



Научно-технологический
университет

Сириус

Структурная биоинформатика | Лекция 3

Вторичная структура

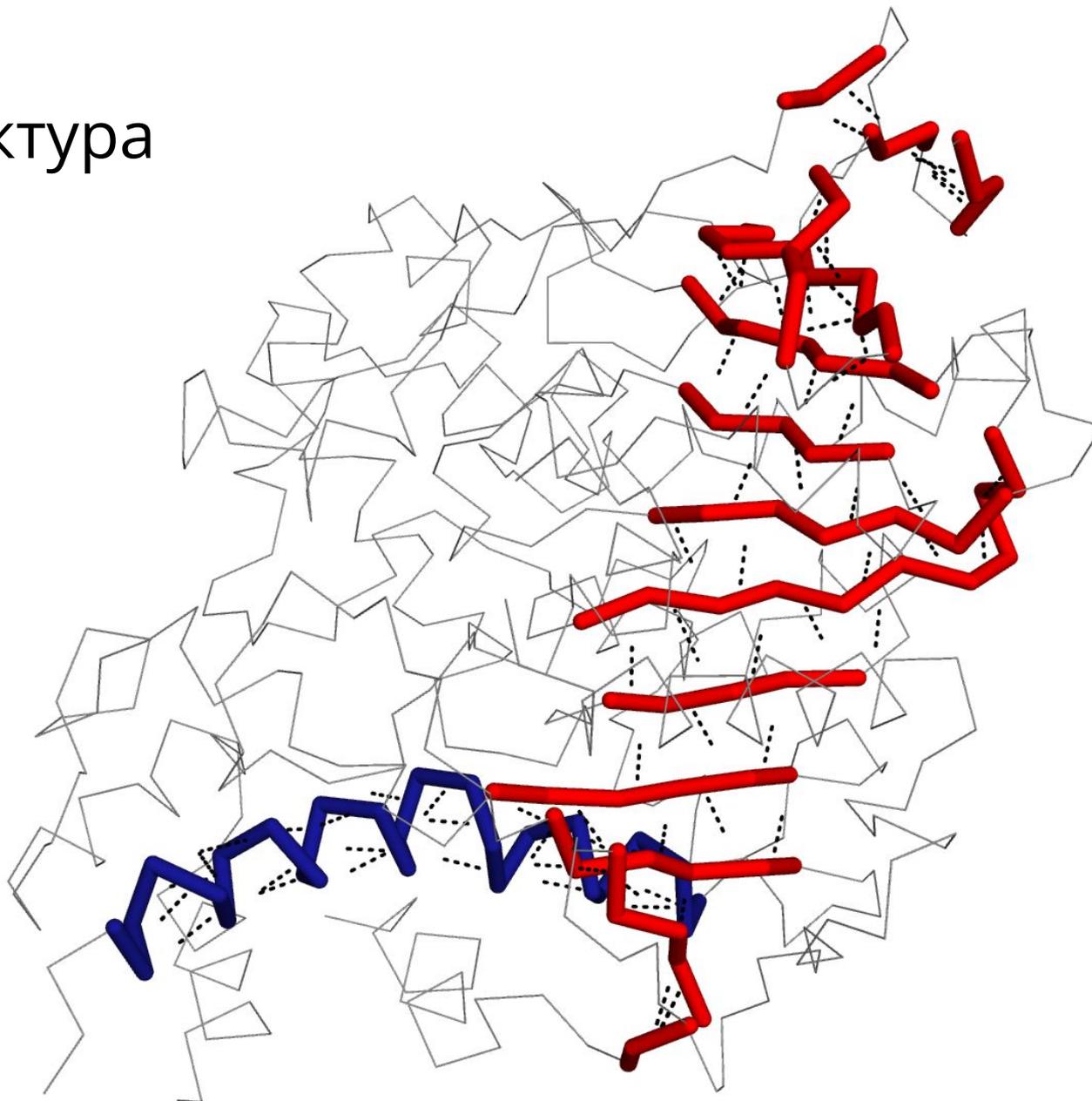
Александр Злобин

Вторичная структура

Вторичная структура -
паттерн во взаимном
расположении атомов
остова

Наслаивается на паттерн
меж-остовных
водородных связей
(определяется им?)

Можно предсказывать
вторичную структуру
отталкиваясь только
от предсказания
водородных связей -
алгоритм **DSSP**

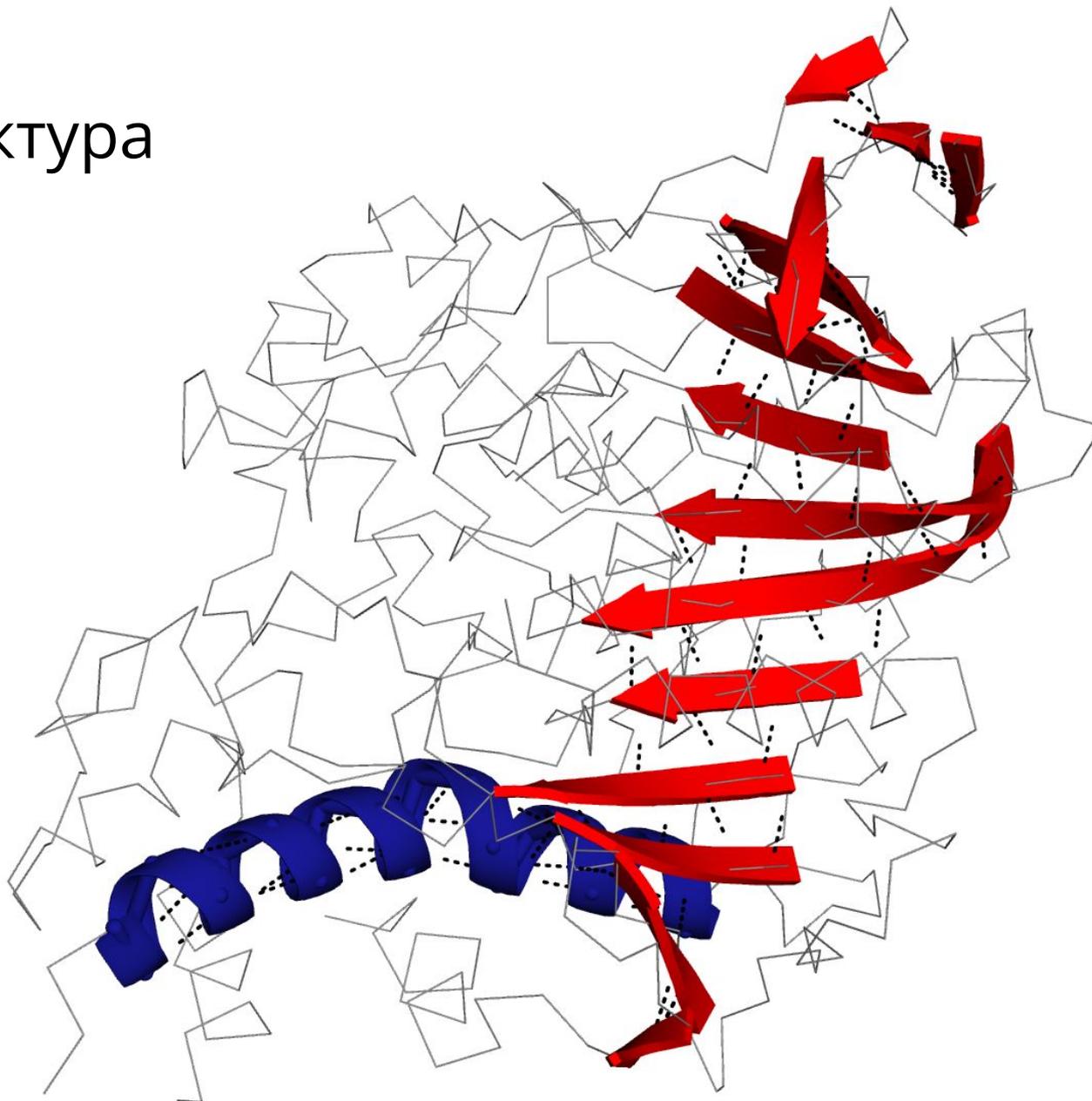


Вторичная структура

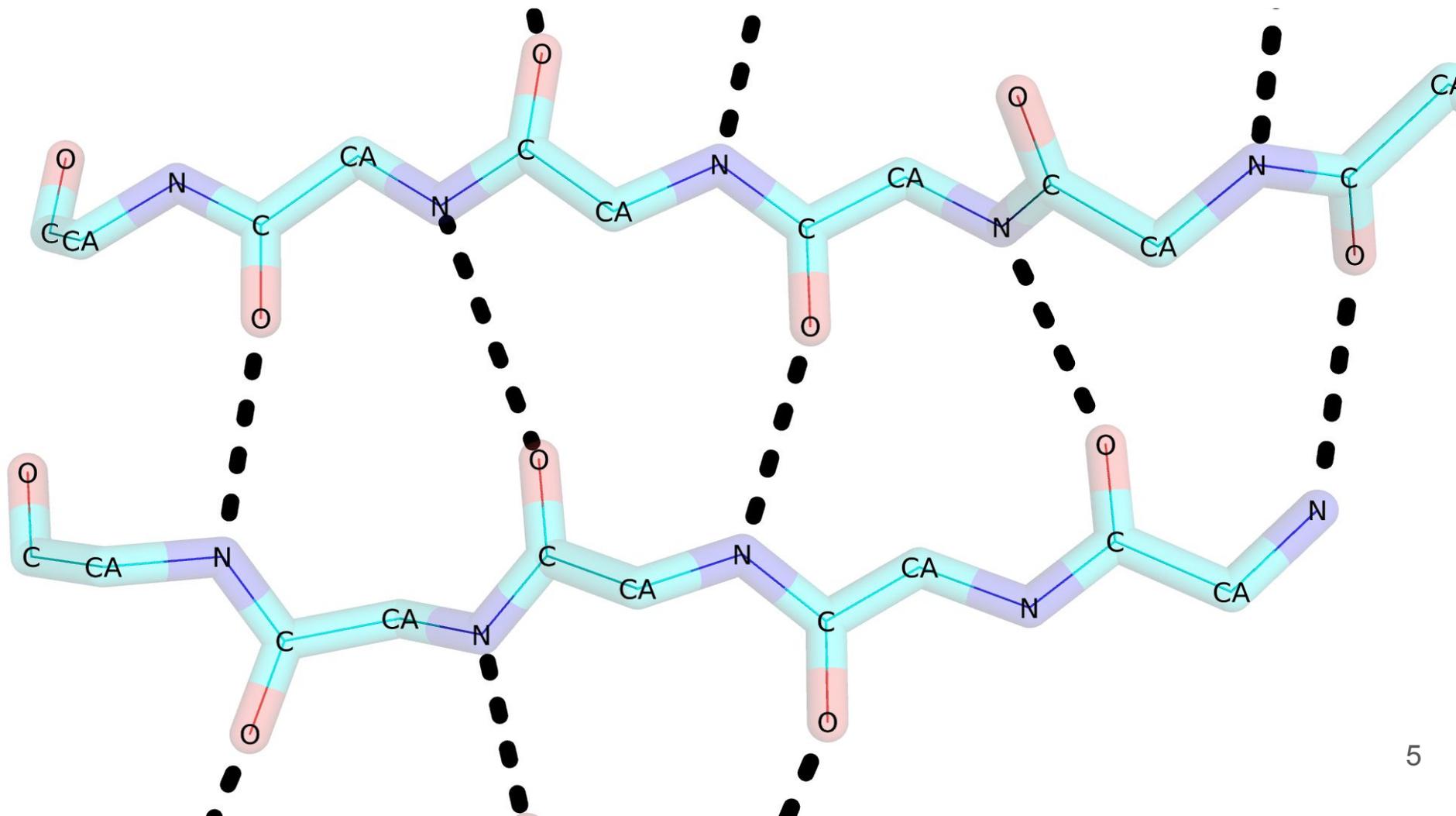
Вторичная структура -
паттерн во взаимном
расположении атомов
остова

Наслаивается на паттерн
меж-остовных
водородных связей
(определяется им?)

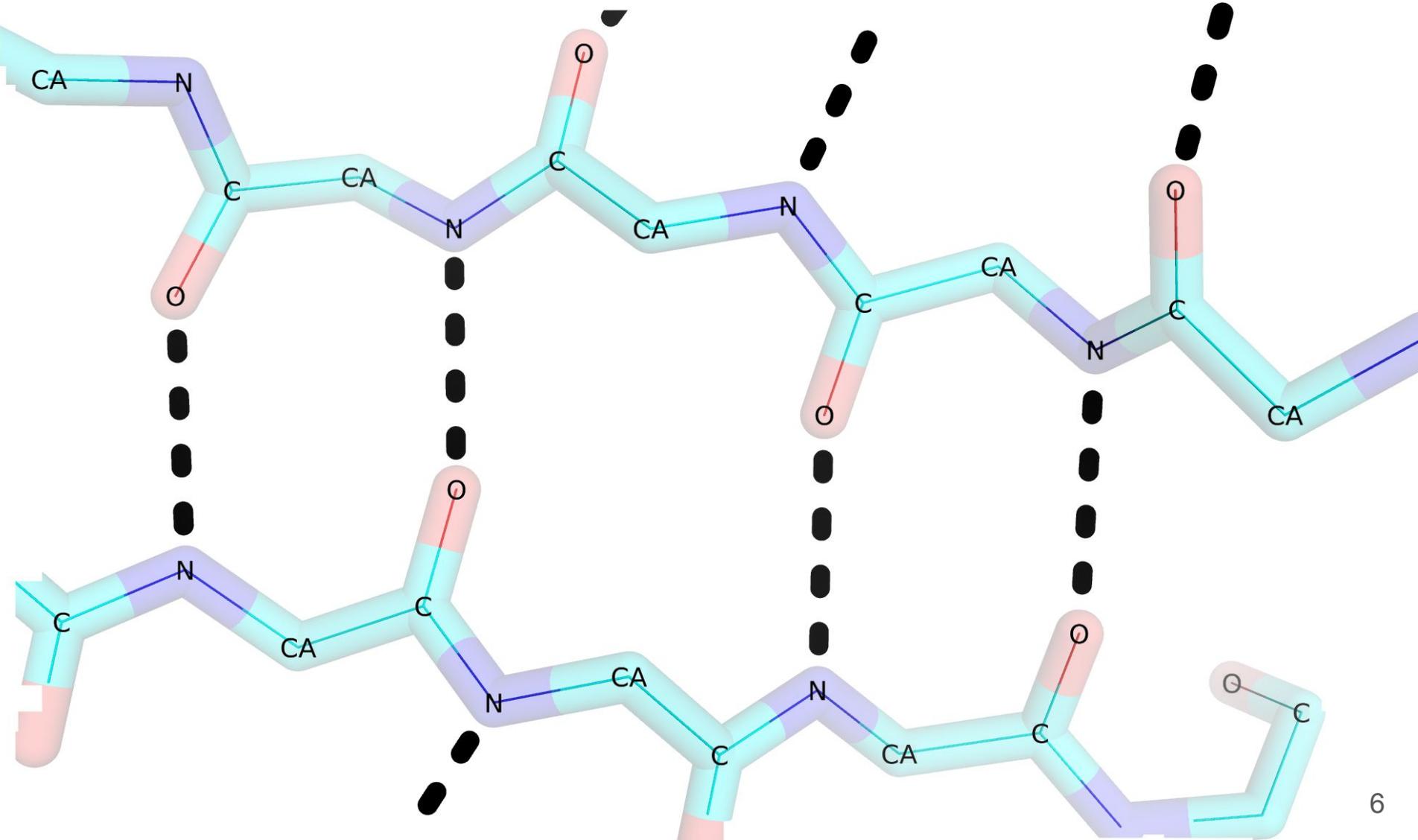
Можно предсказывать
вторичную структуру
отталкиваясь только
от предсказания
водородных связей -
алгоритм **DSSP**



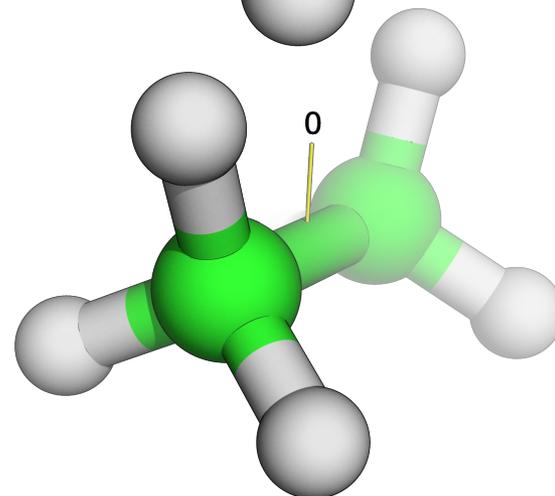
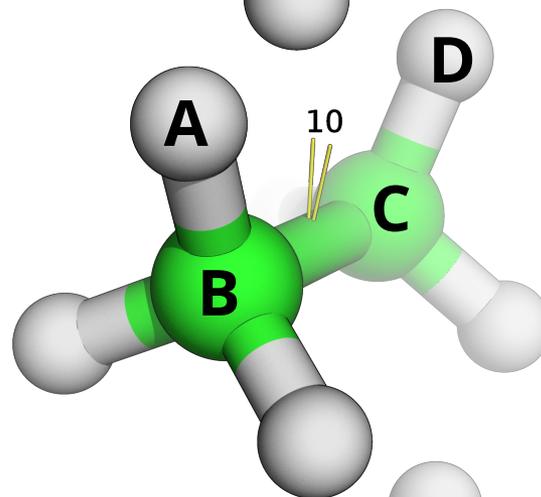
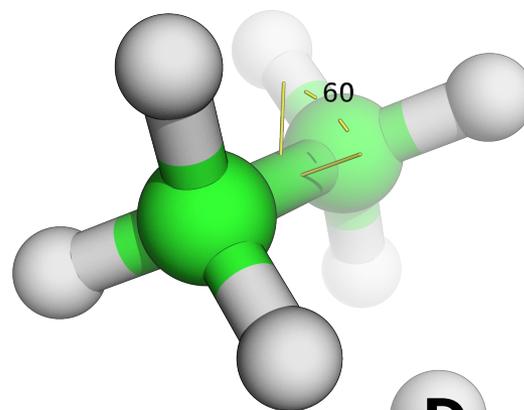
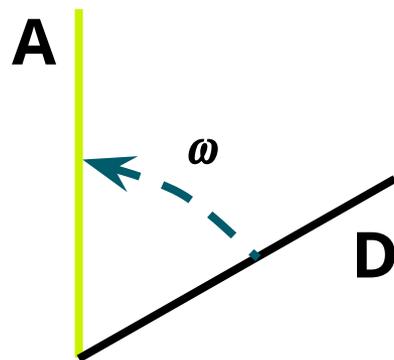
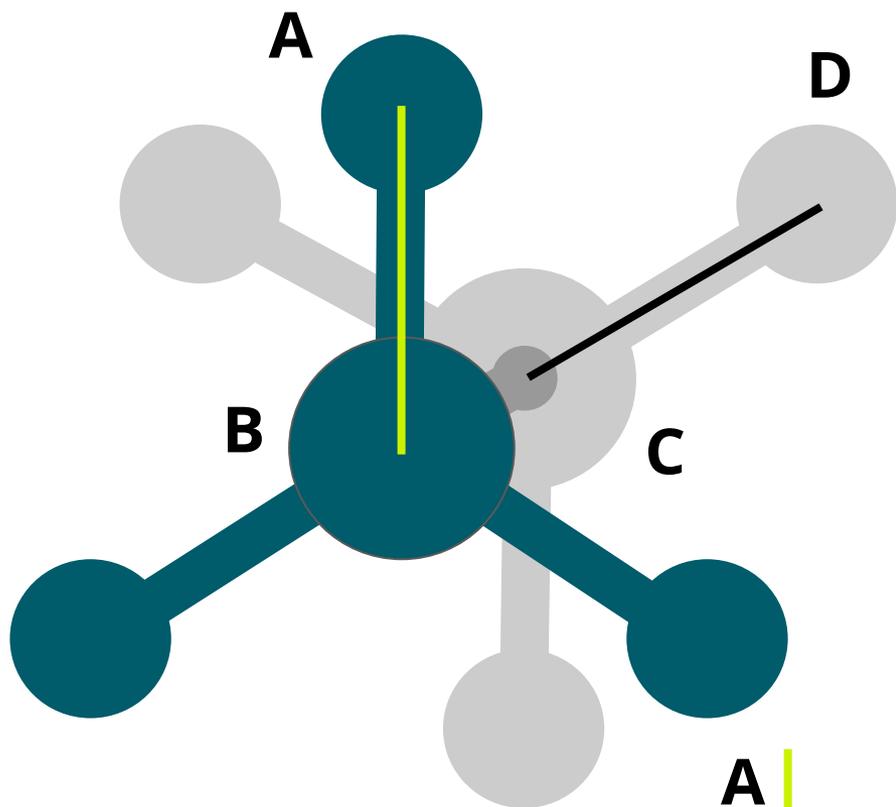
Вторичная структура



Вторичная структура

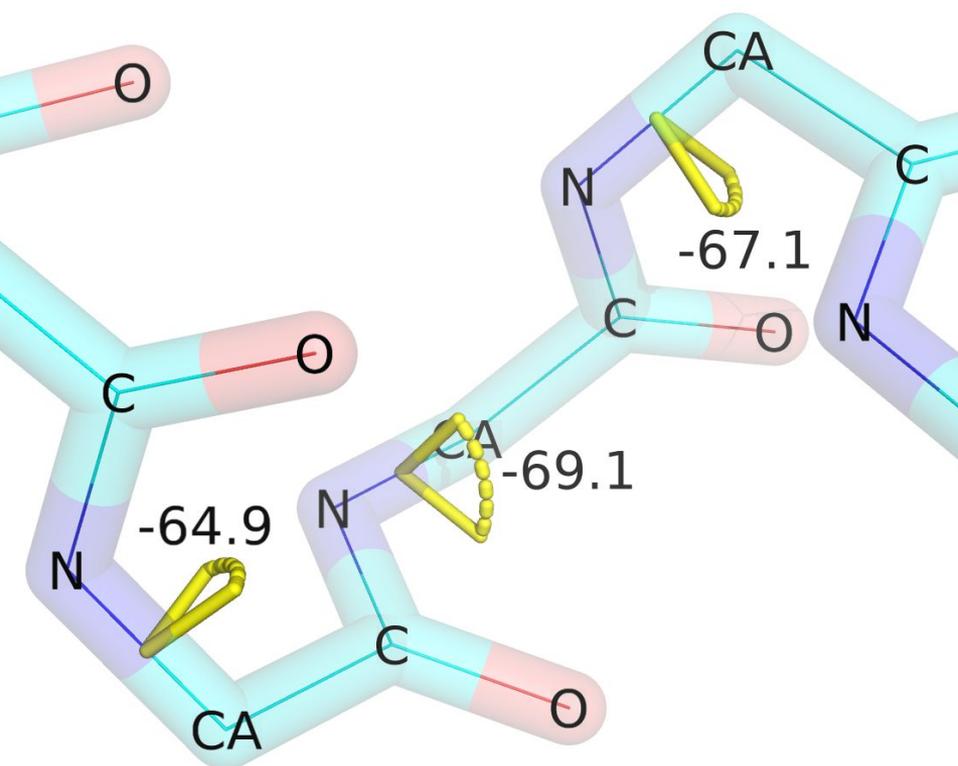


Торсионный угол

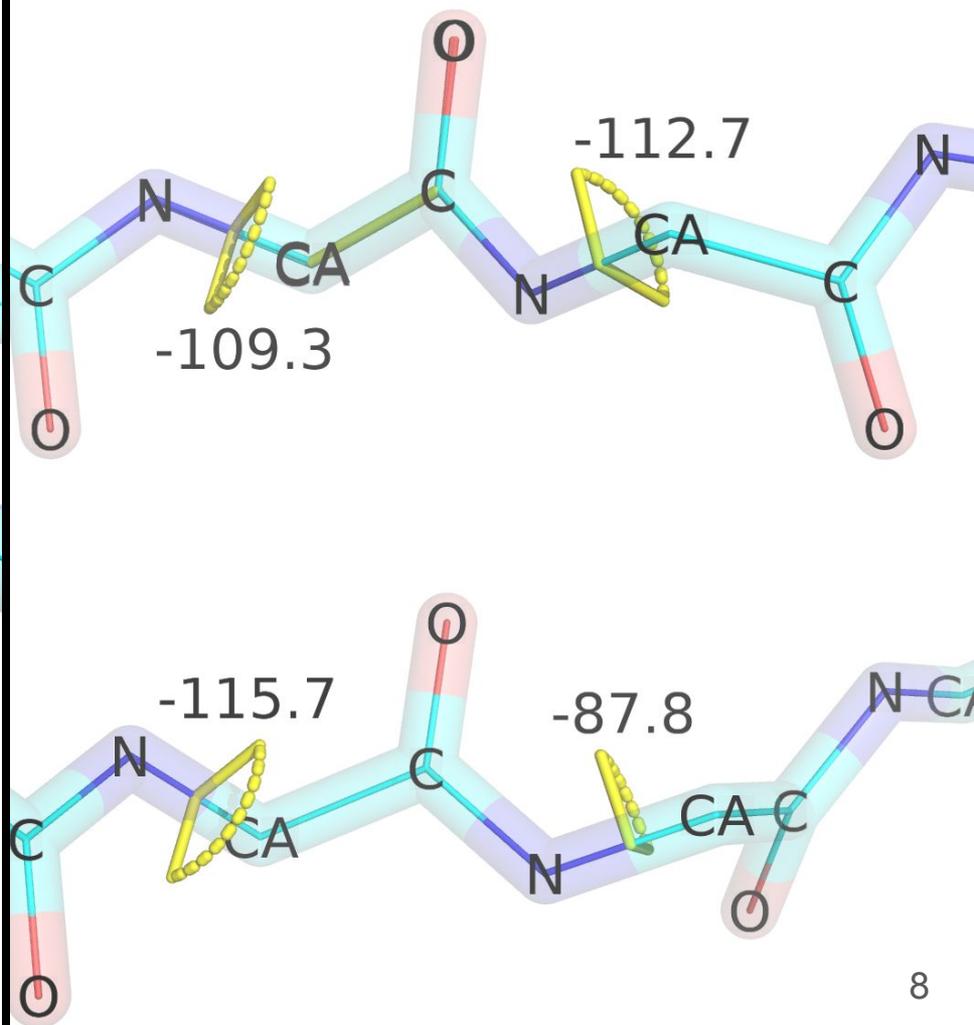


Торсионный угол φ - вращение по N-Ca

в альфа-спиралях

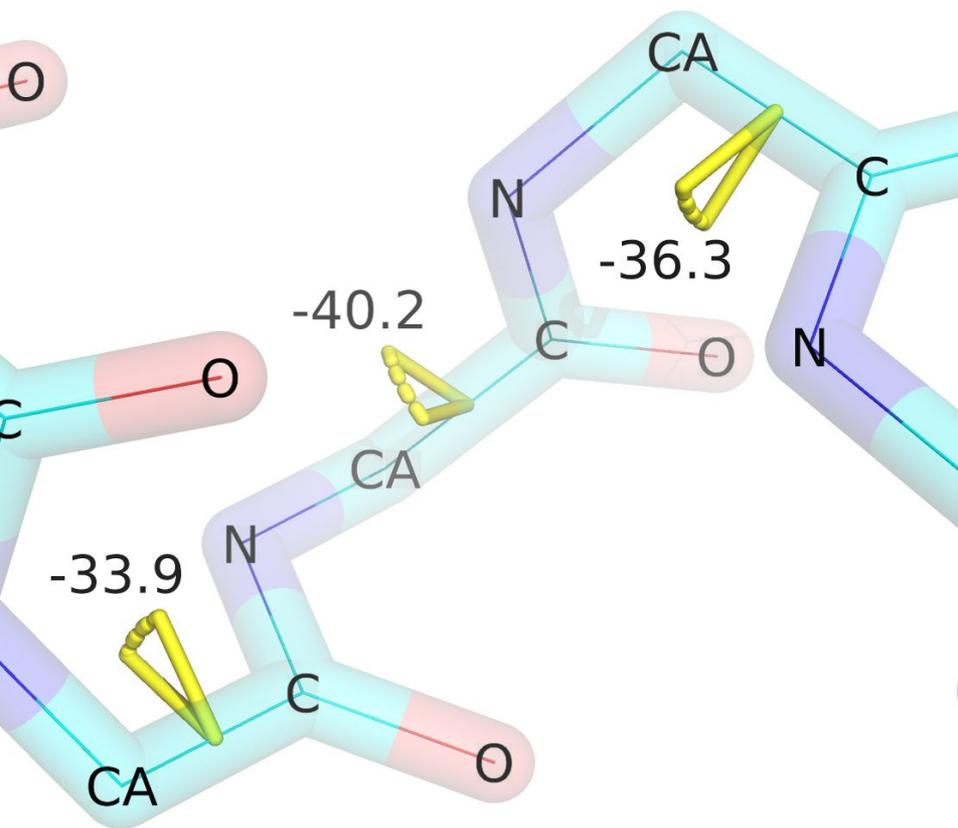


в бета-листах

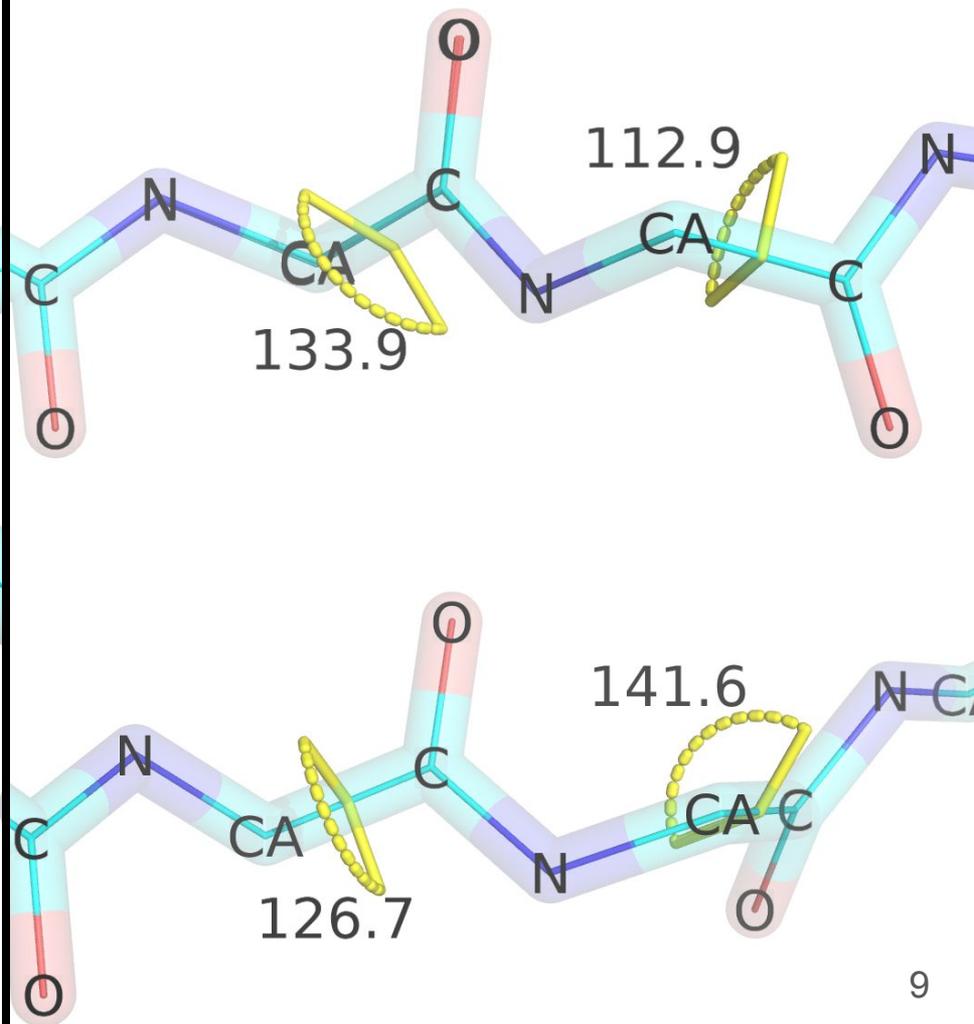


Торсионный угол ψ - вращение по Ca-C

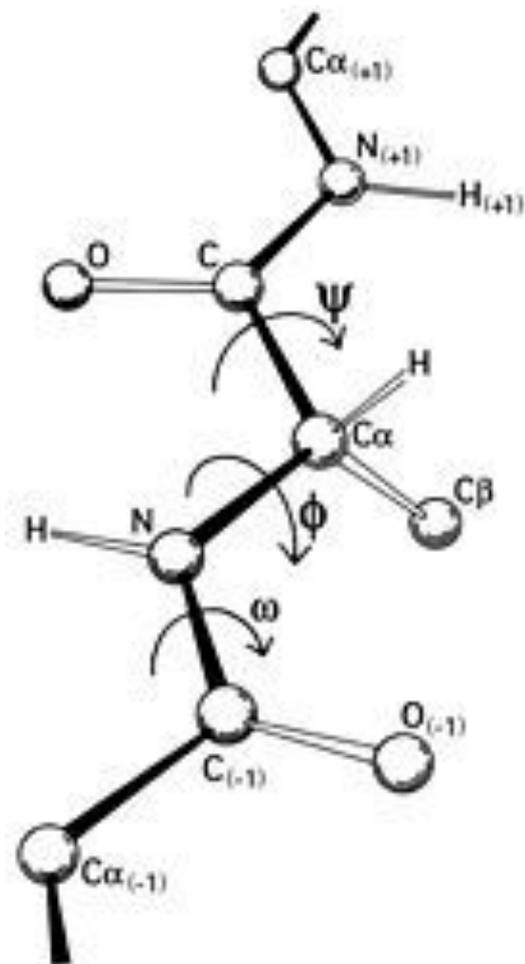
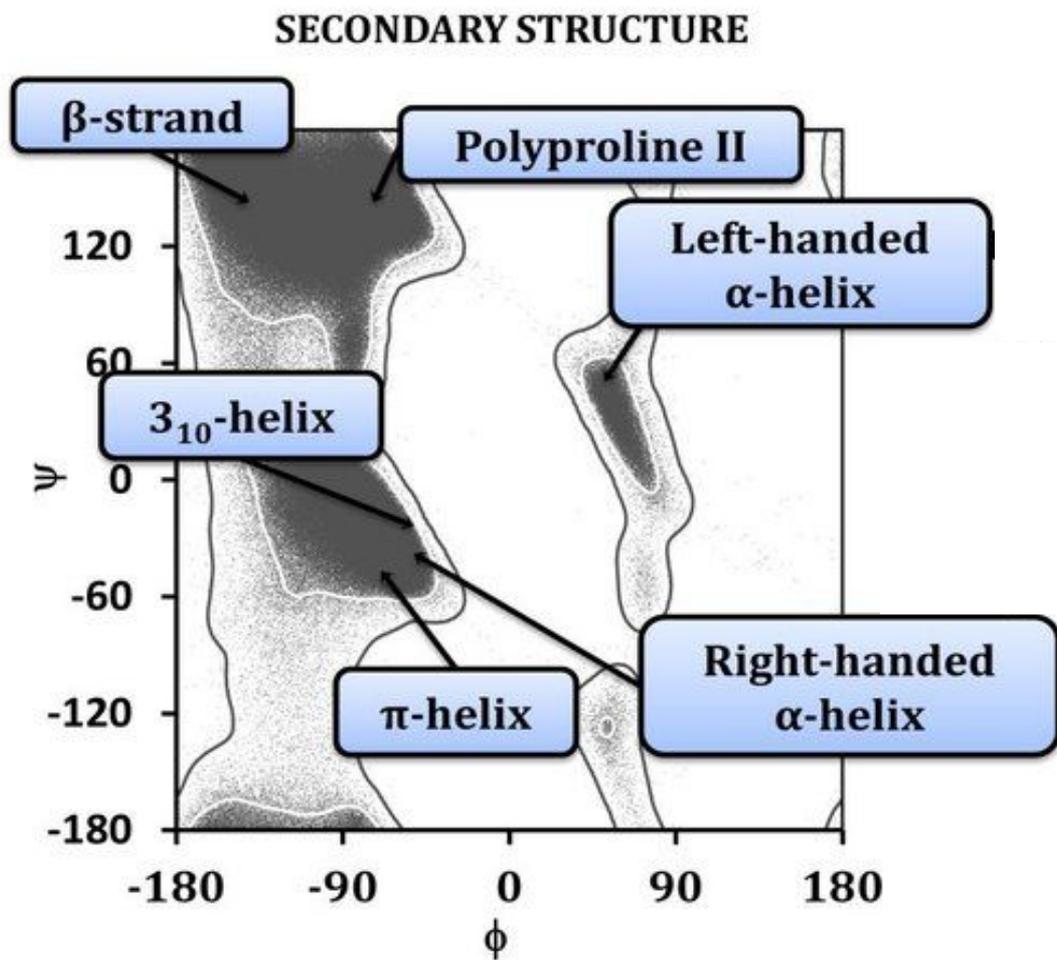
в альфа-спиралях

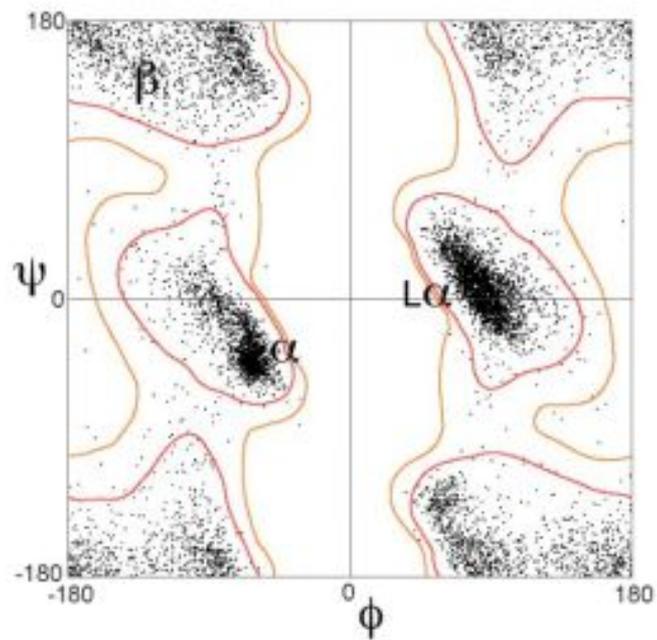


в бета-листах

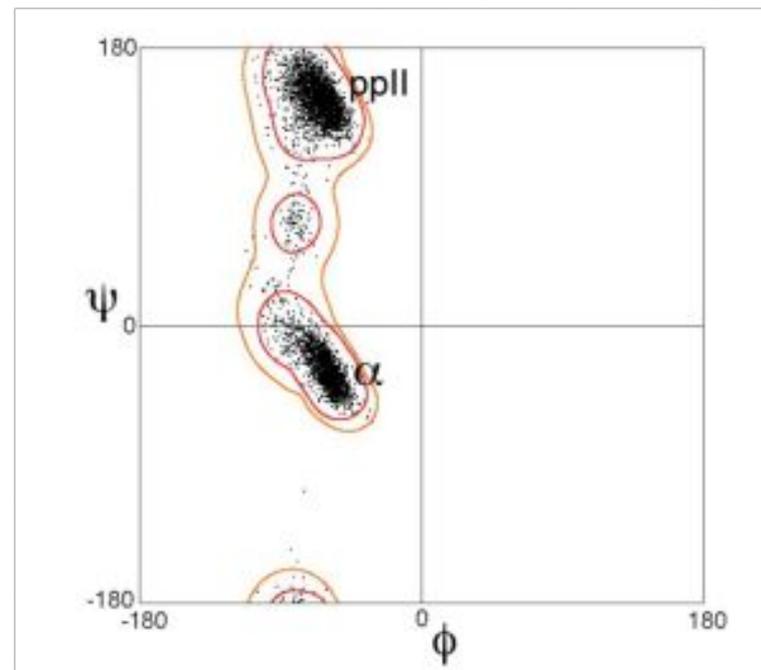


Карты Рамачандрана



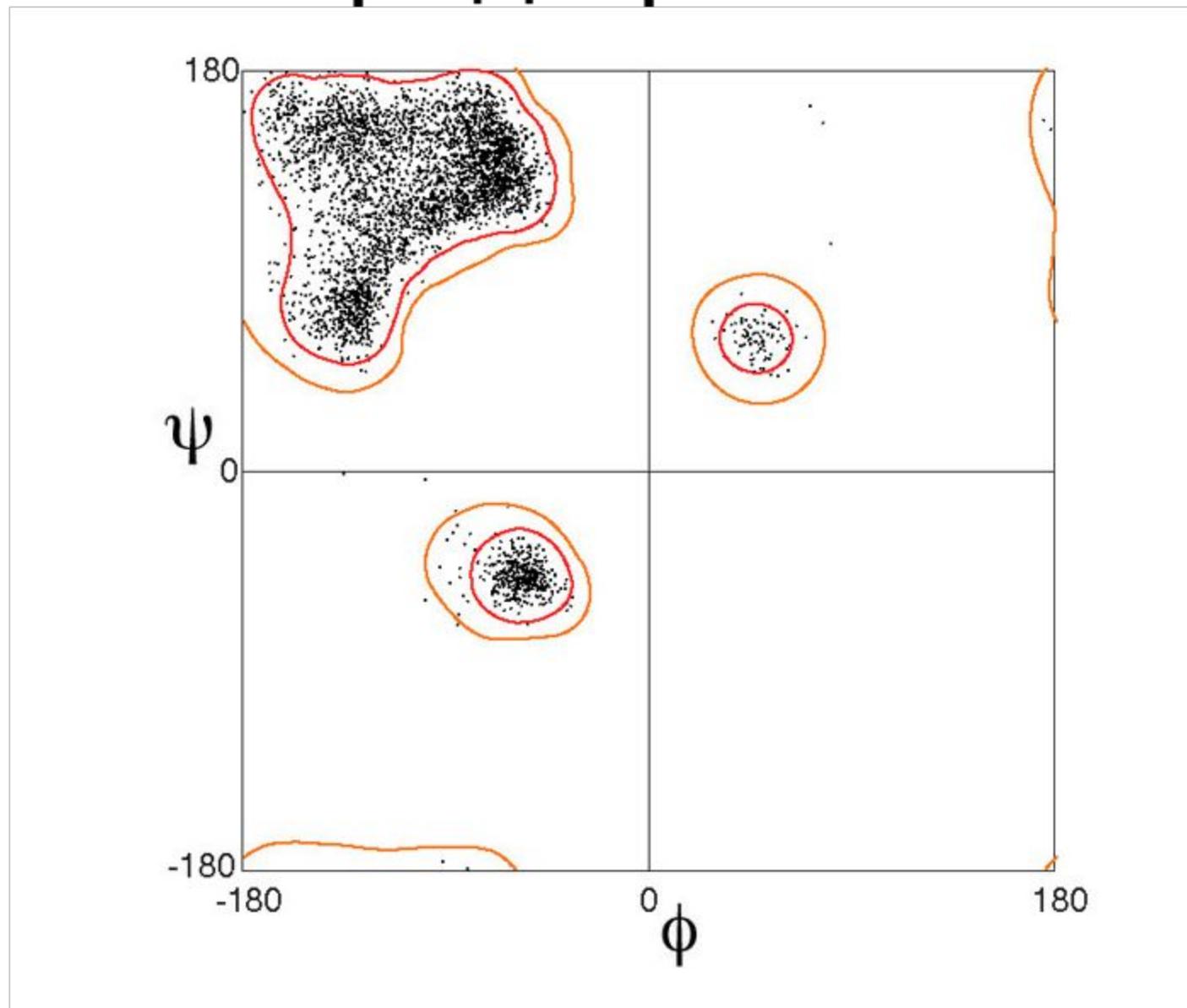


Глицин



Пролин

Перед пролином



DSSP

Definition of the **S**econdary **S**tructure of **P**roteins

Kabsch,W., Sander,C.

Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, no.12 (1983)

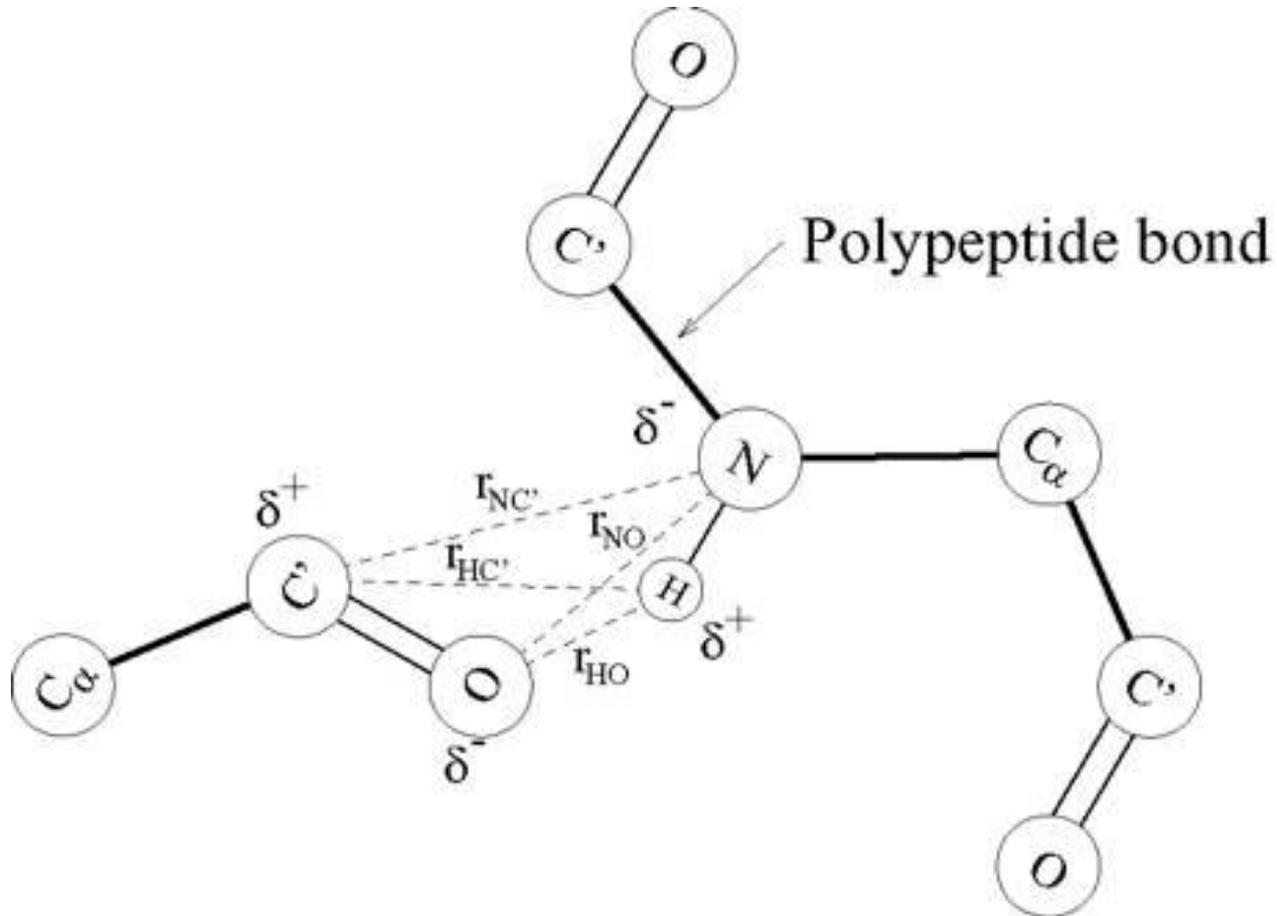
Вход алгоритма: PDB-файл

Выход алгоритма: текстовый файл с информацией о каждом остатке

Этапы алгоритма:

- Нахождение водородных связей между остовными N и O.
- Поиск паттернов водородных связей – “зародышей” элементов вторичной структуры
- Объединение паттернов в элементы вторичной структуры.

DSSP, шаг 1



$$E = f q_1 q_2 (1/r_{NO} + 1/r_{HC'} - 1/r_{HO} - 1/r_{NC'})$$

DSSP, шаг 1

$$E = f q_1 q_2 (1/r_{NO} + 1/r_{HC'} - 1/r_{HO} - 1/r_{NC'})$$

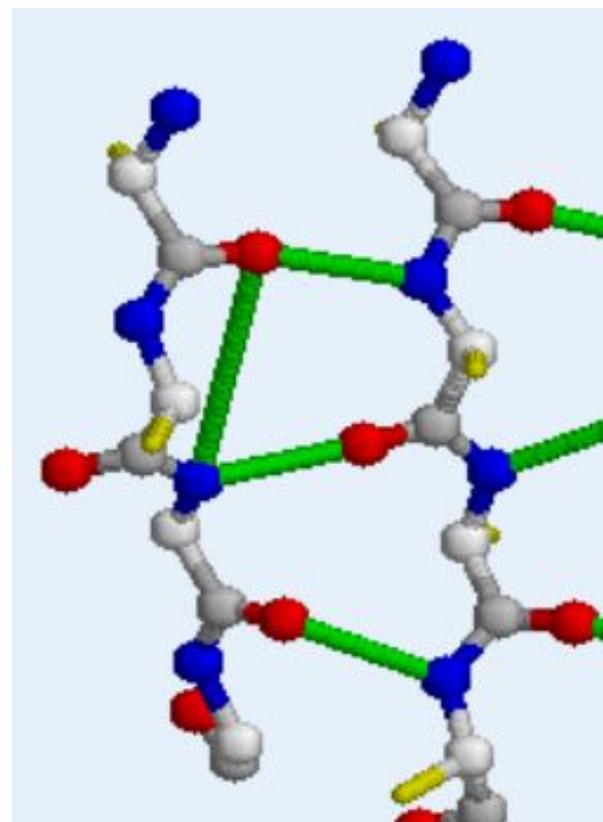
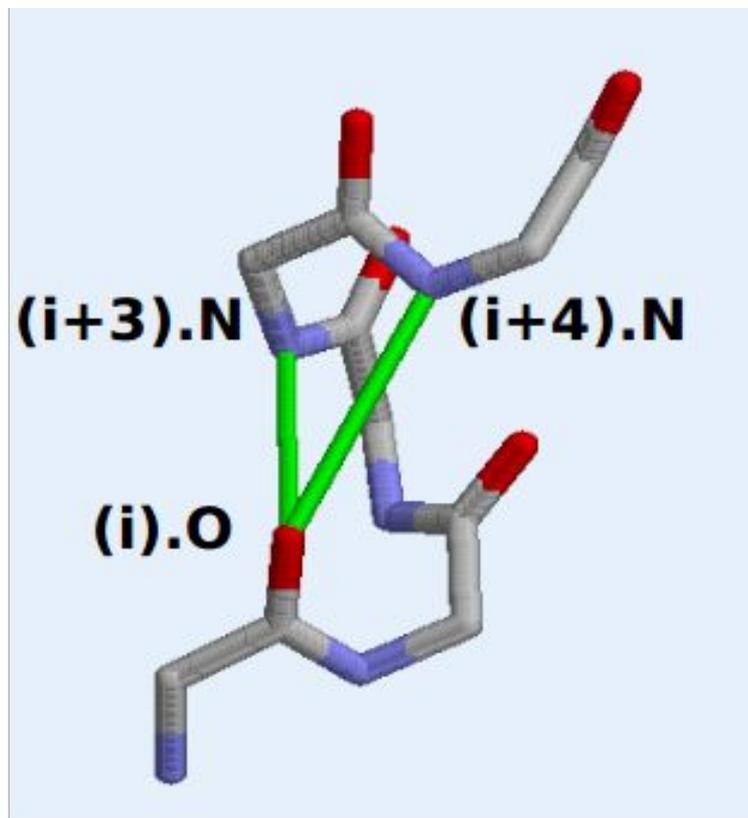
- $q_{1,2}$ – заряды (в единицах заряда электрона):
 $q_1 = 0,42e$; $q_2 = 0,2e$
(считается, что атомы имеют заряды:
C: $+q_1$, O: $-q_1$, N: $-q_2$, H: $+q_2$);
- r_{NO} и др. – межатомные расстояния (в ангстремах)
- f – константа, связанная с размерностью величин,
 $f = 332$ ккал \cdot Å / моль \cdot e
- E исчисляется в ккал/моль
- Положение H восстанавливается
- Критерий водородной связи: $E < -0,5$ ккал/моль
(при «идеальном» расположении атомов
получается примерно -3 ккал/моль)

DSSP может выдать “вилочковые”

СВЯЗИ

в альфа-спиралях

в бета-листах

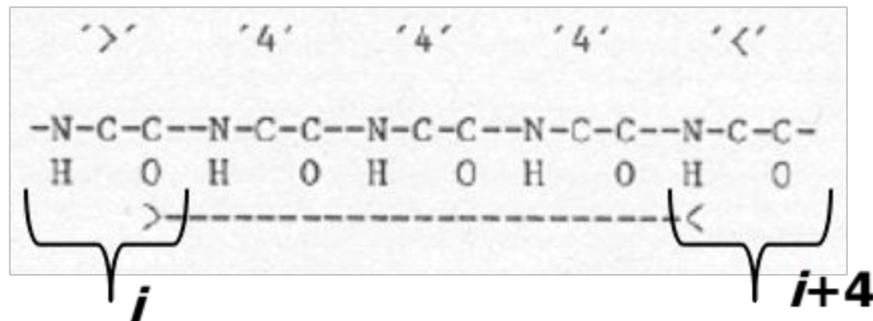


может быть правдой

артефакт алгоритма

Алгоритм DSSP, шаг 2

Поиск элементарных паттернов



4-поворот

n-поворот (*n*-turn)

n=3, 4, 5: *i*.CO---(*i+n*).NH

Детектируются также повороты с *n*= -5, -4, ..., -2, 2

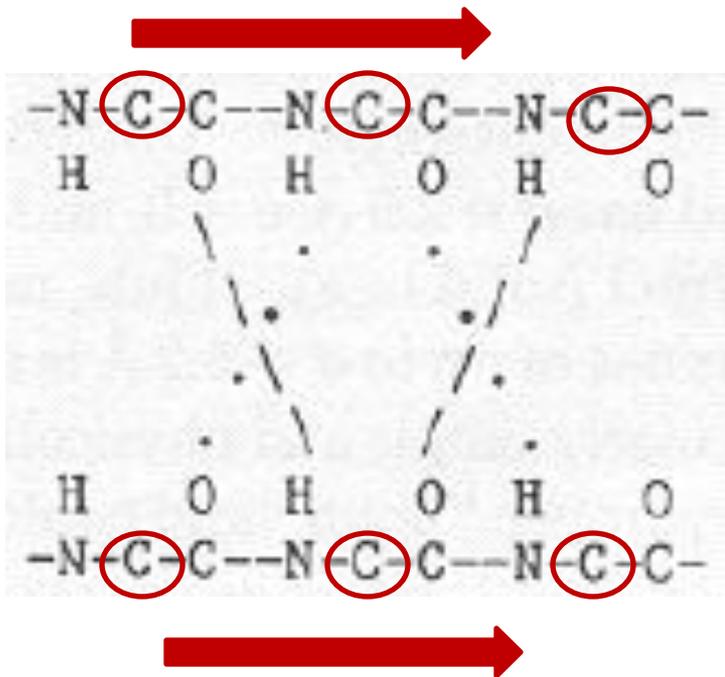
Алгоритм DSSP, шаг 2

Поиск элементарных паттернов

Мостики (bridges)

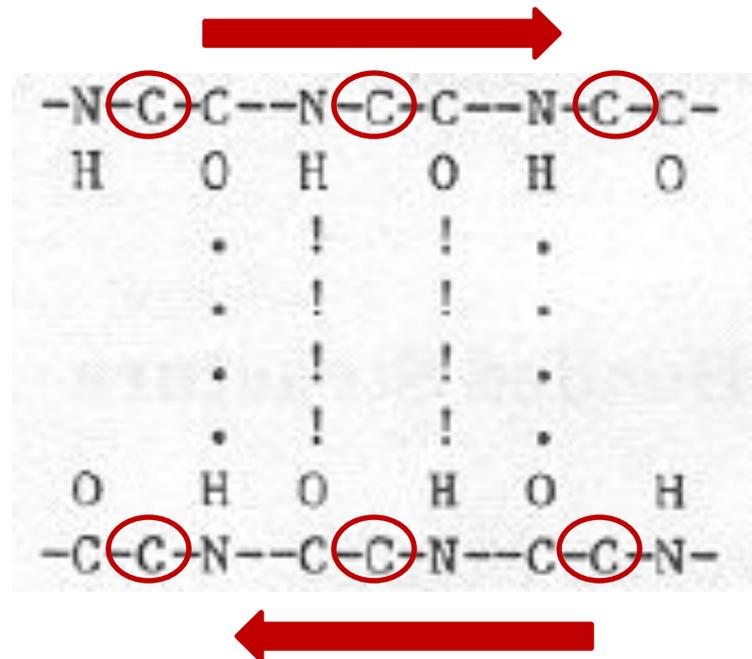
Параллельный (i, j):

Либо $(i-1).O \text{---} (j).N$ и $(j).O \text{---} (i+1).N$,
Либо $(j-1).O \text{---} (i).N$ и $(i).O \text{---} (j+1).N$



Антипараллельный (i, j):

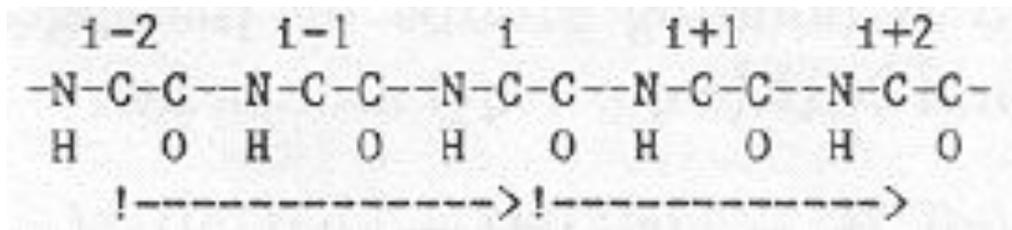
Либо $(i).O \text{---} (j).N$ и $(j).O \text{---} (i).N$,
Либо $(i-1).O \text{---} (j+1).N$ и $(j-1).O \text{---} (i+1).N$



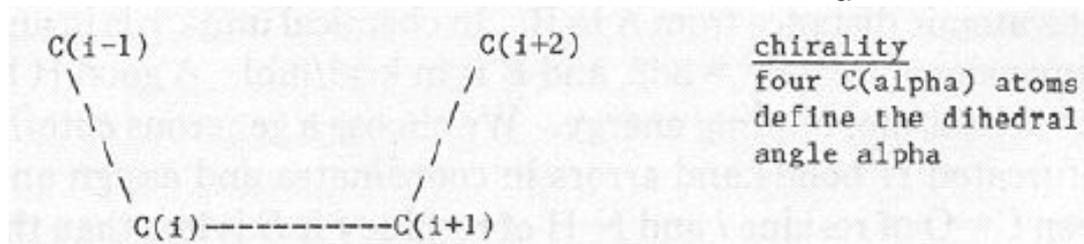
Алгоритм DSSP, шаг 2

Элементарные паттерны, не связанные с водородными связями

Излом (bend): изменение направления более чем на 70°



Хиральность. Остатку i приписывается знак (“+” или “-”) торсионного угла, образованного C_α -атомами $i-1$, i , $i+1$, $i+2$



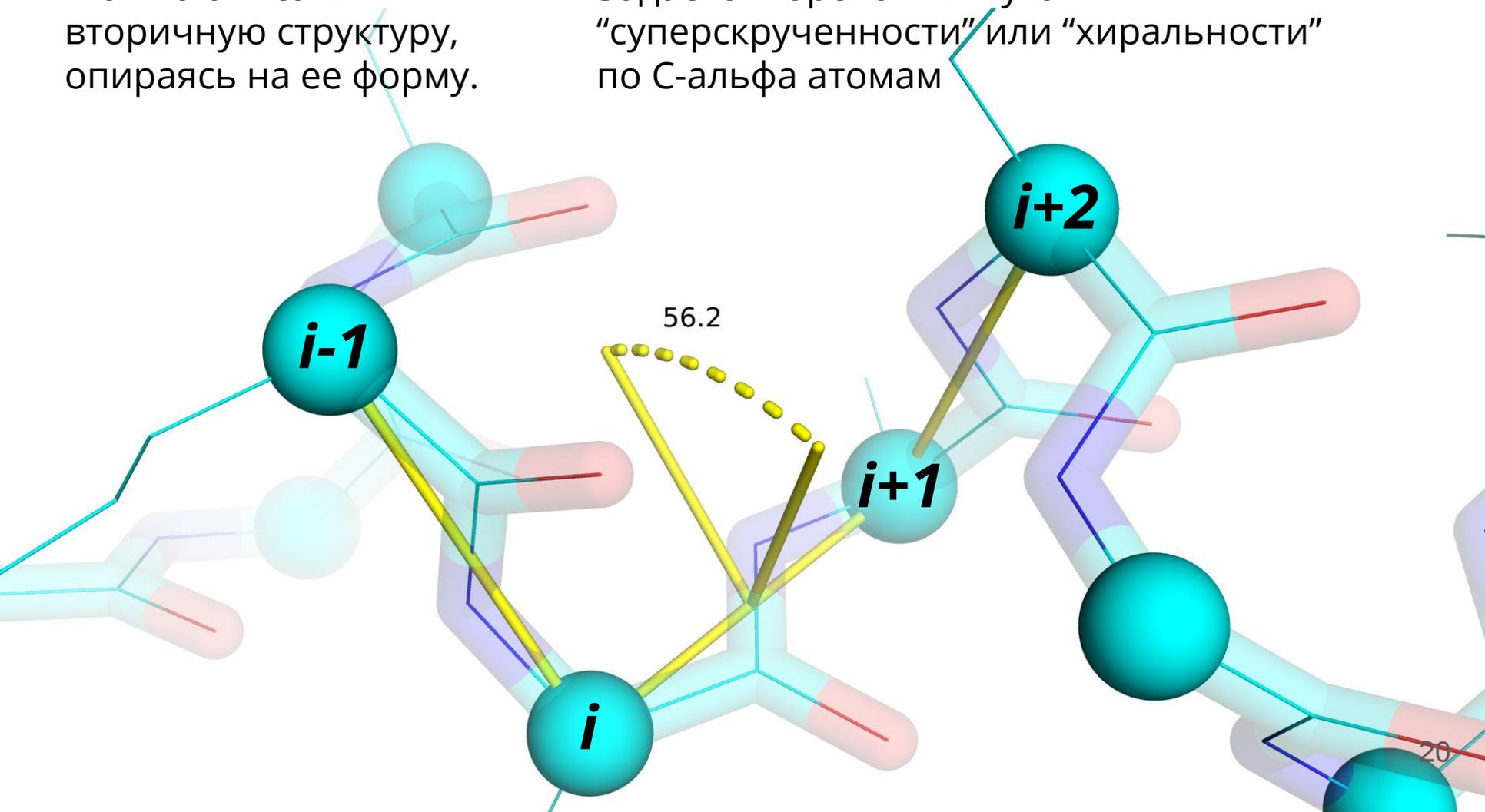
Хиральность остатков из правых спиралей “+”; “левые” спирали исключительно редки

Хиральность остатков из скрученных бета-листов обычно “-”

“Хиральность”

Можно описать
вторичную структуру,
опираясь на ее форму.

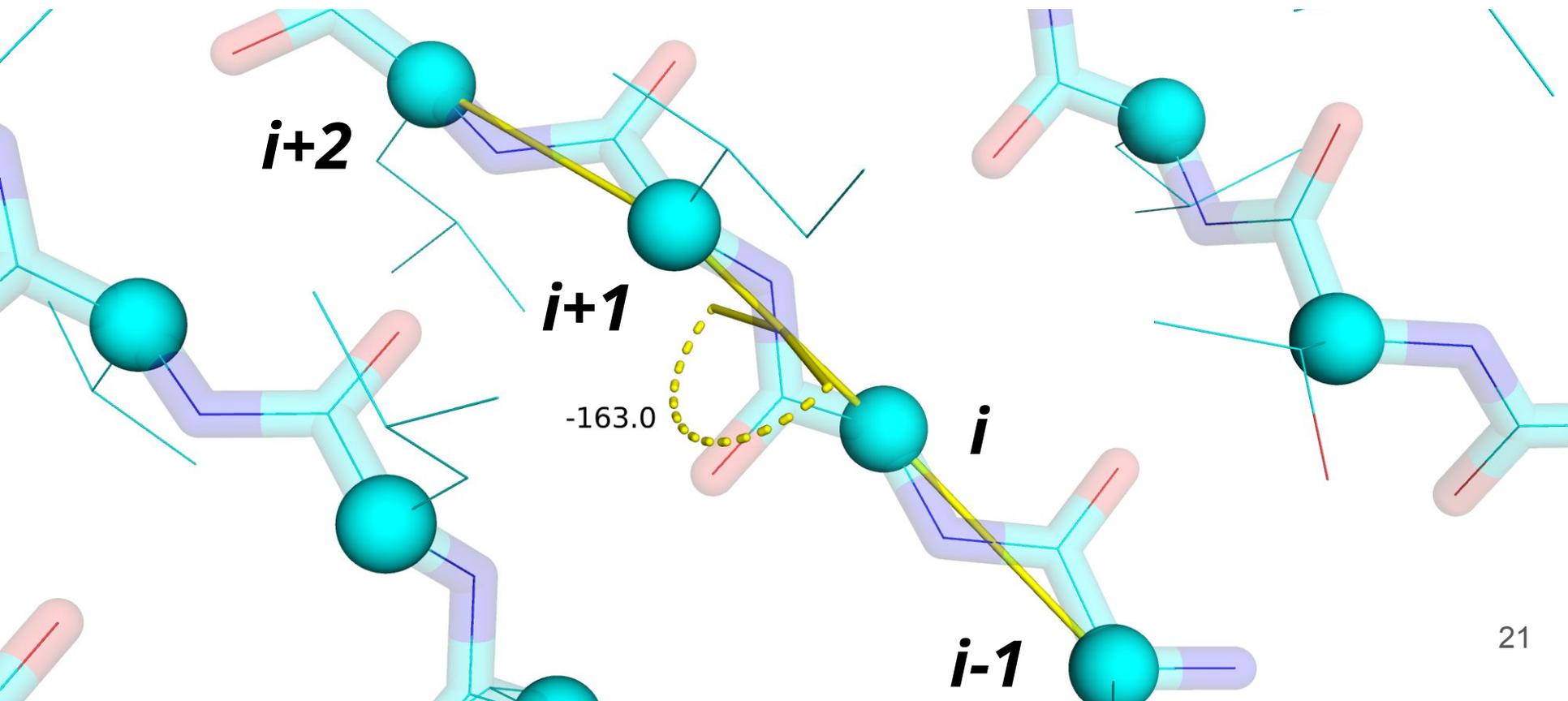
Задается торсионный угол
“суперскрученности” или “хиральности”
по C-альфа атомам



“Хиральность”

Можно описать
вторичную структуру,
опираясь на ее форму.

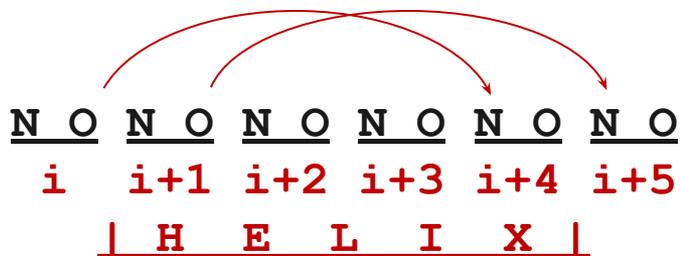
Задается торсионный угол
“суперскрученности” или “хиральности”
по C-альфа атомам



Алгоритм DSSP, шаг 3

Спирали

Минимальная альфа-спираль = два последовательных 4-поворота. Ее длина 4 остатка, крайние не включаются в альфа-спираль!



Альфа-спираль = максимальный набор пересекающихся или следующих вплотную друг за другом минимальных спиралей

Аналогично определены 3_{10} -спирали (из 3-поворотов) и π -спирали (из 5-поворотов)

Алгоритм DSSP, шаг 3

Лестницы и листы

Лестница = один или более мостиков *одного типа*, идущих подряд

Две лестницы связаны, если имеют общий остаток

Бета-лист: максимальный связный набор лестниц

Нерегулярности: допускаются выпячивания (не более одного на одном тяжёлом листе и не более четырёх – на другом)

Обозначения DSSP

- H = alpha helix
- B = residue in isolated beta-bridge
- E = extended strand, participates in beta ladder
- G = 3-helix (3/10 helix)
- I = 5-helix (pi helix)
- T = hydrogen bonded turn
- S = bend

Участки 3-, 4- или 5-поворотов, не составляющие спираль достаточной длины, обозначаются T

Приоритет при детектировании нескольких ситуаций у одного остатка: сверху вниз (от H к S)

Если не детектировано ничего из перечисленного, в соответствующей колонке ставится пробел (хотя некоторые программы обработки заменяют его на букву C, от слова "coil").

Алгоритм Stride

(secondary STRuctural IDEntification)

Frishman,D., Argos,P.

[Knowledge-based protein secondary structure assignment.](#)

Proteins **23** , no, 4 (1995)

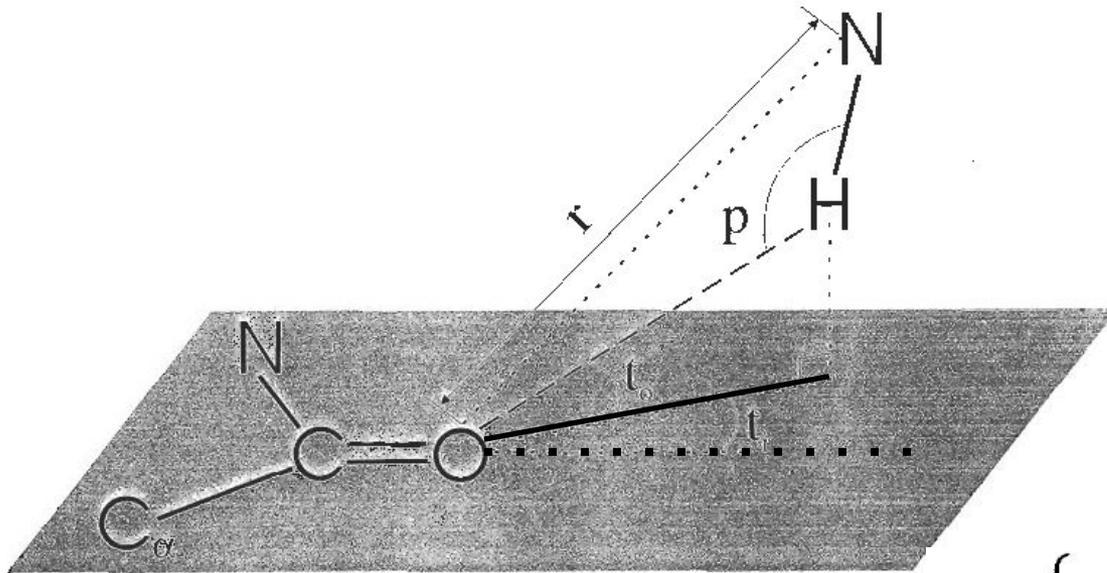
Этапы алгоритма:

- Расчет энергии водородной связи для пар остовных атомов N и O. Способ расчета отличен от используемого в DSSP
- Для каждого остатка: расчет вероятности того, что остаток с данными значениями торсионных углов ϕ и ψ , принадлежит альфа-спирали (бета-тяжу)
- Детекция «зародышей» элементов вторичной структуры
- Объединение зародышей в элементы

Алгоритм основан на подходе “knowledge-based”, то есть использовании статистики экспертных аннотаций.

Алгоритм Stride, шаг 1

“Энергия” водородной связи



$$E_r = \frac{C}{r^8} + \frac{D}{r^6}$$

$$E_p = \cos^2 p$$

$$E = E_r \cdot E_p \cdot E_t$$

$$E_p = \begin{cases} (0.9 + 0.1 \sin 2t_i) \cos t_0, & 0 < t_i < 90^\circ \\ K_1(K_2 - \cos^2 t_i)^3 \cos t_0, & 90^\circ < t_i < 110^\circ \\ 0, & t_i > 110^\circ \end{cases}$$

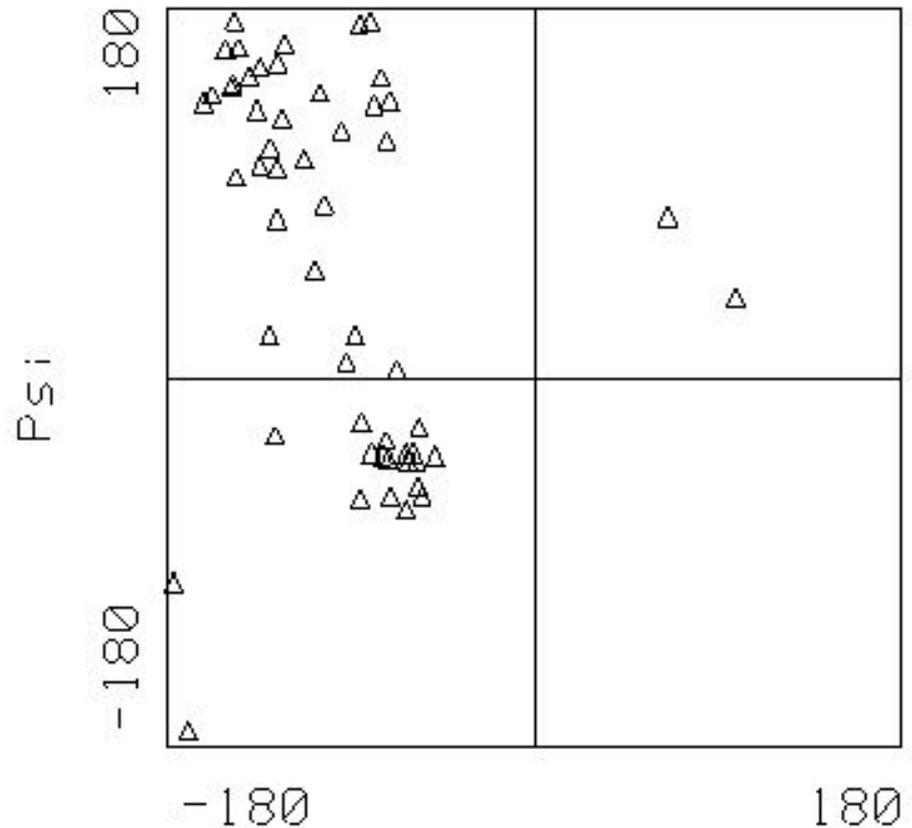
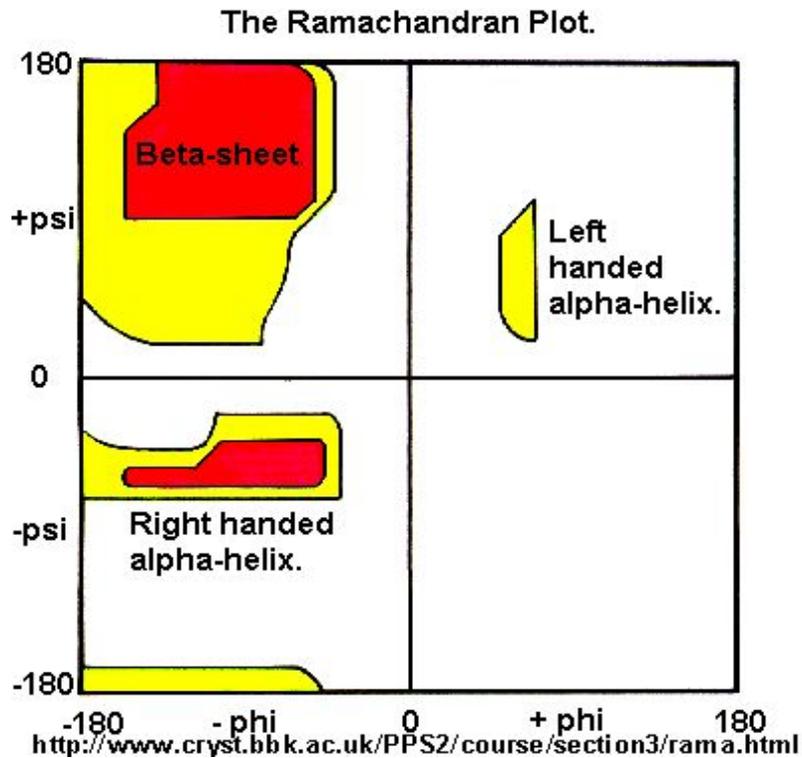
E_r зависит только от расстояния между O и N

E_p зависит от угла p

E_t зависит от углов t_0 и t_1

Алгоритм Stride, шаг 2

Статистика углов ϕ и ψ



ϕ

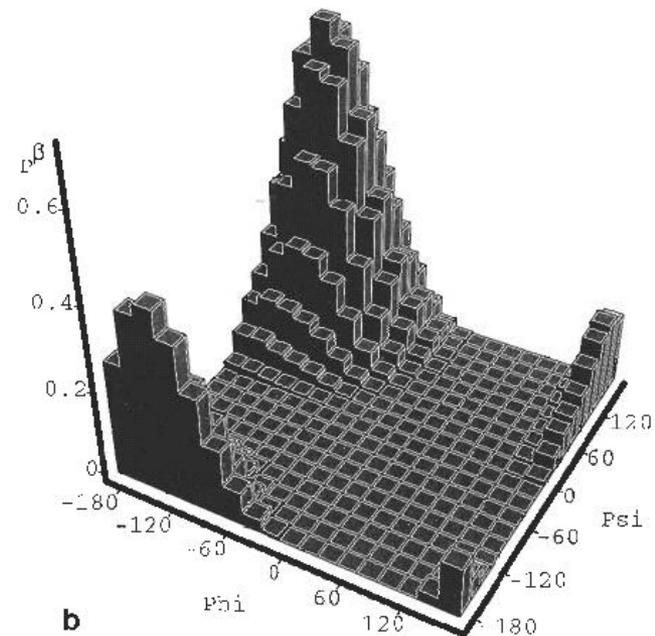
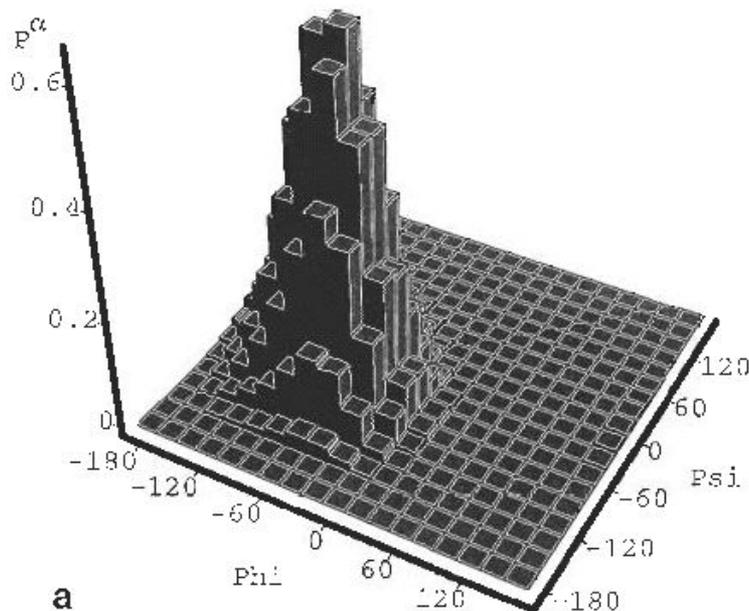
Алгоритм Stride, шаг 2

Статистика углов ϕ и ψ

Собрана статистика по специально составленной выборке из 226 структур (“материал обучения” для уточнения параметров алгоритма). Принадлежность остатка альфа-спирали (бета-тяжу) определялась по аннотации PDB файла.

Карта Рамачандрана разбита на квадратики $20^\circ \times 20^\circ$, построена гистограмма (отдельно для бета- и альфа-).

Удалены точки из областей, заведомо запрещенных для альфа-спиралей (бета-тяжей) – это вероятные ошибки и артефакты.



Алгоритм Stride, шаг 2

Правдоподобие того, что остаток принадлежит спирали или тяжу

Вероятности попасть в квадратик i равны $P_i^\alpha = N_i^\alpha / N^\alpha$, $P_i^\beta = N_i^\beta / N^\beta$,
где N_i^α – число точек из альфа-спиралей в квадратике i ,
 N^α – суммарное число точек из альфа-спиралей.

Аналогично – для бета.

Функции P_i^α и P_i^β сглаживаются одним из стандартных способов

Алгоритм Stride, шаги 3 и 4

Критерий для пары $(i, i+4)$, чтобы считаться зародышем альфа-спирали:

$$E_{\text{hb}}^{i, i+4} \cdot \left(1 + W_1^\alpha + W_2^\alpha \cdot \frac{P_i^\alpha + P_{i+4}^\alpha}{2} \right) < T_1^\alpha$$

Минимальная альфа-спираль – два идущих подряд зародыша $(i, i+4)$, $(i+1, i+5)$; спиралью всегда считаются остатки $i+1, i+2, i+3, i+4$; а остатки i и $i+5$ — при выполнении условий $P_i^\alpha > T_2^\alpha$ и $P_{i+5}^\alpha > T_3^\alpha$, соответственно.

Константы $W_1^\alpha, W_2^\alpha, T_1^\alpha, T_2^\alpha, T_3^\alpha$ подобраны так, чтобы на материале обучения получить наилучшее совпадение с аннотированными в PDB альфа-спиралями.

Для бета-листов процедура аналогична.

Другие алгоритмы

- DEFINE (1988)
- P-CURVE (1989)
- P-SEA (1997)
- XTLSSTR (1999)
- SECSTR (2002)
- VoTAP (2004)
- KAKSI (2005)
- PROSIGN (2008)
- SABA (2011)
- DISICL (2014)
- PCASSO (2014)
- SACF (2016)

Многие из них используют только информацию о положении C α -атомов.

Наиболее популярным по-прежнему остаётся DSSP ☺

В моем белке бета-лист. И что?

- Жесткость листа позволяет ему служить передатчиком механического сигнала на большие расстояния

**nature
communications**

Article | Published: 11 June 2014

Correlated motions are a fundamental property of β -sheets

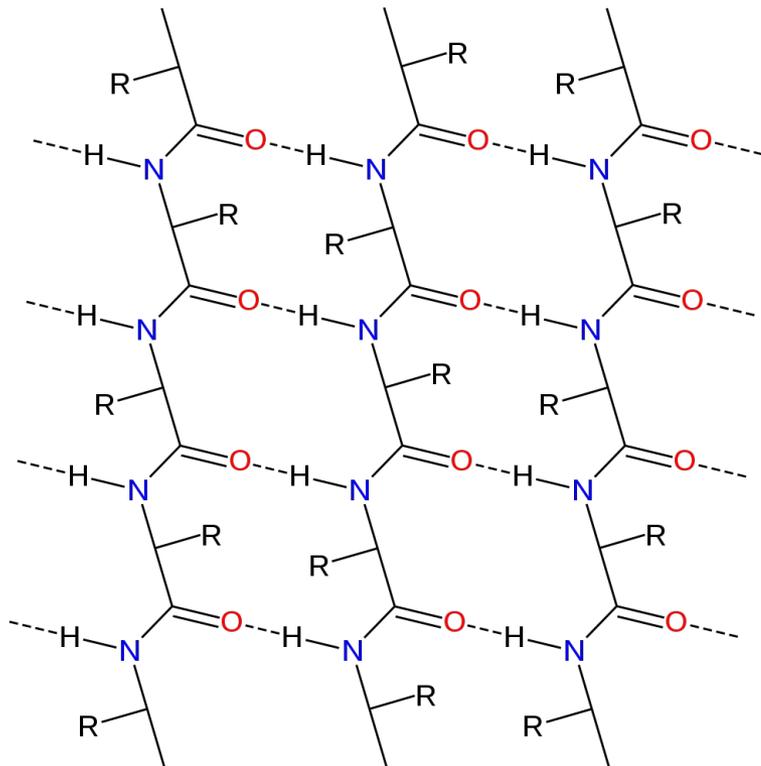
R. Bryn Fenwick, Laura Orellana, Santi Esteban-Martín, Modesto Orozco  & Xavier Salvatella 

Nature Communications 5, Article number: 4070 (2014) | [Cite this article](#)

- Бета-листовые структуры склонны к агрегации



Забытая вторичная структура - альфа-ЛИСТ



npj | Biofilms and Microbiomes

Article | [Open Access](#) | Published: 03 July 2017

Designed α -sheet peptides suppress amyloid formation in *Staphylococcus aureus* biofilms

Alissa Bleem, Robyn Francisco, James D. Bryers & Valerie Daggett [✉](#)

npj Biofilms and Microbiomes 3, Article number: 16 (2017) | [Cite this article](#)

539 Accesses | 8 Citations | 16 Altmetric | [Metrics](#)

α -Sheet secondary structure in amyloid β -peptide drives aggregation and toxicity in Alzheimer's disease



Dylan Shea, Cheng-Chieh Hsu, Timothy M. Bi, Natasha Paranjapye, Matthew Carter Childers, Joshua Cochran, Colson P. Tomberlin, Libo Wang, Daniel Paris, Jeffrey Zonderman, Gabriele Varani, Christopher D. Link, Mike Mullan, and Valerie Daggett

PNAS April 30, 2019 116 (18) 8895-8900; first published April 19, 2019 <https://doi.org/10.1073/pnas.1820585116>

Edited by Angela M. Gronenborn, University of Pittsburgh School of Medicine, Pittsburgh, PA, and approved March 26, 2019 (received for review December 4, 2018)

В моем белке альфа-спираль. И что?

- Спираль - диполь. На N-конце частичный заряд до +0.9. Концы спиралей имеют большую плотность водородных связей. Какой в этом смысл? Конец спирали + примыкающая петля часто являются местом связывания анионных субстратов

nature

Article | Published: 08 June 1978

The α -helix dipole and the properties of proteins

W. G. J. Hol, P. T. van Duijnen & H. J. C. Berendsen

Nature 273, 443–446(1978) | [Cite this article](#)

- Дипольный момент спирали может сдвигать рKa остатков. Играет роль в активации (депротонировании) цистеина в тиоредоксине
- Спираль прямая и жесткая - строительный блок

