

# Лекция 9

## Структурная биоинформатика

апрель 2018

# Поиск белков со схожей структурой

EMBL-EBI  EB-eye Search All Databases   Reset ? Advanced Search

Databases Tools EBI Groups Training Industry About Us Help Site Index   I contact PDBe

Home > Databases > PDBe > Services > SSM

**Match 1 of 24**

[Back to match list](#) [>>](#) [last match](#) [Back to query](#)

---

| Query PDB 1u19:A  |                  |                  |                  | Alignment        |       |                  |                    | Target PDB 1u19:A   |                  |                  |                  |
|---|------------------|------------------|------------------|------------------|-------|------------------|--------------------|---|------------------|------------------|------------------|
| N <sub>res</sub>  | % <sub>res</sub> | N <sub>SSE</sub> | % <sub>SSE</sub> | Q                | P     | RMSD             | N <sub>align</sub> | N <sub>res</sub>  | % <sub>res</sub> | N <sub>SSE</sub> | % <sub>SSE</sub> |
| 351   | 100              | 11               | 100              | 1.000            | 20.85 | 0.000            | 351                | 351   | 100              | 11               | 100              |
| CRYSTAL STRUCTURE OF<br>BOVINE RHODOPSIN AT 2.2<br>ANGSTROMS RESOLUTION |                  |                  |                  | % <sub>seq</sub> | Z     | N <sub>SSE</sub> | N <sub>gaps</sub>  | CRYSTAL STRUCTURE OF<br>BOVINE RHODOPSIN AT 2.2<br>ANGSTROMS RESOLUTION |                  |                  |                  |
|   |                  |                  |                  | 100.0            | 13.59 | 11               | 0                  |   |                  |                  |                  |

[view](#) [download](#)      [view superposed](#)      [view](#) [download](#)

superpose whole entries      Viewer:

---

**Secondary Structure Alignment**

lul9:A    HHHHSSHHHHHH  
lul9:A    HHHHSSHHHHHH

---

**Query PDB 1u19:A**

|      |          |         |     |     |      |          |         |     |
|------|----------|---------|-----|-----|------|----------|---------|-----|
| 1 H1 | 33 A GLU | 33 HIS  | 65  | <-> | 1 H1 | 33 A GLU | 33 HIS  | 65  |
| 2 H1 | 32 A THR | 70 GLY  | 101 | <-> | 2 H1 | 32 A THR | 70 GLY  | 101 |
| 3 H1 | 36 A PHE | 105 CYS | 140 | <-> | 3 H1 | 36 A PHE | 105 CYS | 140 |
| 4 H1 | 21 A GLY | 149 ALA | 169 | <-> | 4 H1 | 21 A GLY | 149 ALA | 169 |
| 5 SD | 3 A ILE  | 179 GLU | 181 | <-> | 5 SD | 3 A ILE  | 179 GLU | 181 |

**Target PDB 1u19:A**

# 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

# Предсказание структуры белков

Сворачивание белка в уникальную конформацию наводит на мысль об алгоритме формирования структуры белка по его последовательности, но доказательством полноты и правильности нашего понимания могла бы стать его реализация в виде компьютерной программы...

Методы предсказания структуры по последовательности:

- Предсказание вторичной структуры;
- Моделирование по гомологии;
- Распознавание типов укладки (по известной библиотеке фолдов);
- Априорное предсказание новых типов укладки.



# Предсказание вторичной структуры

В настоящее время предсказание может быть выполнено с помощью различных алгоритмов и является относительно точным - ~ 80% элементов вторичной структуры выявляется правильно.

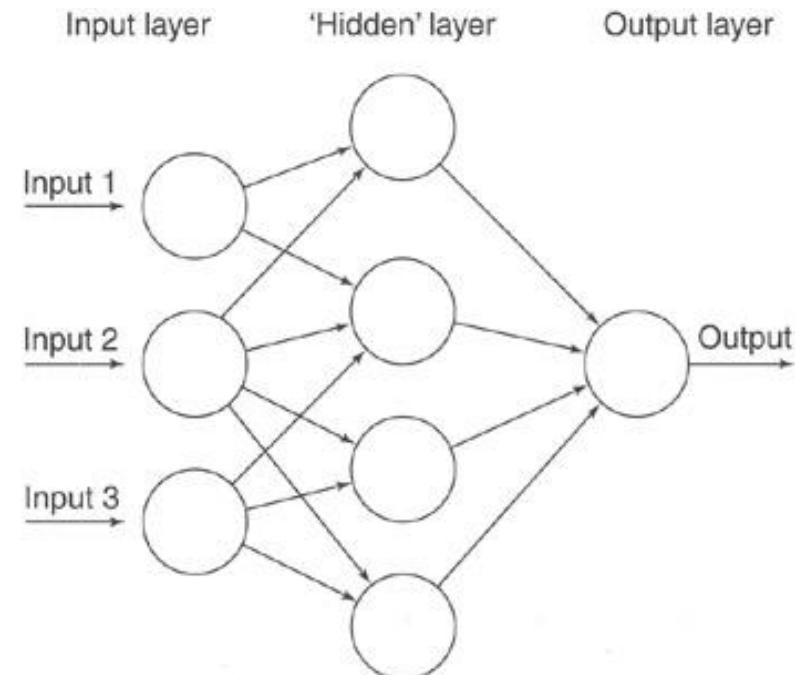
|             | 10   | 20  | 30 | 40 | 50  |
|-------------|--|-----|----|----|-----|
| AA sequence | ALVEDPPLKVSEGLLIREGYDPDLDALRAAHREGVAYFLELEERERERTG     |     |    |    |     |
| Prediction  | HH-----EEE-----HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH    |     |    |    |     |
| Experiment  | -E-----E-----HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH-   |     |    |    |     |
|             | 60   | 70  | 80 | 90 | 100 |
| AA sequence | IPTLKVGYMAVFGYYLEVTRPYYERVPKEYRPVQLKDRQRYTLPEMKEK      |     |    |    |     |
| Prediction  | --EEEEEEEEE-----EEEEEEEEE--EEEE-HHHHHH                 |     |    |    |     |
| Experiment  | -----EEEEEE---EEEEEEEEEHHHHHH-----EEEE---EEEEEE-HHHHHH |     |    |    |     |
|             | 110  | 120 |    |    |     |
| AA sequence | EREVYRLEALIRRREEEVFLEVLERAKRQ                          |     |    |    |     |
| Prediction  | HH   |     |    |    |     |
| Experiment  | HH-- |     |    |    |     |

Наиболее мощные методы предсказания вторичной структуры основаны на нейронных сетях.

# Нейронные сети

**Искусственные нейронные сети (ИНС)** — математические модели, а также их программные или аппаратные реализации, построенные по принципу организации и функционирования сетей нервных клеток живого организма.

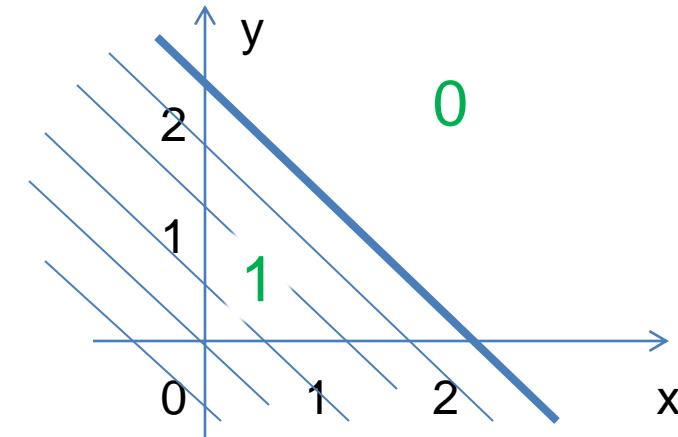
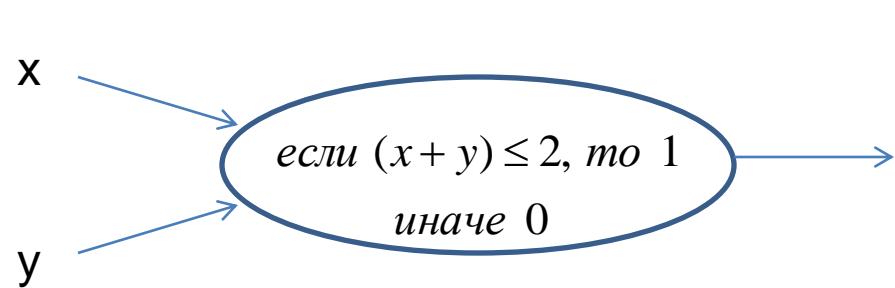
В вычислительной схеме одиночный нейрон является вершиной графа с несколькими входящими ребрами и одним исходящим. Для формирования сети необходимо соединить выходы одних нейронов со входами других.



При этом некоторые нейроны содержат входы для всей сети, некоторые – выходы наружу, а некоторые с внешним миром не связаны (скрытые нейроны).

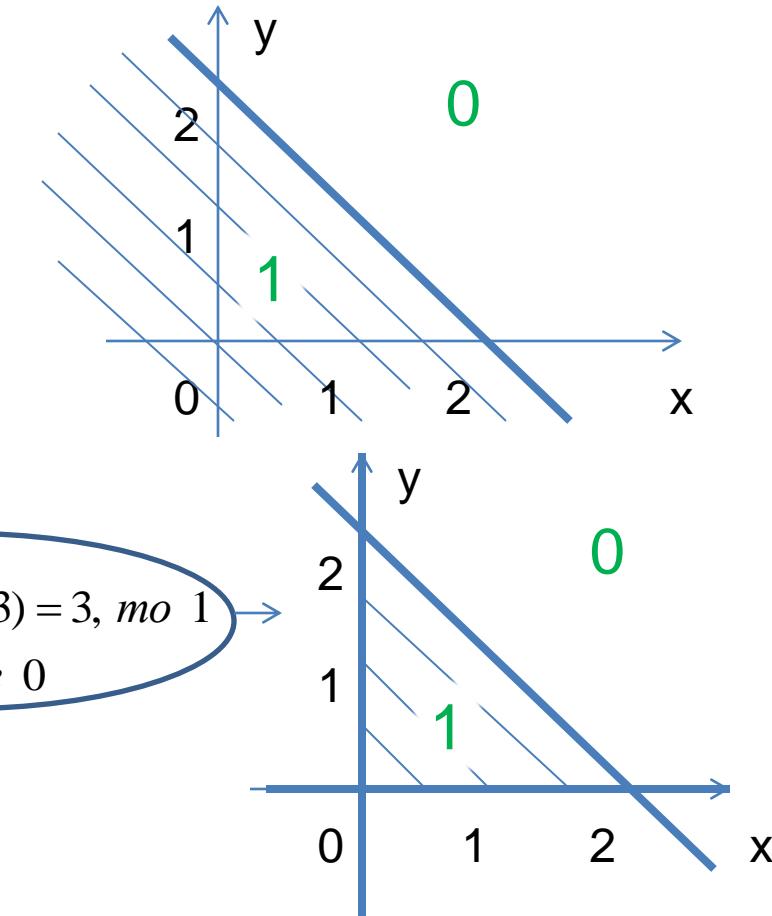
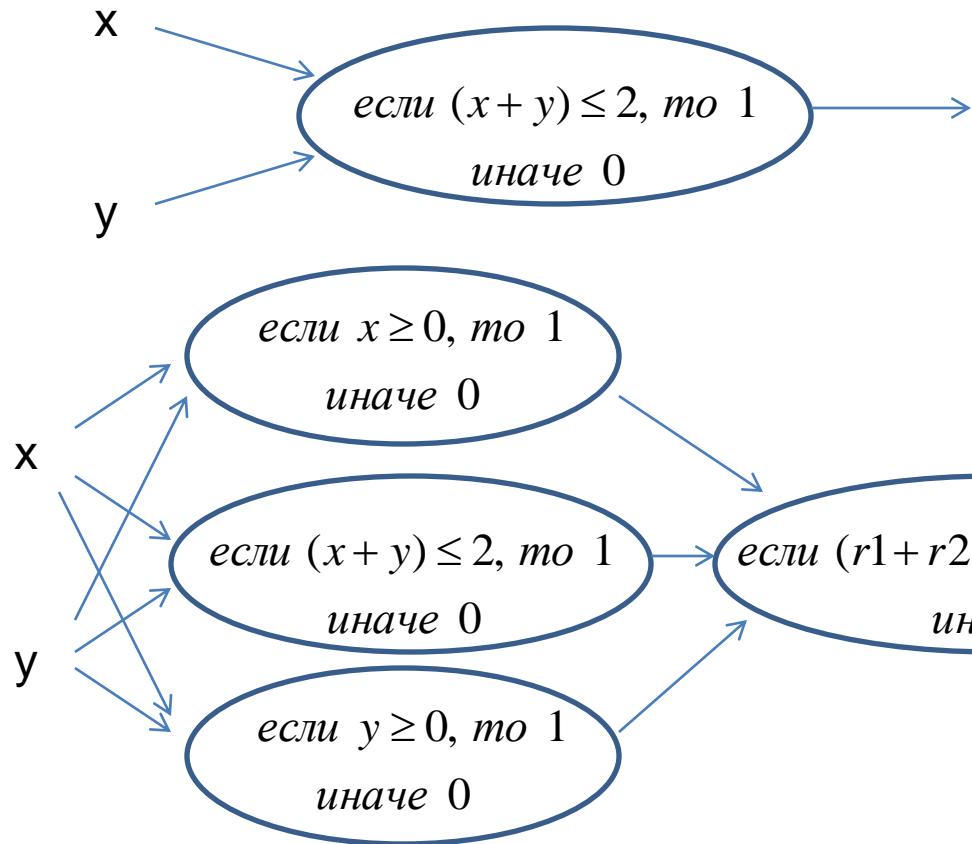
# Нейронные сети. Геометрическая интерпретация

Если интерпретировать пару чисел  $(x, y)$  на входе как точку на плоскости, то данный нейрон принимает решение, на какой стороне от линии находится вход.



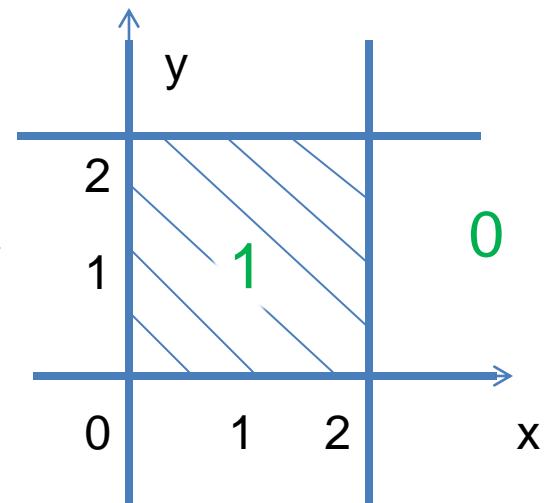
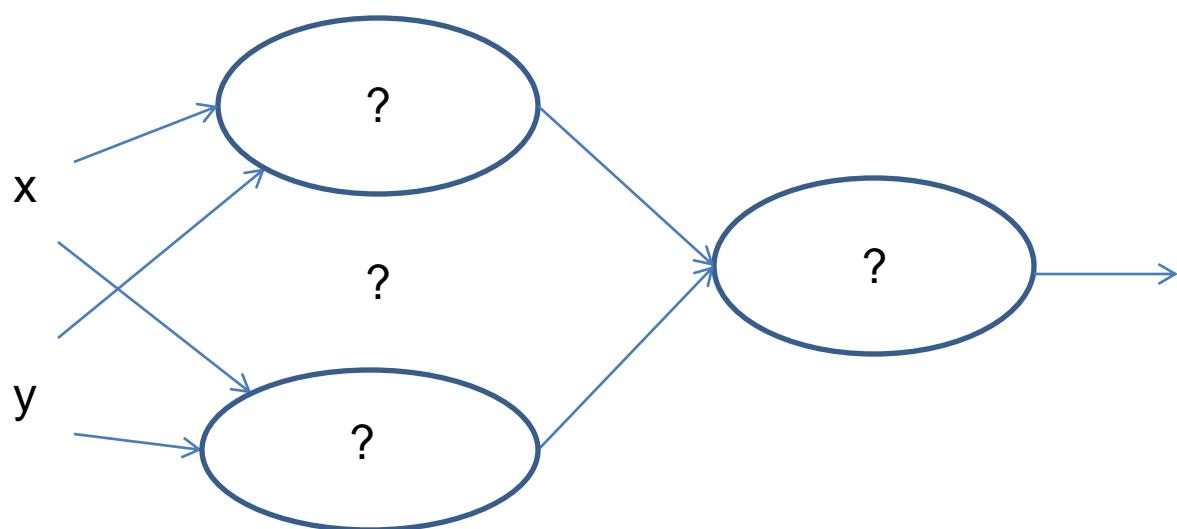
# Нейронные сети. Геометрическая интерпретация

Если интерпретировать пару чисел  $(x, y)$  на входе как точку на плоскости, то данный нейрон принимает решение, на какой стороне от линии находится вход.

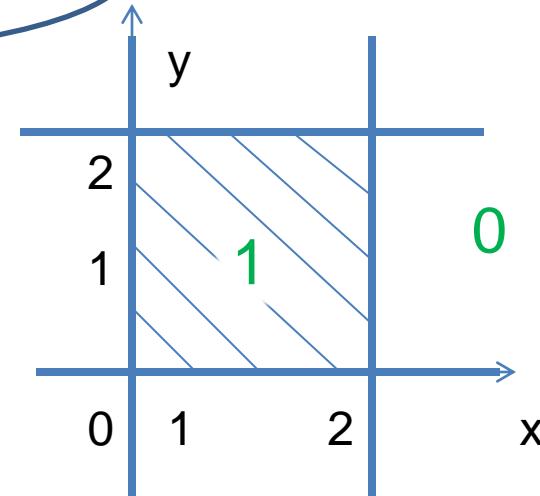
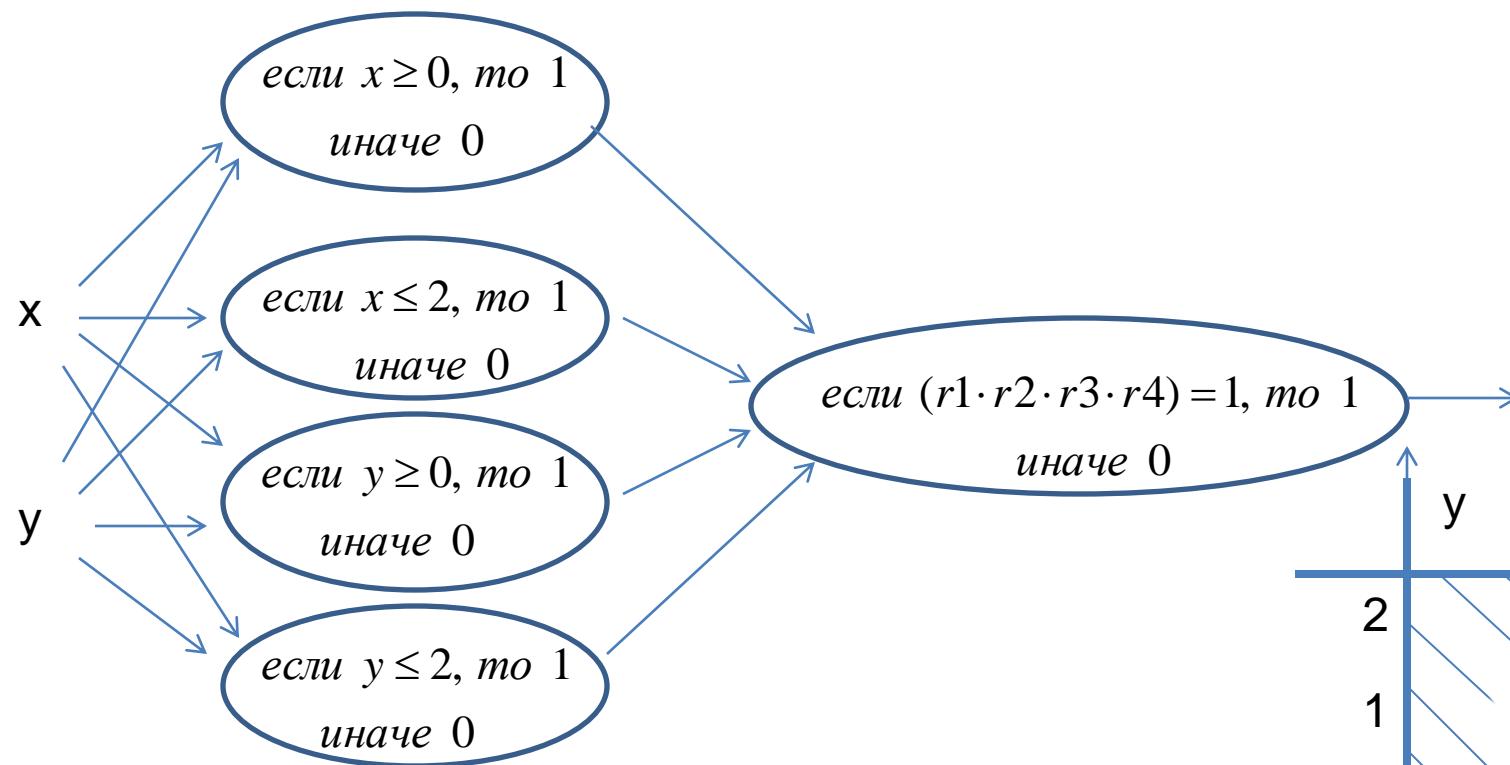


Нейронная сеть определяется топологией связей, весами и формулой принятия решения в узлах. Очевидно, сеть может принимать более сложные решения, чем один нейрон.

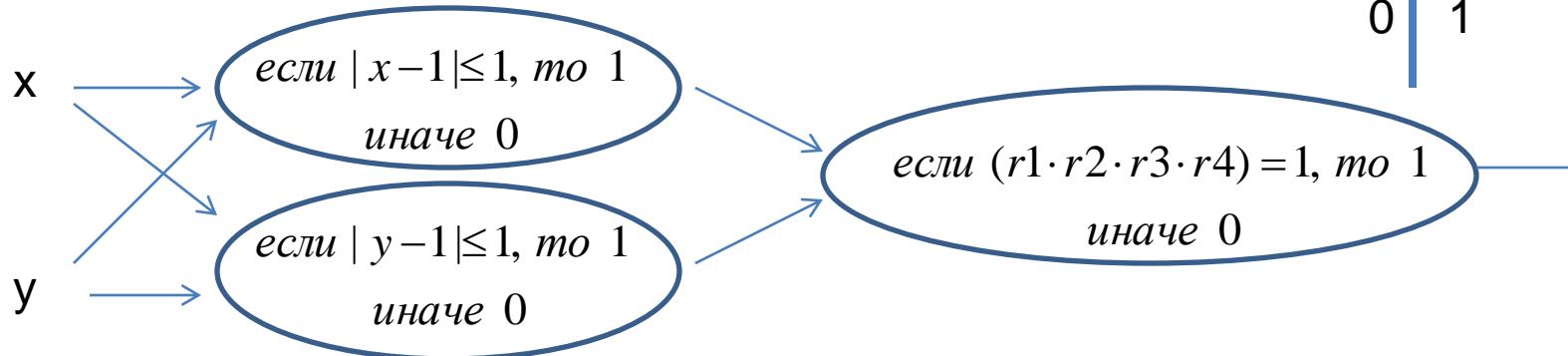
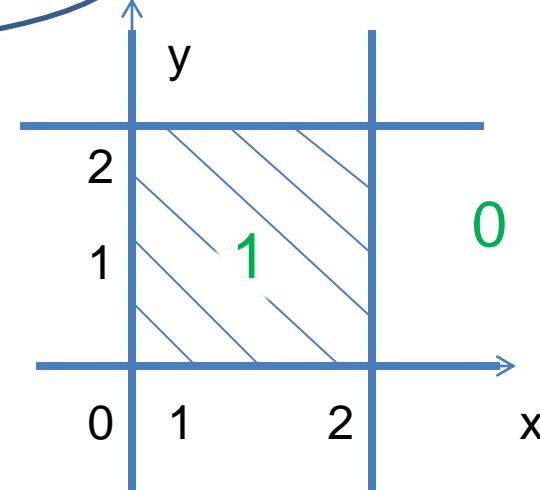
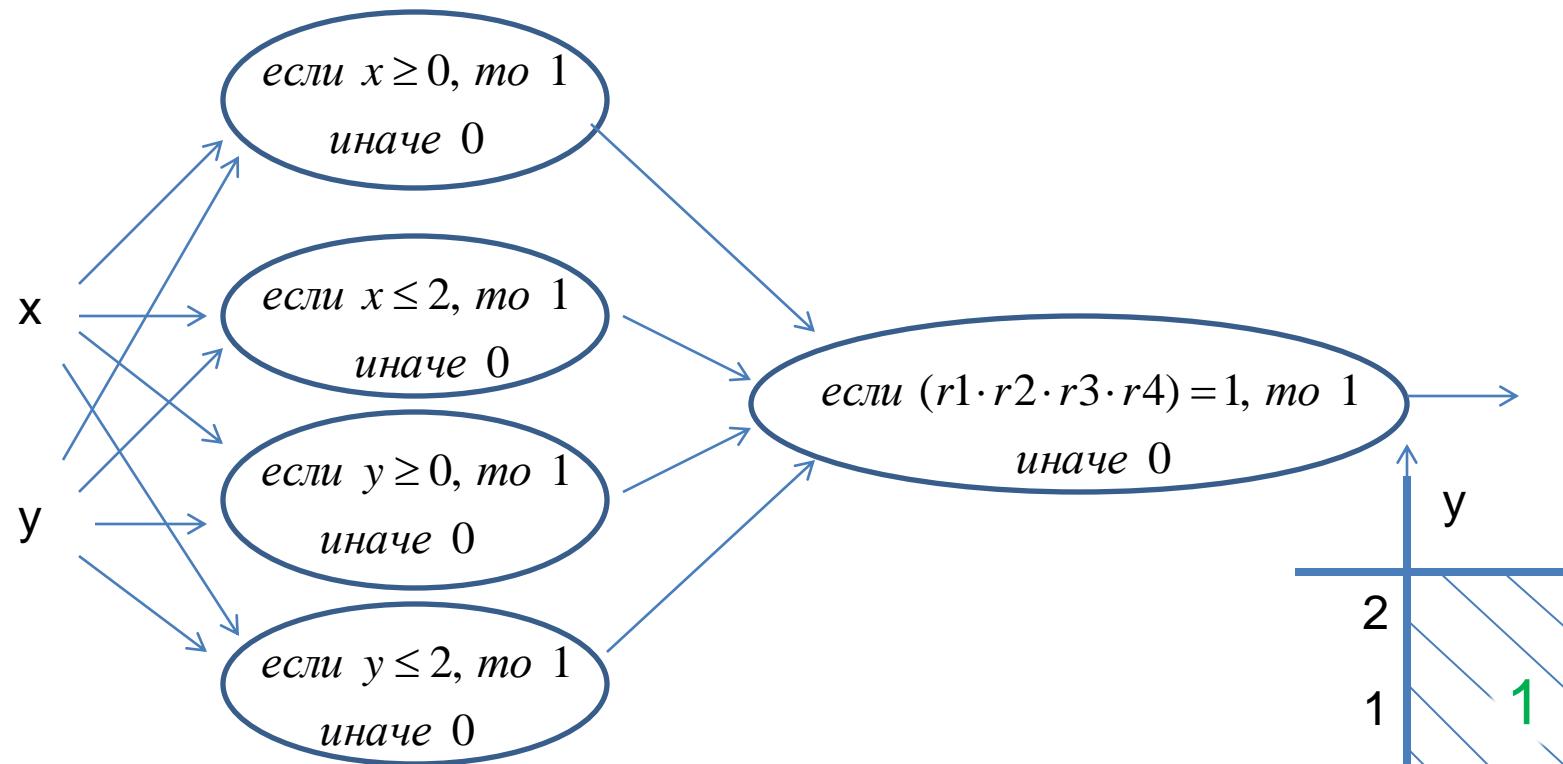
# Нейронные сети. Геометрическая интерпретация



# Нейронные сети. Геометрическая интерпретация

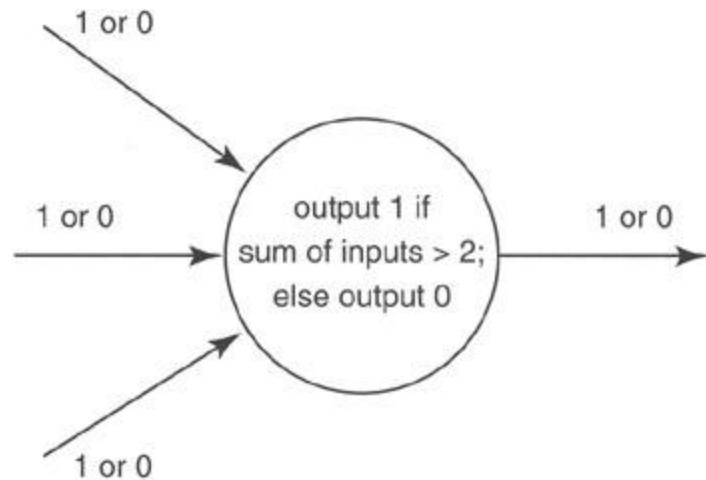


# Нейронные сети. Геометрическая интерпретация



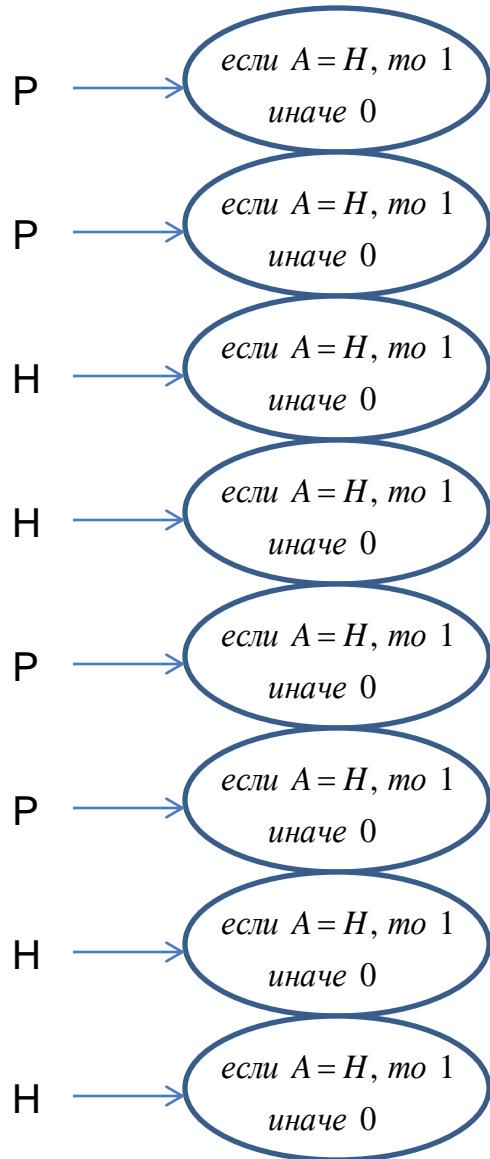
# Нейронные сети

Неограниченная сложность возможна как при создании и соединении нейронов, так и при определении строгости связей. Если вместо того, чтобы просто просуммировать входные сигналы  $i_1 + i_2 + i_3$ , использовать взвешенную сумму входов  $10*i_1 + i_2 + 0,5*i_3$ , то сеть станет более чувствительной ко входу 1 и менее ко входу 3.



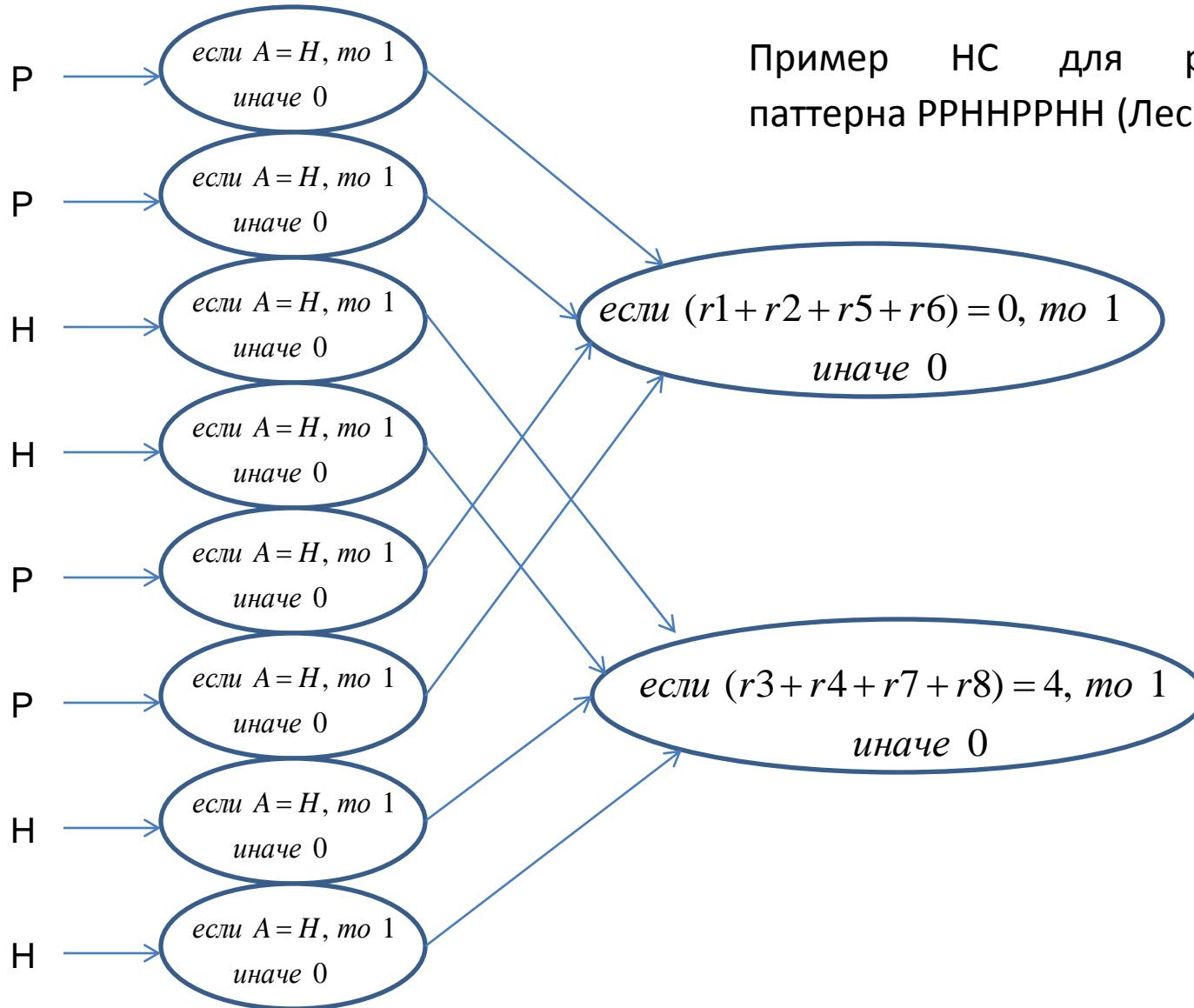
В процессе обучения происходит подбор параметров при неизменной топологии сети. Для этого применяют сеть с начальными параметрами к различным примерам и сравнивают ответ с правильным. При несовпадении производят уточнение параметров.

# Нейронные сети



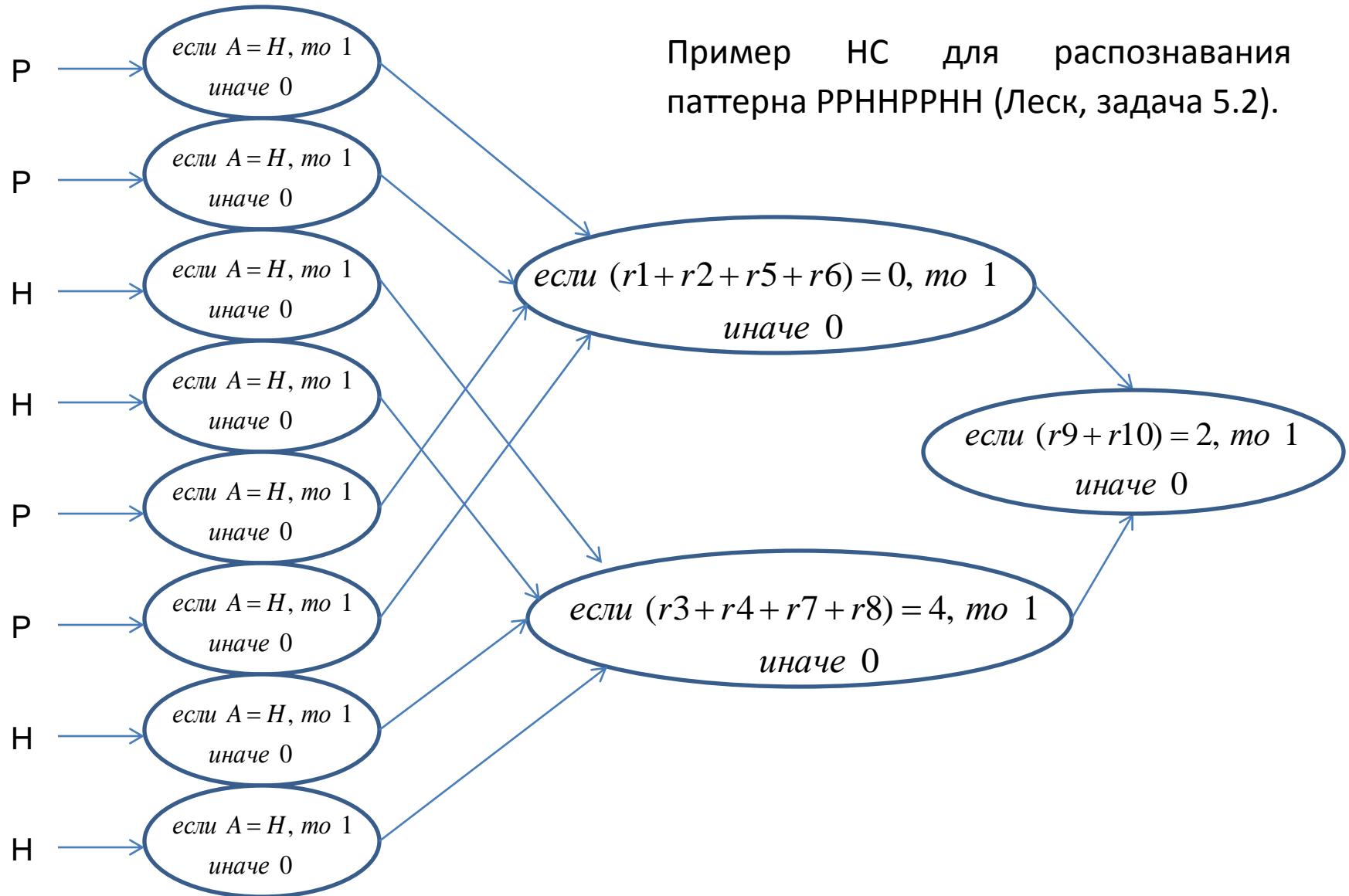
Пример НС для распознавания паттерна РРННРРНН (Леск, задача 5.2).

# Нейронные сети



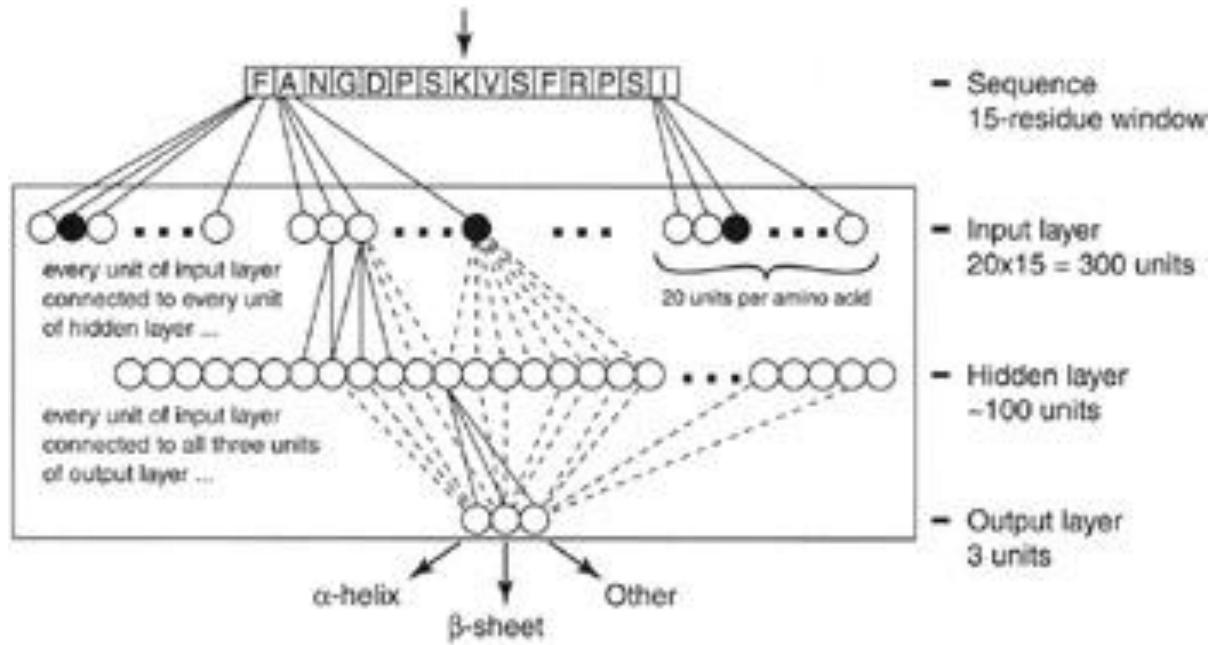
Пример НС для распознавания паттерна РРННРРНН (Леск, задача 5.2).

# Нейронные сети



# Предсказание вторичной структуры. Нейронные сети

Пример нейронной сети для предсказания вторичной структуры (PSIPRED, 1999).



Входная область сканирует последовательность окном шириной в 15 остатков, при этом предсказание делается для центрального остатка. Каждому из остатков соответствует 20 входных нейронов, один из которых активен.

Скрытая область состоит из ~ 100 нейронов, соединенных с каждым нейроном ввода и вывода.

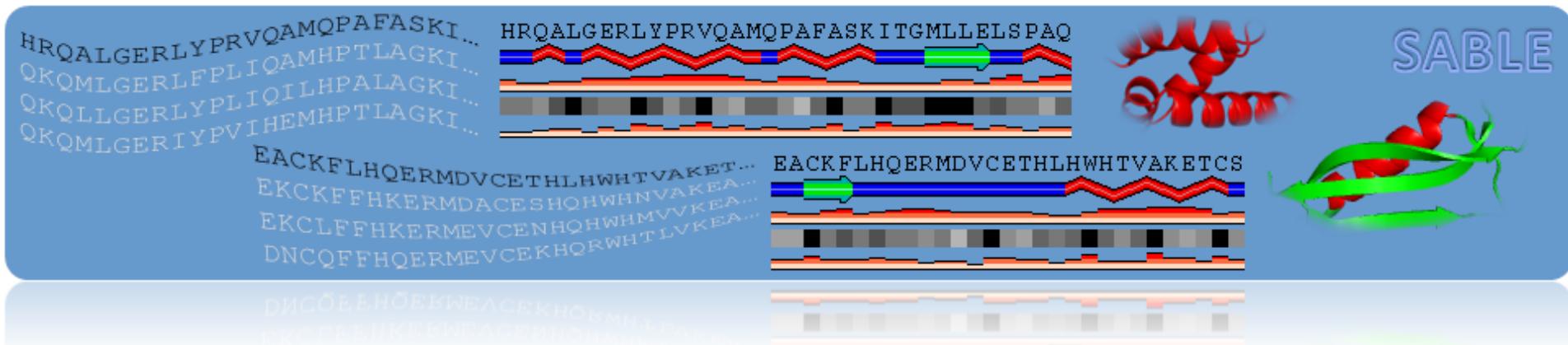
Область вывода состоит из трех нейронов, которые делают предсказание: спираль, лист или ни то, ни другое.

# Предсказание вторичной структуры. PSIPRED и SABLE

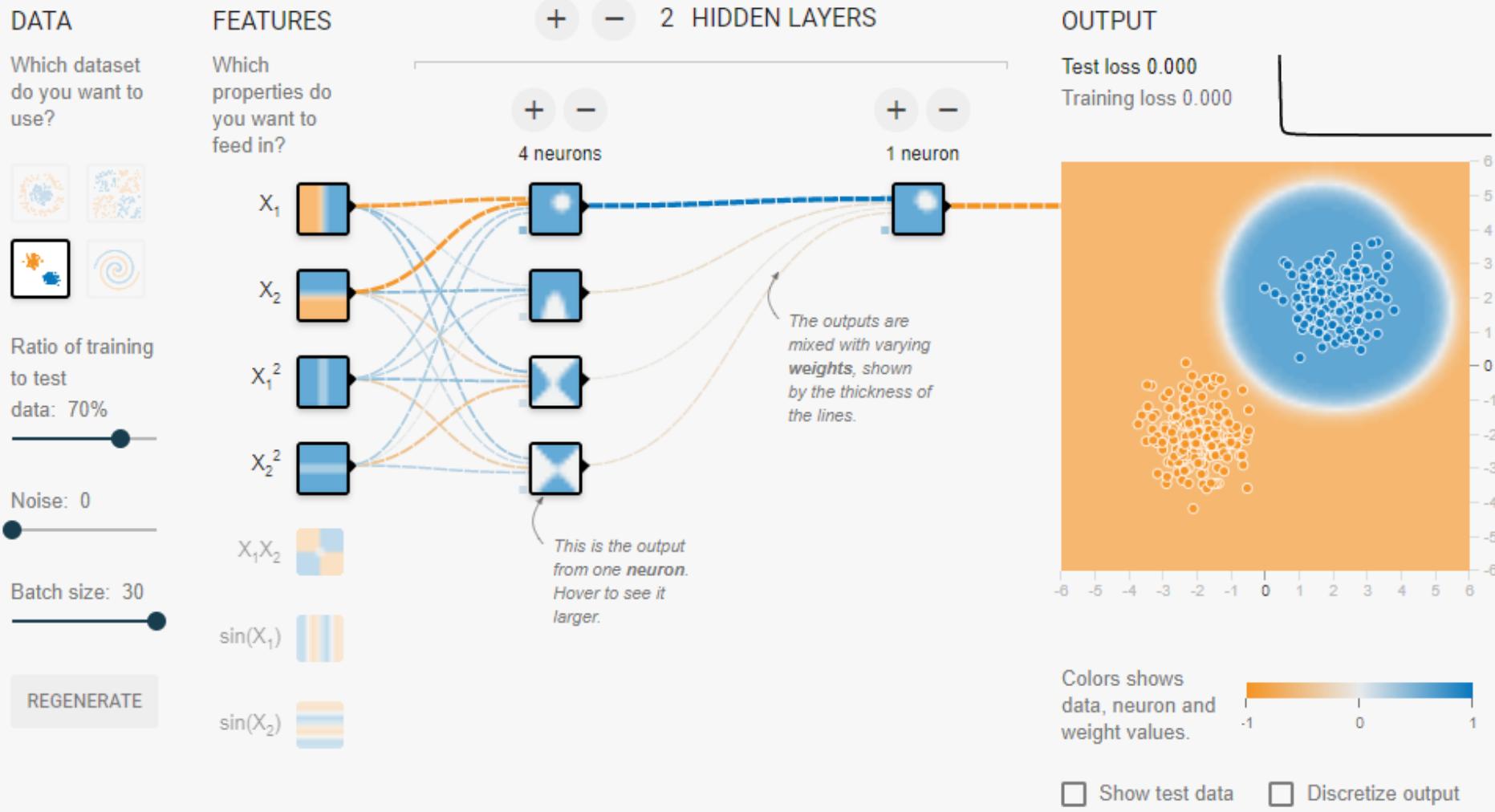
## The PSIPRED Protein Sequence Analysis Workbench

The PSIPRED Protein Sequence Analysis Workbench aggregates several UCL structure prediction methods into one location. Users can submit a protein sequence, perform the predictions of their choice and receive the results of the prediction via e-mail or the web.

For a summary of the available methods you can read [More...](#)



# A Neural Network Playground



# Контроль посещаемости сегодня

Заполнить форму через QR-код



3 мин

2 мин

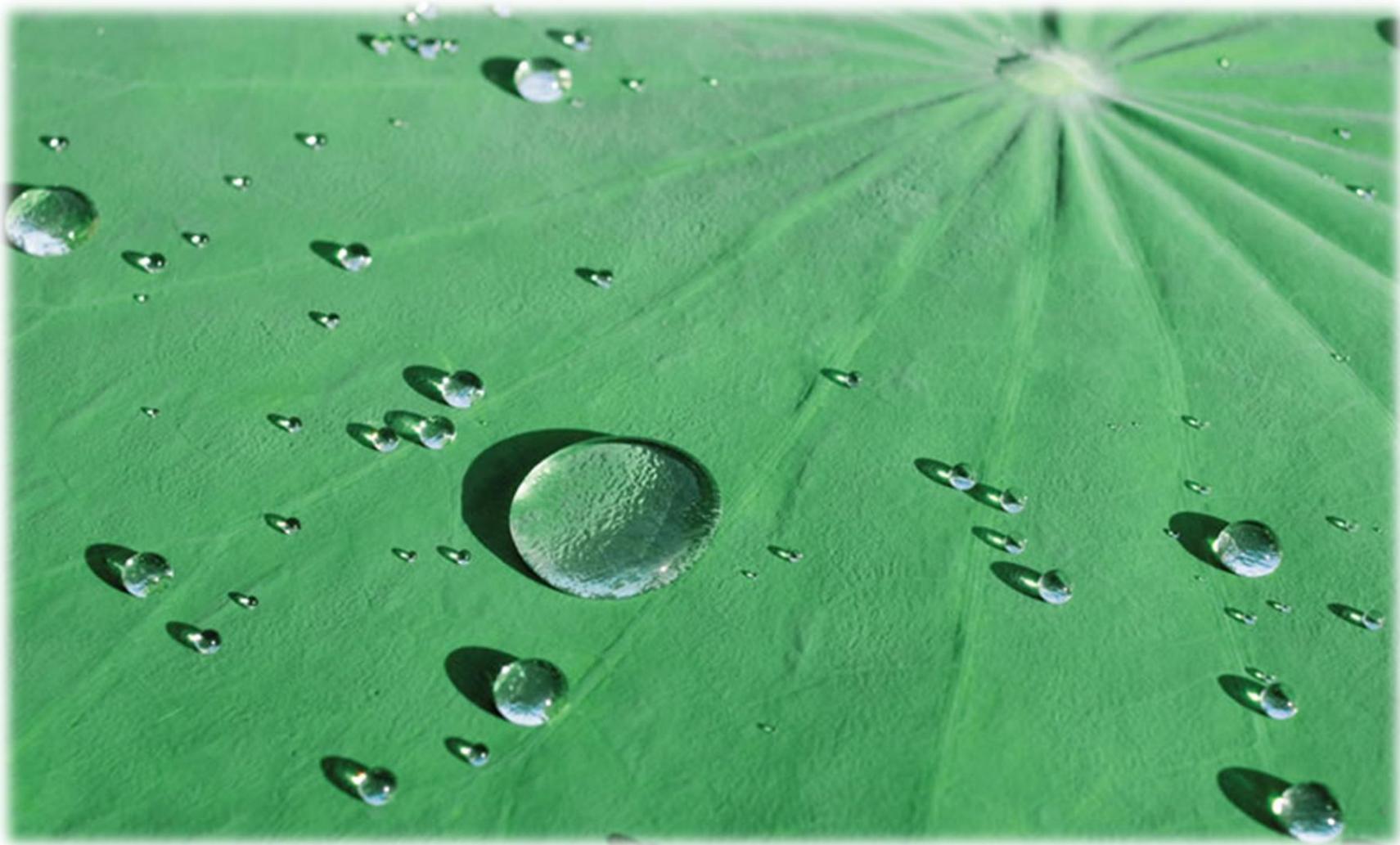
1 мин

Всё!



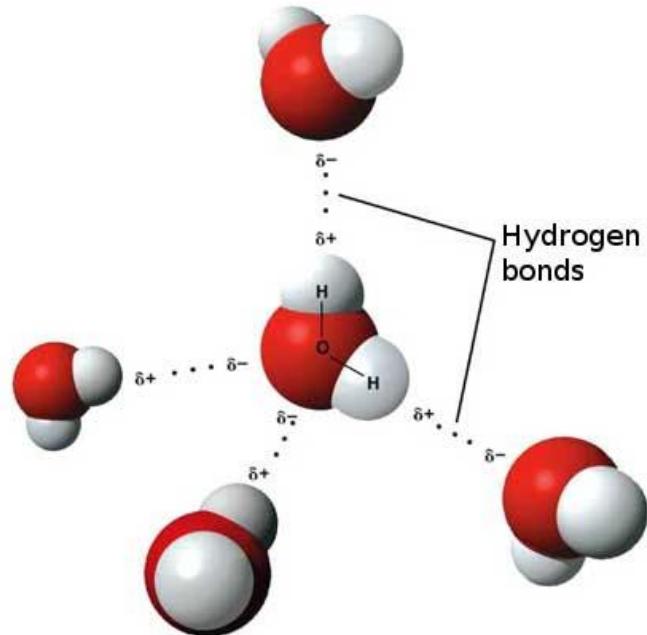
либо записаться в список на перерыве

# Гидрофобность



# Гидрофобность

Гидрофобный эффект – следствие большей упорядоченности молекул воды вокруг неполярной молекулы.

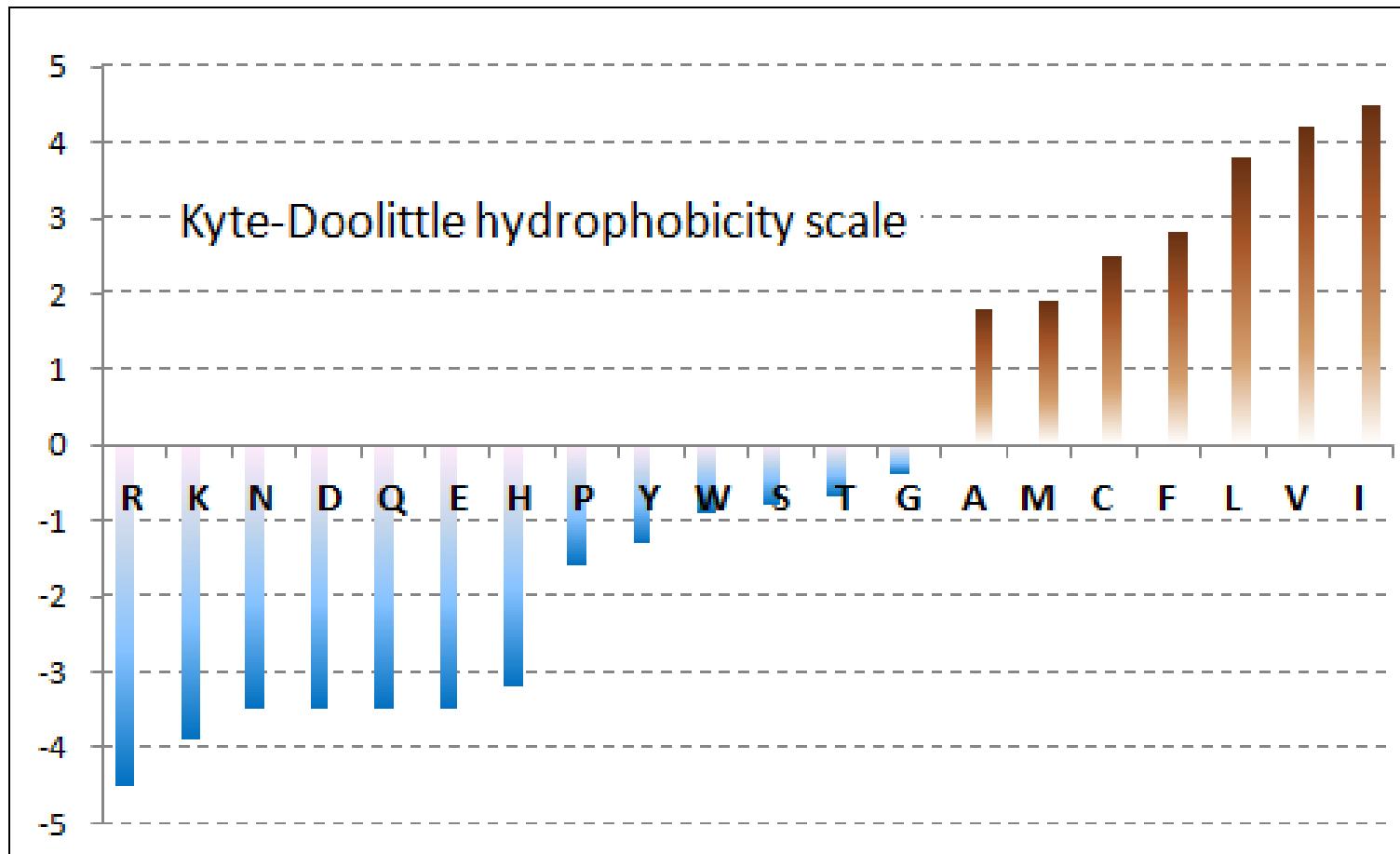


Мерой гидрофобности молекул может служить коэффициент разделения – равновесное отношение концентраций вещества в двух фазах в случае несмешивающихся растворителей:

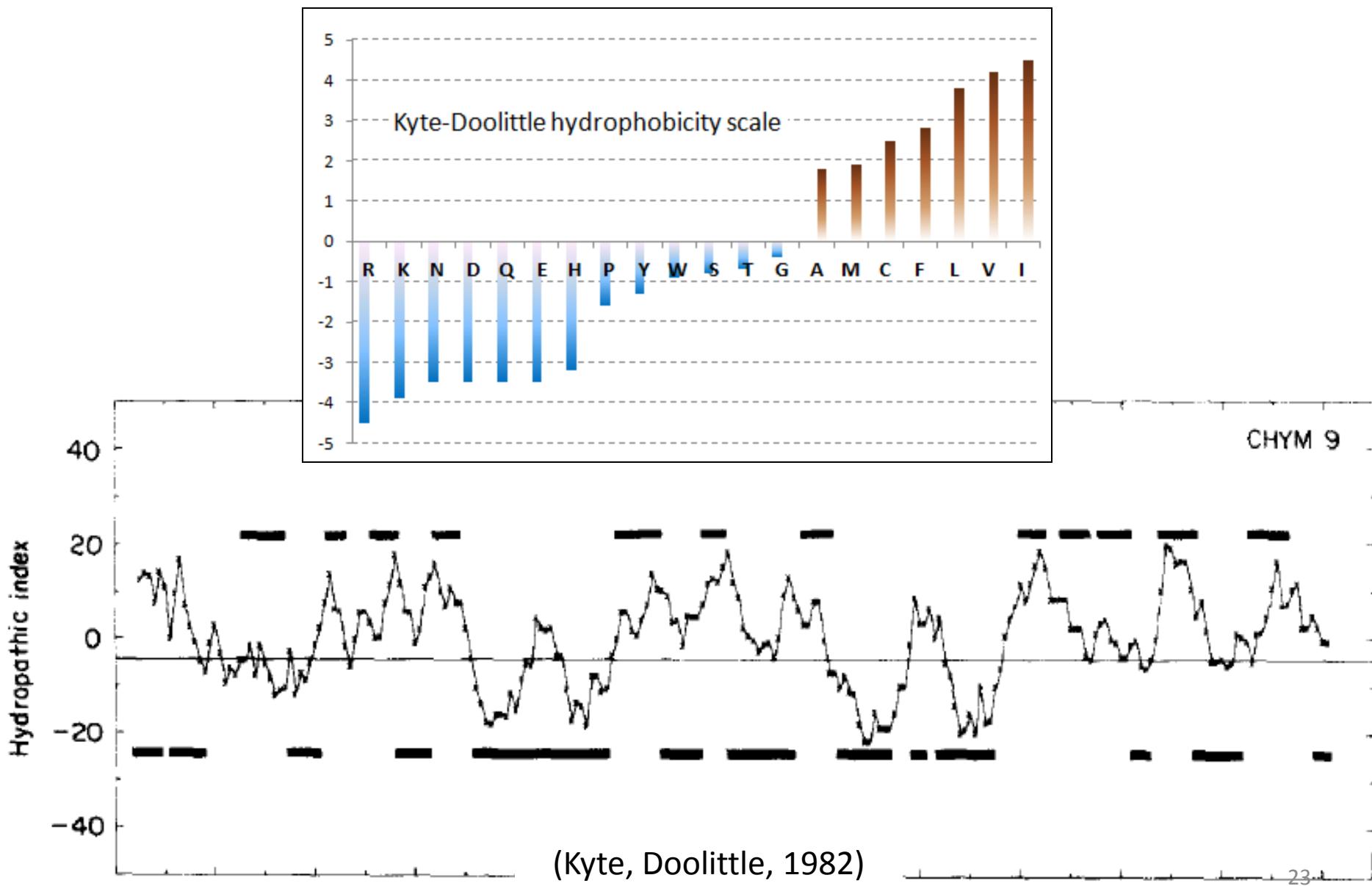
$$\log P_{oct/wat} = \log \left( \frac{[solute]_{octanol}}{[solute]_{un-ionized}_{water}} \right)$$

# Гидрофобность

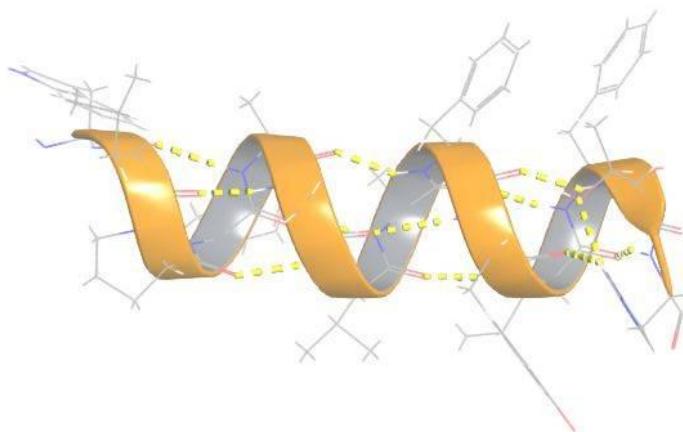
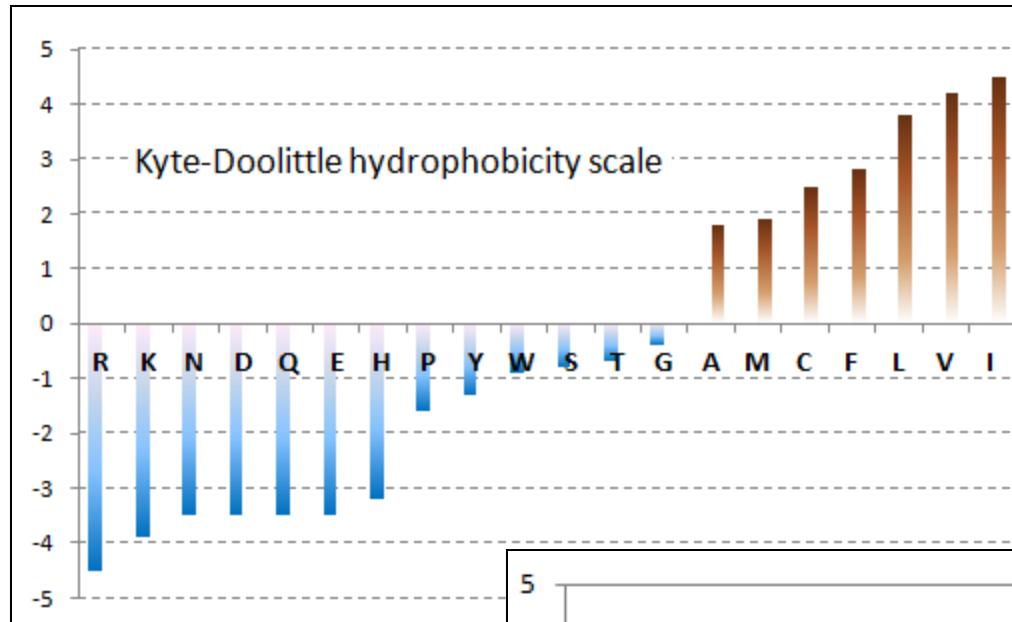
(Kyte, Doolittle, 1982) – распространенная шкала туманного происхождения



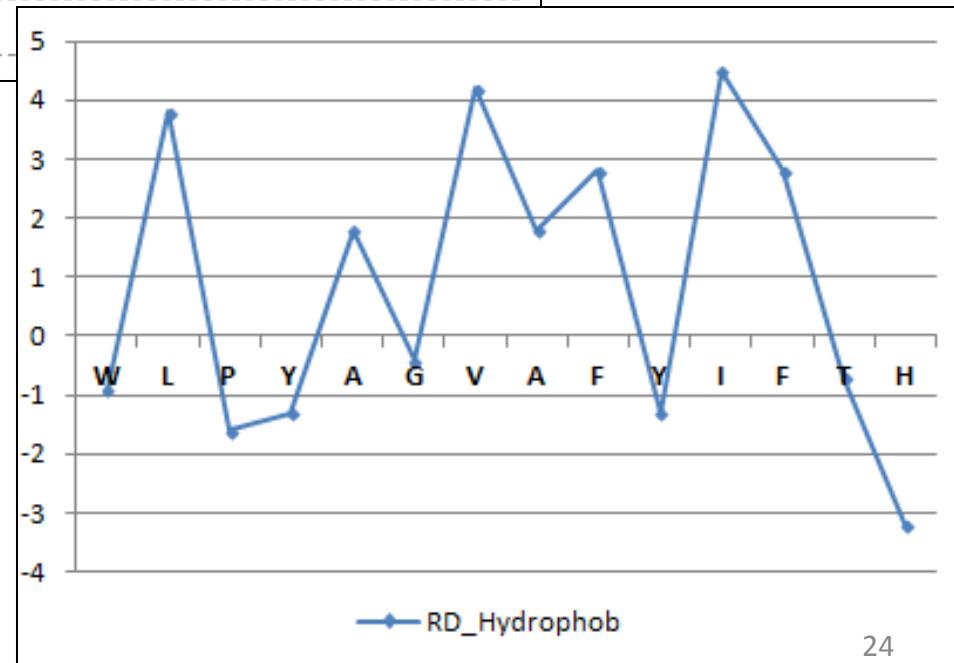
# Гидрофобность



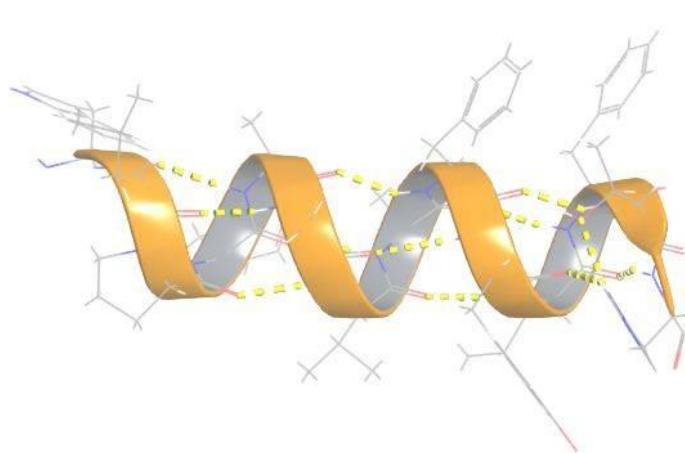
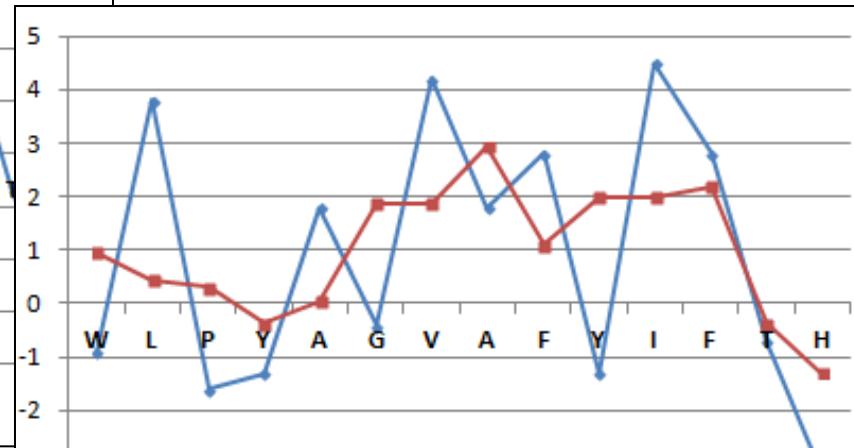
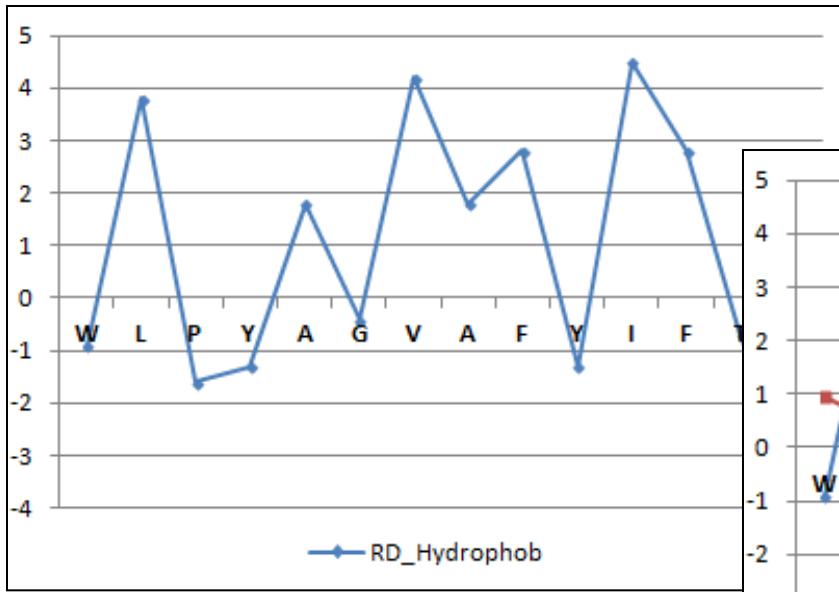
# Гидрофобность



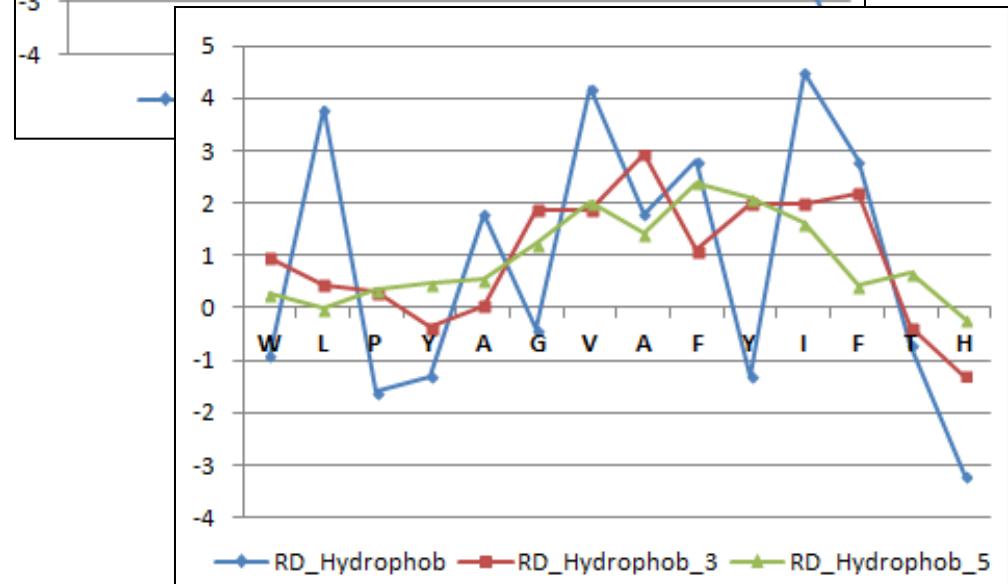
...WLPY**A**GVA**F**Y**I**FTH...



# Предсказание топологии. Профили гидрофобности

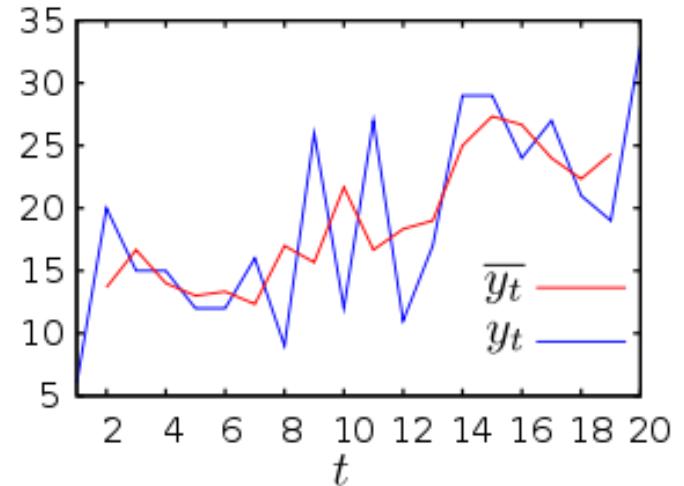


. . . W L P Y A G V A F Y I F T H . . .



## Скользящее среднее

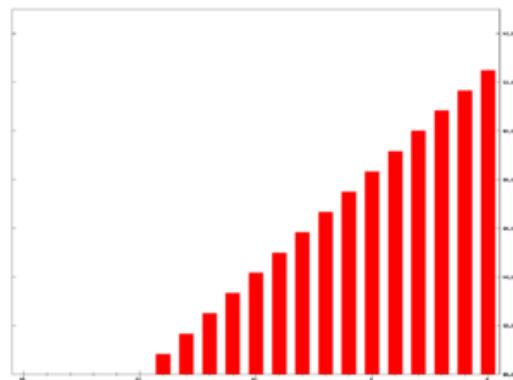
$$\bar{y}_t = \frac{y_t + y_{t-1}}{2}$$



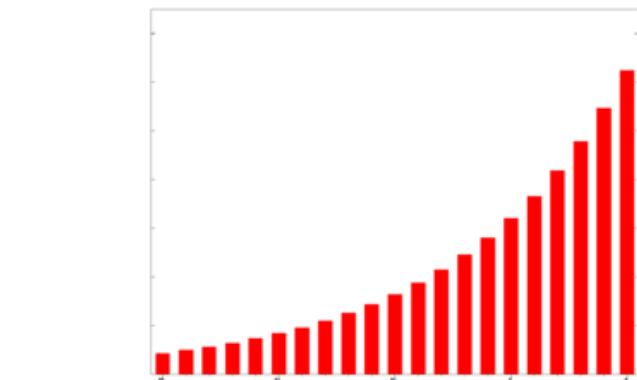
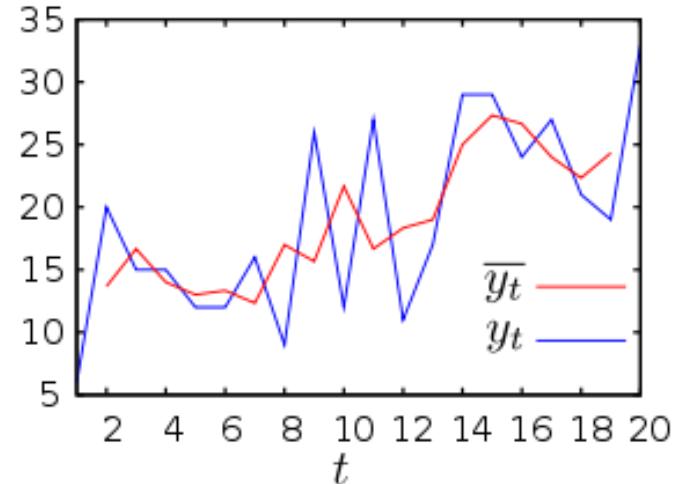
# Скользящее среднее

$$\bar{y}_t = \frac{y_t + y_{t-1}}{2}$$

$$WMA_t = \frac{\sum_{i=0}^{n-1} W_{t-i} \cdot p_{t-i}}{\sum_{i=0}^{n-1} W_{t-i}}$$

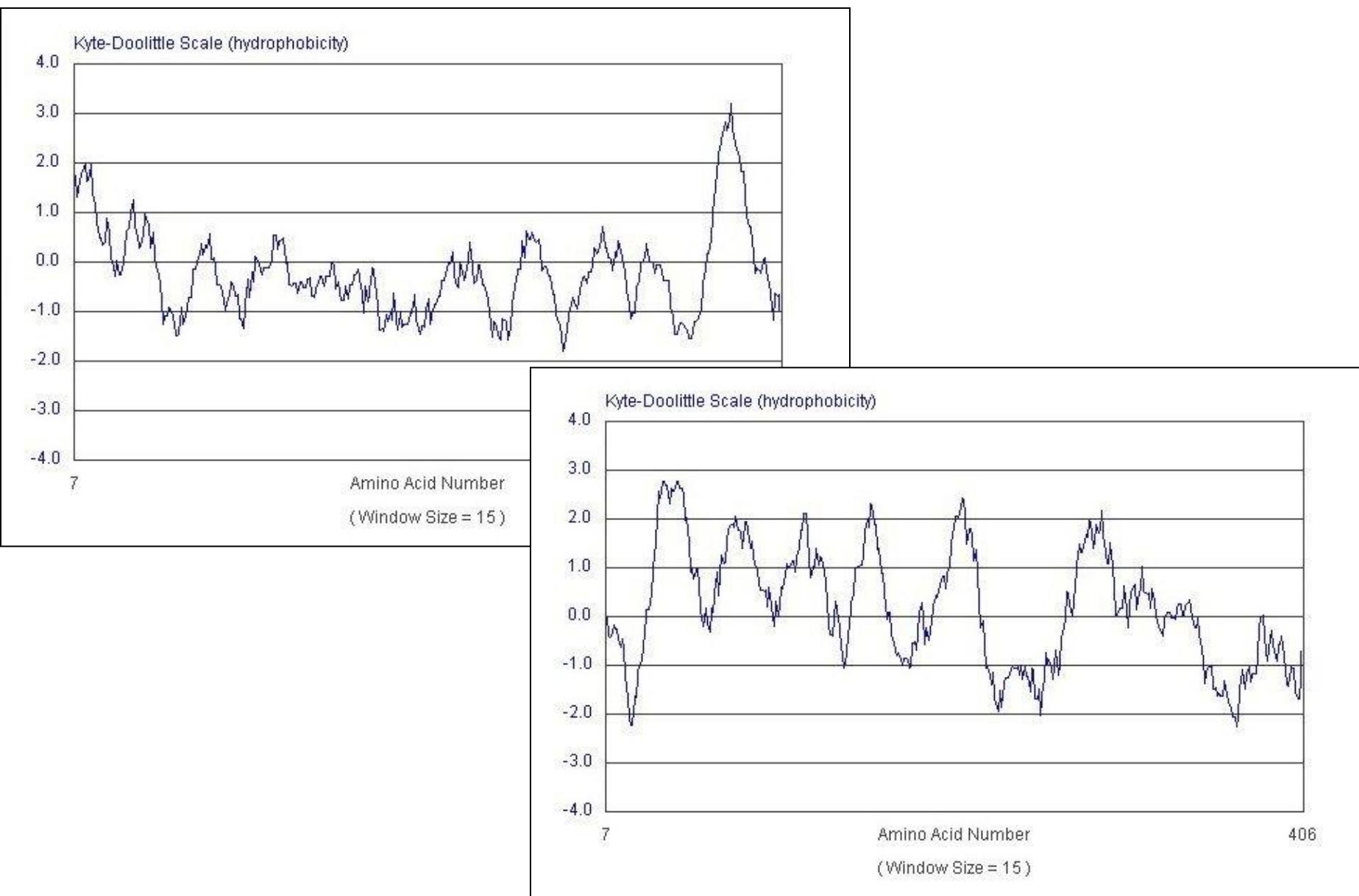


Линейное взвешивание



Экспоненциальное взвешивание <sub>27</sub>

# Предсказание топологии. Профили гидрофобности



# Профили гидрофобности

<http://www.vivo.colostate.edu/molkit/hydropathy/>



«Apologies, but this site was taken down. The programs were written in an older version of Java which is not compatible with most modern browsers and I do not have time to re-code.»

<http://gcat.davidson.edu/DGPB/kd/kyte-doolittle.htm>



Last Modified: Wednesday, 27 February 2002

<https://web.expasy.org/protscale/>

# Предсказание структуры белков

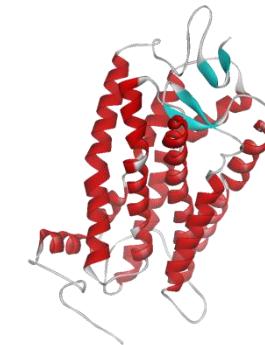
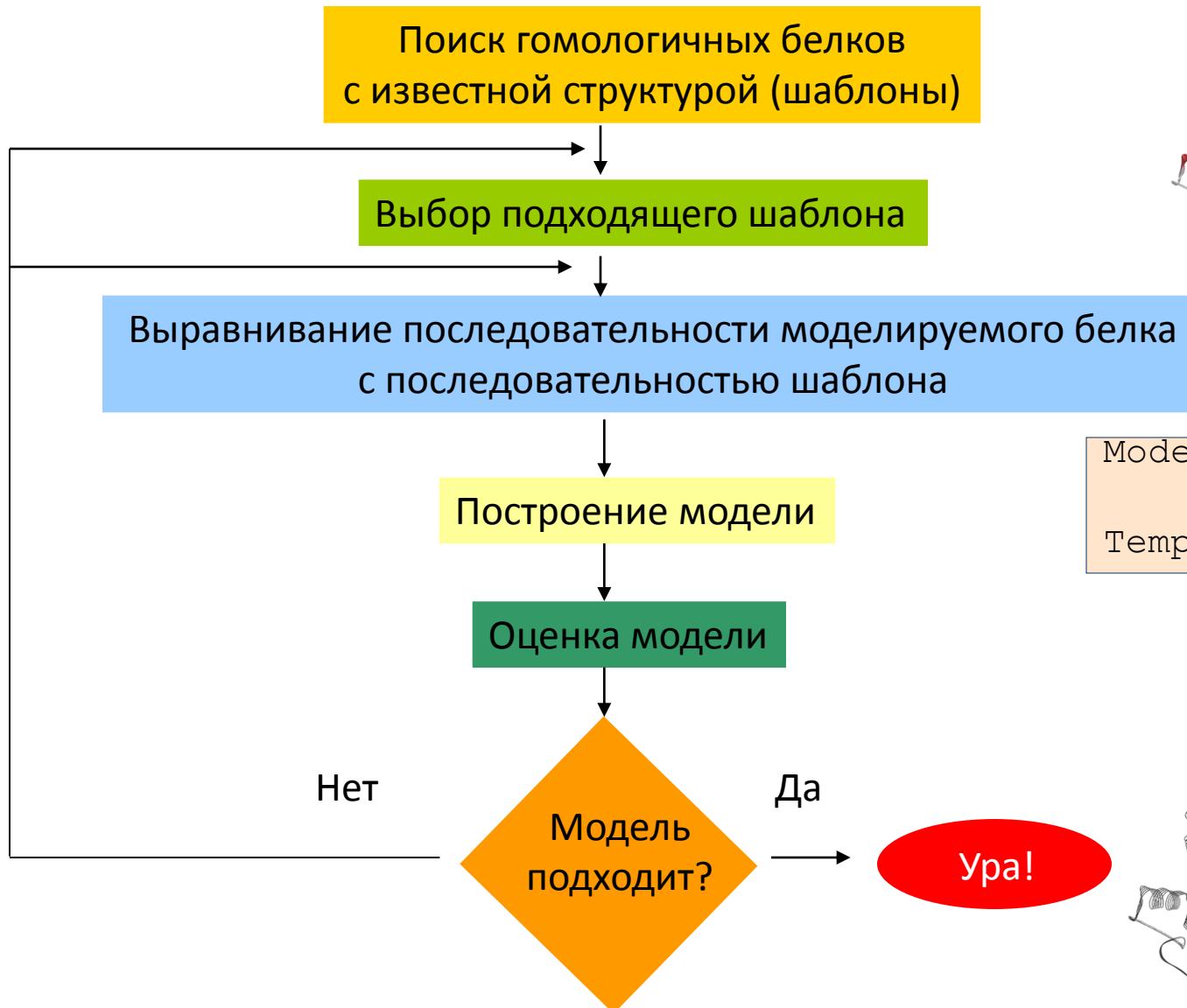
Сворачивание белка в уникальную конформацию наводит на мысль об алгоритме формирования структуры белка по его последовательности, но доказательством полноты и правильности нашего понимания могла бы стать его реализация в виде компьютерной программы...

Методы предсказания структуры по последовательности:

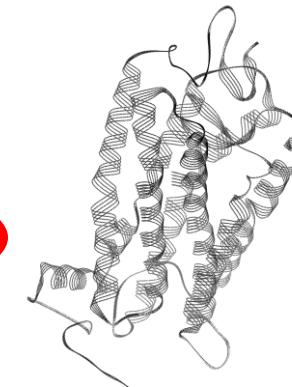
- Предсказание вторичной структуры;
- Моделирование по гомологии;
- Распознавание типов укладки (по известной библиотеке фолдов);
- Априорное предсказание новых типов укладки.



# Моделирование на основании гомологии. Алгоритм



|           |             |
|-----------|-------------|
| Model:    | FVVFVL.FAIC |
|           | ::  :       |
| Template: | VIIMVIAFLIC |

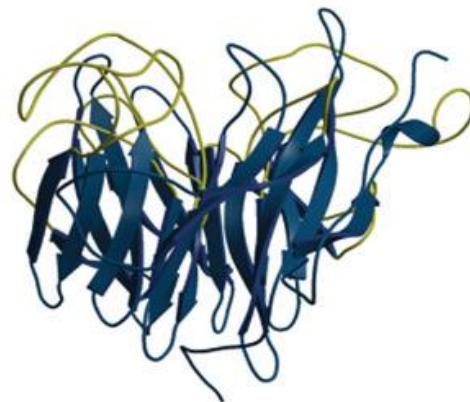
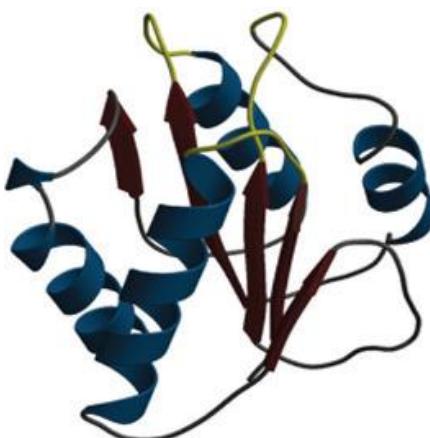


# Методы построения модели

- Сборка модели из «жестких фрагментов»
- Моделирование на основе пространственных ограничений
- Моделирование путем сопоставления сегментов и другие методы

## Моделирование петель

- *ab initio*
- путем поиска в базах данных



# Сборка модели из «жестких фрагментов»

**COMPOSER:** исторически первый подход к моделированию

- моделирование в декартовых координатах (Sutcliffe, ..., Blundell, 1987)

- Поиск белковых структур с последовательностями, гомологичными моделируемой. Выполнение выравнивание последовательностей, определение положения Сα-атомов консервативных остатков.
- Составление общего шаблона из перекрывающихся структурно консервативных фрагментов (при необходимости).
- Достройка боковых цепей **с учетом библиотек ротамеров**.
- Достройка петель путем подбора подходящих по геометрии гомологичных фрагментов среди белковых структур.
- Общая оптимизация геометрии.



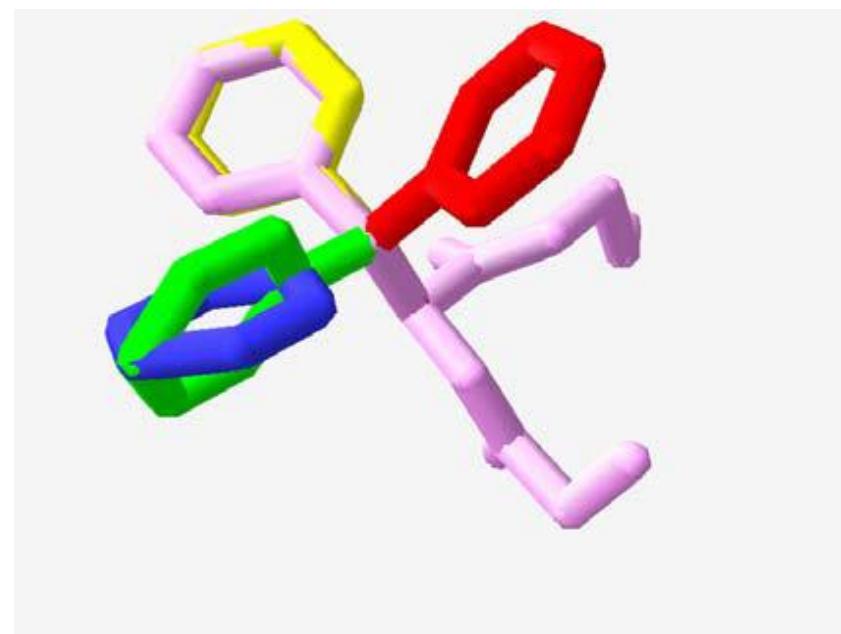
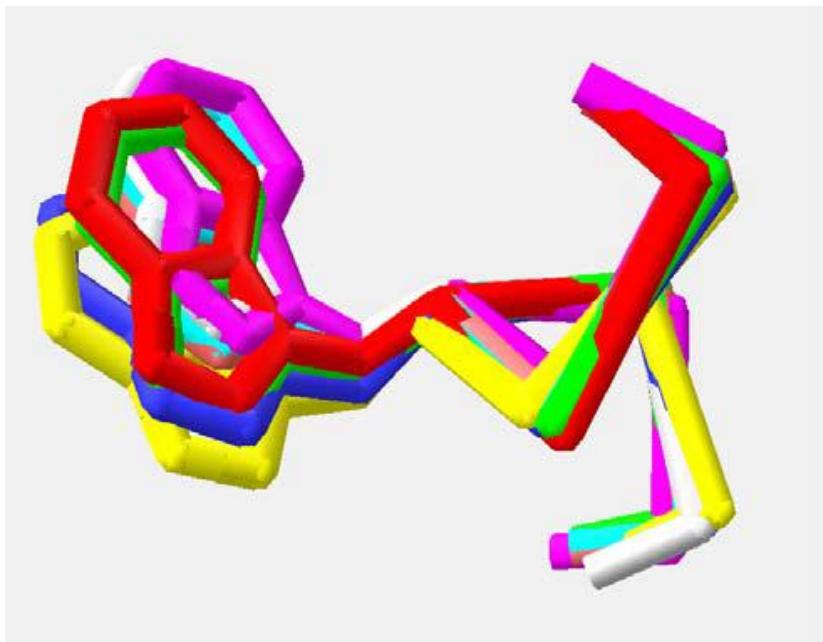
BIOZENTRUM

University of Basel  
The Center for Molecular Life Sciences

SWISS-MODEL

# Библиотеки ротамеров

- Лишь небольшая доля всех возможных конформаций боковых цепей реально наблюдается в экспериментальных структурах
- Конформация боковой цепи зависит от геометрии основной цепи
- Библиотеки ротамеров содержат наборы вероятных конформаций



# Моделирование на основе пространственных ограничений. MODELLER

Наиболее распространенный подход к моделированию (Sali & Blundell, 1993) – моделирование во внутренних координатах

- Поиск белковых структур с последовательностями, гомологичными моделируемой.
- Выполнение выравнивания последовательностей.
- Извлечение пространственных ограничений из шаблонов.
- Построение модели с учетом этих ограничений
- Общая оптимизация геометрии.

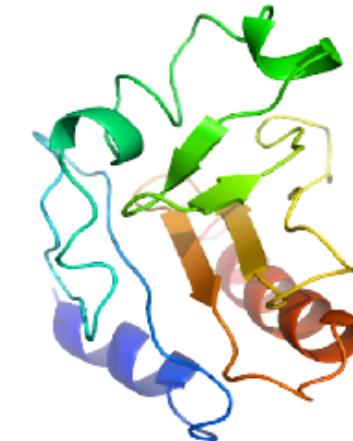
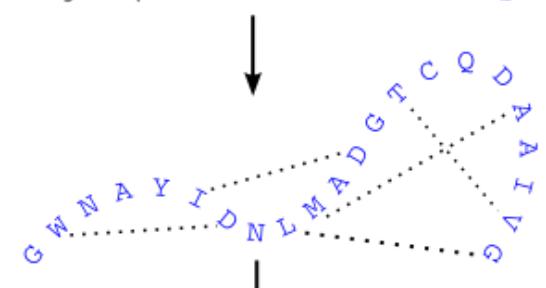
1. Align sequence with structures

Template structure(s)  
Target sequence

SWQTYVDTNLVGTGAVTQA--AI  
-GWNAYIDNLMADGTCQDAAIVG

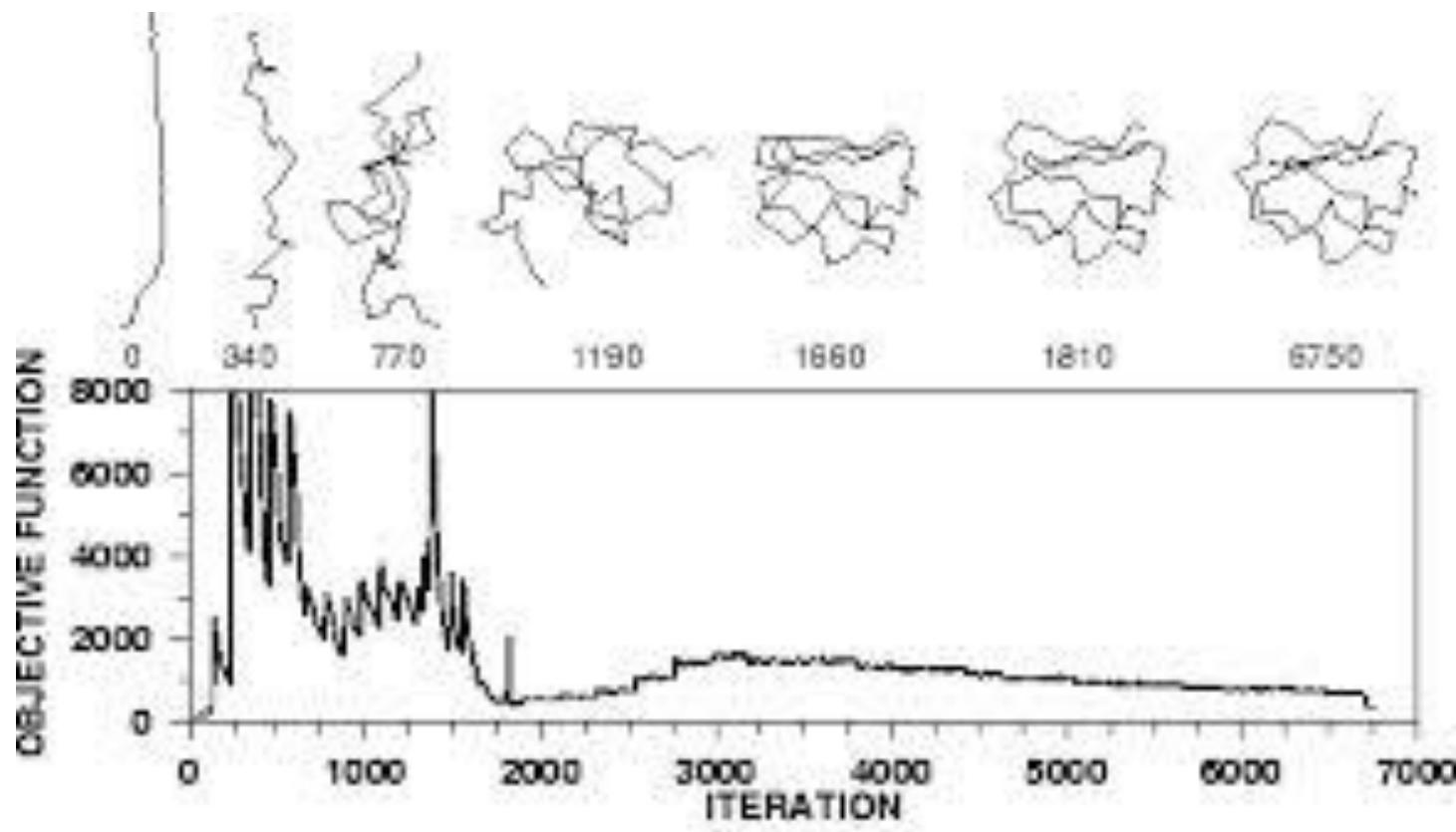
2. Extract spatial restraints

3. Satisfy spatial restraints



# Моделирование на основе пространственных ограничений. MODELLER

Начиная с распрямленной конформации или конформации шаблона, выполняется учет всех более далеких ограничений, чередующийся с минимизацией энергии методом сопряженных градиентов.



# Выбор шаблона

Методы, применяемые для сравнения последовательностей:

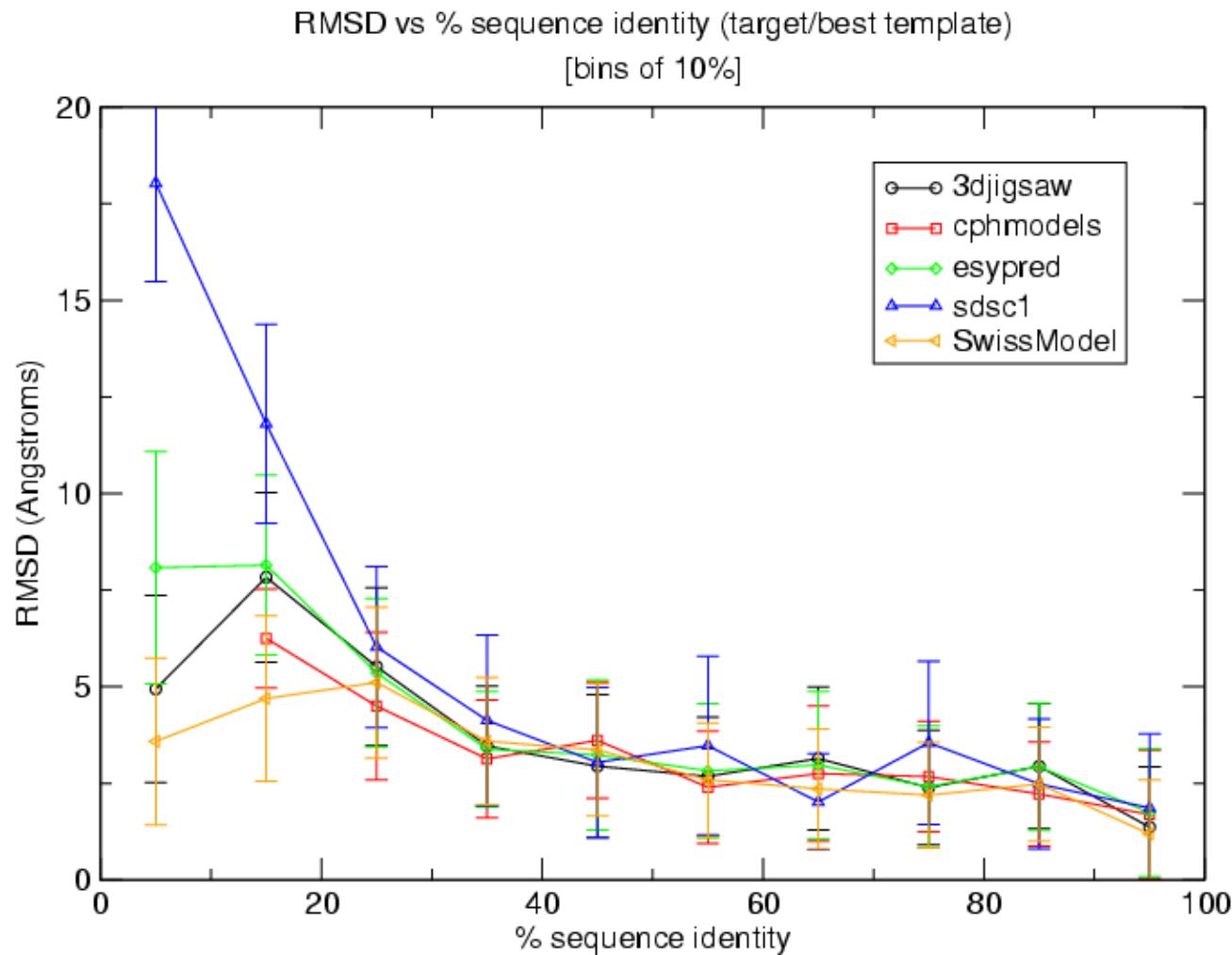
1. Попарное сравнивание моделируемой последовательности с каждой последовательностью из базы данных (FASTA, BLAST).
2. Сравнивание сразу нескольких последовательностей (Clustal, Muscle).
3. Протягивание последовательности через библиотеку пространственных структур.

Факторы, влияющие на выбор шаблона:

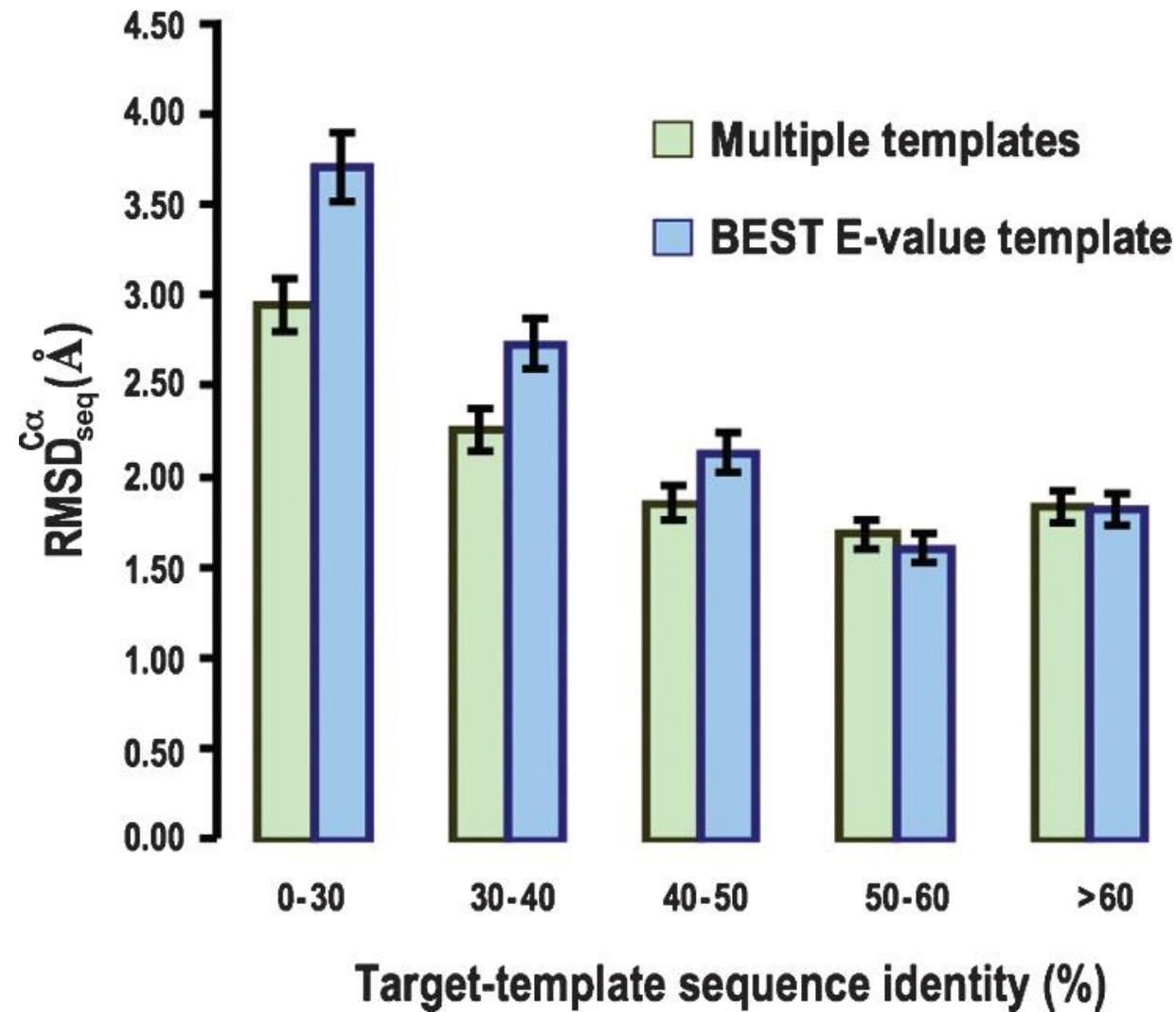
1. Высокая идентичность последовательностей.
2. Белки принадлежат к одному подсемейству.
3. Качество экспериментальной структуры (разрешение или количество ограничений на аминокислотный остаток).

| Процент идентичности | Качество выравнивания        |
|----------------------|------------------------------|
| > 40%                | почти всегда высокое         |
| > 30%                | ошибочно выровненные участки |
| < 30%                | «сумеречная зона»            |

# Качество модели: идентичность последовательностей



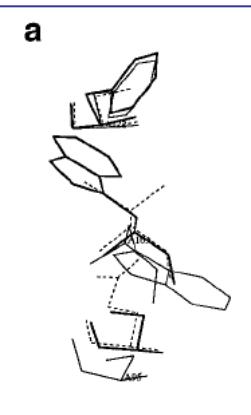
## Качество модели: несколько шаблонов



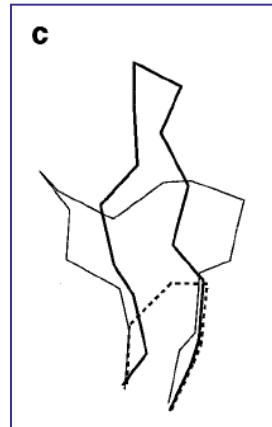
# Ошибки построения модели

1. Ошибки в ориентации боковых цепей.

Мышиный белок, связывающий ретиноевую кислоту. Тонкая линия – кристалл, толстая линия – модель, пунктир – шаблон (мышиный липидсвязывающий белок).



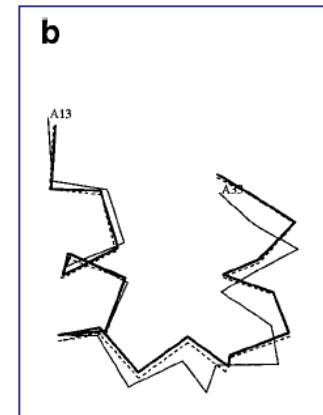
3. Ошибки в участках, для которых отсутствует шаблон. Показан контур С $\alpha$  атомов остатков 112-117 кристаллографической структуры человеческого эозинофильного нейротоксина (тонкая линия), его модели (толстая линия), и шаблона – рибонуклеазы А (пунктир).



5. Неправильно выбранный шаблон.

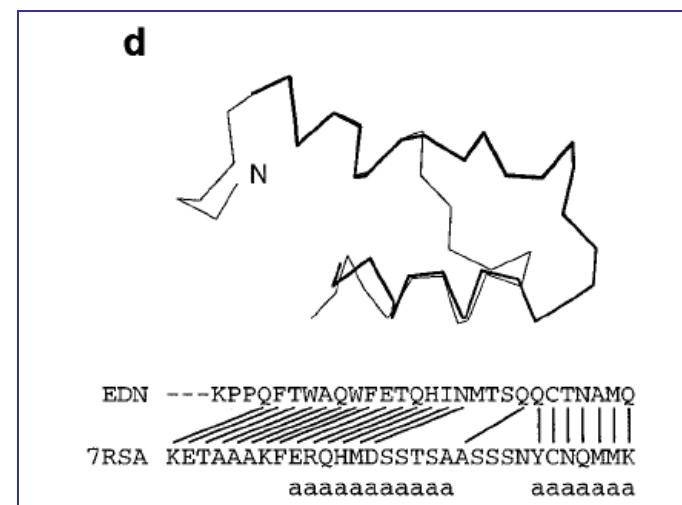
2. Сдвиги в корректно выровненных участках.

Сравнение участка кристаллической структуры мышного белка, связывающего ретиноевую кислоту, с его моделью и с шаблоном.



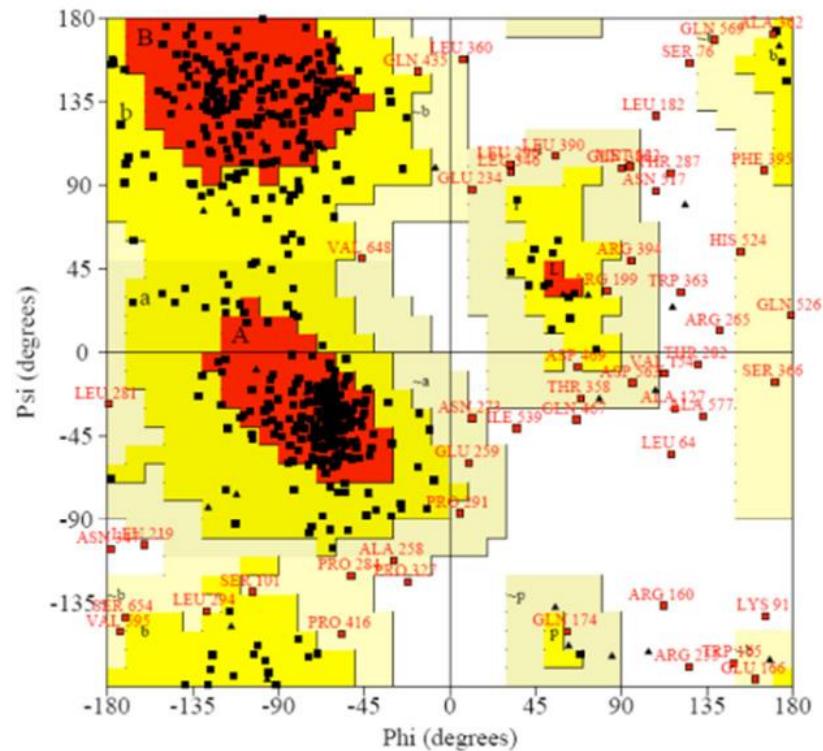
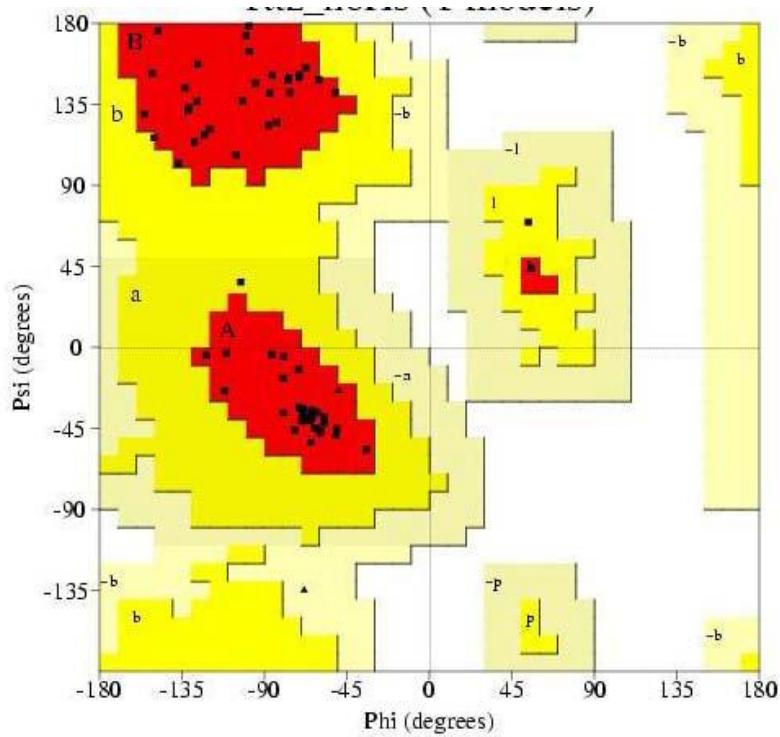
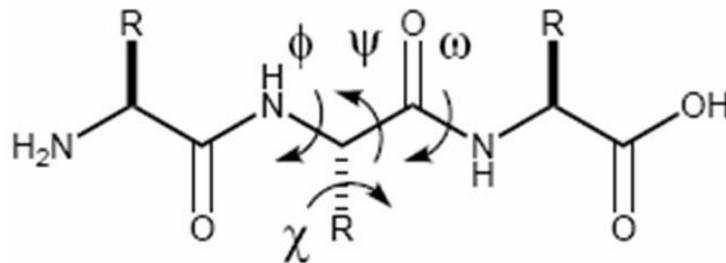
4. Ошибки из-за неправильного выравнивания.

N-концевой участок токсина сравнивается с его моделью. Показан соответствующий участок выравнивания, линии показывают эквивалентные остатки.



# Оценка модели. Проверка стереохимии

Карты Рамачандрана



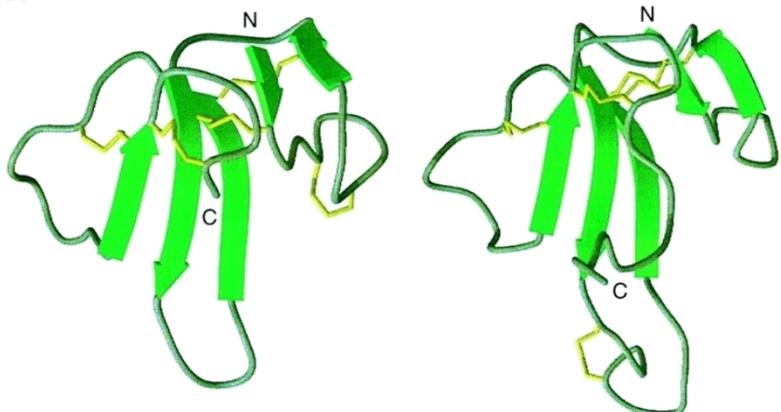
# Пример моделирования

Построение модели *lynx1*, эндогенного модулятора никотиновых ацетилхолиновых рецепторов.

Открытая рамка считывания (ORF) содержала последовательность, кодирующую белок 11 кДа.

Поиск в базах данных выявил гомологов среди секретируемых нейротоксинов и молекул иммунной системы Ly-6. Характерная черта этих белков – присутствие богатого цистеинами мотива, который также присутствует в *lynx1*. (Цистеиновые остатки выделены черным в выравнивании).

|               | ...C.....C.....C.....C.....C.....C.....C.....CC.....CN                            | 49/37 |
|---------------|---|-------|
| <i>lynx1</i>  | LECHVCAYNGDN CFKPMRCPAMATYCMTTRTYFTPYRM KV RKSCVP SCFETVYDG YSKHASATSCCQYYLCN     |       |
| E48           | LRCHVCTSSSN CKHSVVCPASSRFCKTTNTVEPLRGN LVK KDCAE SCTPSYTLQQQ VSSGTSSTOCCQEDLCN    | 37/16 |
| Ly-6AE        | LECYQCYGVPFETSCPSC ITCPYPDGVCVTQEAAVIVDSQTRKVKNNLCL PICPPNIESMEILGTVNVKTSCCQEDLCN |       |
| CD59          | LTCYHCFQPVVS SCNMMNSTCSPDQDSCLYAVAGMQVYQ RCWKQSDCHGEII MDQLEETK LKFRCCQFNLCN      | 27/19 |
| $\alpha$ Cbtx | LECHNQQSSQ TPTTTGCSGGENCYKKRWRDHRGYR TERGC GCPS VKNGIEINCCTTDRCN                  | 21/19 |
| $\alpha$ Bgtx | IVCHTTATSP ISAVTCPGENLCYRKMWCAFSSRGKVVELGCAA TCPS KKPYEEVTCCSTDKCN                | 30/18 |
| M3tx          | LTGVTTSKSI GITTENCAGQNLCFKRRHYVIPRYT EITRGCAA TCPI PENYDSIHCCKTDKCN               | 31/20 |



Модель *lynx1*

Структура  $\alpha$ -бунгаротоксина

Структурные данные известны для CD59,  $\alpha$ -бунгаротоксина и кобратоксина. Структуры этих молекул очень схожи несмотря на невысокую идентичность последовательностей. Это объясняется тем, что консервативные цистеиновые остатки определяют структуру молекул, создавая стабильное ядро из  $\beta$ -тяжей благодаря дисульфидным связям.

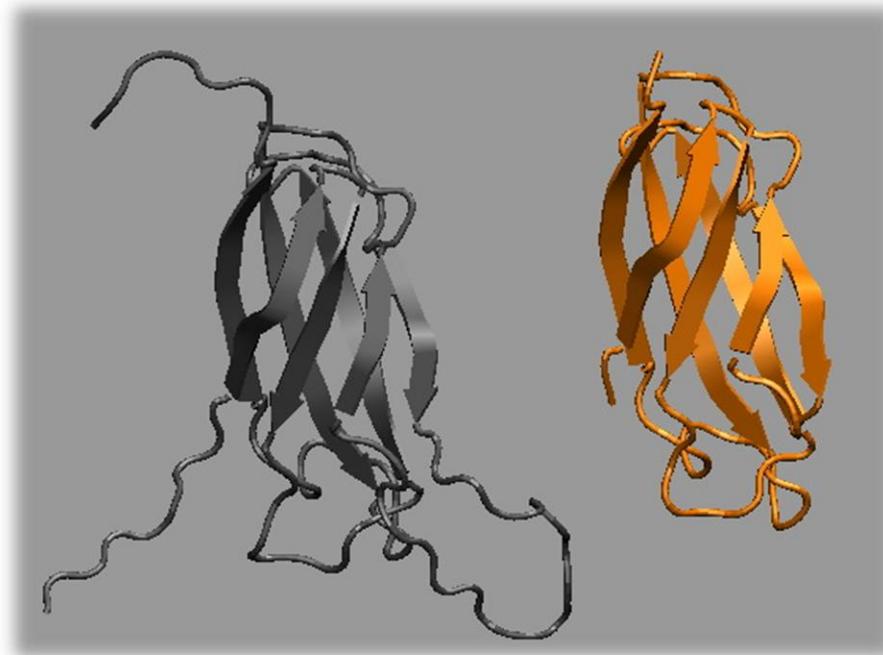
Все три структуры были использованы в качестве шаблонов независимо и все вместе, было построено 4 модели. Оценка моделей показала, что все они достоверны с вероятностью 0.84 (CD59), 0.91 (кобратоксин), 0.93 ( $\alpha$ -бунгаротоксина) и 0.94 (все три шаблона).

## Еще один пример моделирования

FN3H -MQVSDVPTNLEVVAATPTSLLISWYTFHYG--MNRYYRITYGETGGNS 47  
1FNA\_A -----RDLEVVAATPTSLLISWDAP-AVT---VRYYYRITYGETGGNS 38  
:\*\*\*\*\* . \*\*\*\*\*

FN3H PVQEFTVPWINTYTGEPTYADDFKGRFTATISGLKPGVDYTITVYAVTEF 97  
1FNA\_A PVQEFTVP-----GSKSTATISGLKPGVDYTITVYAVTGR 73  
\*\*\*\*\*

FN3H SGTGDFDYPISINYRTLEHHHHH 121  
1FNA\_A GDSPASSKPISINYRTEI---- 91  
\*\*\*\*\*



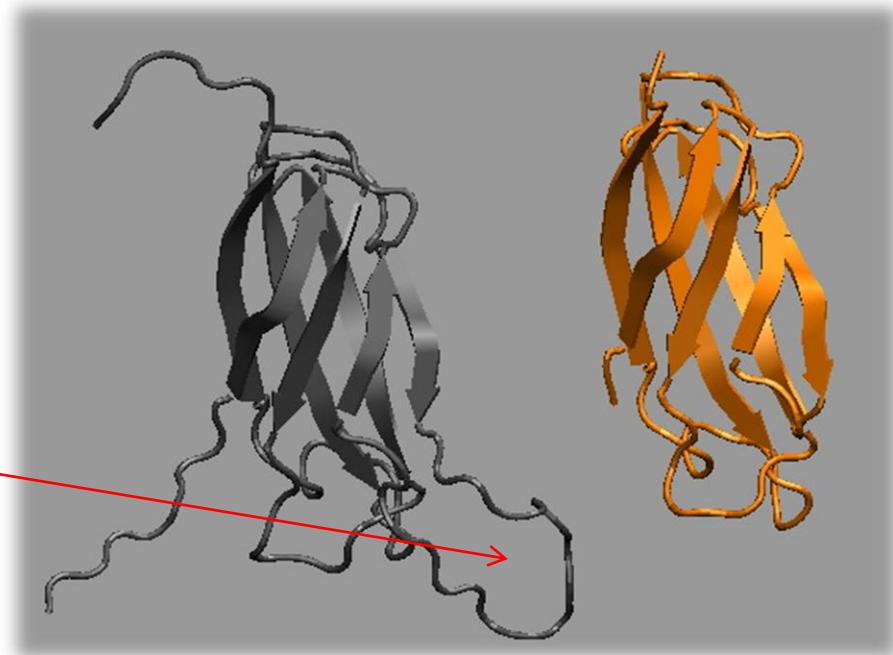
# Еще один пример моделирования

|        |  |       |
|--------|--|-------|
| FN3H   | -MQVSDVPTNLEVVAATPTSLLISWYTFTHYG--MNRYYRITYGETGGNS | 47    |
| 1FNA_A | -----RDLEVVAATPTSLLISWDAP-AVT---VRYYYRITYGETGGNS   | 38    |
|        | :***** .   | ***** |

|        |  |       |
|--------|--|-------|
| FN3H   | PVQEFTVPWINTYTGEPTYADDFKGRFTATISGLKPGVDYTITVYAVTEF | 97    |
| 1FNA_A | PVQEFTVP-----GSKSTATISGLKPGVDYTITVYAVTGR           | 73    |
|        | *****  | ***** |

|        |                         |     |
|--------|-------------------------|-----|
| FN3H   | SGTGDFDYPISINYRTLEHHHHH | 121 |
| 1FNA_A | GDSPASSKPISINYRTEI----- | 91  |
|        | *****                   |     |

Вставка



# Welcome to ROSIE

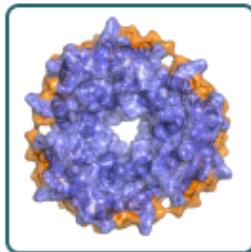
## Rosetta Online Server that Includes Everyone

Welcome Queue About ChangeLog Documentation Support

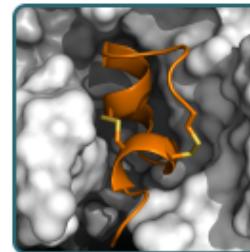
Login Create an account

Recommend Share 5 G+ 26

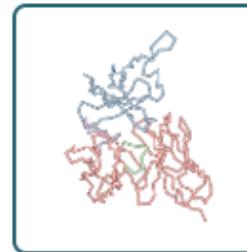
### Rosetta Protocols opened for academic users:



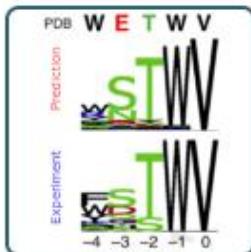
[\[Mp\\_lipid\\_acc\]](#)



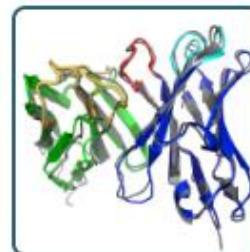
[\[Tox\\_dock\]](#)



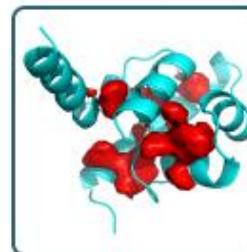
[\[Snug\\_dock\]](#)



[\[Sequence\\_tolerance\]](#)



[\[Antibody\]](#)



[\[Vip\]](#)

### ROSIE stats (24hrs):

Users: 4,815 +1

Jobs: 31,163 +28

CPU hours: 3,521,032 +5,956

See more info at our [About](#) page.

### Get Started with ROSIE

- [ROSIE Documentation](#) – Server related documentation and info.
- [Rosetta Forums](#) This is a list of forums for Rosetta users to discuss problems with running Rosetta and is monitored by Rosetta developers.

# Предсказание структуры белков

Сворачивание белка в уникальную конформацию наводит на мысль об алгоритме формирования структуры белка по его последовательности, но доказательством полноты и правильности нашего понимания могла бы стать его реализация в виде компьютерной программы...

Методы предсказания структуры по последовательности:

- Предсказание вторичной структуры;
- Моделирование по гомологии;
- Распознавание фолда;
- Априорное предсказание новых типов укладки.



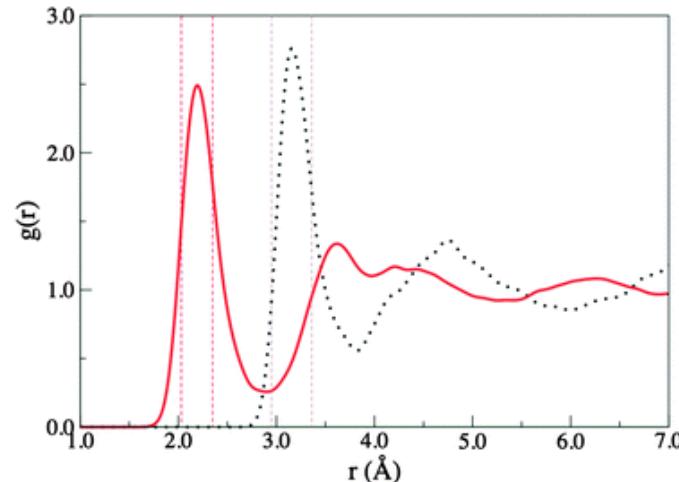
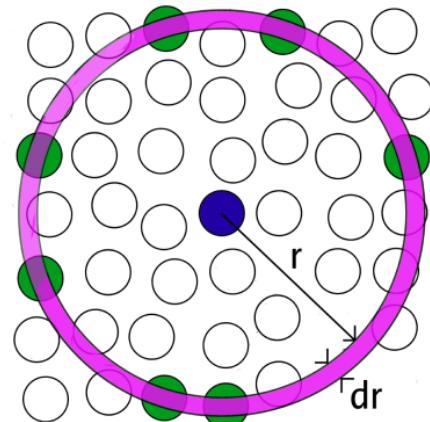
# Распознавание фолда

Если для последовательности нет гомолога с известной структурой, то, возможно, есть хотя бы структура, подходящая для данной последовательности?

Фактически, нужно примерить данную последовательность на все типы укладки и выбрать наиболее подходящую - **метод «протягивания»** (threading) состоит в построении большого числа грубых моделей для данной последовательности и их последующей экспресс-оценки.

⇒ **нужна функция оценки соответствия последовательности и фолда.**

Например, функции распределения вероятности парных расстояний между остатками (например, по Сβ-атомам) ( $20 \times 20$  штук). Соотнося расстояния в моделях с этими функциями, можно оценить насколько вероятны как эти расстояния, так и модели в целом.

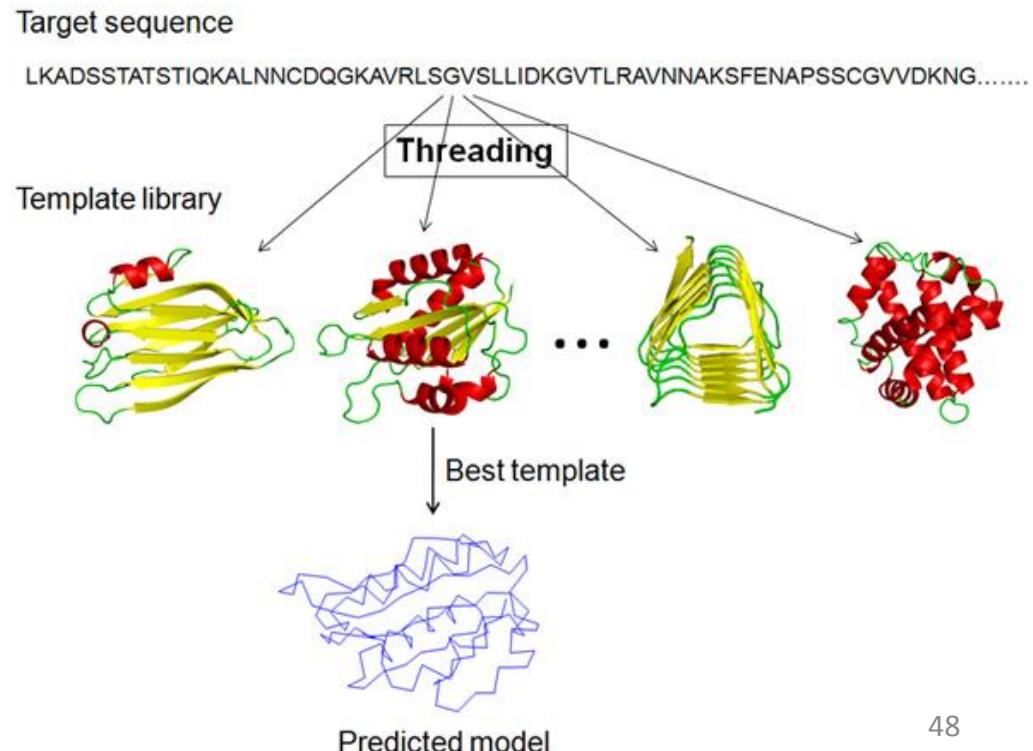


# Распознавание фолда. Алгоритм «протягивания»

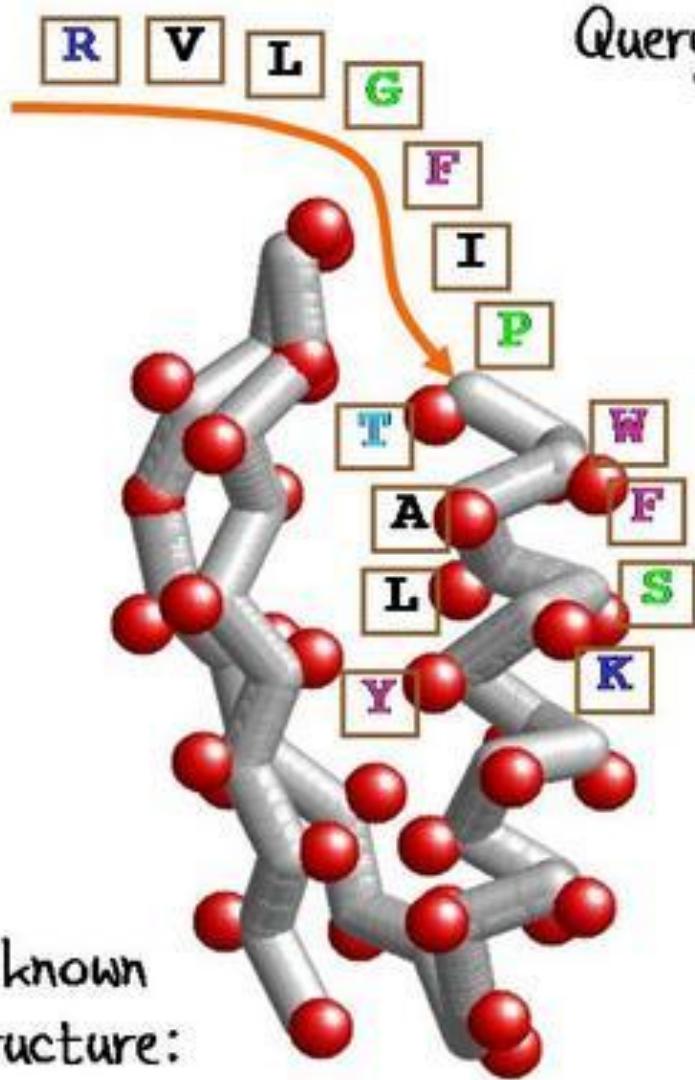
1. Выбор некой известной структуры в качестве потенциального шаблона
2. Генерация всевозможных выравниваний последовательности шаблона с новой последовательностью

...IIAWLVKEKKVDVIV...   ...IIAWLVK-EKKVDVIV...   ...IIAWLVKEKKVDVIV...  
...NGLELVLDSDLVLDATF...   ...NGLELVLDSDLVLDATF...   ...NGLELVLD-SVLDATEF...  
    \*\*              \*                        \*\*                          \*\*              \*

3. Построение и оценка моделей
4. Переход к следующему шаблону (п.1)
5. Сопоставление моделей, построенных по различным шаблонам, и выбор оптимальной



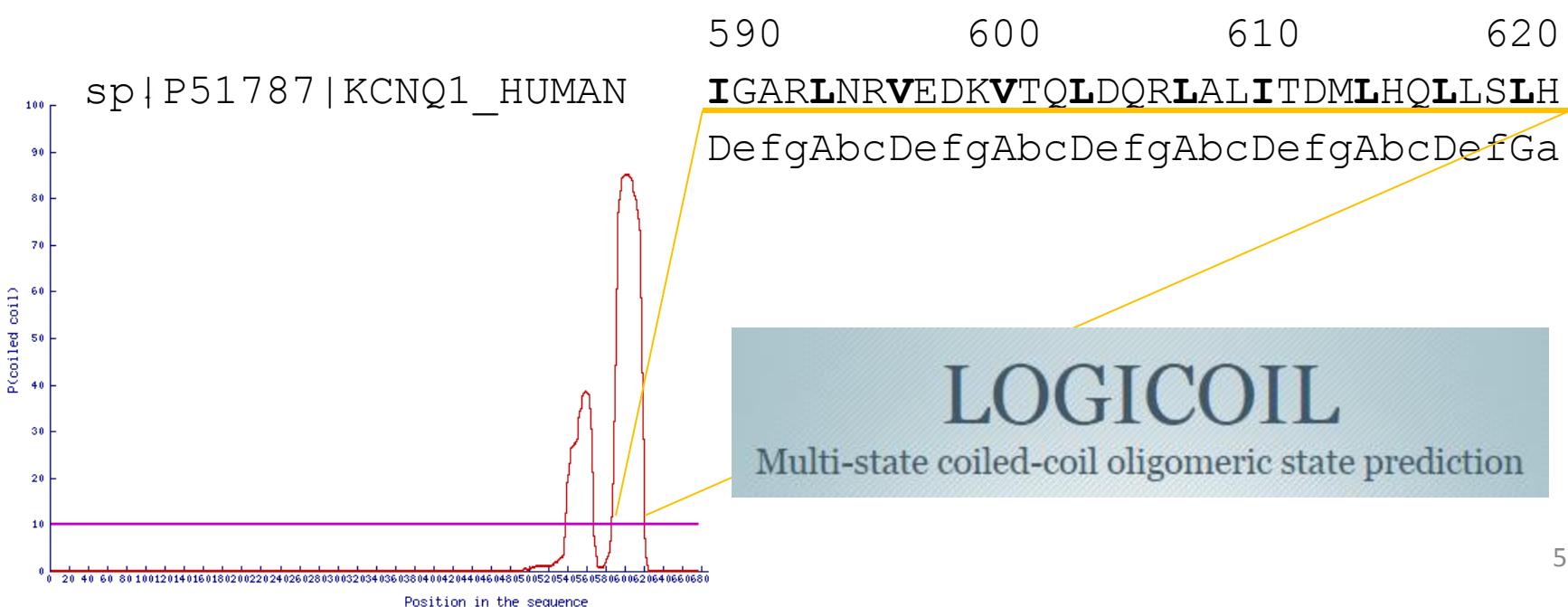
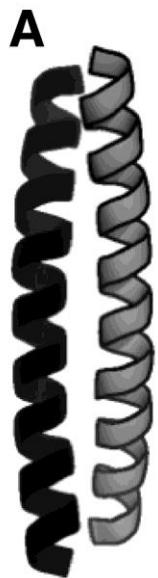
# WHAT IS THREADING



Query Sequence: RVLGVFIPTWFALSKY

- Thread the sequence onto the structure.
- Use structural properties to evaluate the fit:
  - Local structure
  - Environment
  - Pairwise interactions.

# Распознавание фолда. Суперспирали



# Неупорядоченные белки

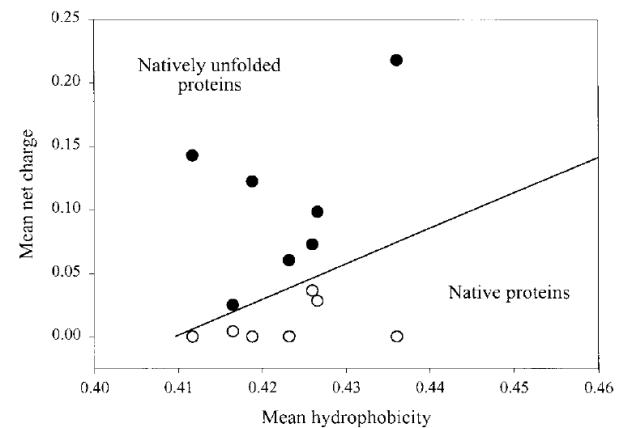
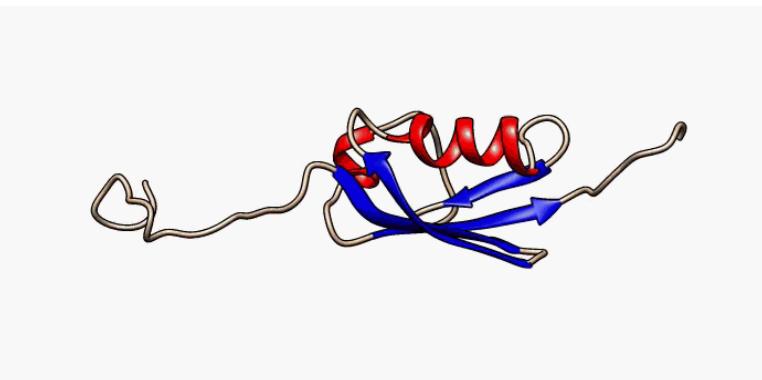
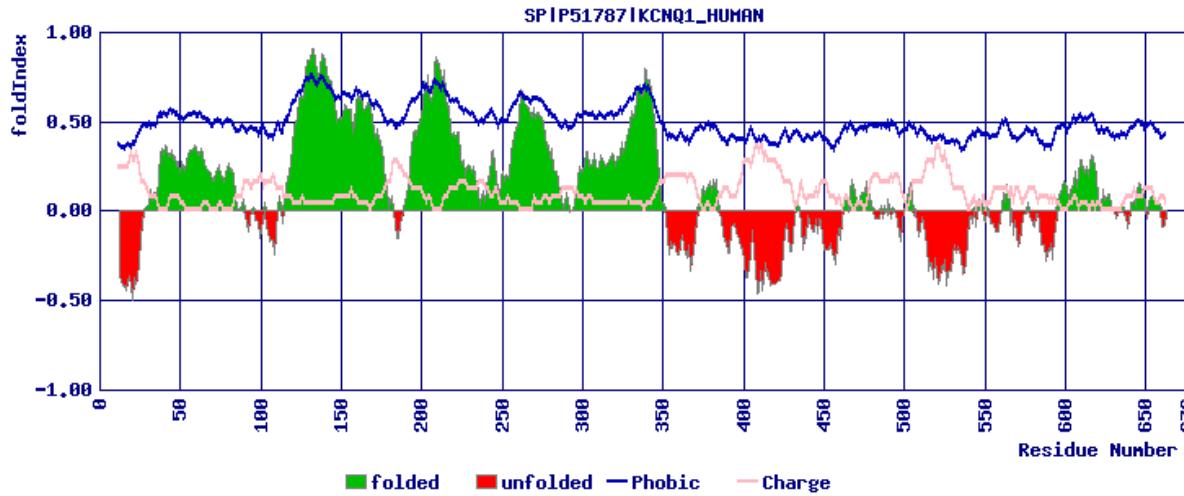
Intrinsically disordered proteins - «нарушители»  
догмы «структура определяет функцию»

Предсказание неупорядоченных участков по  
последовательности:

- использование структурных данных  
(нейронные сети)
- использование свойств аминокислот

$$\langle R \rangle = 2.785 \langle H \rangle - 1.151 \quad (\text{Uversky VN, et al. 2000})$$

FoldIndex <http://bip.weizmann.ac.il/fldbin/findex>





# I-TASSER

Protein Structure & Function Predictions

(The server completed predictions for [390328 proteins](#) submitted by [94188 users](#) from [138 countries](#))

([The template library](#) was updated on [2018/04/02](#))

