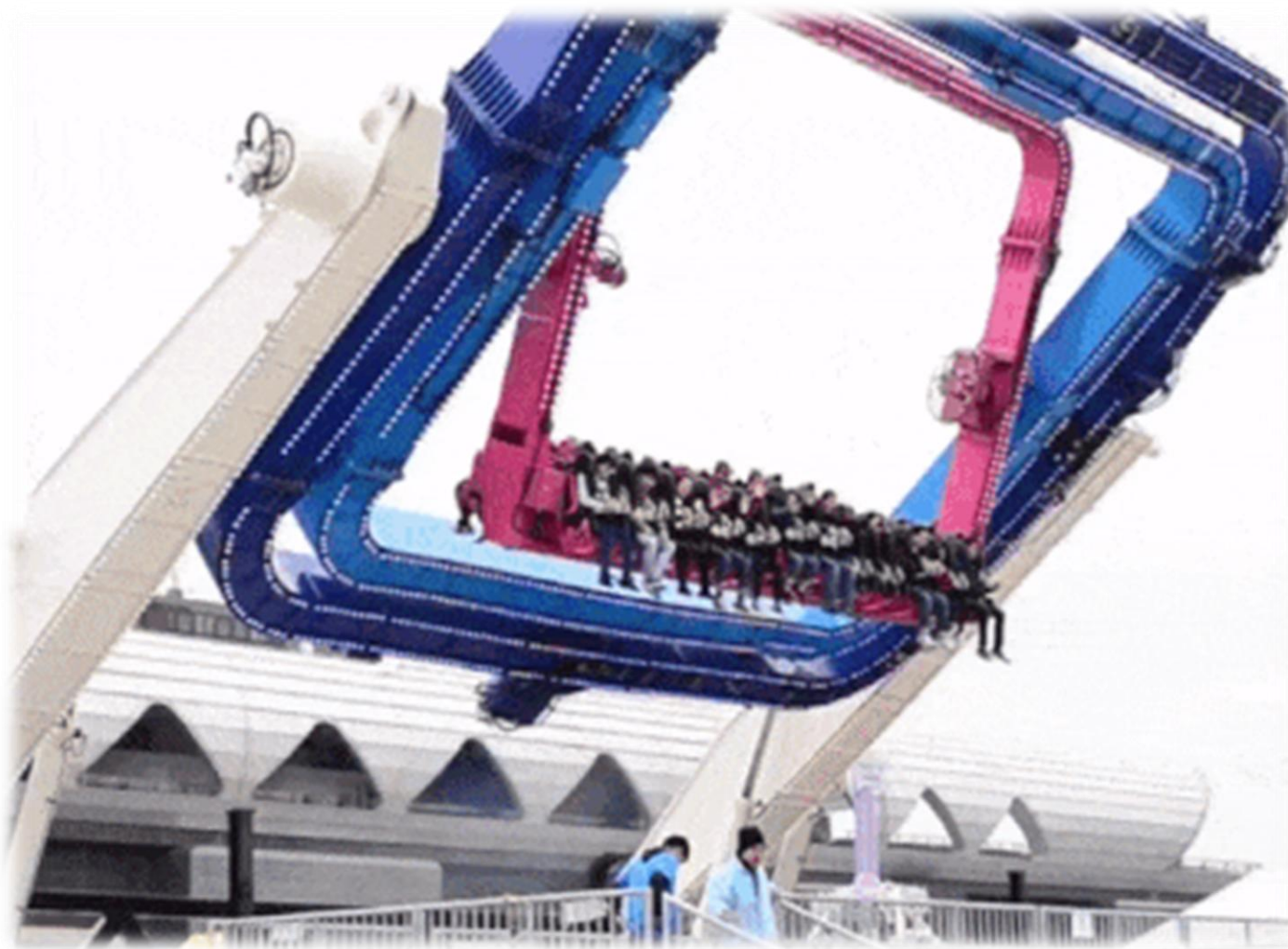
Такие повороты **некоммутативны** и конечное положение системы зависит от порядка, в котором совершаются повороты.

Недостаток: сложность вычислений при комбинации вращений, «складывание рамок карданова подвеса» ($\beta=0$).

Используются в ЭМ, ЯМР, ЭПР.

Задание поворота в пространстве

Карданов подвес



Задание поворота в пространстве

Кватернионы (1843): система гиперкомплексных чисел, образующая векторное пространство размерностью четыре над полем вещественных чисел.

$$q = a + bi + cj + dk$$

$$\bar{q} = a - bi - cj - dk$$

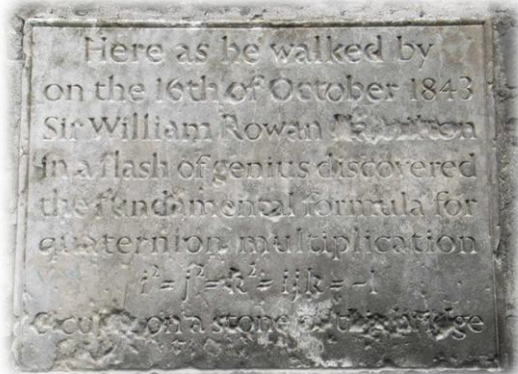
$$|q|^2 = a^2 + b^2 + c^2 + d^2$$

$$q^{-1} = \frac{\bar{q}}{|q|^2}$$

×	1	i	j	k
1	1	<i>i</i>	<i>j</i>	<i>k</i>
i	<i>i</i>	-1	<i>k</i>	- <i>j</i>
j	<i>j</i>	- <i>k</i>	-1	<i>i</i>
k	<i>k</i>	<i>j</i>	- <i>i</i>	-1



Уильям Роуэн Гамильтон
(1805 — 1865)



Задание поворота в пространстве

Описание поворота через кватернион:

$$q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = w + (x, y, z) = \cos(\alpha/2) + \mathbf{u} \sin(\alpha/2)$$

$$\vec{t} = q\vec{v}q^{-1} = \frac{q\vec{v}\bar{q}}{|q|^2} \quad \begin{array}{l} \text{-результат поворота вектора } \mathbf{v} \\ \text{вокруг вектора } \mathbf{u} \text{ на угол } \alpha \end{array}$$

$$T = QVQ^{-1} = \frac{QVQ^T}{\sqrt{\det(Q)}} \quad \text{- он же в матричном виде}$$

Преимущества: а) интуитивно понятный смысл и отсутствие «складывания рамок»
б) возможность использования векторной алгебры
в) легкость выполнения интерполяции

$$q(t) = (1-t)q_1 + tq_2, \quad 0 < t < 1$$

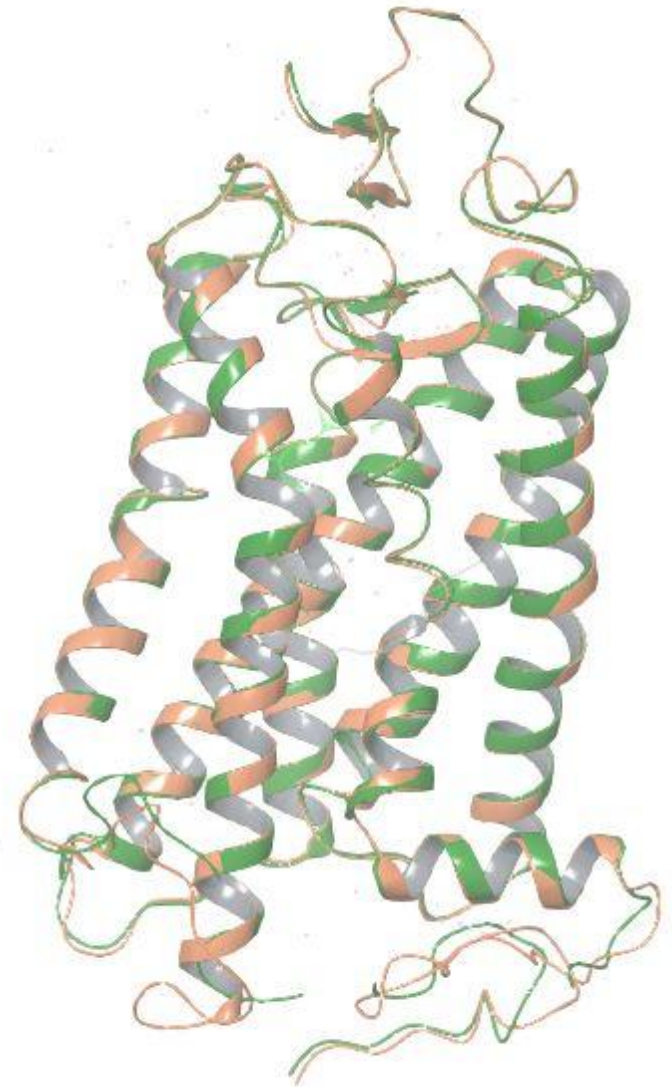
Структурные выравнивания

Варианты задачи:

1) Взаимное соответствие атомов
известно заранее:

$$p_i \leftrightarrow q_i, i = 1, \dots, N$$

(различные структуры или модели одной
и той же последовательности или близко
гомологичных последовательностей
одинаковой длины);



Совмещение структур зрительного родопсина (pdb-коды 1u19 и 1f88)

Структурные выравнивания

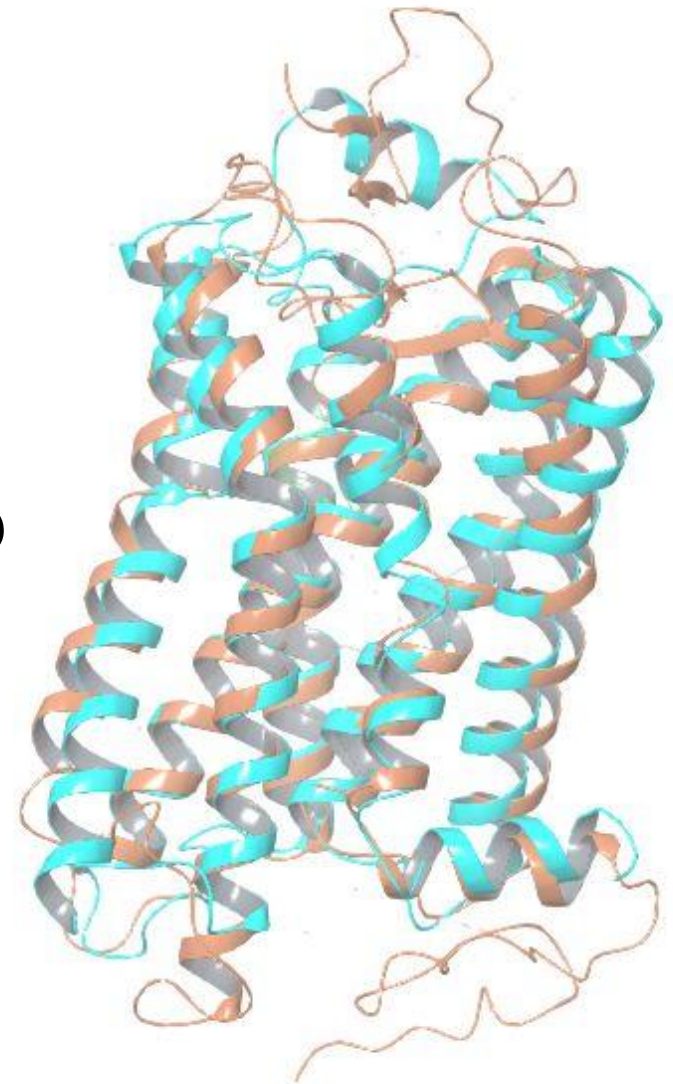
Варианты задачи:

2) Соответствие неизвестно, но известен порядок следования:

$$P_{i(k)} \leftrightarrow q_{j(k)}, i = 1, \dots, K \leq N, M,$$

$$\text{причем } k_1 > k_2 \Leftrightarrow i(k_1) > i(k_2), j(k_1) > j(k_2)$$

(структуры последовательностей разной длины, имеющие вставки в выравнивании последовательностей);



Совмещение структур зрительного родопсина (pdb-код 1u19) и β 2-адренэргического рецептора (pdb-код 2rh1)

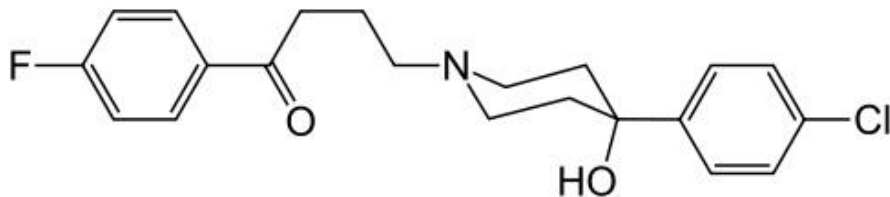
Структурные выравнивания

Варианты задачи:

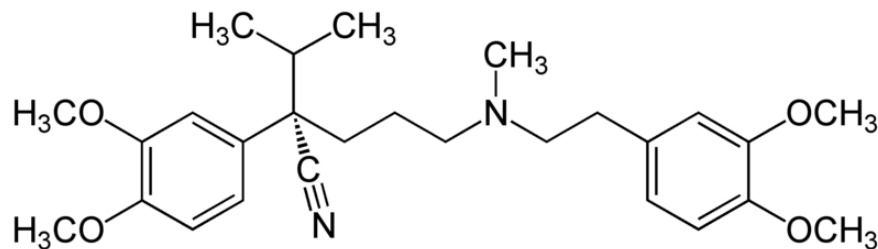
3) Неизвестно ни соответствие, ни порядок следования:

$$P_{i(k)} \leftrightarrow q_{j(k)}, i = 1, \dots, K \leq N, M$$

(фармакофоры)



Галоперидол
(блокатор дофаминовых
рецепторов)



Верапамил
(блокатор потенциалзависимых
кальциевых каналов)

Оба проявляют активность в
отношении калиевых каналов

Структурные выравнивания

Меры близости (для структур с заранее известным взаимным соответствием атомов):

-Среднеквадратичное отклонение атомов (СКО)

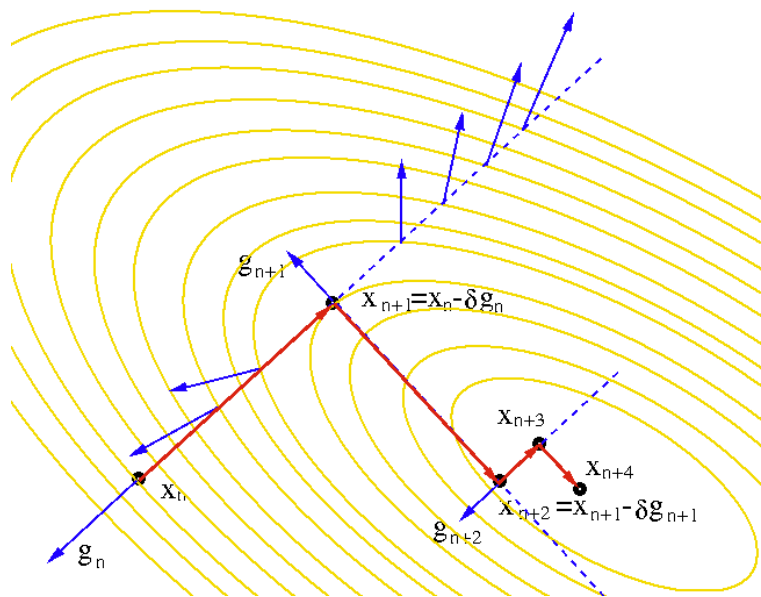
$$СКО = \sqrt{\frac{\sum_i \left((x_i - x_i^0)^2 + (y_i - y_i^0)^2 + (z_i - z_i^0)^2 \right)}{n}}$$

(обычно рассчитывается для Са-атомов, но в случае очень близких гомологов – для всех тяжелых атомов)

Методы минимизации. Градиентный спуск

Метод нахождения локального минимума функции с помощью движения вдоль градиента.

Пусть есть некая функция $F(\mathbf{x})$, для которой мы хотим найти такое значение \mathbf{x} , что функция принимает минимальное значение. Основная идея метода заключается в том, чтобы идти в направлении наискорейшего спуска, а это направление задаётся антиградиентом $-\nabla F$:



$$\vec{x}^{[j+1]} = \vec{x}^{[j]} - \lambda^{[j]} \nabla F(\vec{x}^{[j]})$$

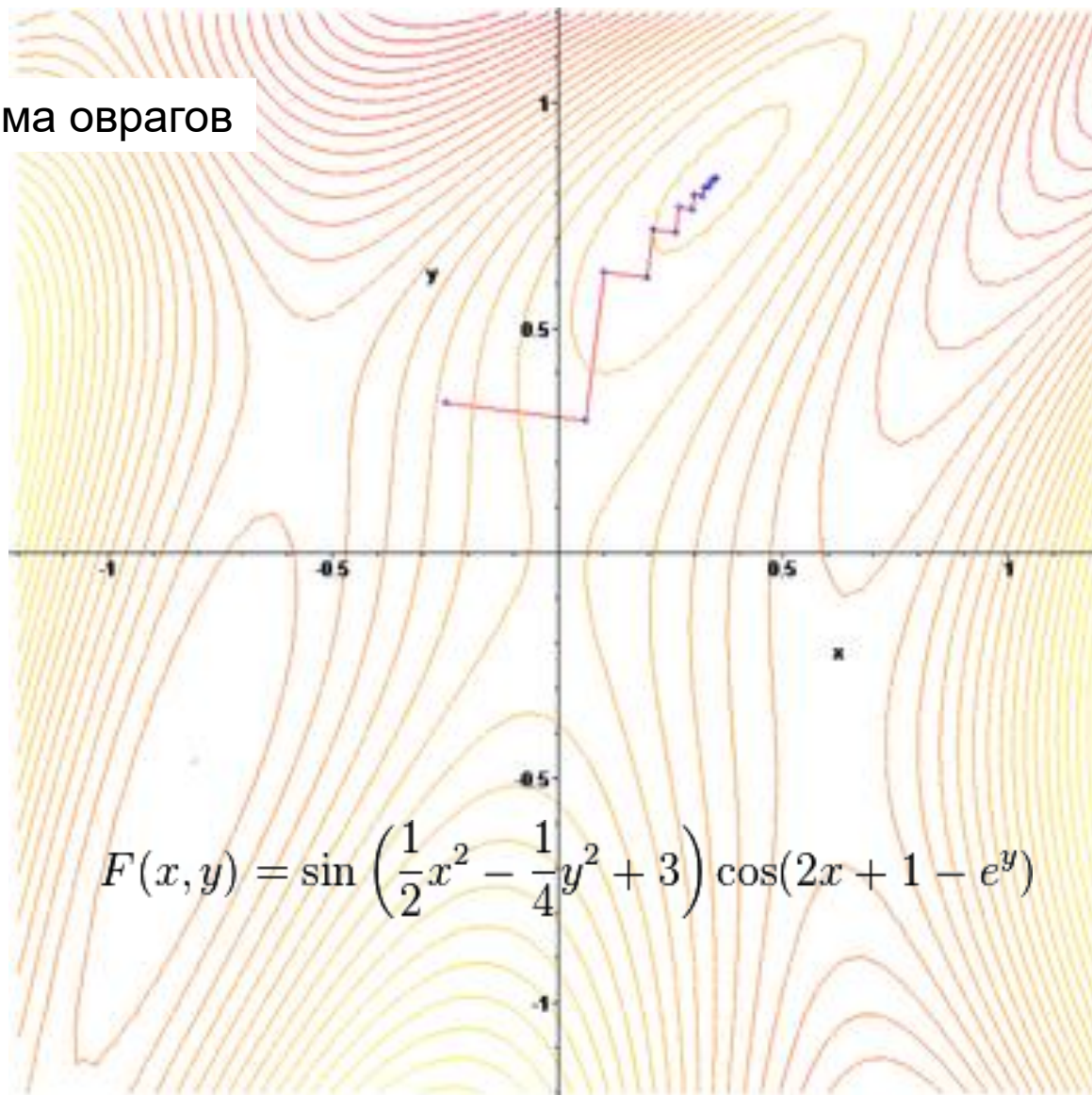
Величина шага может быть

- постоянной
- дробной
- **определяемой на каждой итерации заново**

(метод наискорейшего спуска, steepest descent)

Методы минимизации. Градиентный спуск

Проблема оврагов

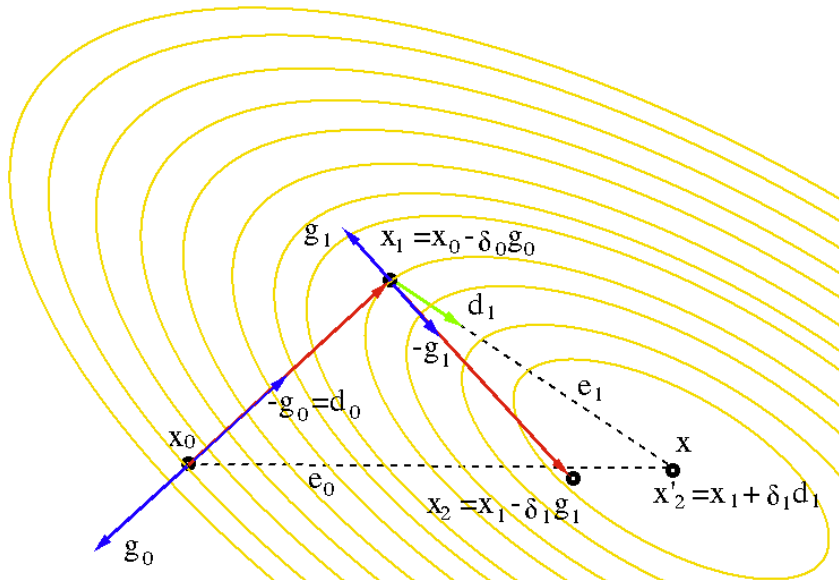


Методы минимизации. Метод сопряженных градиентов

(conjugated gradients)

Метод нахождения локального минимума функции, при котором направление и величина шага рассчитываются с учетом градиента в текущей точке и информации о предыдущем шаге.

$$\vec{S}_k = -\nabla f(\vec{x}_k) + \omega_k \vec{S}_{k-1}, \quad \omega_i = \frac{\|\nabla f(\vec{x}_i)\|^2}{\|\nabla f(\vec{x}_{i-1})\|^2}$$



В случае квадратичной функции от n переменных

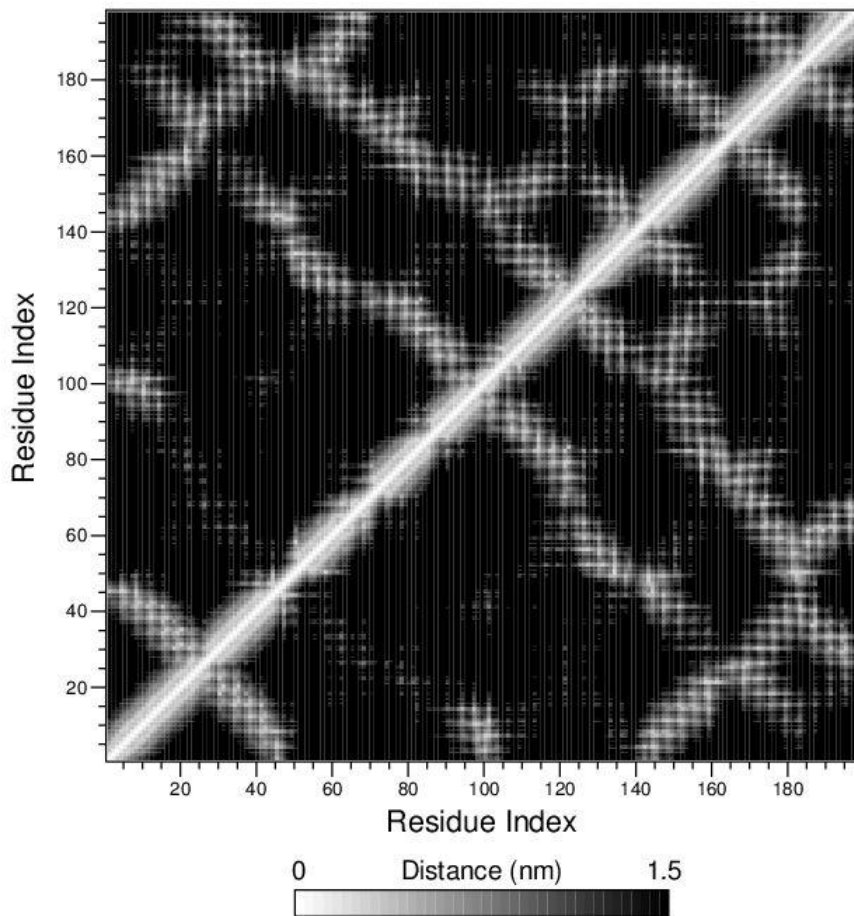
$$f(\vec{x}) = \vec{x}^T A \vec{x} + \vec{b} \cdot \vec{x} + c.$$

метод гарантированно сходится за n шагов.

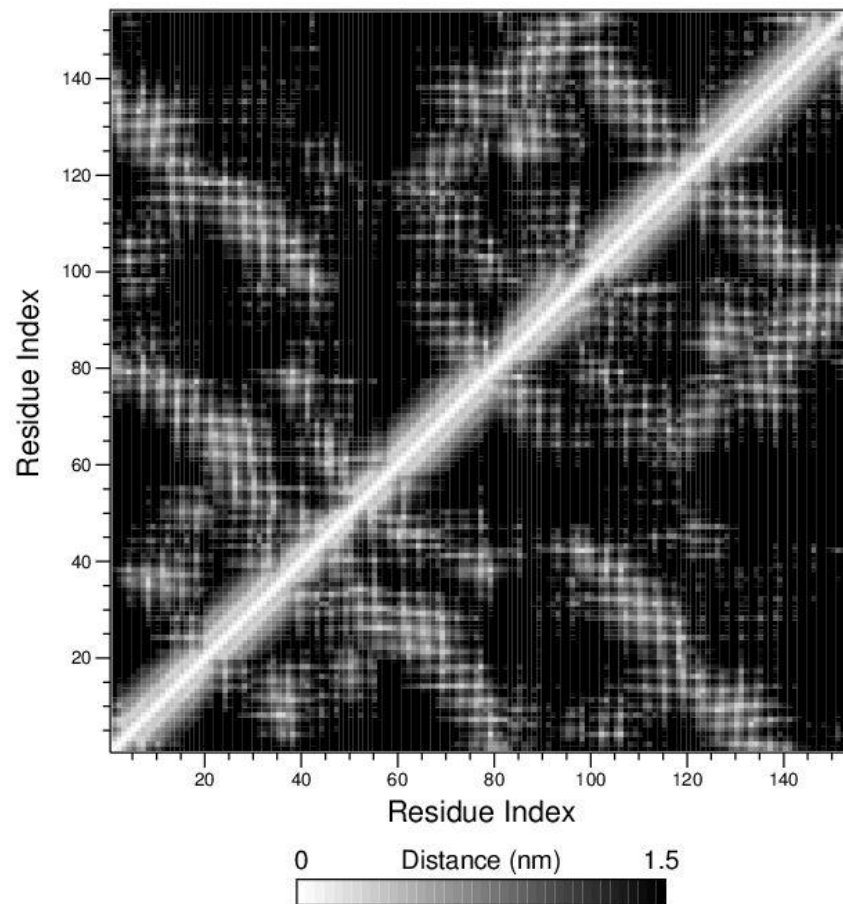
Сравнение структур

Если структуры схожи, то сохраняются паттерны контактов между остатками
⇒ анализируя матрицы расстояний (Ca-Ca), можно распознать схожие структуры

colicin

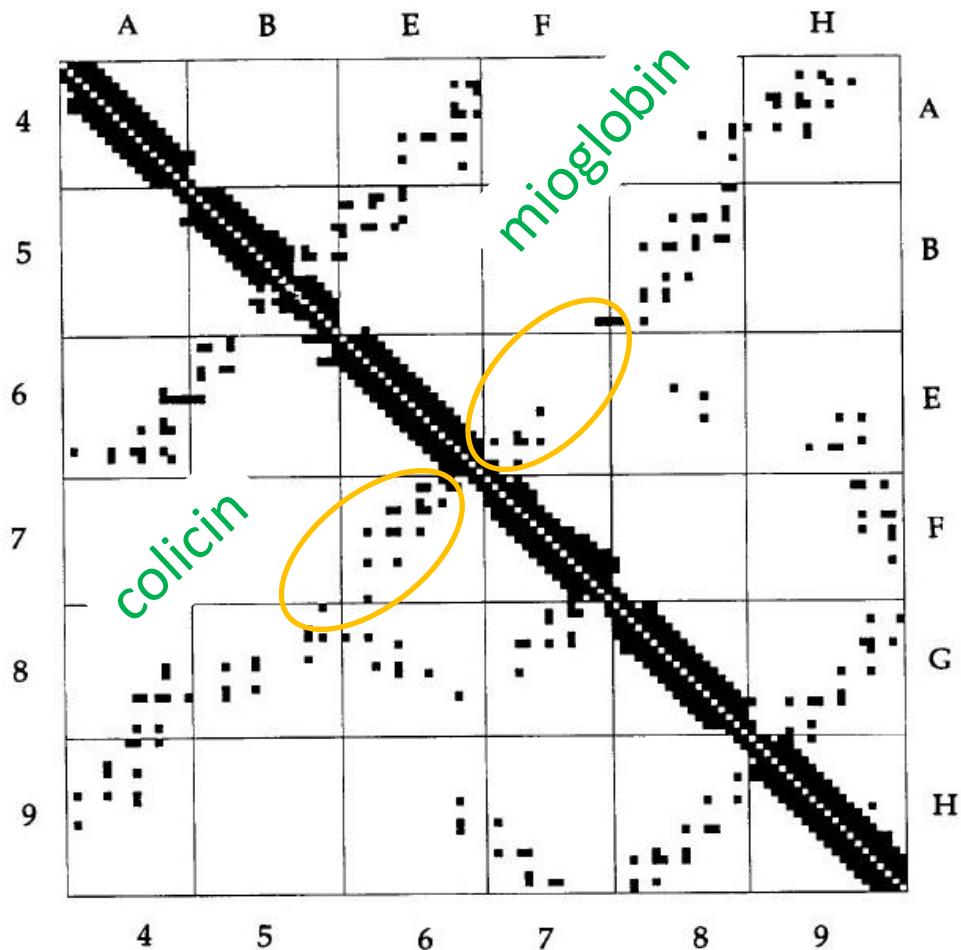


mioglobin



Сравнение структур

Выравнивание матриц расстояний – программа **DALI** (Distance-matrix ALIngment) (1993)

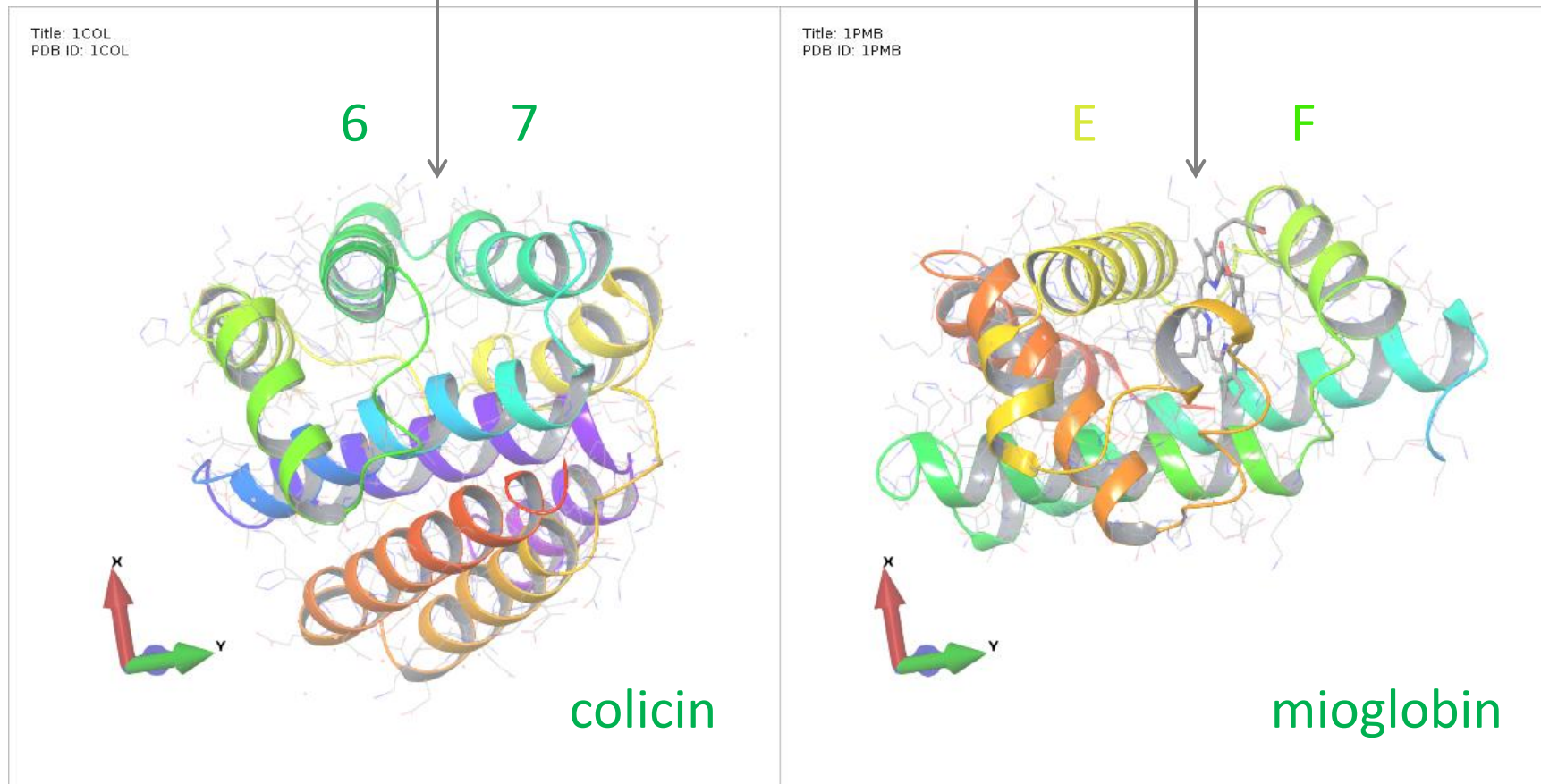


Особенности: свободная топология обнаруженных сходств в структуре, в т.ч. и «обратных» фрагментов.

Сравнение структур

Есть контакт между спиралями

Гем; нет контакта между спиралями




Search for proteins with a similar structure

Structure Alignment Results.

Query: pdb entry 1u19

CRYSTAL STRUCTURE OF BOVINE RHODOPSIN AT 2.2 ANGSTROMS RESOLUTION

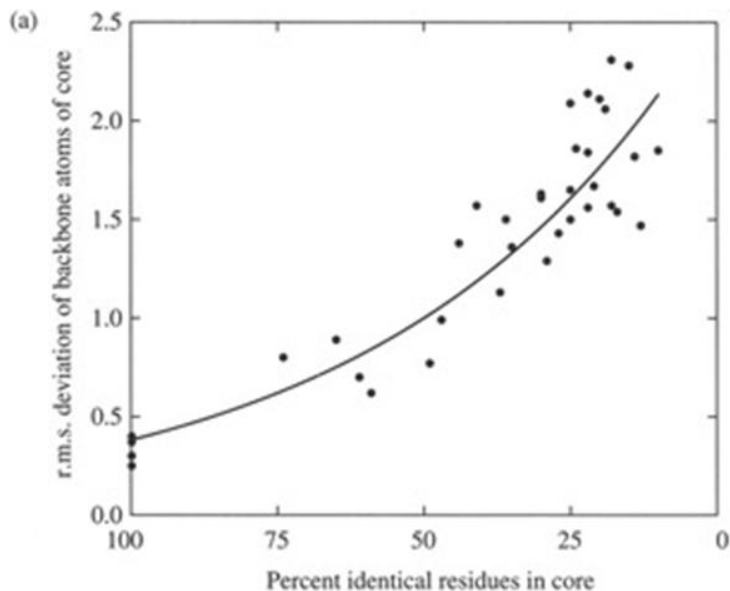
Examined 1 entry, (1 chain). Displaying Matches 1-2 of 2.

##	Scoring 			RMSD	N _{align}	N _g	%seq	Query					Target (PDB entry)				
	Q	P	Z					Ch	N _{res}	%sse	Match	%sse	N _{res}	×	Title		
1	0.23	1.4	7.1	2.20	235	11	22	A	351	55	2rh1:A	27	443	<input checked="" type="checkbox"/>	HIGH RESOLUTION CRYSTAL STRUCTURE OF HUMAN B2-ADRENERGIC G PROTEIN- COUPLED RECEPTOR.		
2	0.23	1.3	7.1	2.23	236	9	21	B	351	73	2rh1:A	36	443	<input checked="" type="checkbox"/>	HIGH RESOLUTION CRYSTAL STRUCTURE OF HUMAN B2-ADRENERGIC G PROTEIN- COUPLED RECEPTOR.		

Examined 1 entry, (1 chain). Displaying Matches 1-2 of 2.

Эволюция белковых структур

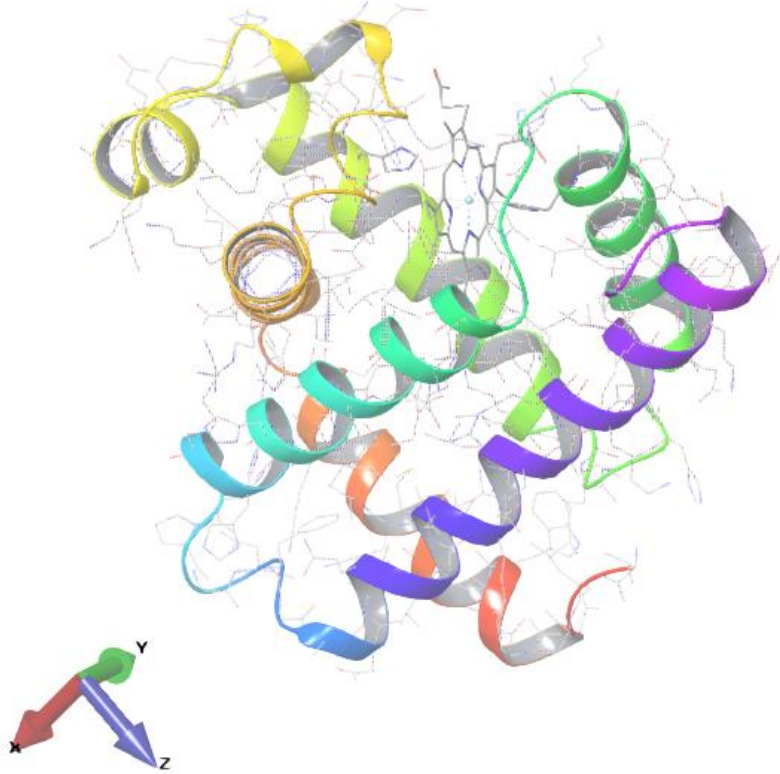
Вариации, встречающиеся в семействах гомологичных белков с одинаковой функцией, показывают, как структура приспособляется к изменениям в последовательности: структура устойчива к мутациям.



Свободно могут мутировать участки на поверхности белка, не влияющие на функцию. В частности, внешние петли легко адаптируются к изменению количества остатков, в то время как мутации, изменяющие число внутренних остатков, приводят к изменению взаимной ориентации спиралей и листов, но не их конформации.

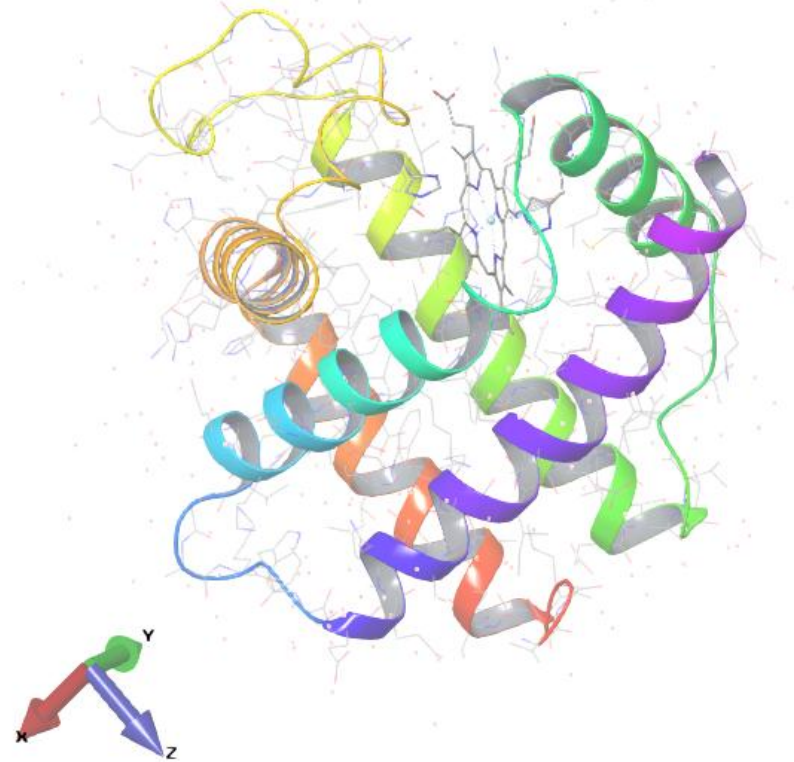
Эволюция белковых структур

Title: 1MBN
PDB ID: 1MBN



Миоглобин кашалота (1mbn, 1969)

Title: 1GDJ
PDB ID: 1GDJ



Леггемоглобин люпина (1gdj, 1995) (ИК РАН)

Эволюция белковых структур

миоглобин кашалота и леггемоглобин люпина



Идентичность 18%*, но родство подтверждается сходством структуры и функции

Классификация структур белков. SCOP

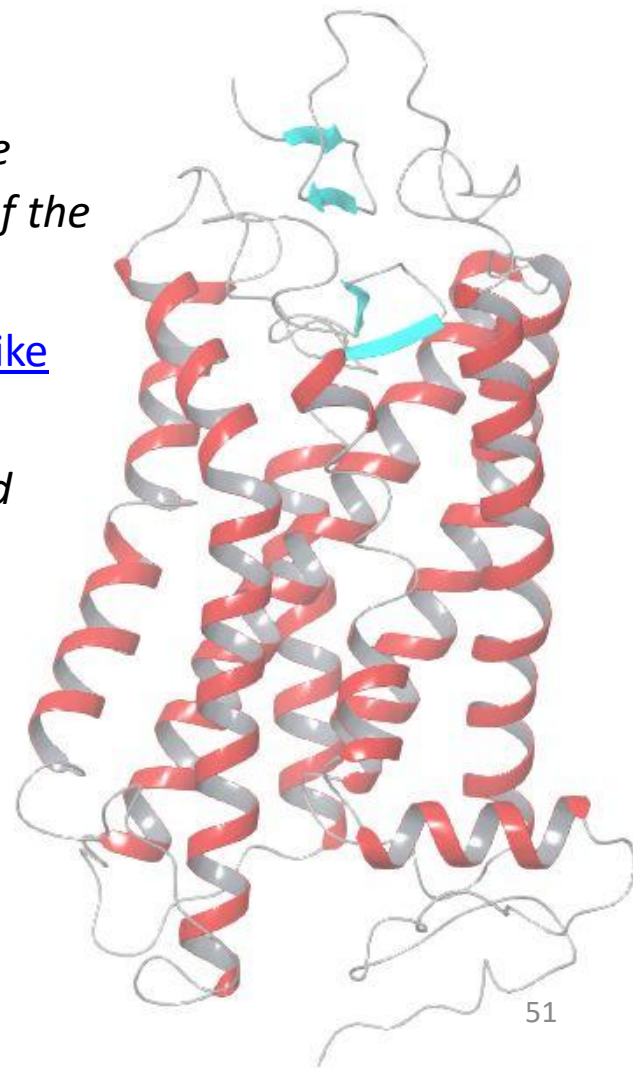


SCOP (Structural classification of proteins) (Murzin et al., 1994 - 2009) – организация структур для отображения их эволюционного происхождения и структурного сходства. Основные уровни организации:

- **Классы:** α , β , $\alpha+\beta$, α/β и разнообразные «малые белки», зачастую имеющие слабую вторичную структуру.
- **Фолды** – надсемейства, имеющие общую топологию укладки (одинаковые элементы вторичной структуры с одинаковым чередованием и одинаковым расположением, по крайней мере в «ядре»), наличие эволюционного предка маловероятно;
- **Надсемейства** – вероятно эволюционно близкие белки с низкой идентичностью, но функции и структуры которых позволяют предположить наличие общего предка (например, актин, АТФ-азный домен HSP и гексакиназы);
- **Семейства** – очевидно эволюционно близкие белки с идентичностью остатков, как правило, не менее 30% (глобины – 15%);

Классификация структур белков. SCOP

- Root: [scop](#)
- Class: [Membrane and cell surface proteins and peptides](#)
Does not include proteins in the immune system
- Fold: [Family A G protein-coupled receptor-like](#)
core: up-and-down bundle of seven transmembrane helices tilted 20 degrees with respect to the plane of the membrane
- Superfamily: [Family A G protein-coupled receptor-like](#)
- Family: [Rhodopsin-like](#)
Individual TM segments have a number of kinks and distortions



<http://scop2.mrc-lmb.cam.ac.uk/> - что-то странное ☹

Классификация структур белков. CATH

CATH (Orengo et al., 1997) – полуавтоматическая иерархическая классификация белковых доменов. Основные уровни организации:

Class – эквивалентно уровню «класс» в SCOP

Architecture – эквивалентно уровню «фолд» в SCOP

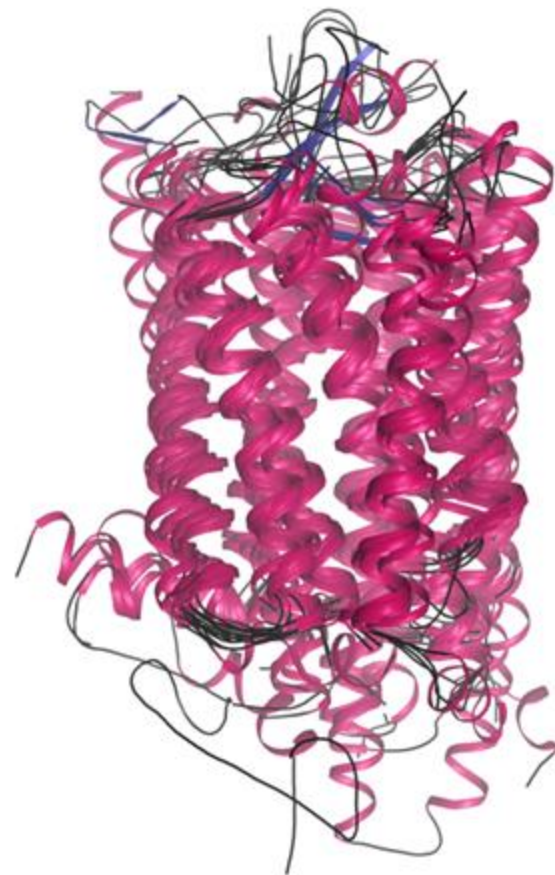
Topology – нечеткий уровень, объединяющий фолды с характерными особенностями

Homologous superfamily – эквивалентно уровню «надсемейство» в SCOP

CATH Superfamily 1.20.1070.10

Rhodopsin 7-helix transmembrane proteins

Хорошо заметно «ядро»



Структурная геномика

Цель: получение максимального количества разнообразных типов укладки (фолдов) белковых структур.

Методы: РСА и спектроскопия ЯМР.

Сроки: 2000 - 2015



Правила отбора мишеней:

- 1) **Новизна**, т.е. в идеале последовательность не должна иметь сходства с белками с уже известной структурой;
- 2) **Актуальность**, т.е. наличие перспективы практического использования полученной структуры, а не только академический интерес;
- 3) **Удобство в работе**, т.е. желательно, чтобы белки были растворимыми, имели повышенное содержание метионина (для решения фазовой проблемы в РСА) и т.п.
- поиск «под фонарем».

Функции расшифрованных структур, как правило, неизвестны и становятся предметом для отдельных исследований

Продолжение следует...