



Научно-технологический  
университет

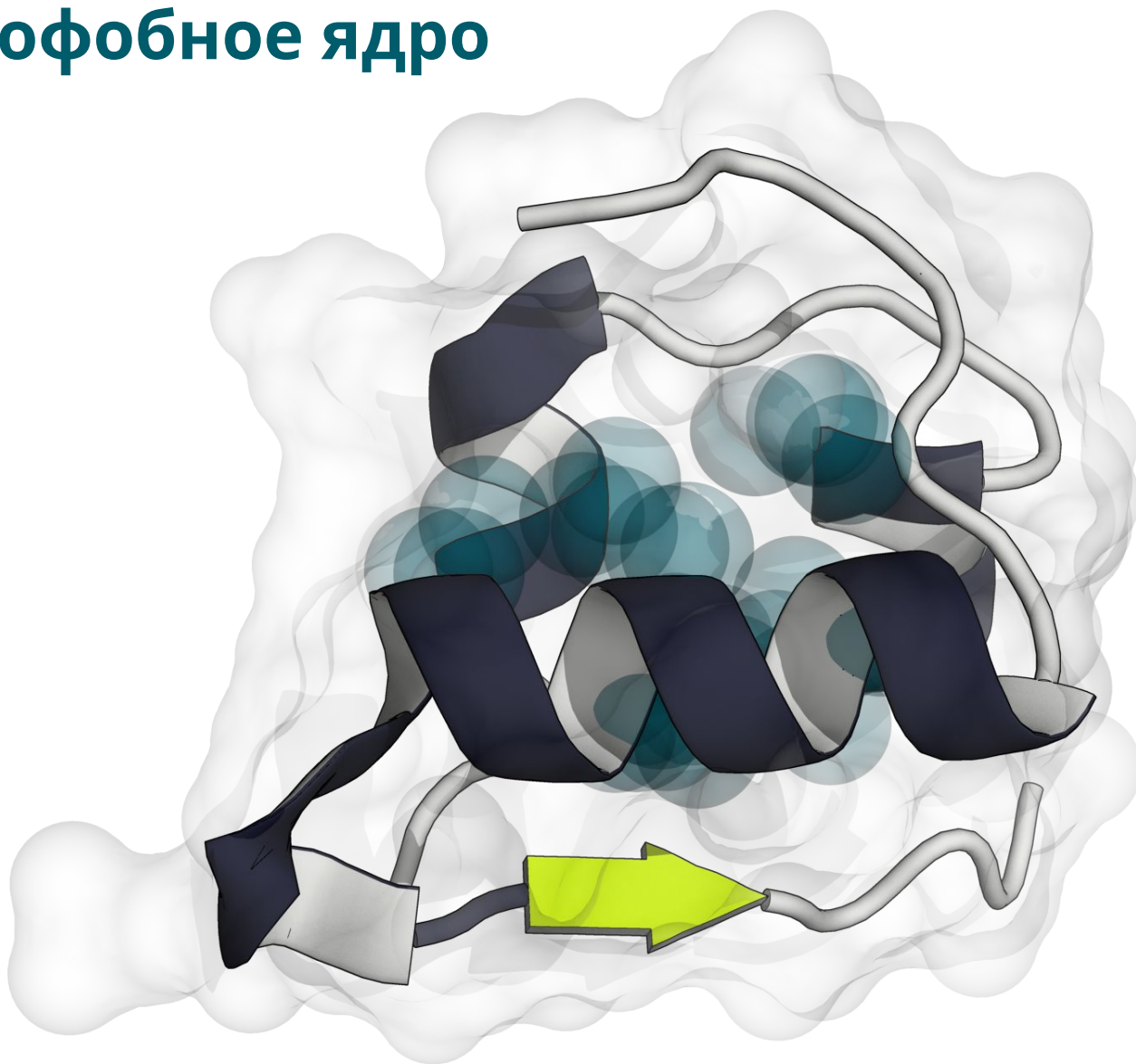
Сириус

Структурная биоинформатика | Лекция 4

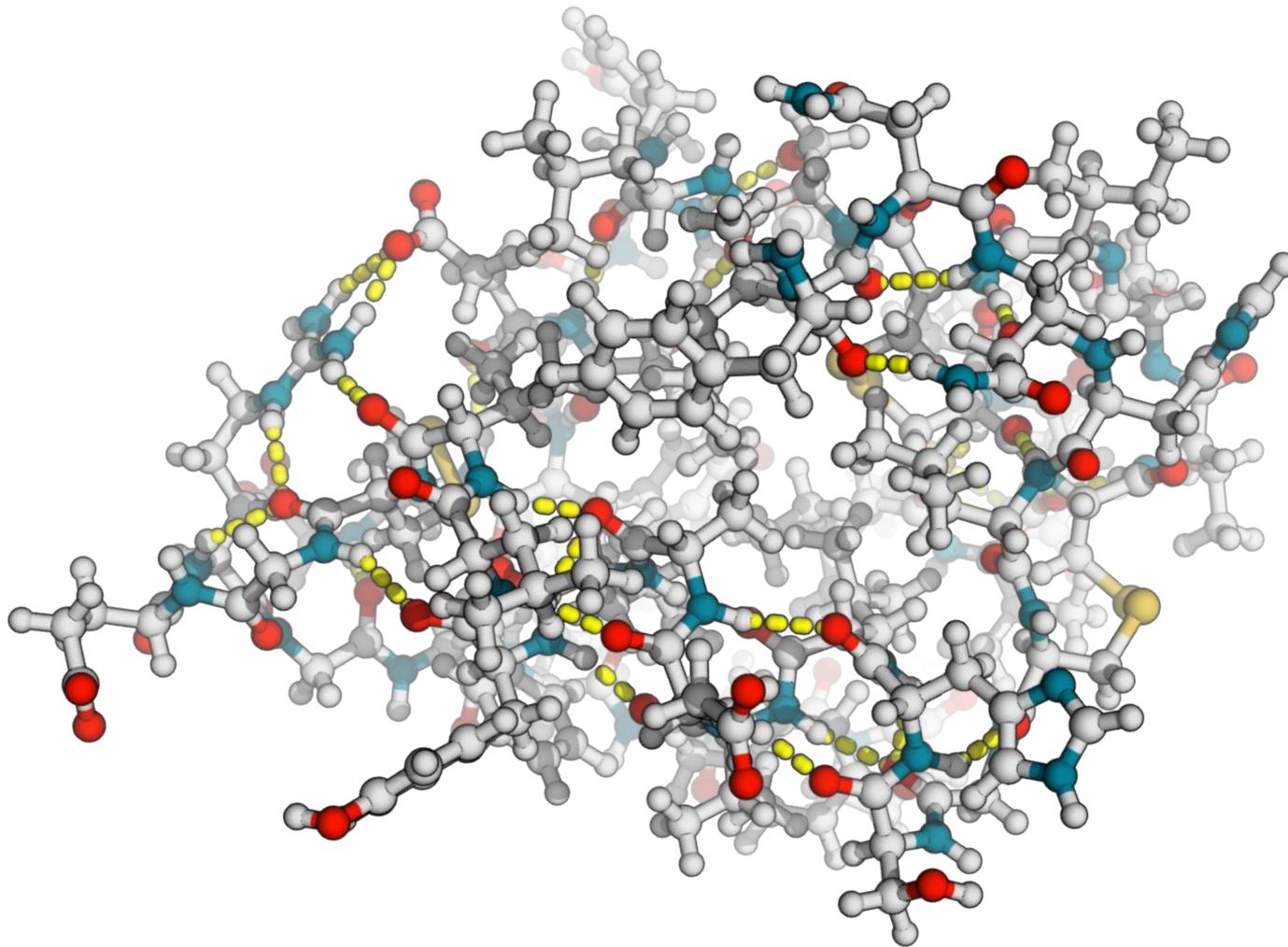
# Третичная структура. Пространство укладок

Александр Злобин

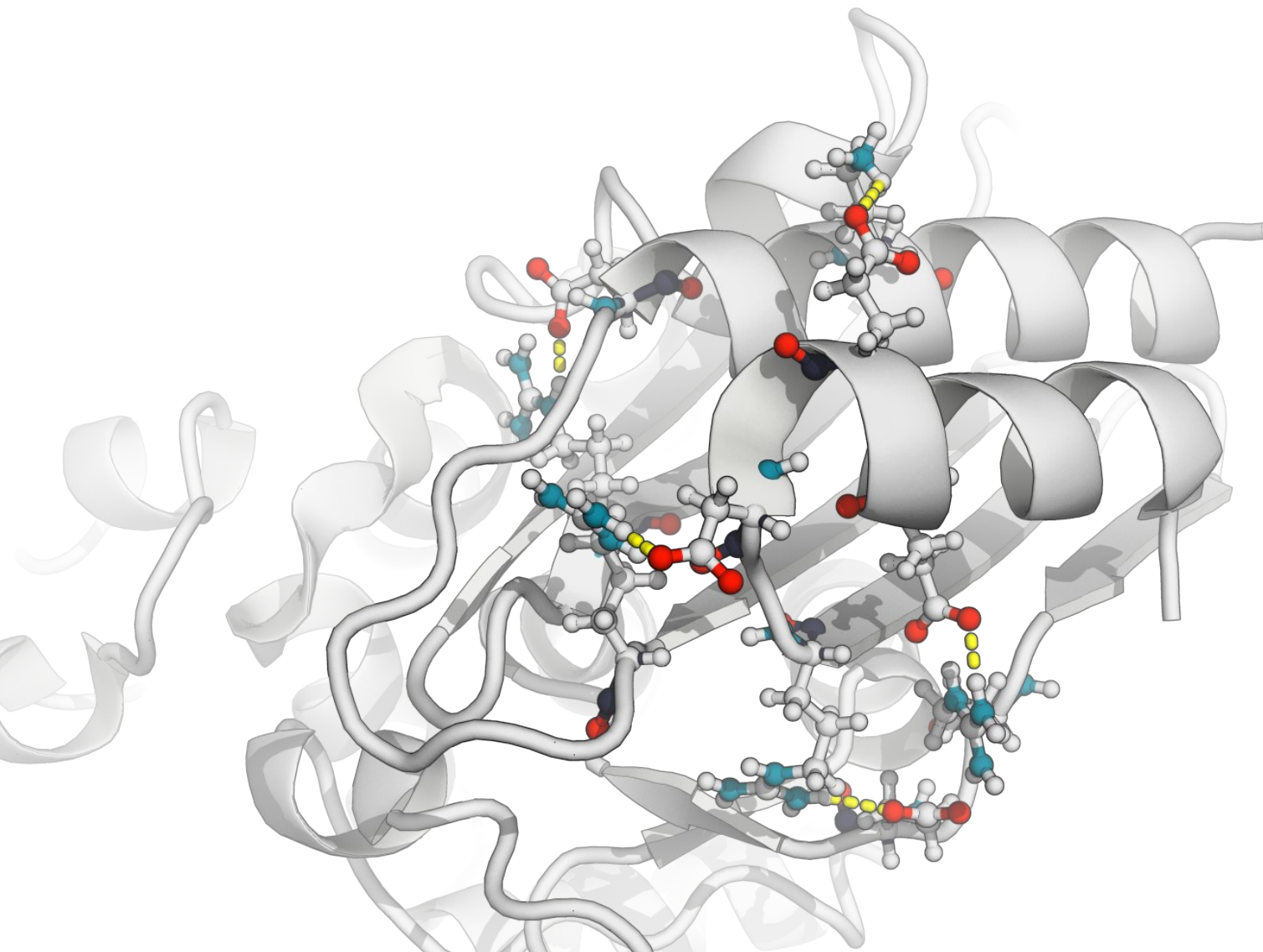
# Гидрофобное ядро



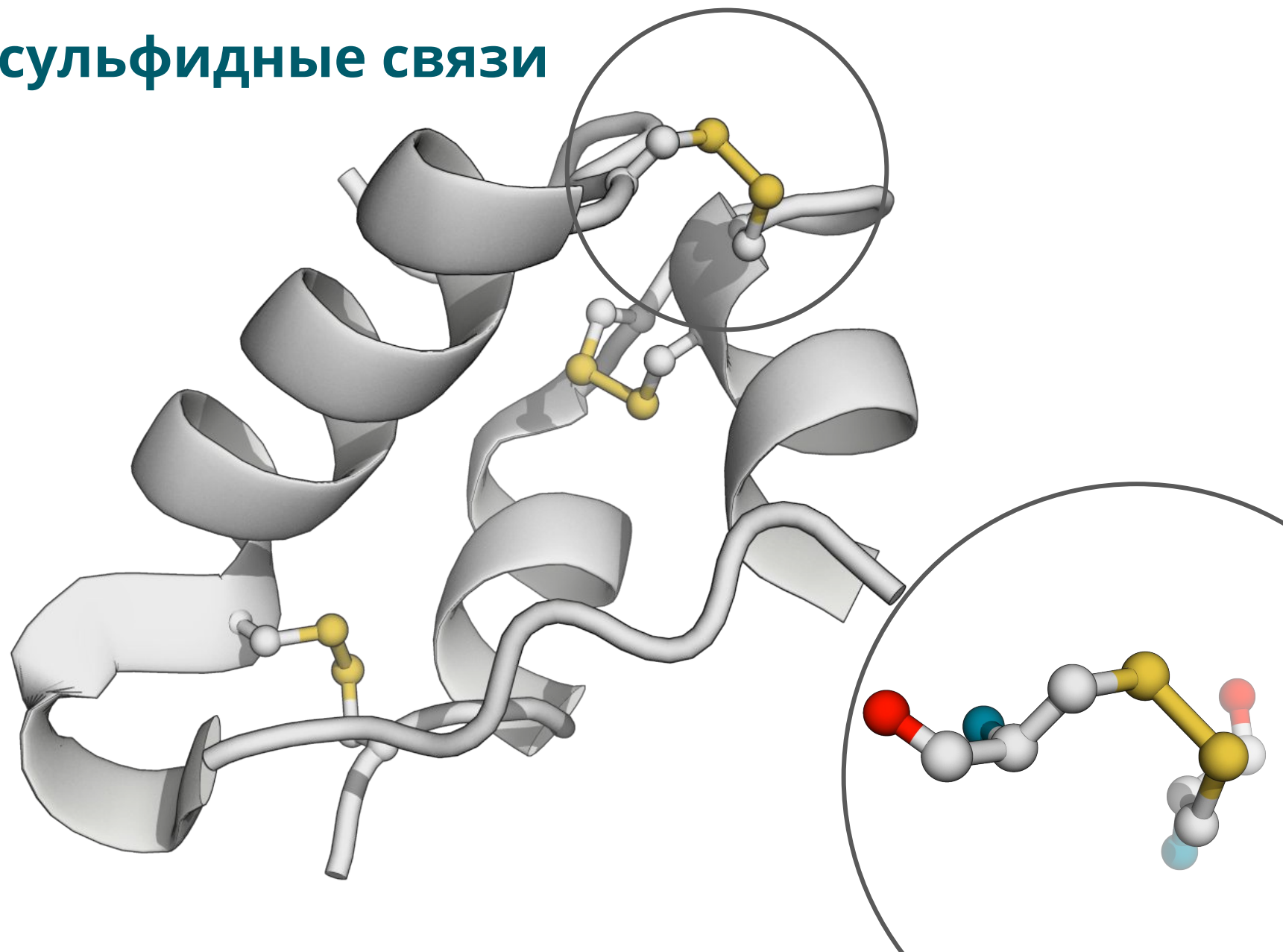
# Сеть водородных связей внутри белка



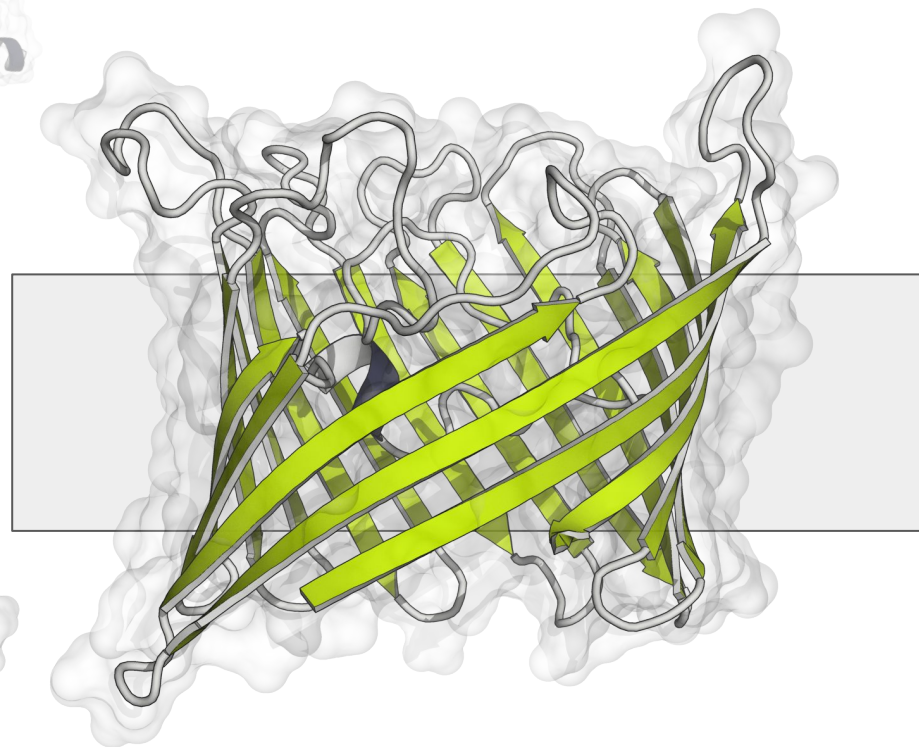
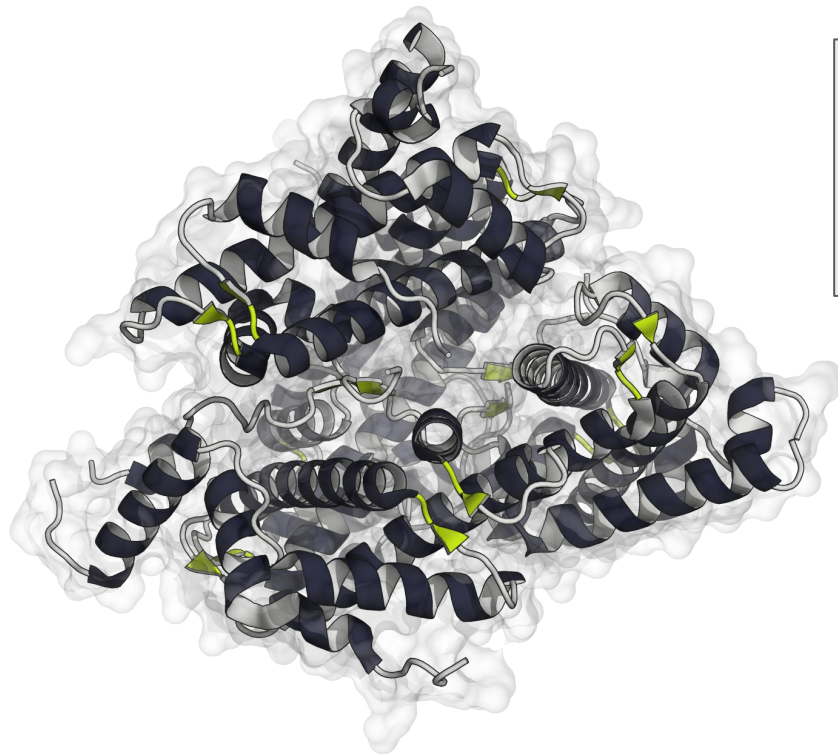
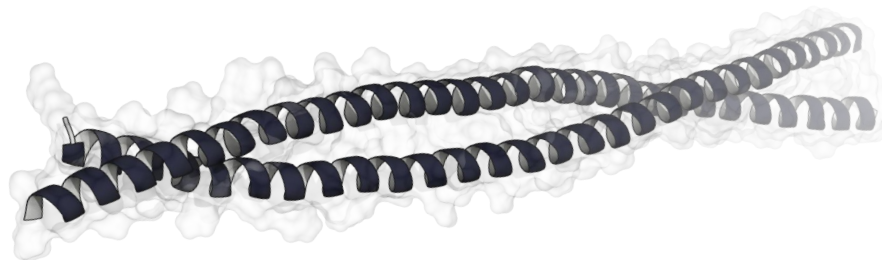
# Соляные мостики



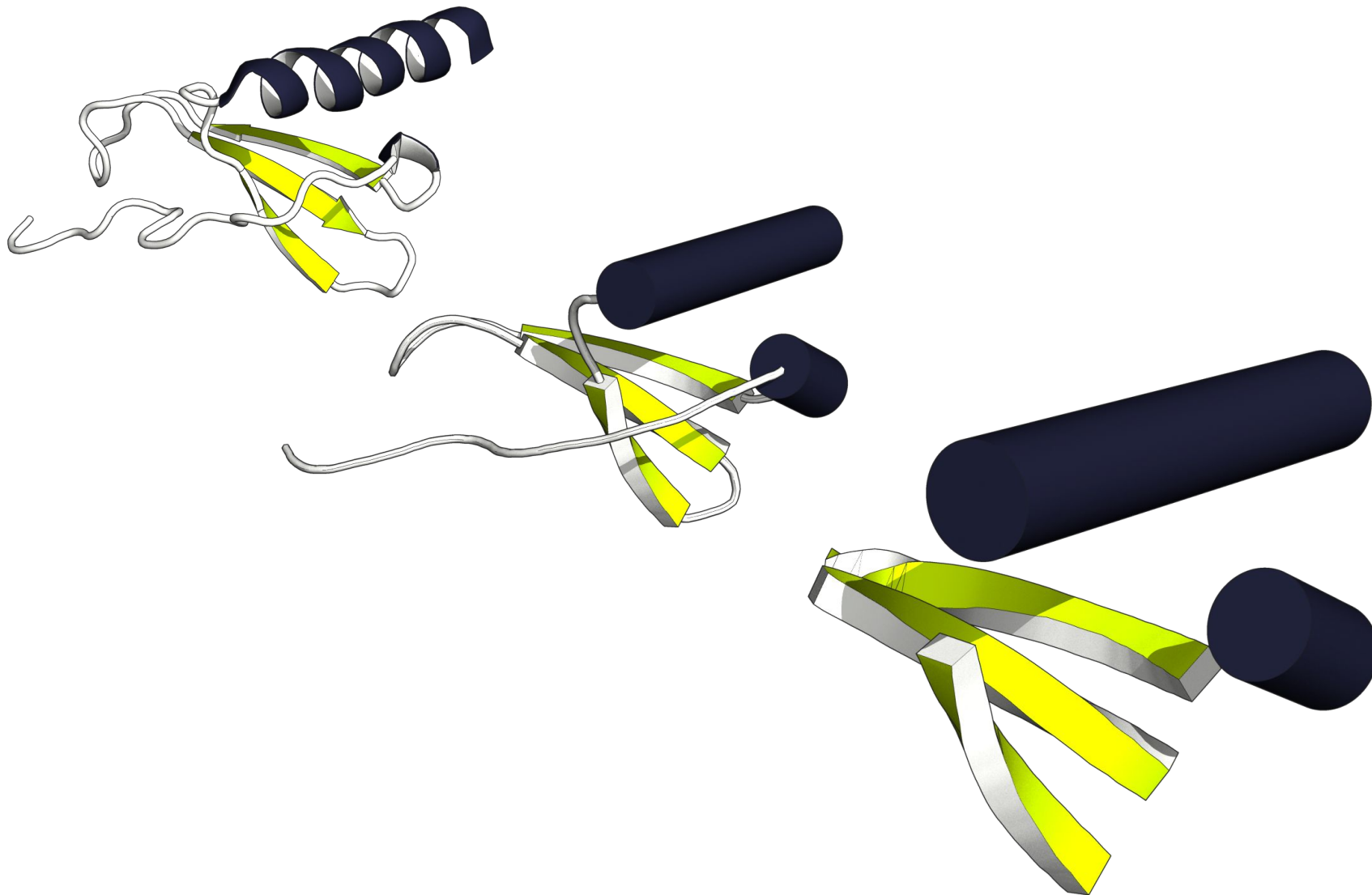
# Дисульфидные связи



# Классификация третичных структур



# Укладка глобулярного белка



# Структурная классификация: укладка

Абстрактный термин, подразумевающий композицию и способ взаиморасположения элементов вторичной структуры внутри домена.

## **Fold**

1. Мы знаем много укладок. И что?
2. Почему разные белки свернуты по-разному?  
Зачем такой-то белок свернут именно так?
3. Зачем нужна большая часть белка (кроме активного центра)?
4. Почему одни укладки распространены больше других?
5. Как связана укладка с процессом сворачивания (фолдингом)?
6. Как возникают и изменяются укладки в ходе эволюции?
7. и т.д.



# Структурные классификации

- SCOP – полуавтоматическая иерархическая классификация, не обновляется с 2009.
- SCOP2 – замена SCOP, under construction
- SCOPe – SCOP extended (оживший SCOP)
- ECOD – независимо оживший SCOP
- CATH – полуавтоматическая иерархическая классификация, обновляется.

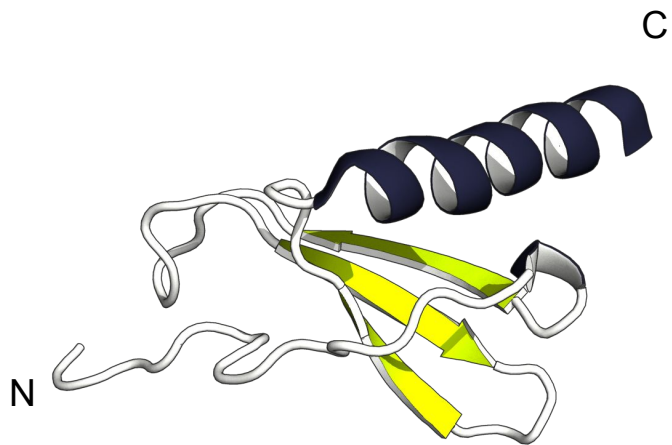
# CATH: уровни классификации

**Class** - structures are classified according to their secondary structure composition (mostly alpha, mostly beta, mixed alpha/beta or few secondary structures).

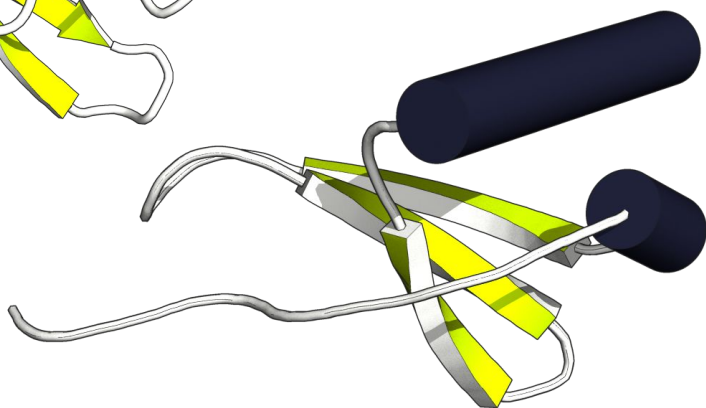
**Architecture** - structures are classified according to their overall shape as determined by the orientations of the secondary structures in 3D space but ignores the connectivity between them. ~ *Fold*

**Topology (fold family)** - structures are grouped into fold groups at this level depending on both the overall shape and connectivity of the secondary structures.

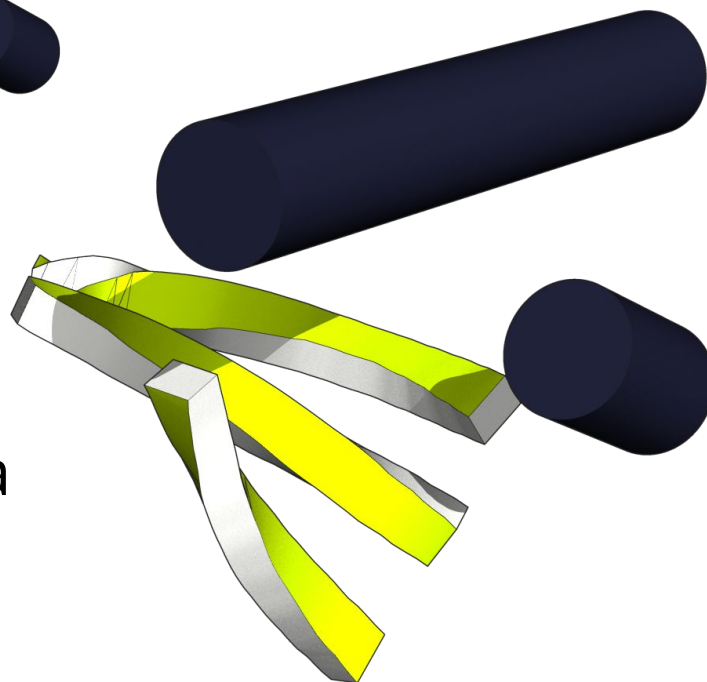
**Homologous superfamily** - this level groups together protein domains which are thought to share a common ancestor and can therefore be described as homologous.

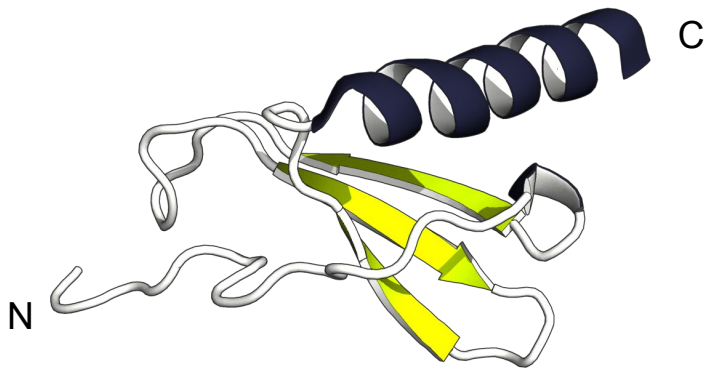


Топология

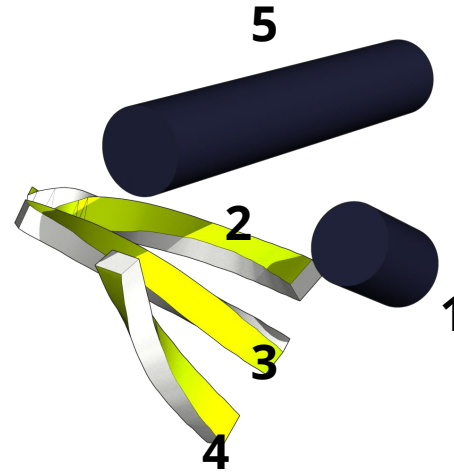


Архитектура





Топология

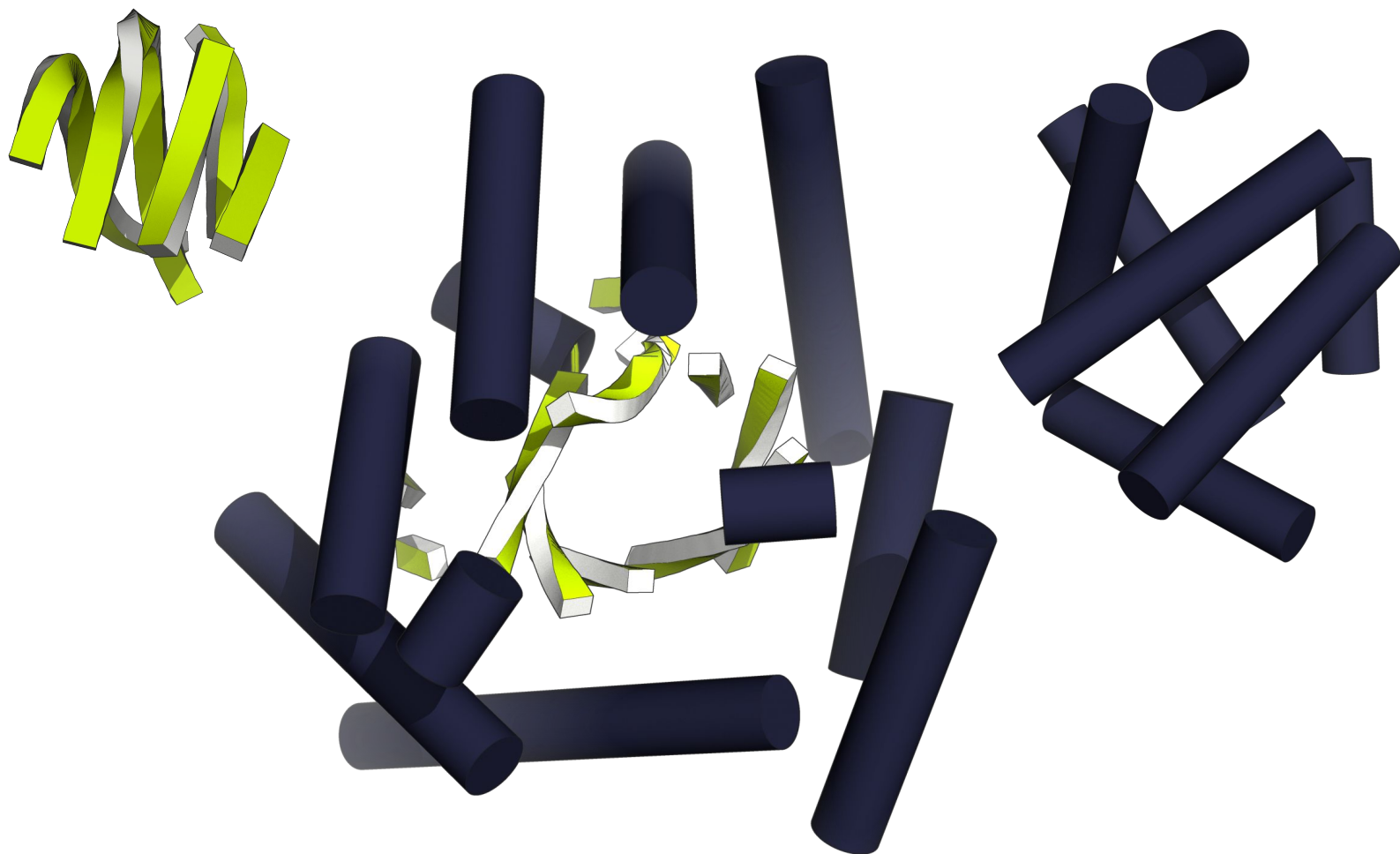


Архитектура

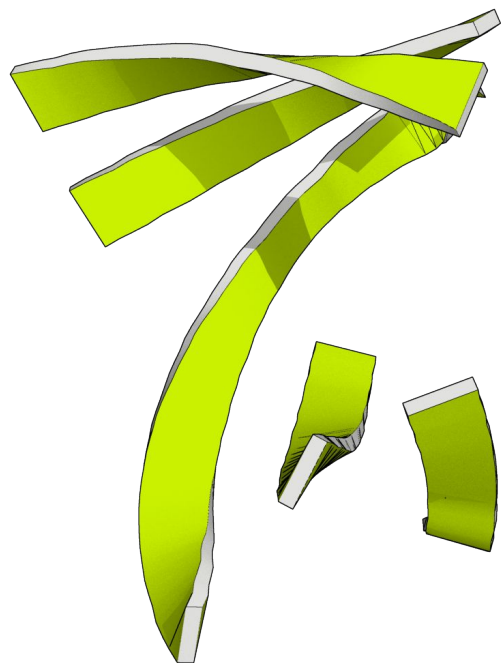
Данная топология: 1-2-3-4-5

Однако мы вполне можем представить себе случай, когда она будет 1-4-3-2-5

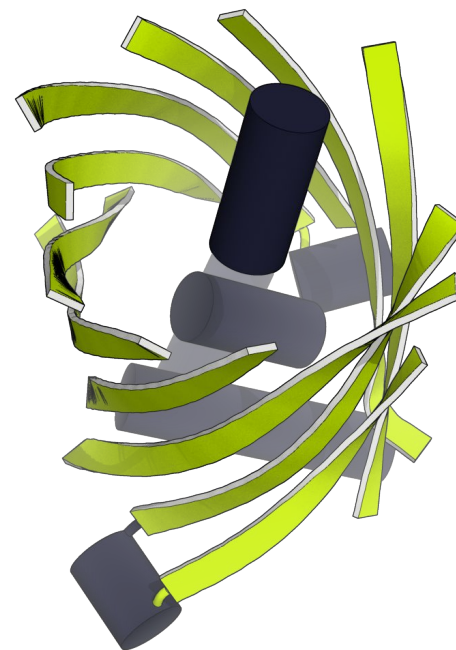
# Классификация упадок по архитектуре



# Разнообразие фолдов: преимущественно $\beta$

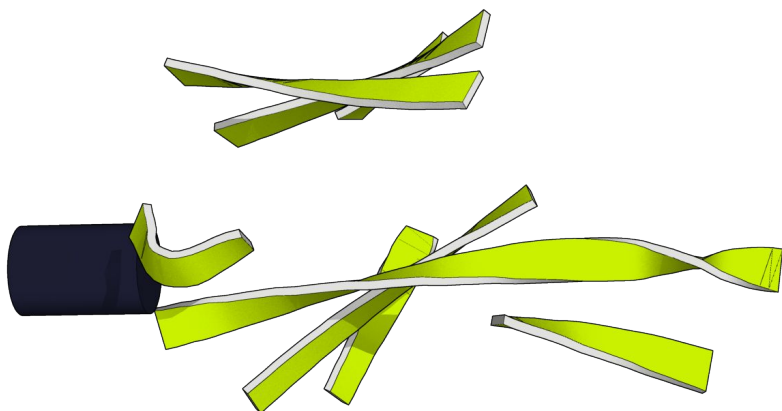


$\beta$ -сверток  
 *$\beta$ -roll*

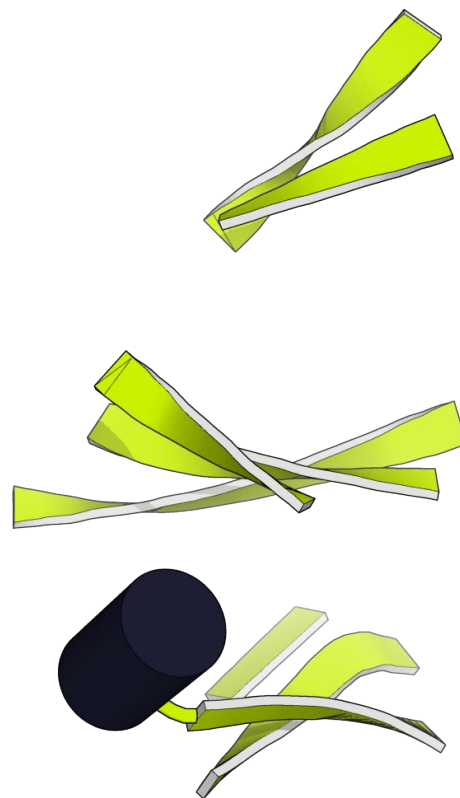


$\beta$ -бочка  
 *$\beta$ -barrel*

# Разнообразие фолдов: преимущественно $\beta$

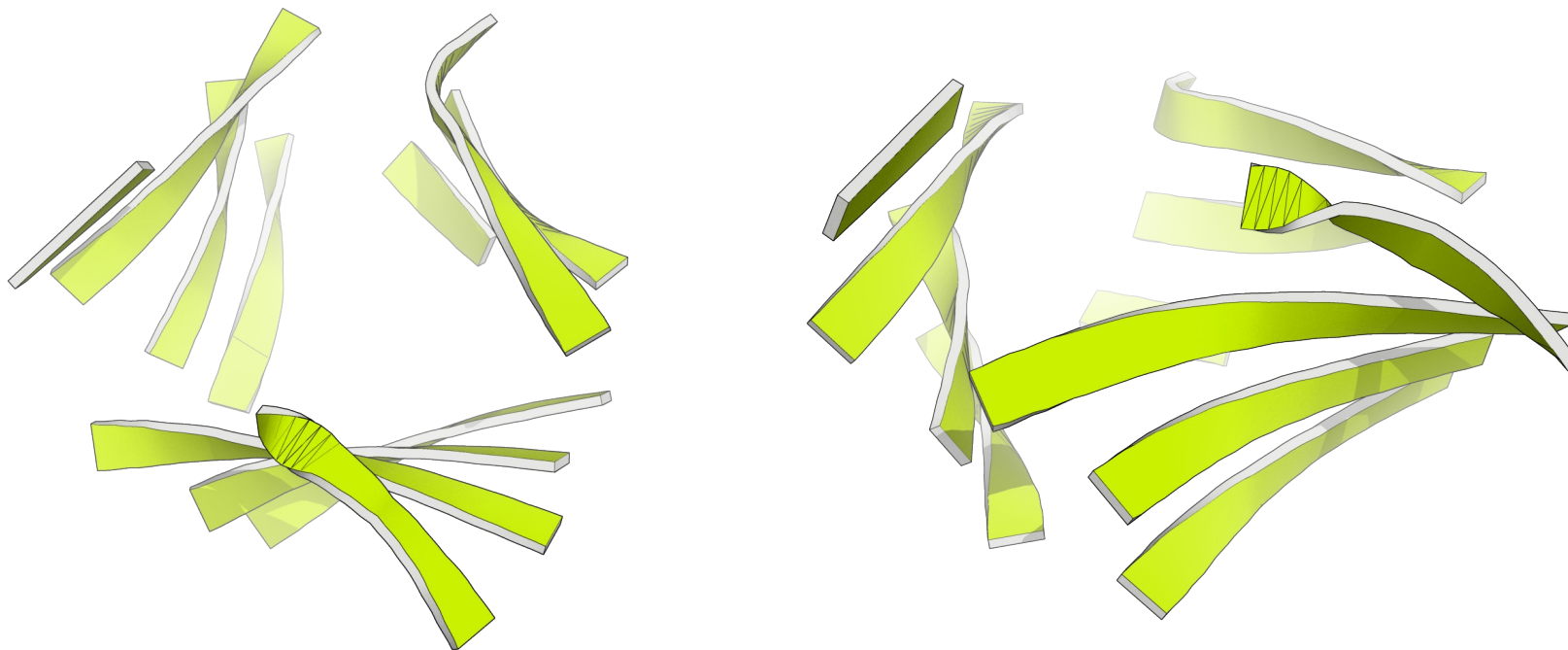


$\beta$ -сендвич  
 *$\beta$ -sandwich*



Тройной  $\beta$ -сендвич  
*3-layer sandwich*

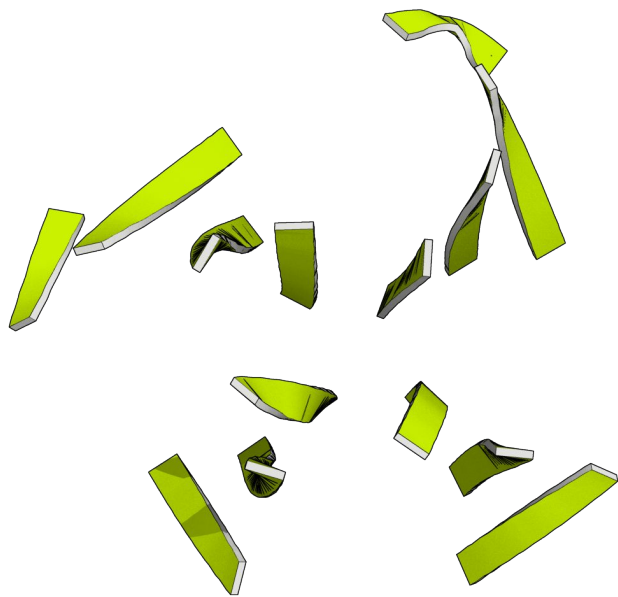
# Разнообразие фолдов: преимущественно $\beta$



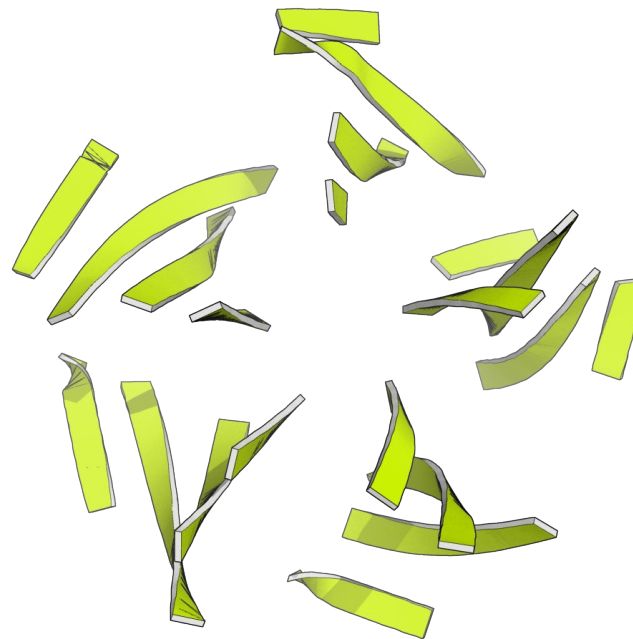
$\beta$ -призма  
 *$\beta$ -prism*



# Разнообразие фолдов: преимущественно $\beta$

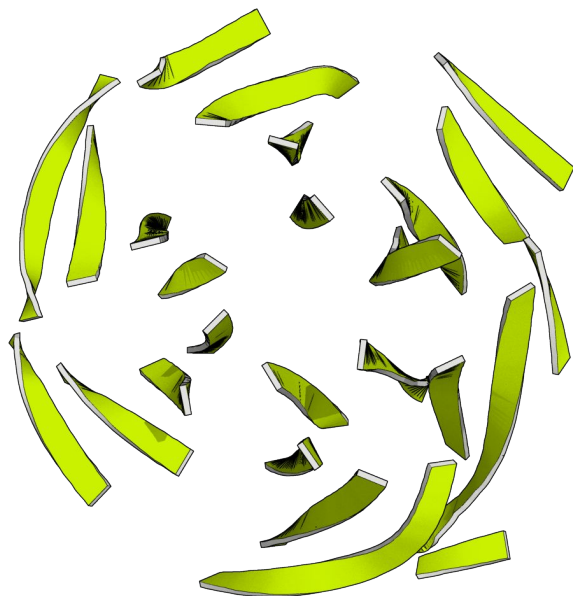


4-пропеллер  
*4-propeller*



5-пропеллер  
*5-propeller*

# Разнообразие фолдов: преимущественно $\beta$

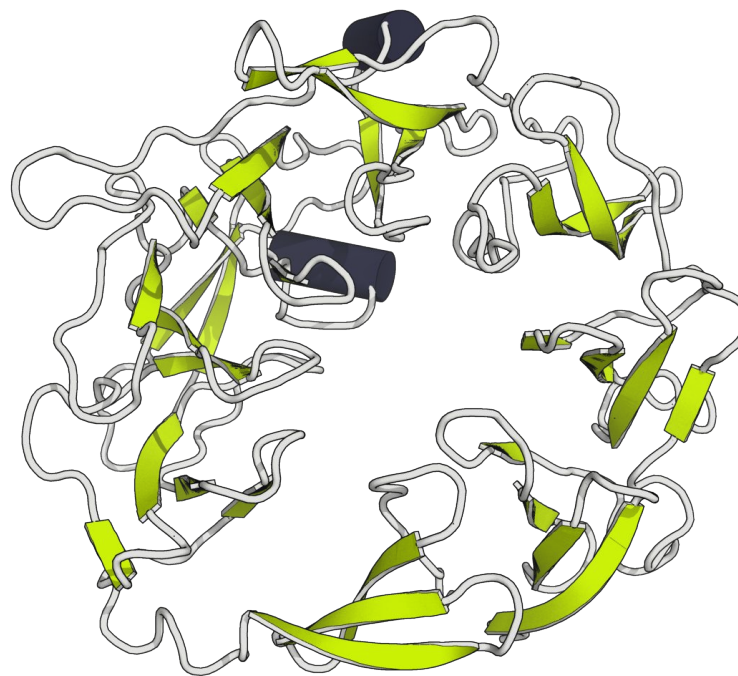
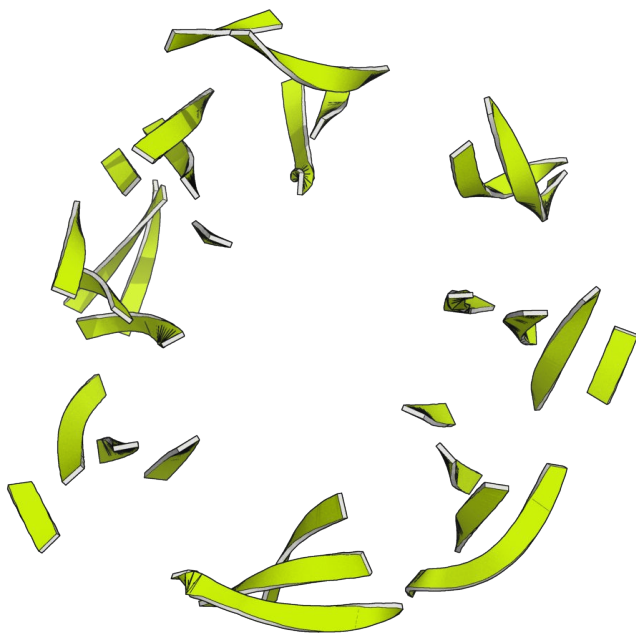


6-пропеллер  
*6-propeller*



7-пропеллер  
*7-propeller*

# Разнообразие фолдов: преимущественно $\beta$

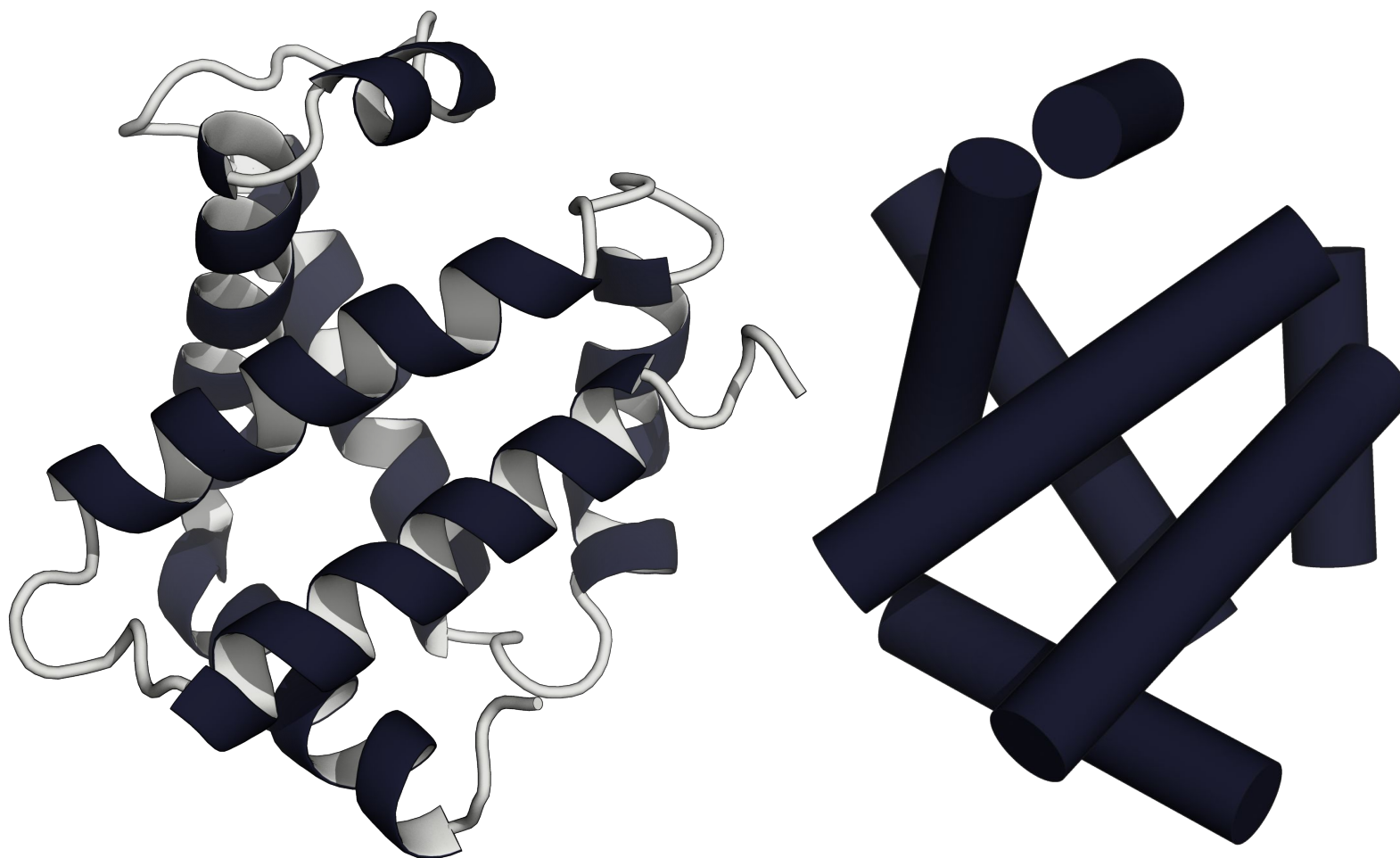


8-пропеллер  
*8-propeller*

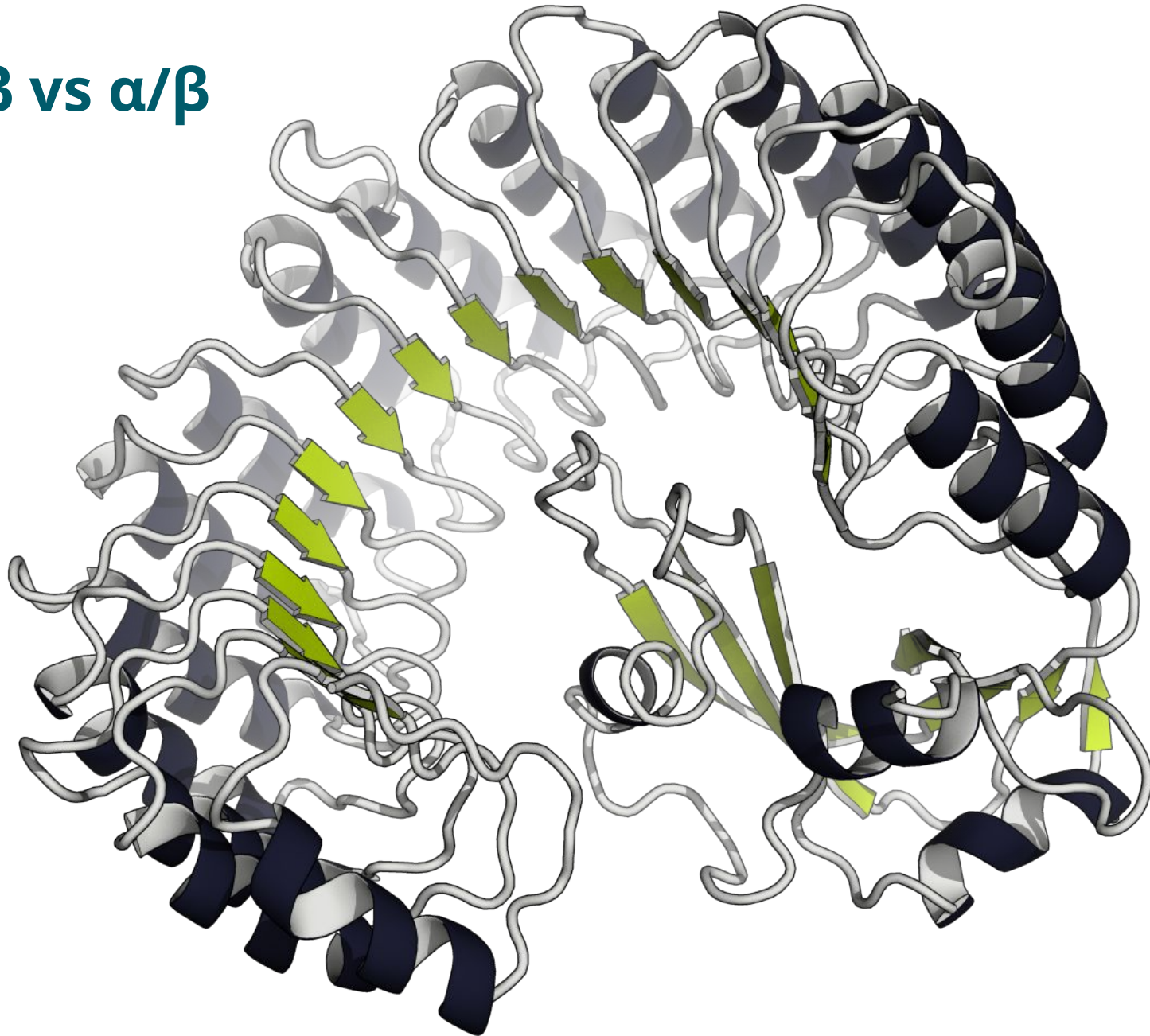
# $\alpha$ -архитектуры: $\alpha$ -пучки



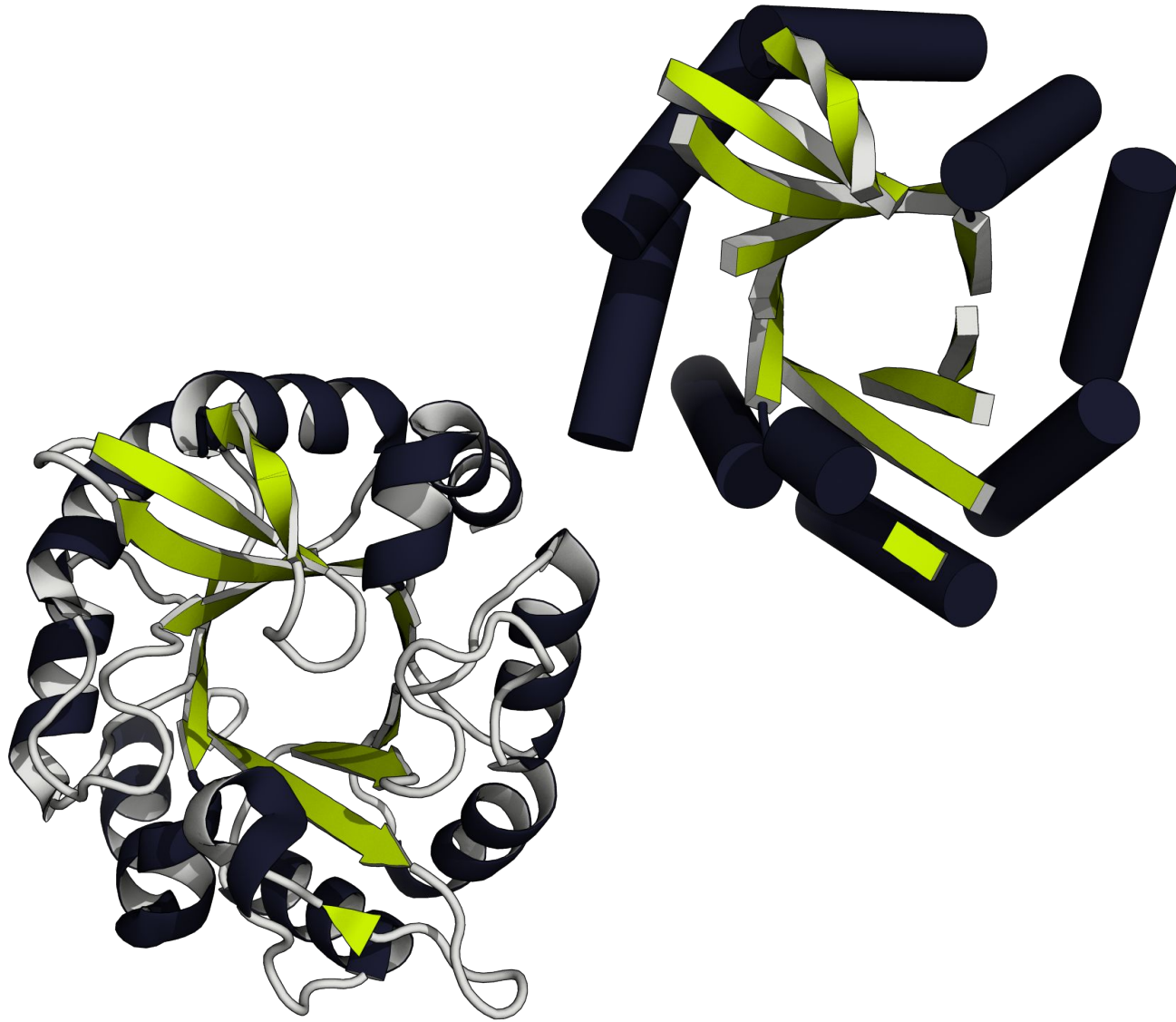
# $\alpha$ -архитектуры: $\alpha$ -многогранники



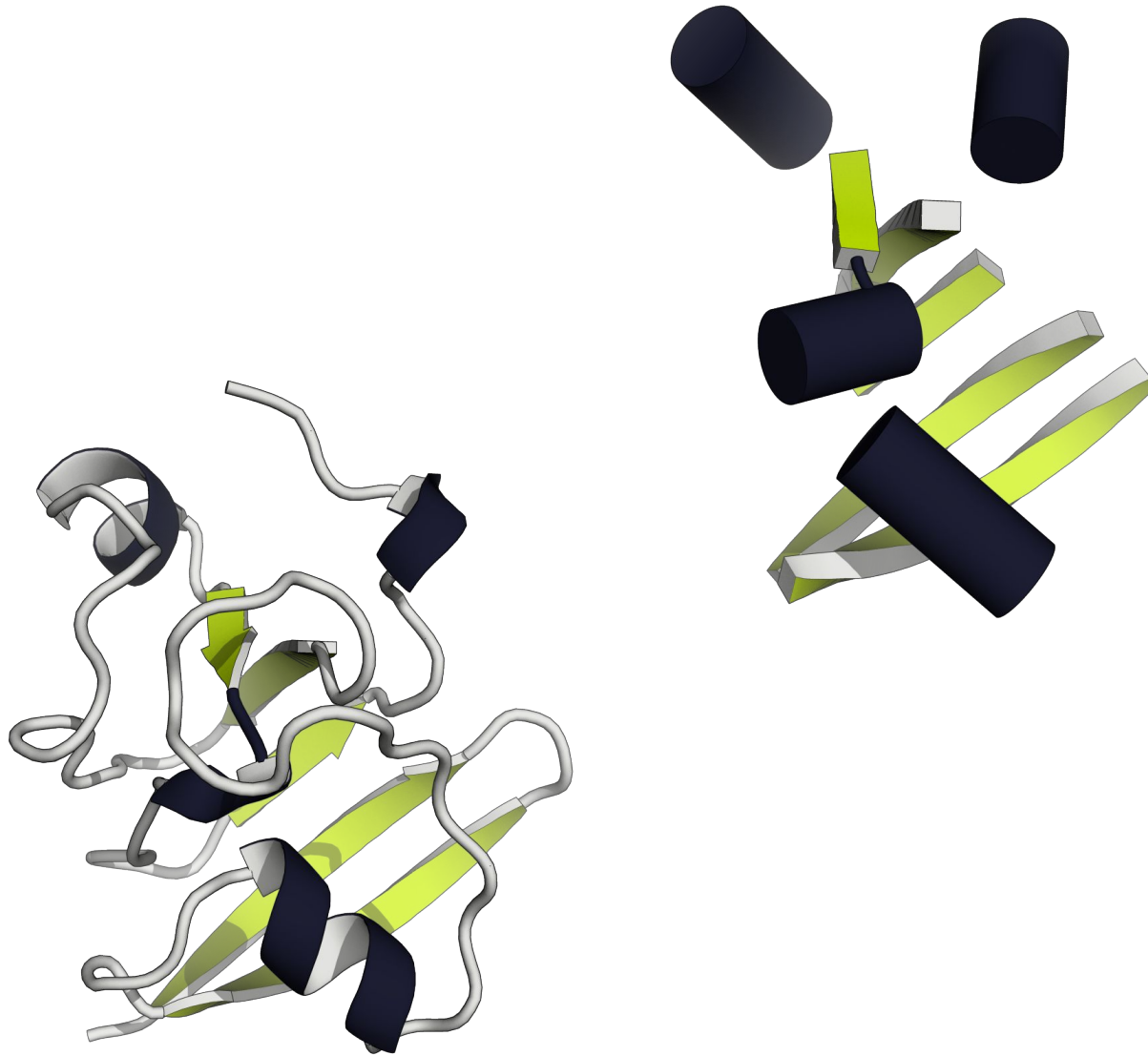
# $\alpha+\beta$ vs $\alpha/\beta$



$\alpha/\beta$

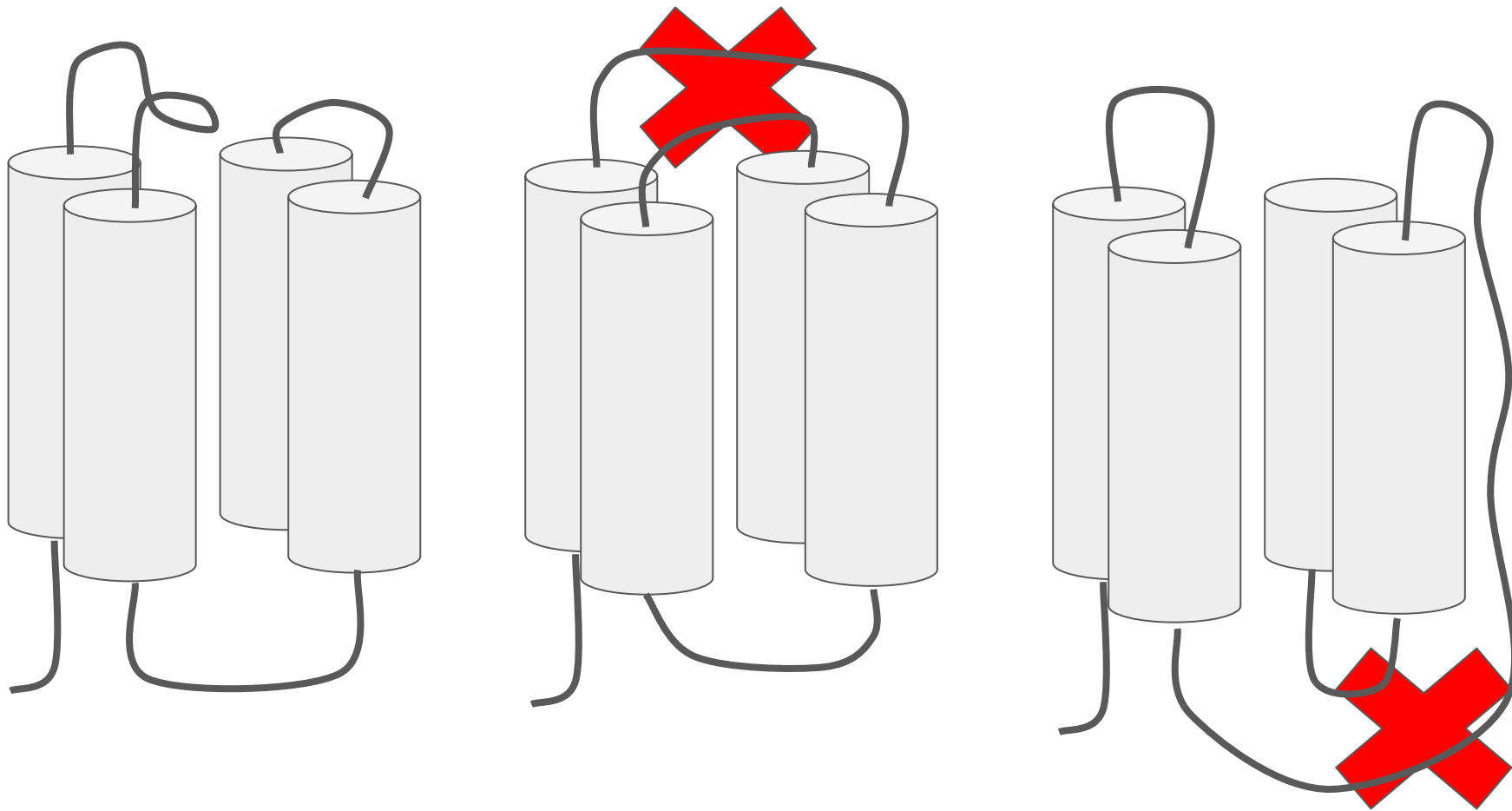


$\alpha+\beta$

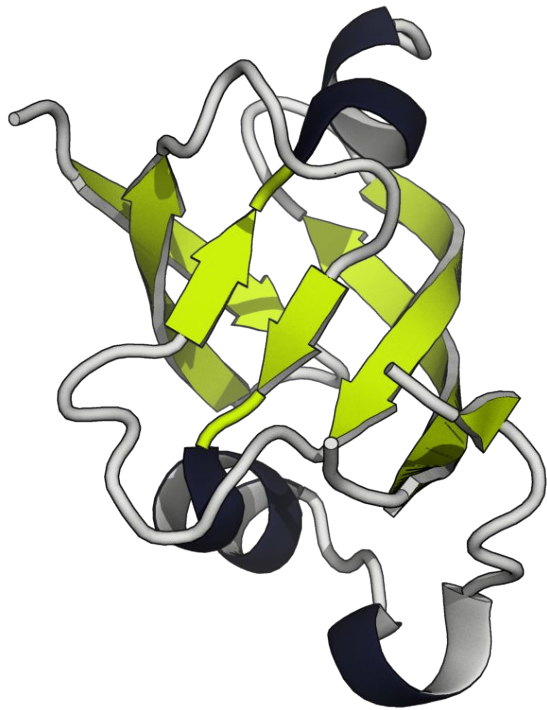




# Топология белковой цепи



# Топология белковой цепи

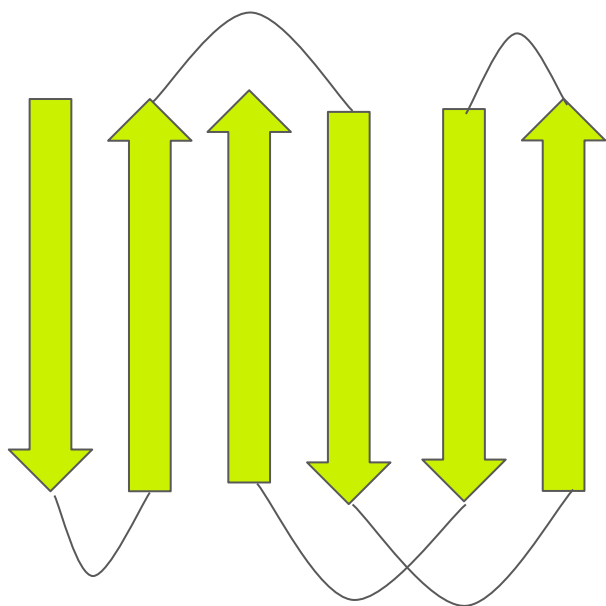


Кислая протеаза



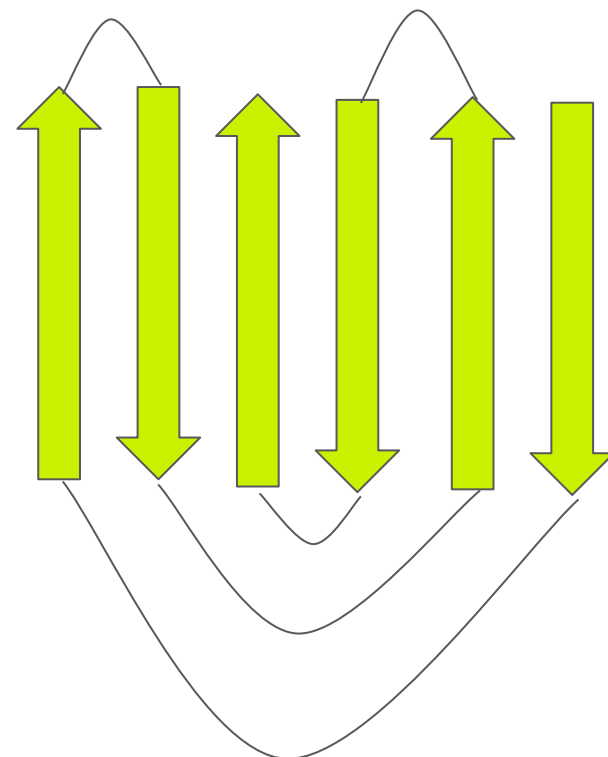
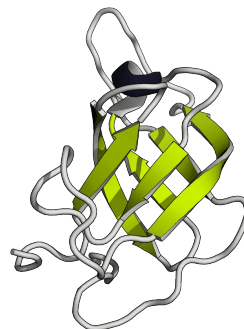
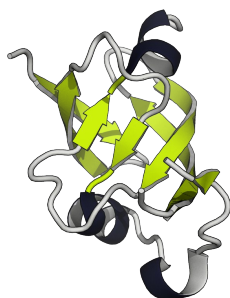
Сериновая протеаза

# Топология белковой цепи



126354

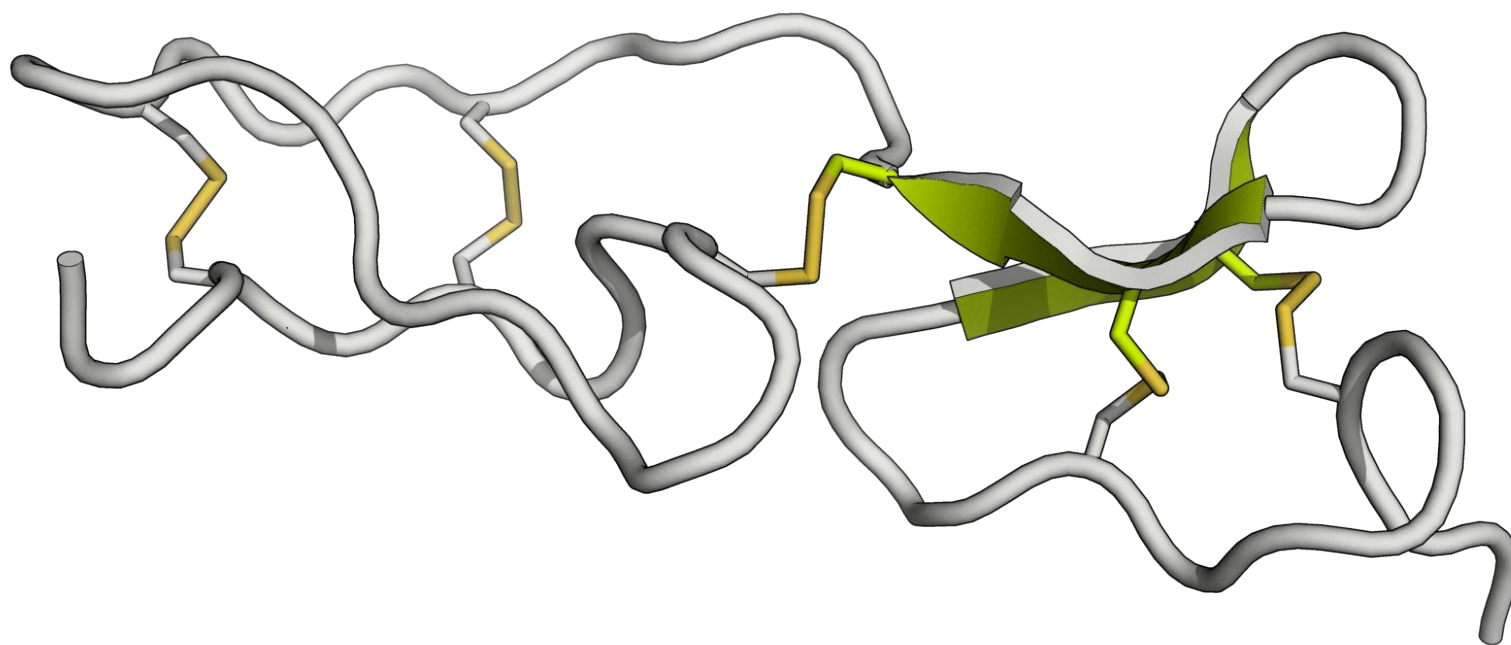
Кислая протеаза



236541

Сериновая протеаза

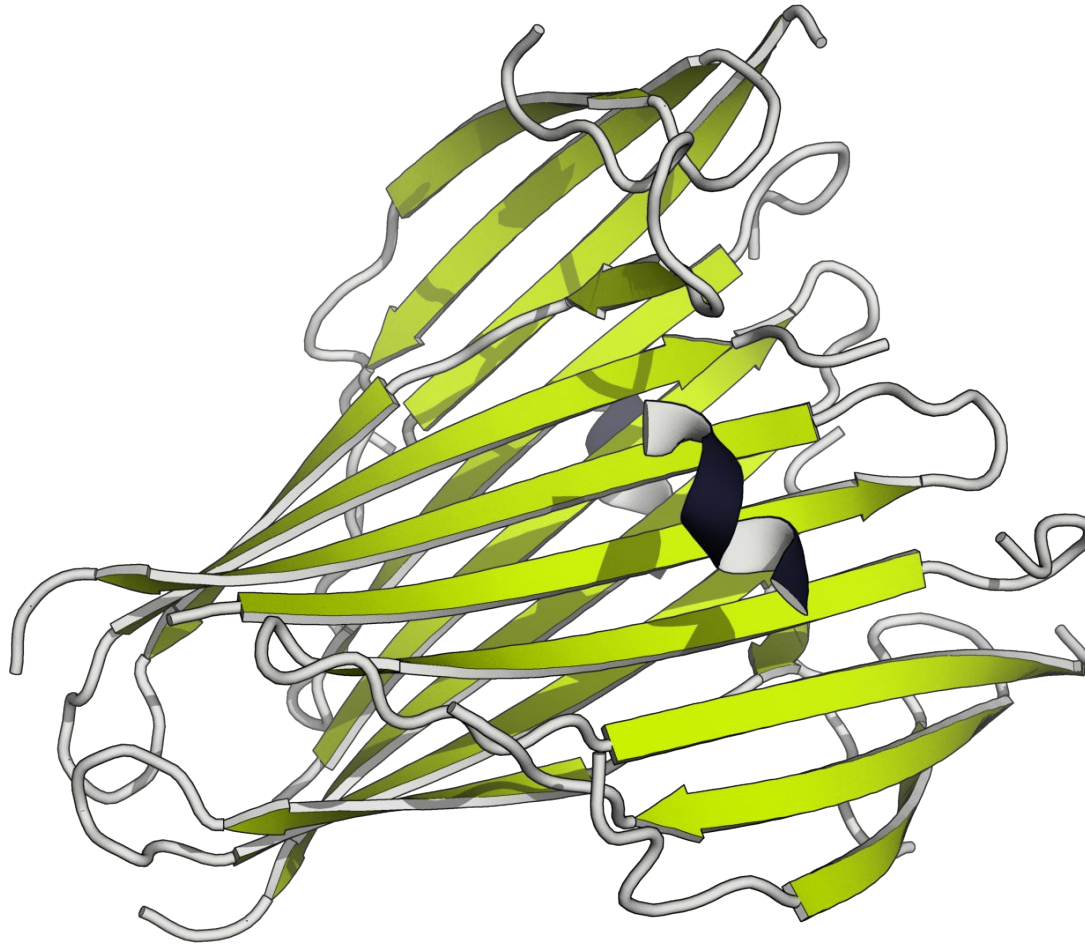
# Глобулярные белки с нестандартной укладкой



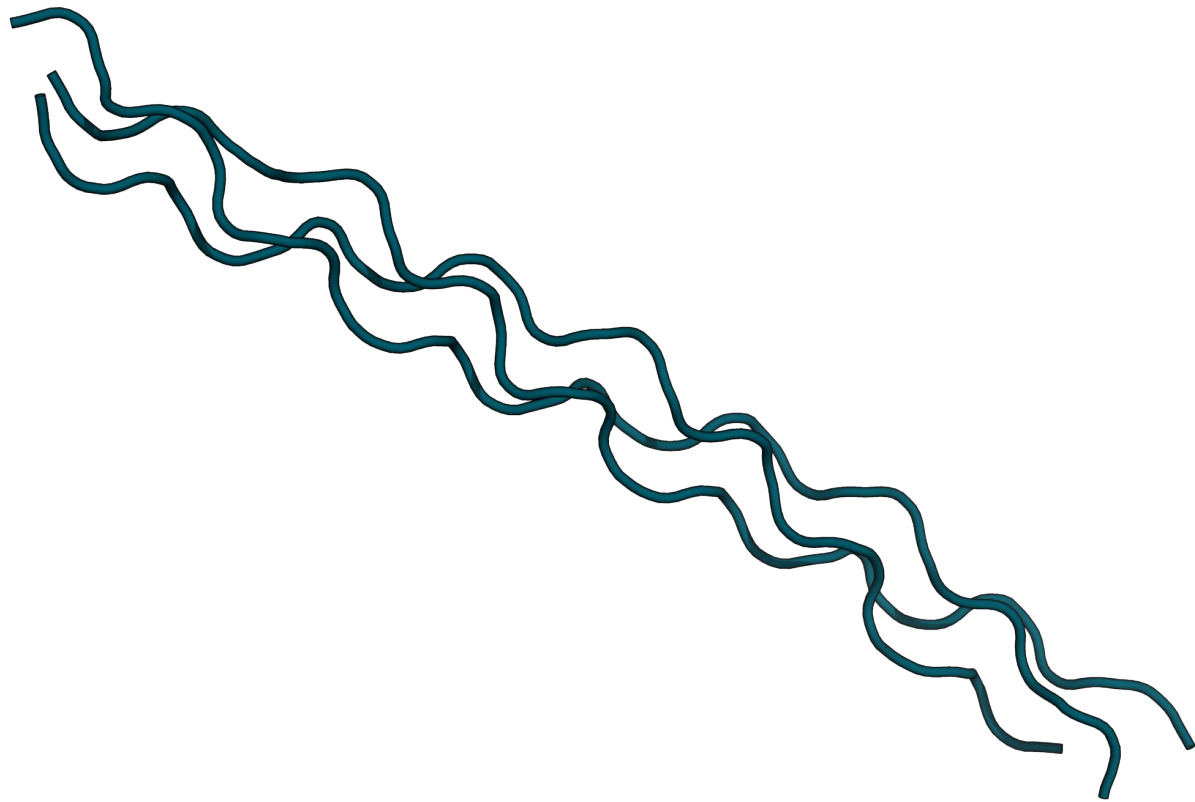
# Фибриллярные белки: $\alpha$ -белки



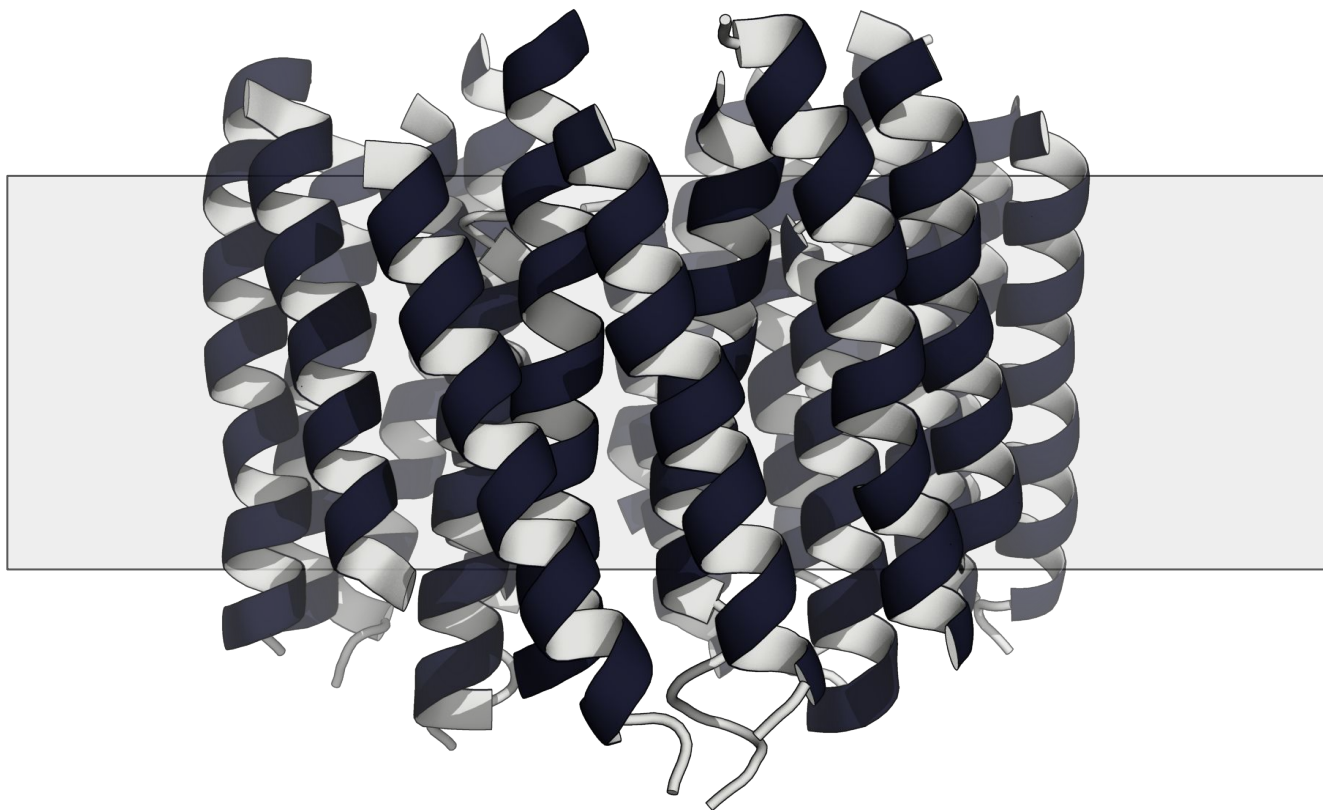
# Фибриллярные белки: $\beta$ -белки



# Фибриллярные белки: полипролиновые спирали

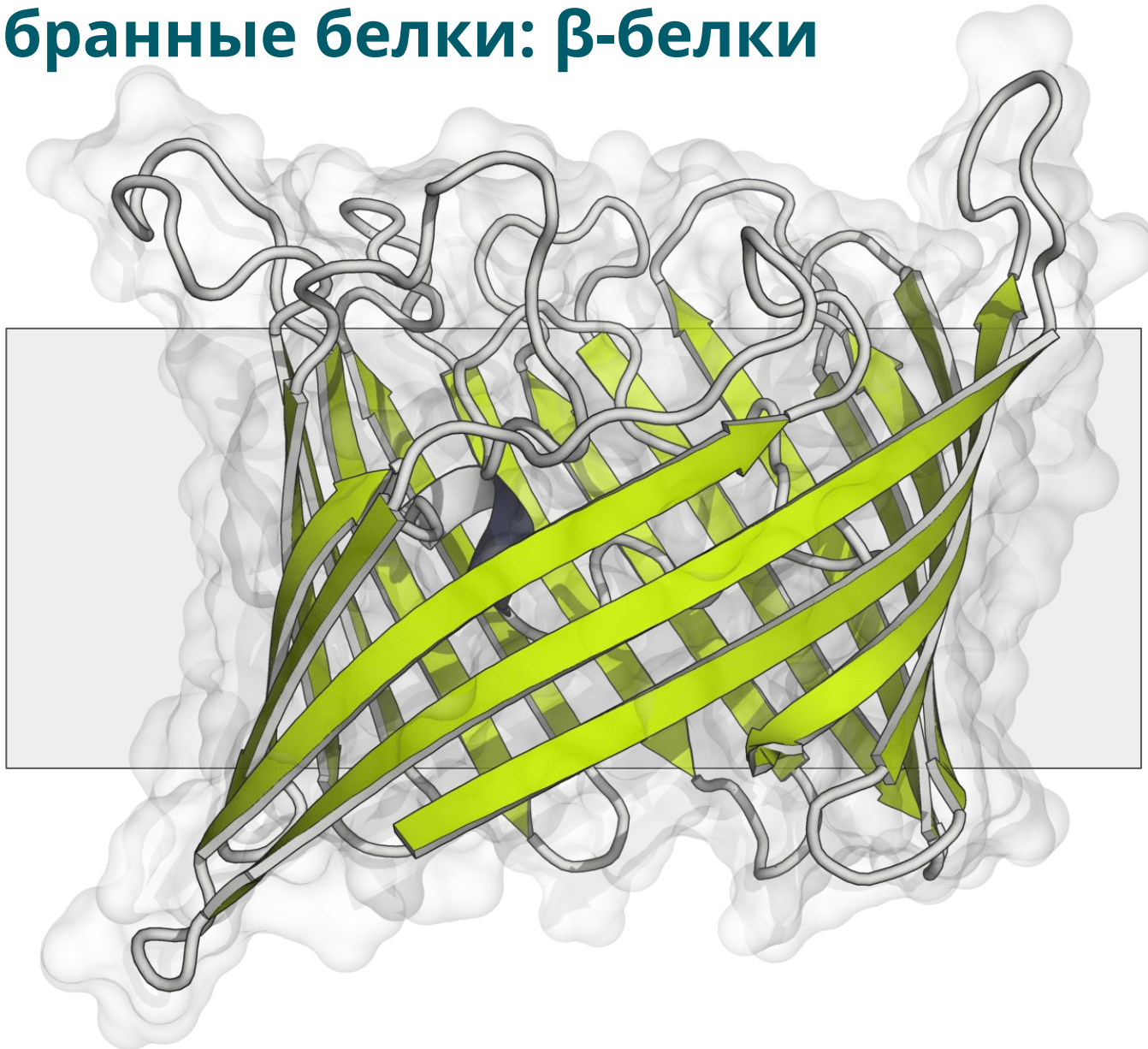


# Мембранные белки: $\alpha$ -белки





# Мембранные белки: $\beta$ -белки

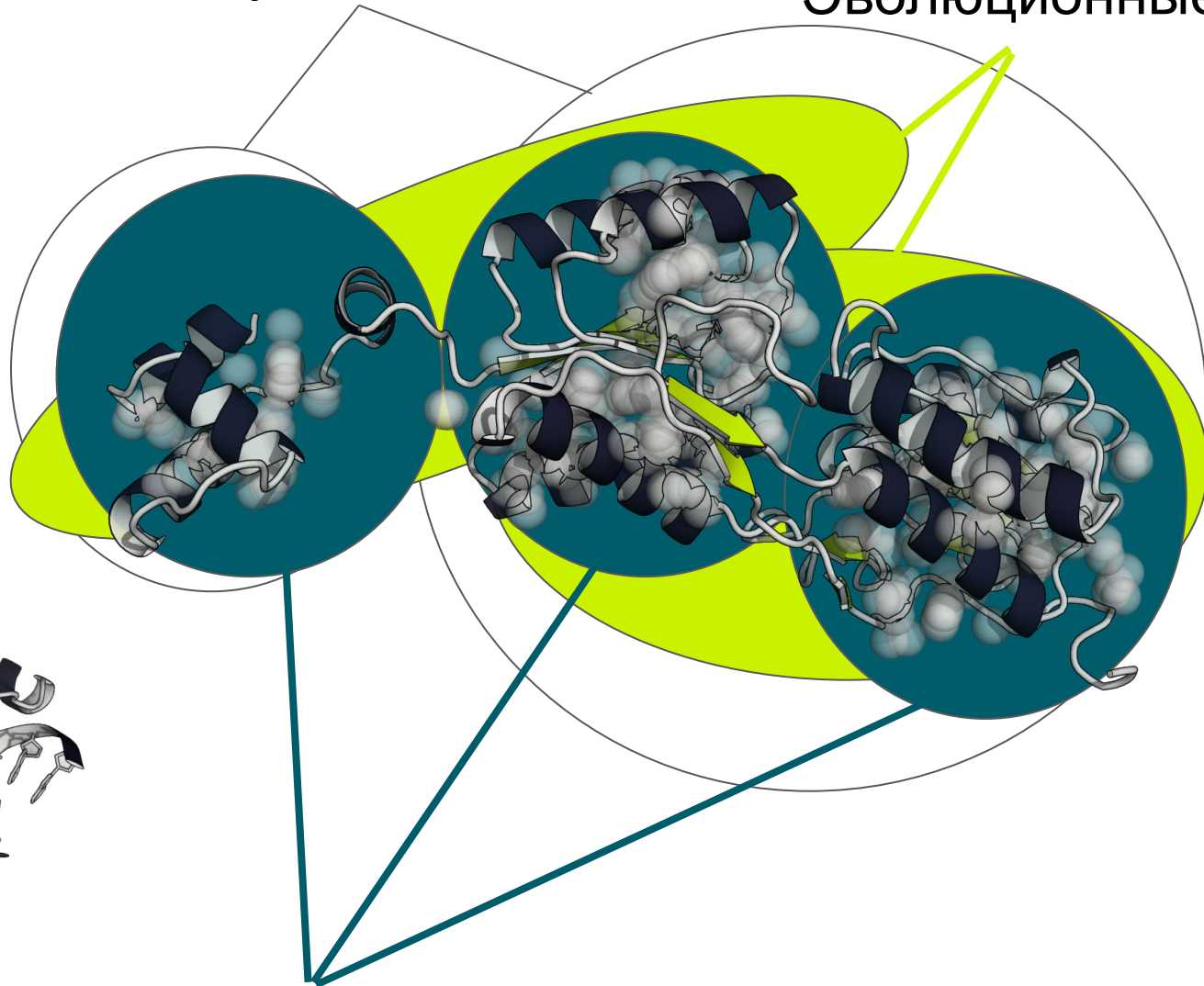


# Домены



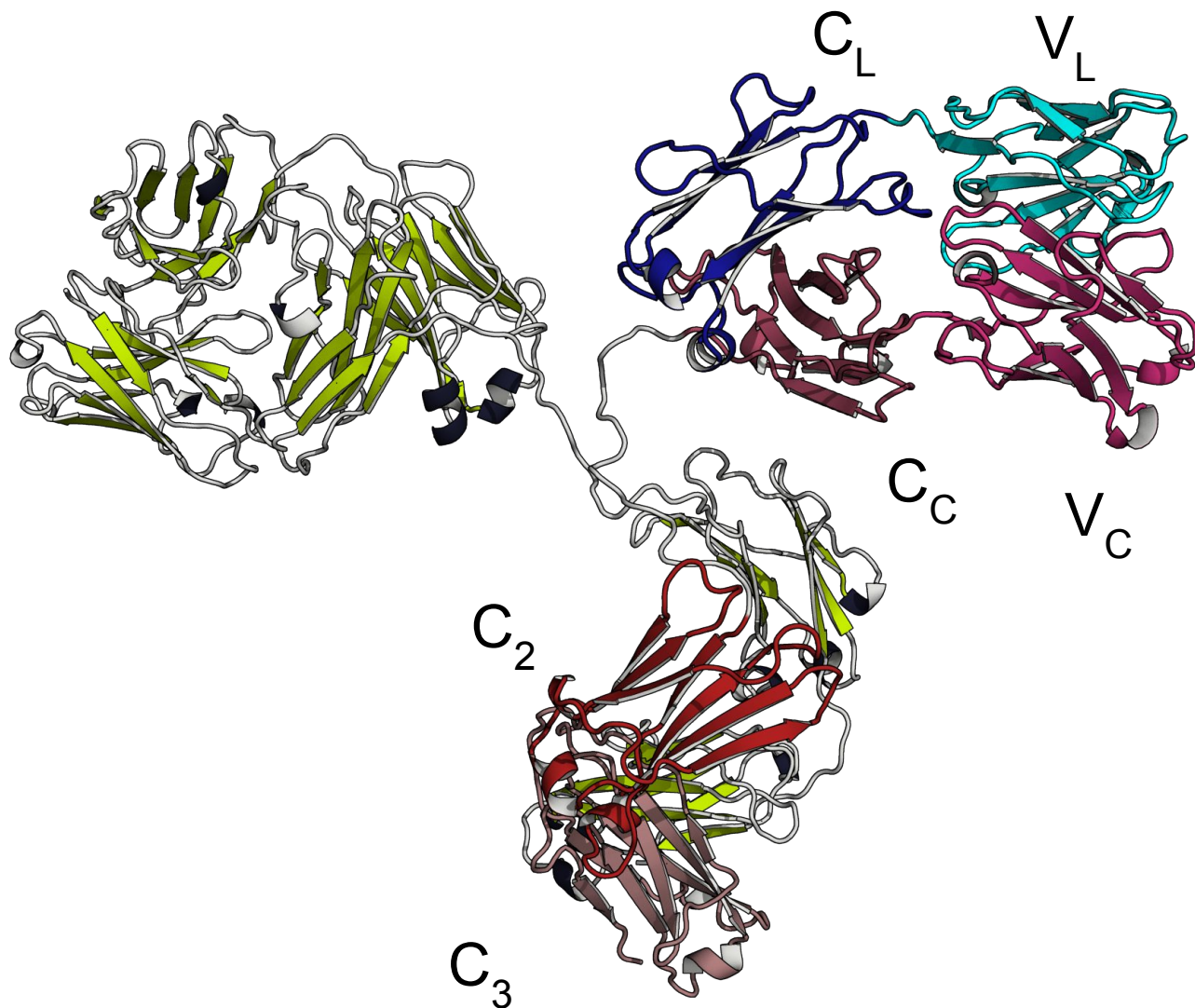
Функциональные

Эволюционные



Структурные

# Примеры многодоменных белков



Superfamily Comparison

Select a CATH node...

**A** 3-Layer(aba) Sandwich

3.40

CATH ID	3.40
Topologies	126
Superfamilies	577
Domains	86984
Example Domain	<a href="#">2hbaA00 [PDB]</a>



## Top of CATH Hierarchy (4 Classes)

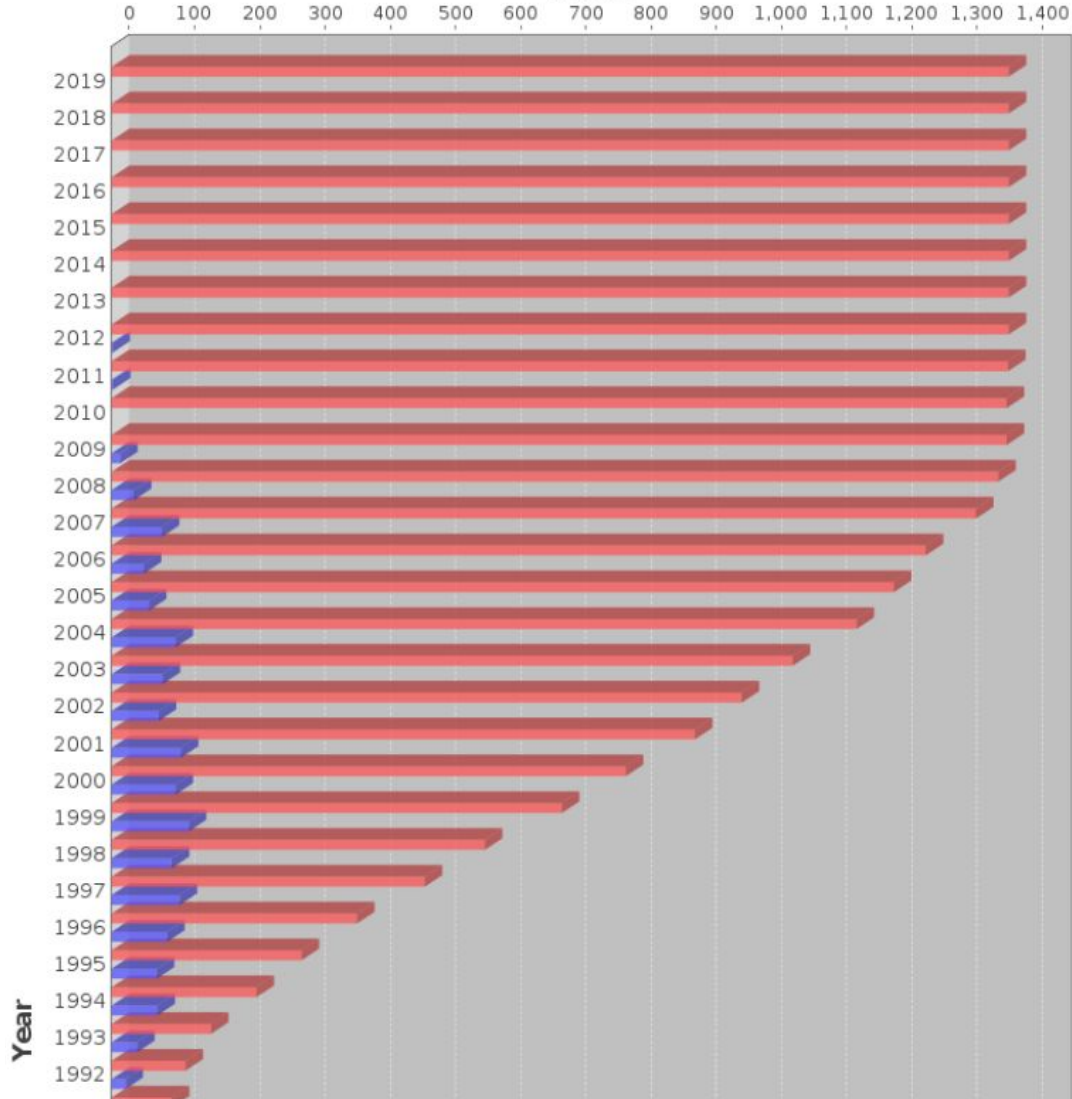
- ▶ **C** 1 Mainly Alpha *5 Architectures, 405 Folds, 2174 Superfamilies, 90302 Domains*
- ▶ **C** 2 Mainly Beta *21 Architectures, 244 Folds, 1395 Superfamilies, 110260 Domains*
- ▶ **C** 3 Alpha Beta *14 Architectures, 634 Folds, 2428 Superfamilies, 229776 Domains*
  - ▶ **A** 3.10 Roll *60 Folds, 248 Superfamilies, 16507 Domains*
  - ▶ **A** 3.15 Super Roll *3 Folds, 7 Superfamilies, 56 Domains*
  - ▶ **A** 3.20 Alpha-Beta Barrel *18 Folds, 63 Superfamilies, 16668 Domains*
  - ▶ **A** 3.30 2-Layer Sandwich *224 Folds, 1178 Superfamilies, 66583 Domains*
  - ▶ **A** 3.40 3-Layer(aba) Sandwich *126 Folds, 577 Superfamilies, 86984 Domains*
    - ▶ **T** 3.40.5 Ribosomal Protein L9; domain 1 *10 Superfamilies, 292 Domains*
    - ▶ **T** 3.40.10 DNA Methylphosphotriester Repair Domain *1 Superfamilies, 5 Domains*
    - ▶ **T** 3.40.20 Severin *2 Superfamilies, 290 Domains*
    - ▶ **T** 3.40.30 Glutaredoxin *12 Superfamilies, 4013 Domains*
    - ▶ **T** 3.40.33 Pathogenesis-related Protein p14a *1 Superfamilies, 62 Domains*
    - ▶ **T** 3.40.35 Fructose Permease *1 Superfamilies, 15 Domains*
    - ▶ **T** 3.40.47 Peroxisomal Thiolase; Chain A, domain 1 *2 Superfamilies, 1877 Domains*
    - ▶ **T** 3.40.50 Rossmann fold *298 Superfamilies, 52880 Domains*
    - ▶ **T** 3.40.80 Lysozyme-like *1 Superfamilies, 257 Domains*
    - ▶ **T** 3.40.91 Restriction Endonuclease *9 Superfamilies, 265 Domains*
    - ▶ **T** 3.40.109 NADH Oxidase *3 Superfamilies, 288 Domains*
    - ▶ **T** 3.40.120 Alpha-D-Glucose-1,6-Bisphosphate; Chain A, domain 3 *1 Superfamilies, 221 Domains*
    - ▶ **T** 3.40.140 Cytidine Deaminase; domain 2 *7 Superfamilies, 420 Domains*
    - ▶ **T** 3.40.190 D-Maltodextrin-Binding Protein; domain 2 *18 Superfamilies, 5295 Domains*
    - ▶ **T** 3.40.198 Delta-endotoxin CytB *1 Superfamilies, 16 Domains*
    - ▶ **T** 3.40.210 PvuII Endonuclease; Chain A *3 Superfamilies, 35 Domains*
    - ▶ **T** 3.40.220 Leucine Aminopeptidase, subunit E; domain 1 *3 Superfamilies, 484 Domains*
    - ▶ **T** 3.40.225 L-fucose-1-phosphate Aldolase *1 Superfamilies, 110 Domains*
    - ▶ **T** 3.40.228 Dimethylsulfoxide Reductase; domain 2 *1 Superfamilies, 104 Domains*



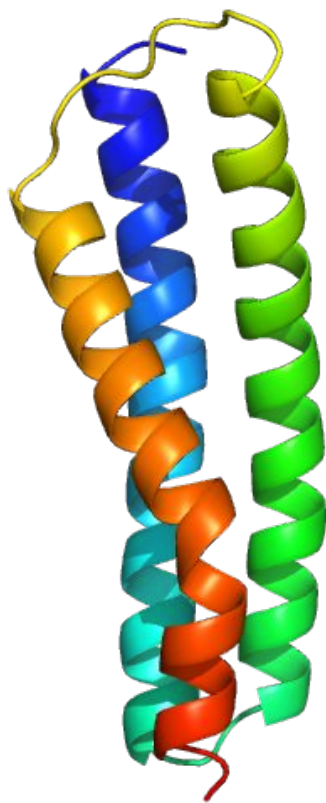
# Growth Of Unique Folds (Topologies) Per Year As Defined By CATH (v4.0.0)

number of folds can be viewed by hovering mouse over the bar

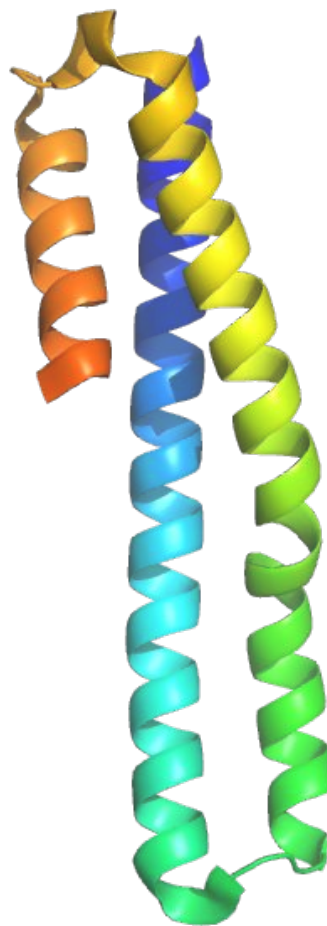
Number



# Разные белки - одинаковые укладки

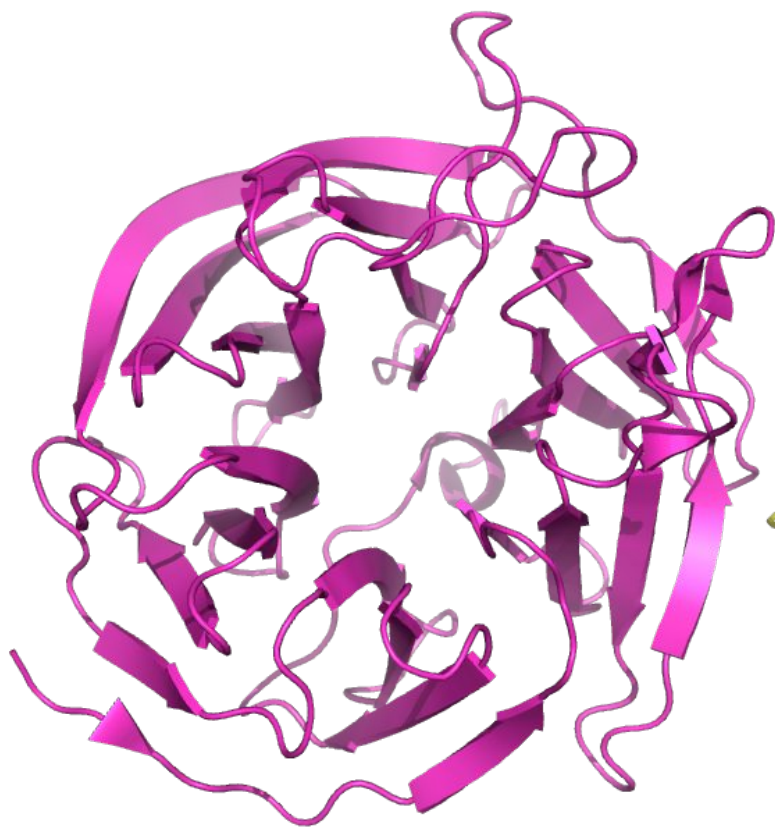


Лактаза бактерии

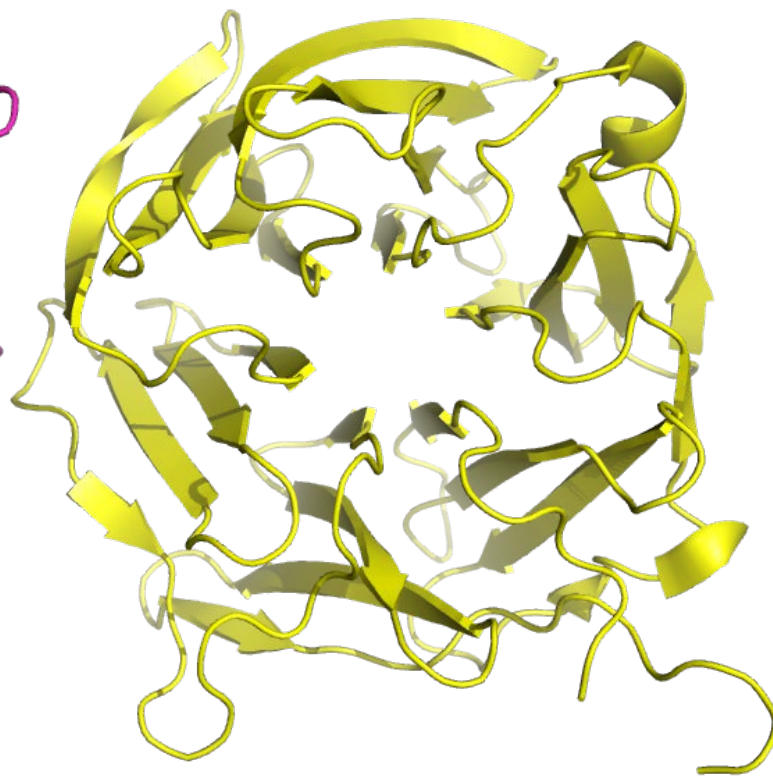


Шаперон человека

# Разные белки - одинаковые укладки



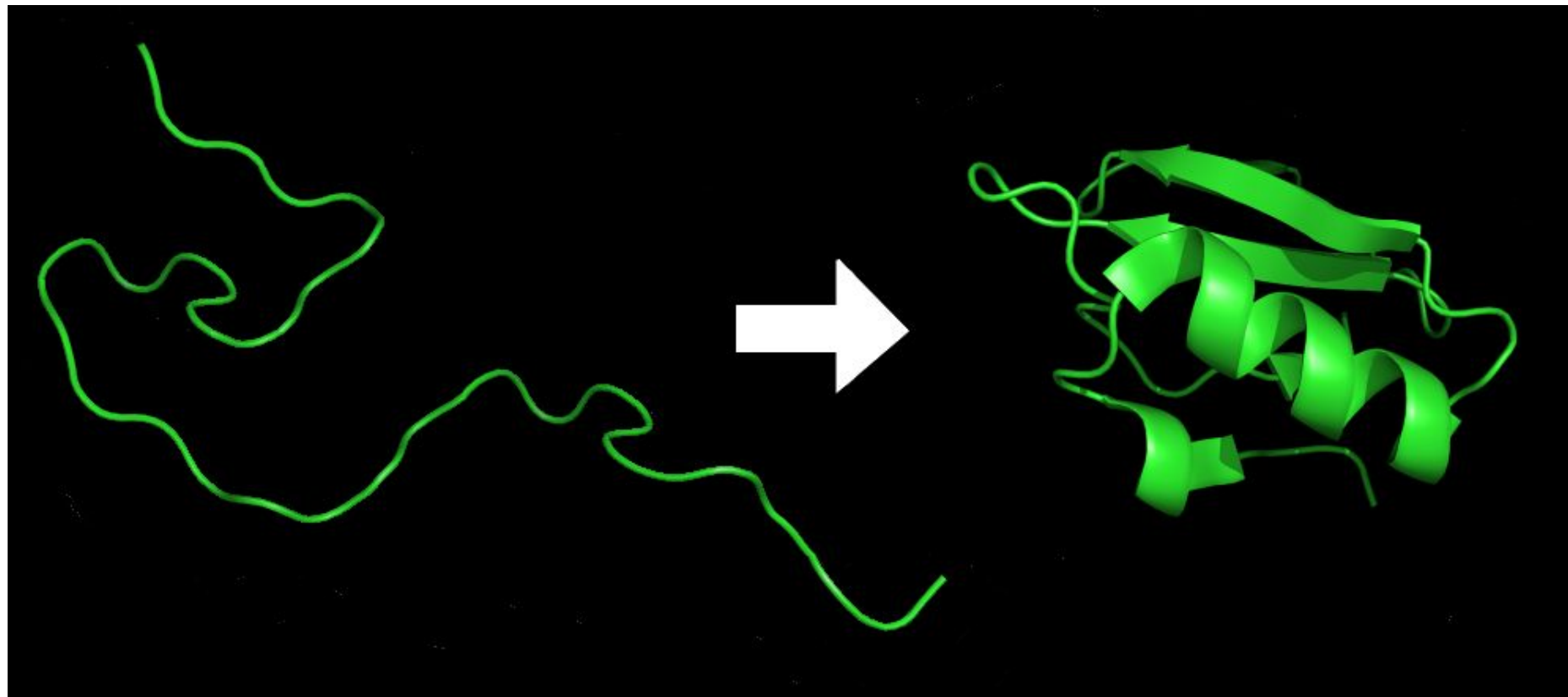
DFP-аза каракатицы



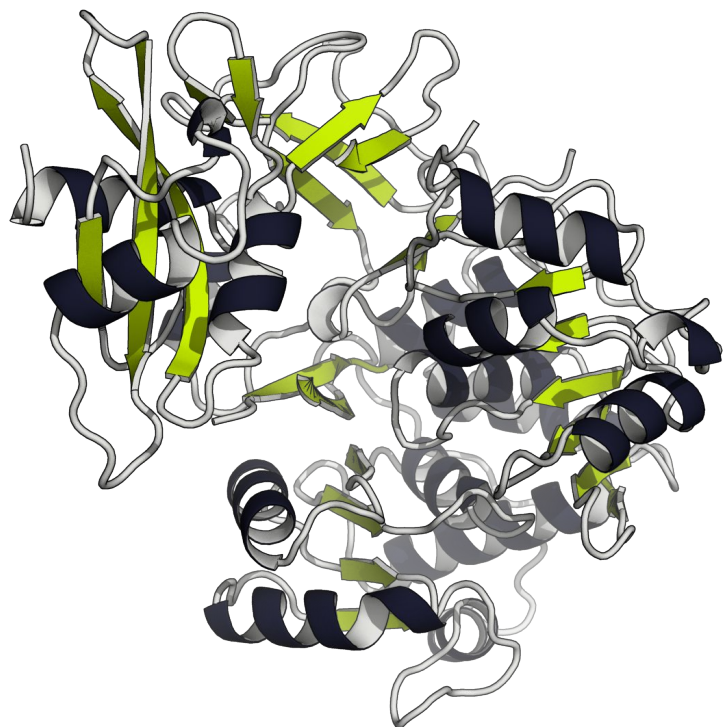
Лектин гриба



# Фолдинг - процесс принятия укладки

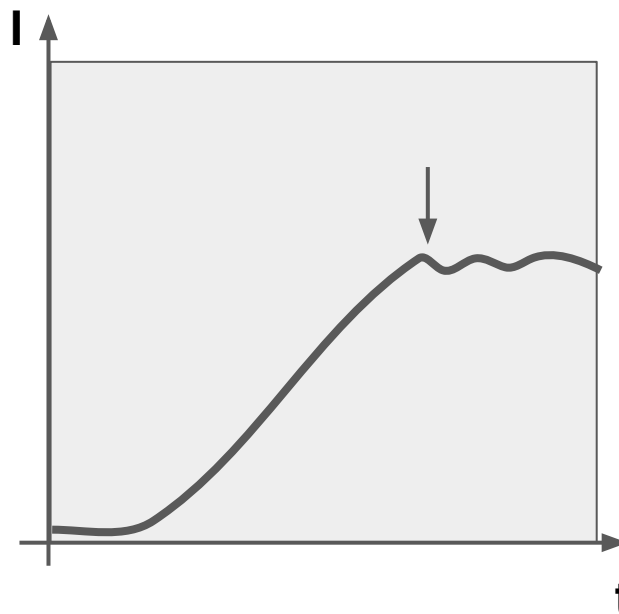


# Как быстро сворачиваются белки?

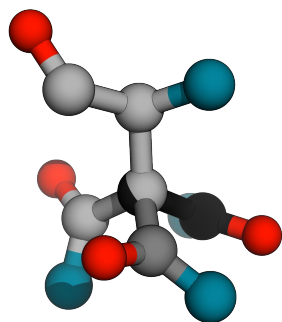


Люцифераза

люциферин + O<sub>2</sub> = окслюциферин + свет

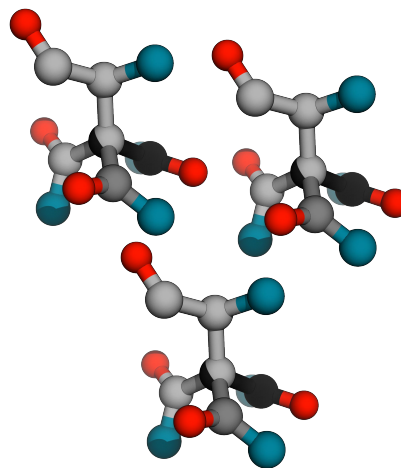
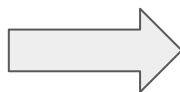


# Парадокс Левинтала



1 АК ~ 10 конформаций

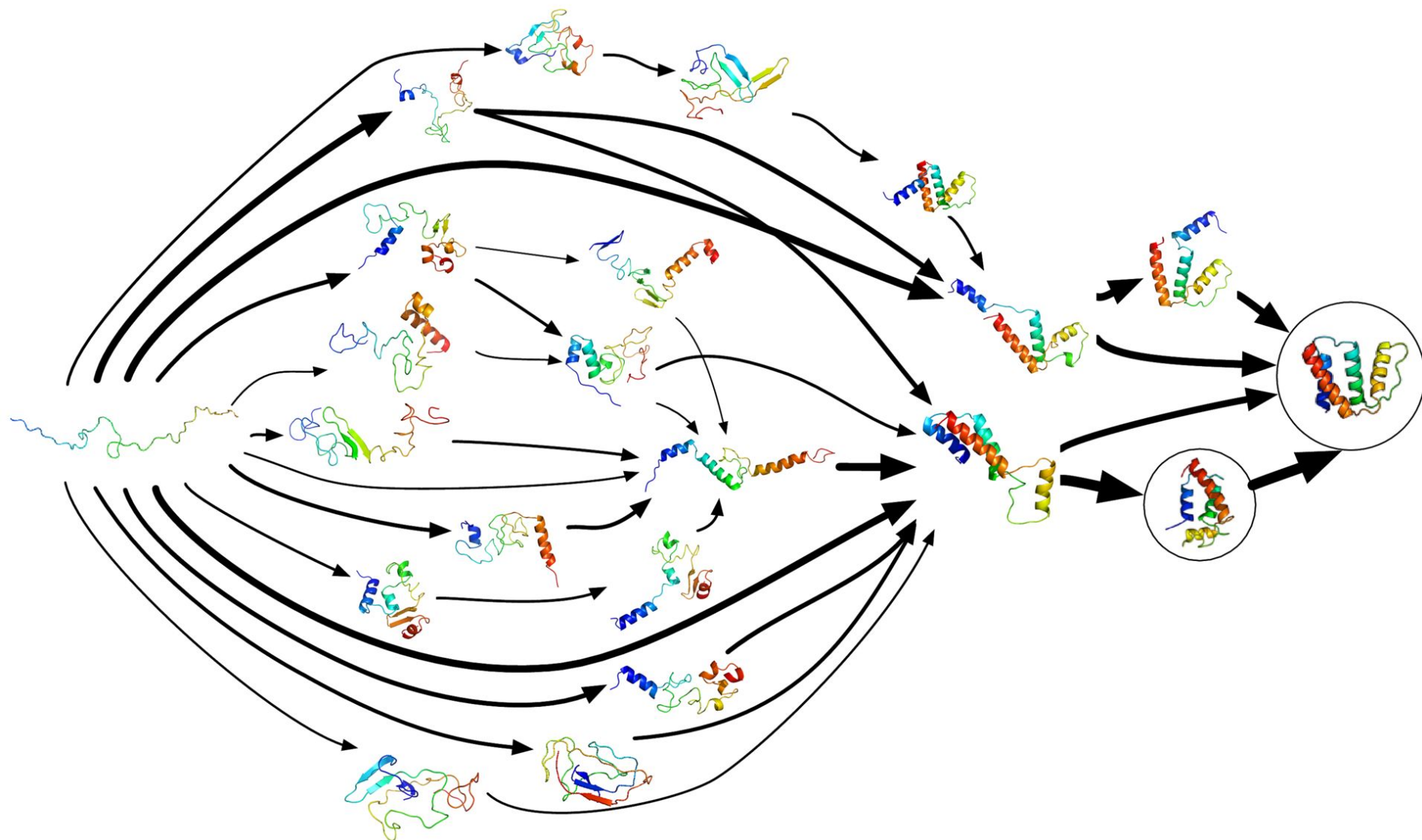
1 конформационный  
переход ~  $10^{-13}$  с



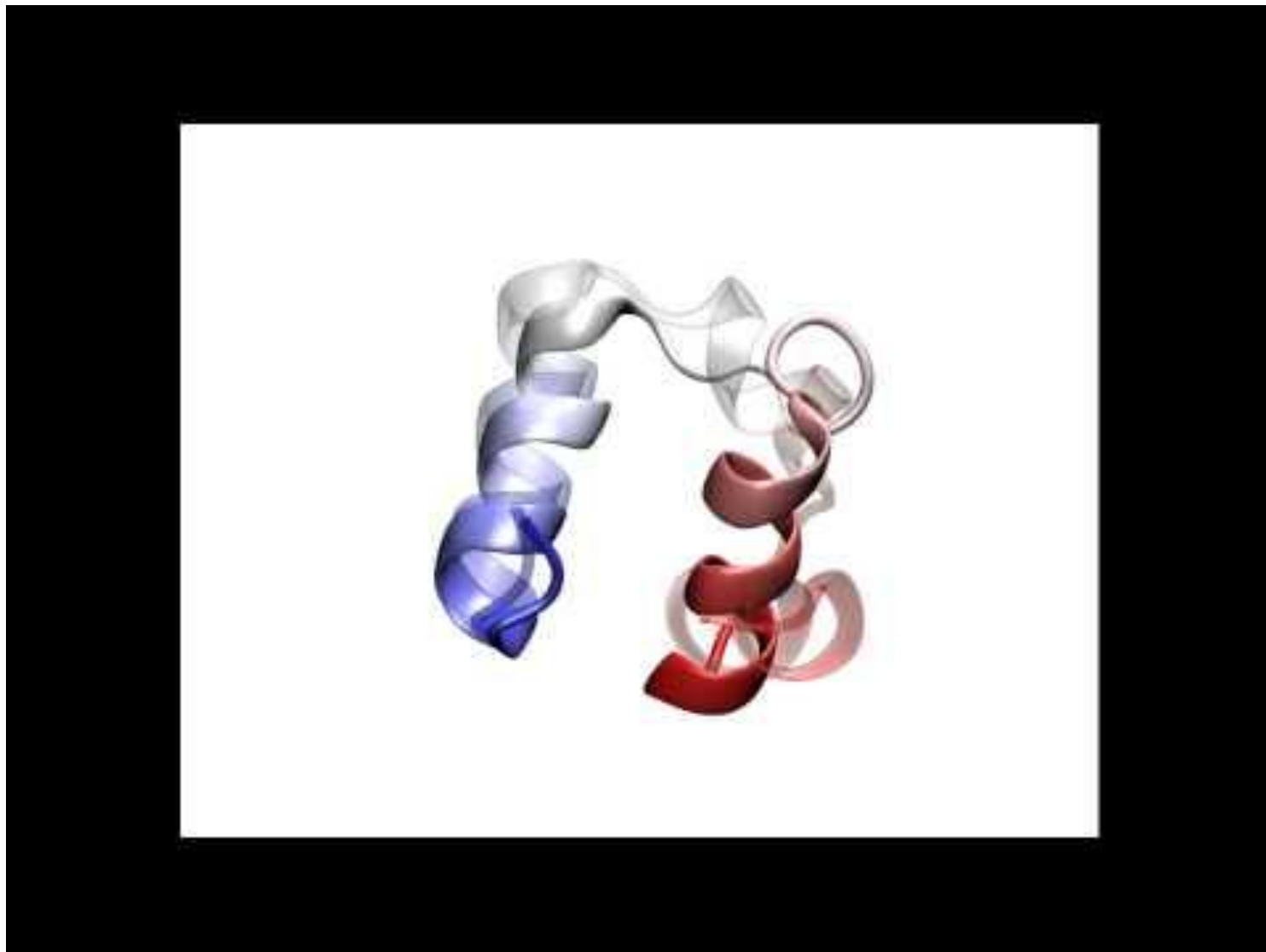
100 АК ~  $10^{100}$  конформаций

Поиск лучшей конформации  
~  $10^{80}$  лет

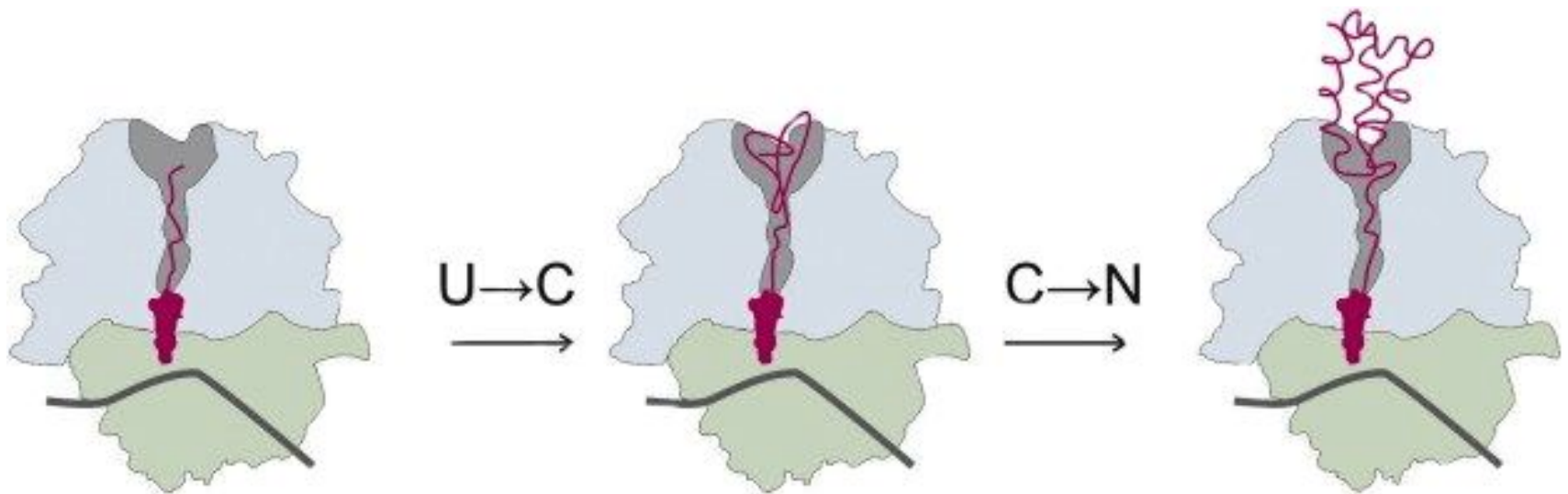
# Путь фолдинга



# Путь фолдинга



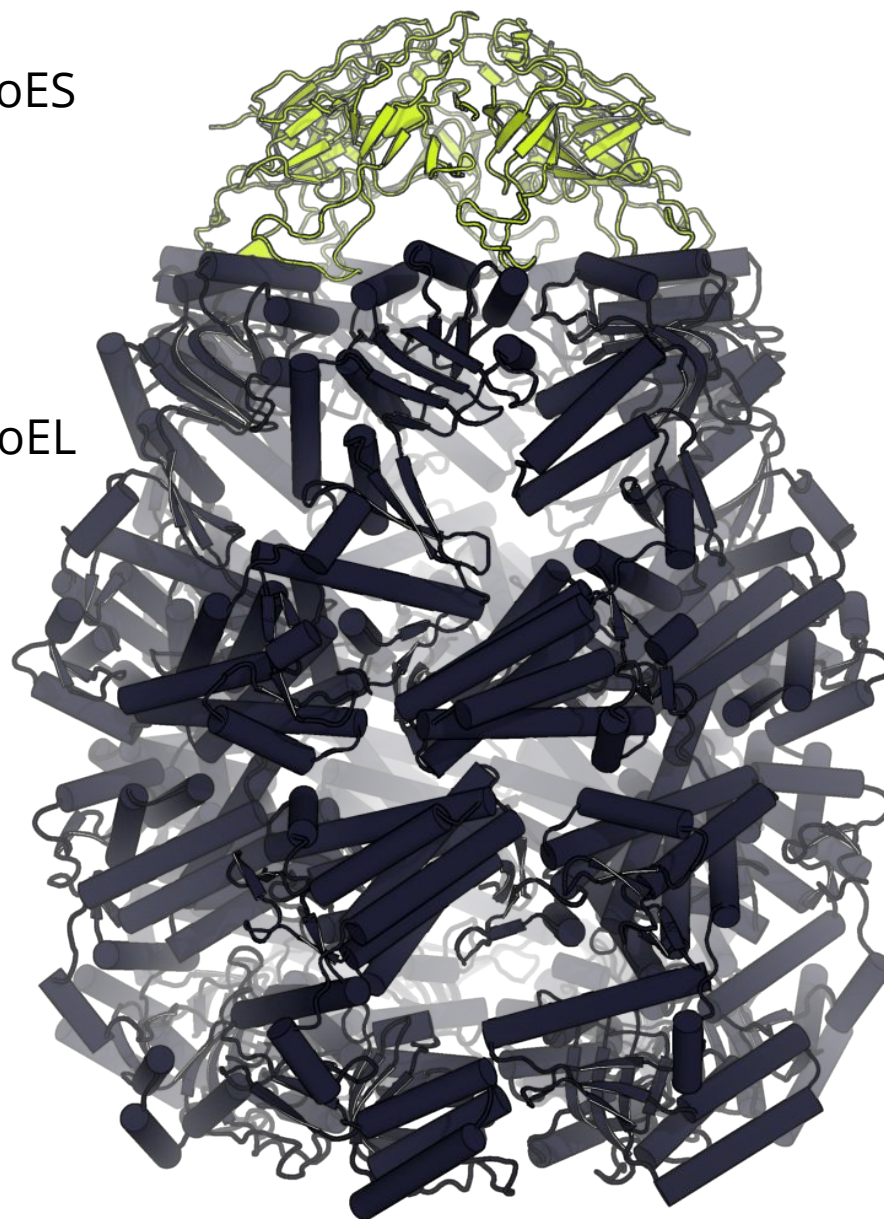
# Трансляция и фолдинг



# Шапероны

GroES

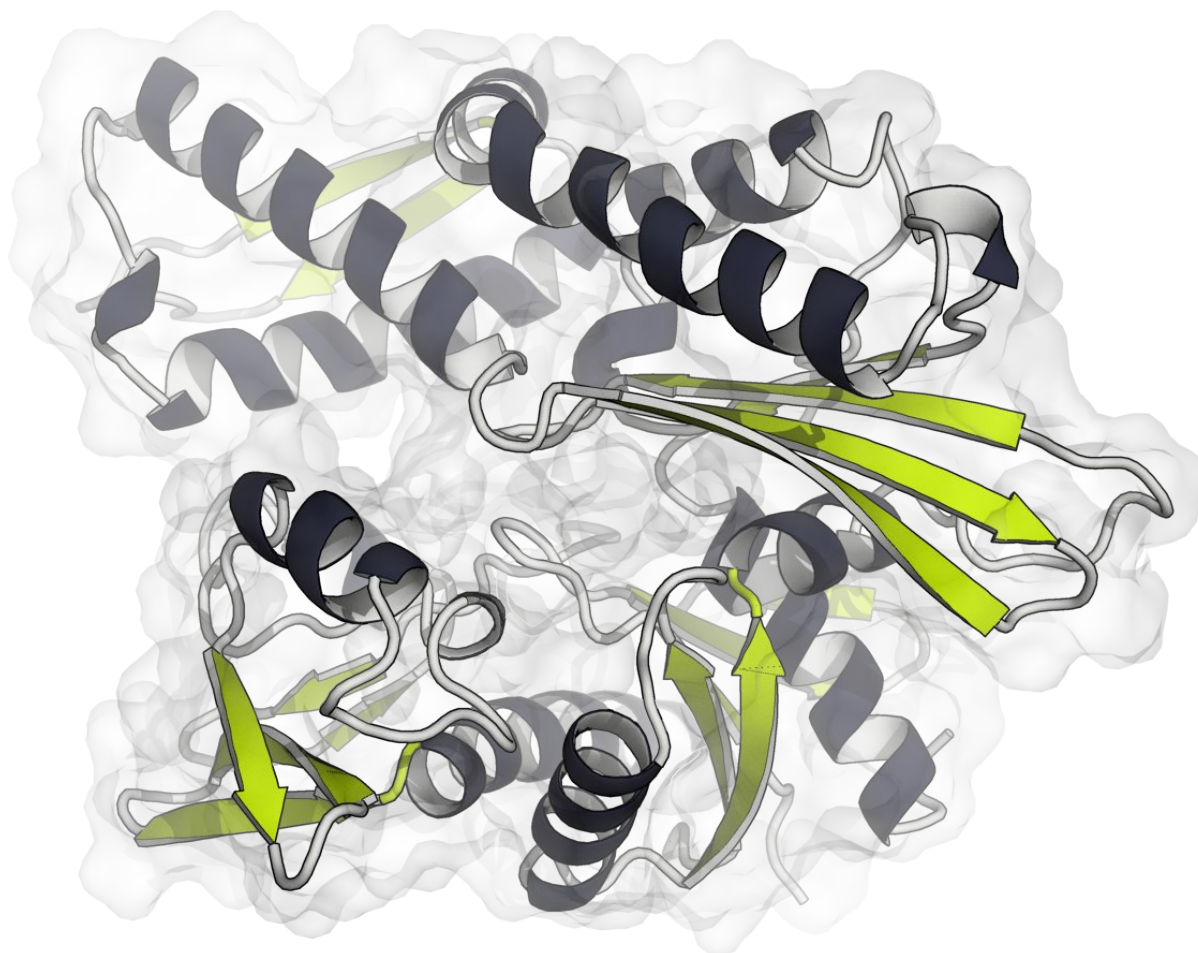
GroEL



## GroEL/GroES

Создание изолированной гидрофобной среды для полного фолдинга целой молекулы белка.

# Шапероны

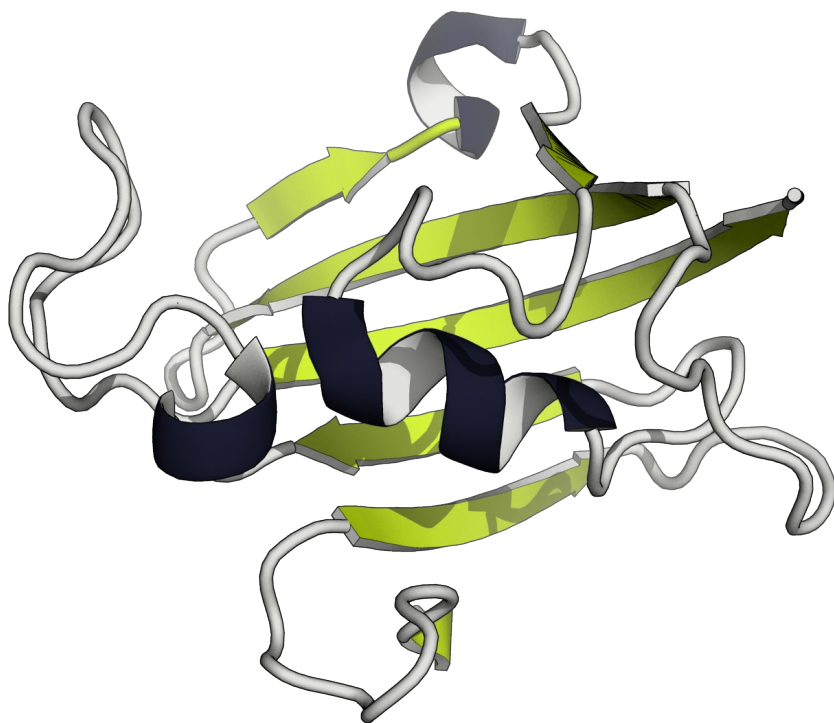


## Hsp70

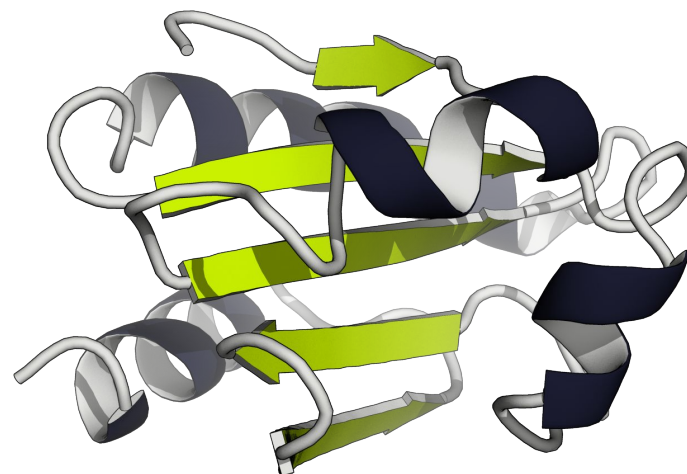
Защита  
свежесинтезированных  
петель белка от агрегации  
до окончания его синтеза.



# Другие ферменты

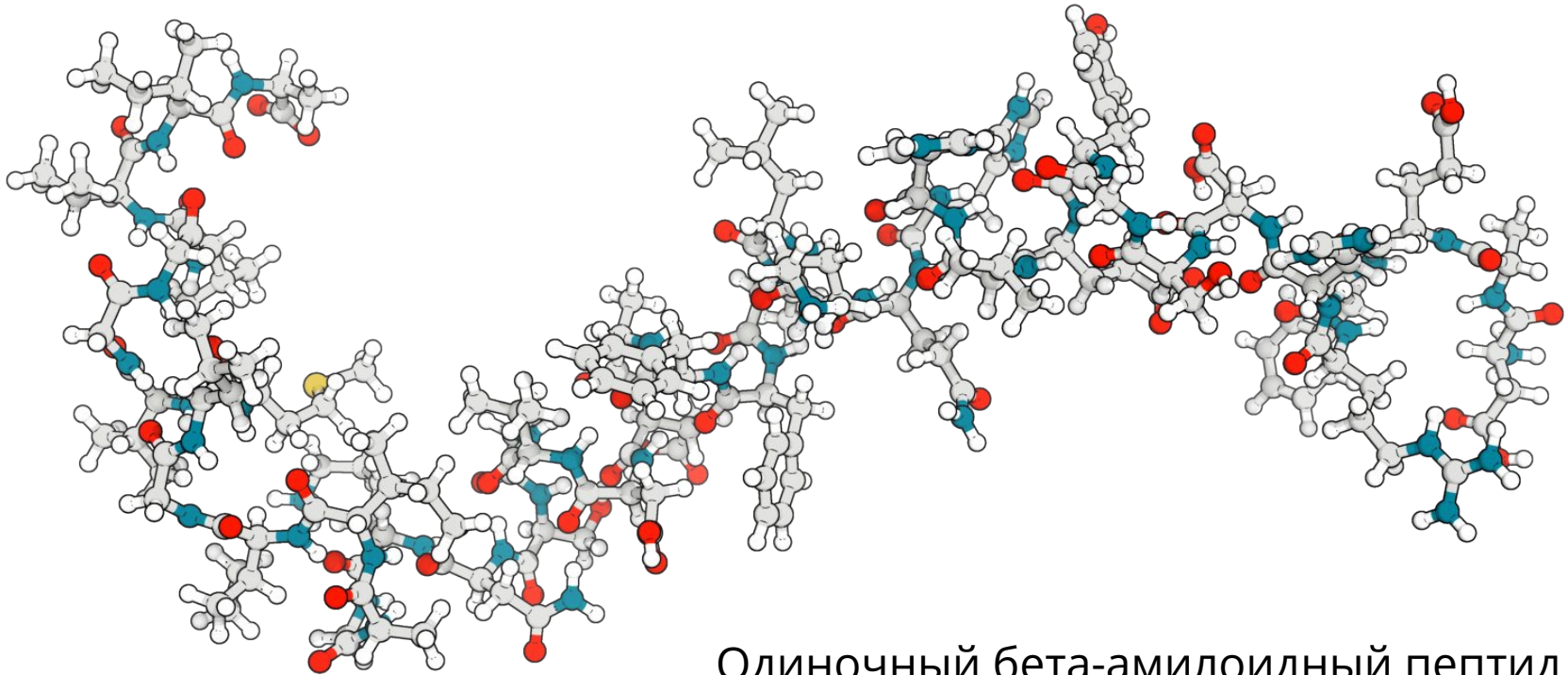


Пролил-изомераза А



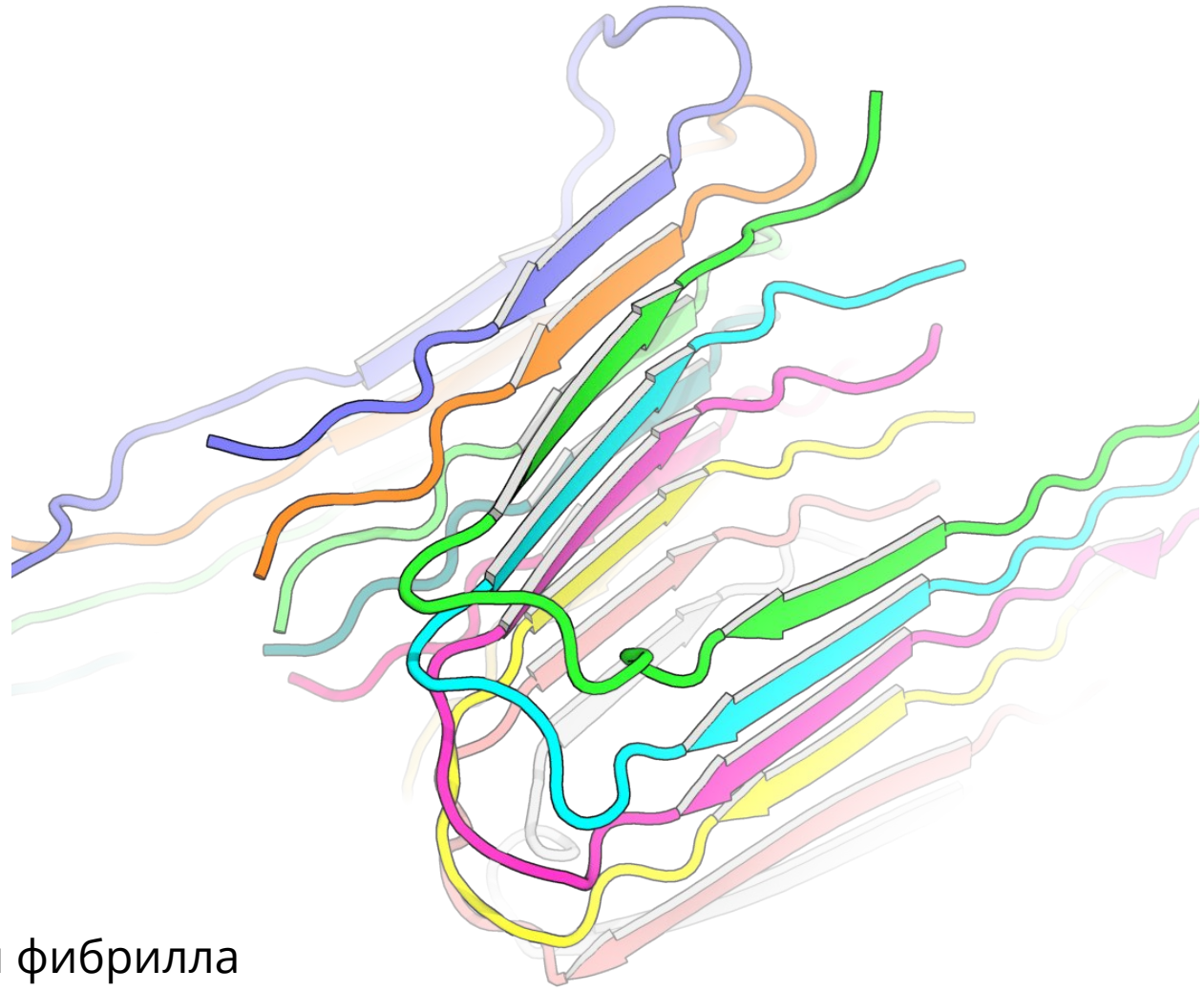
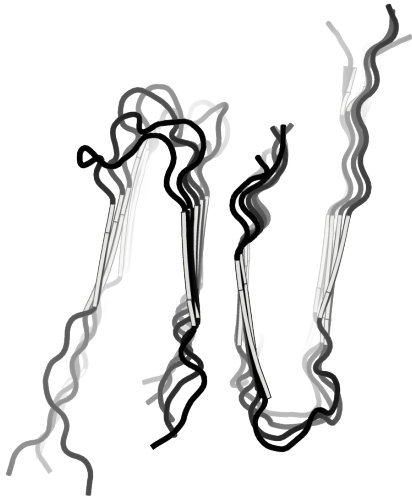
Тиоредоксин

# Амилоиды



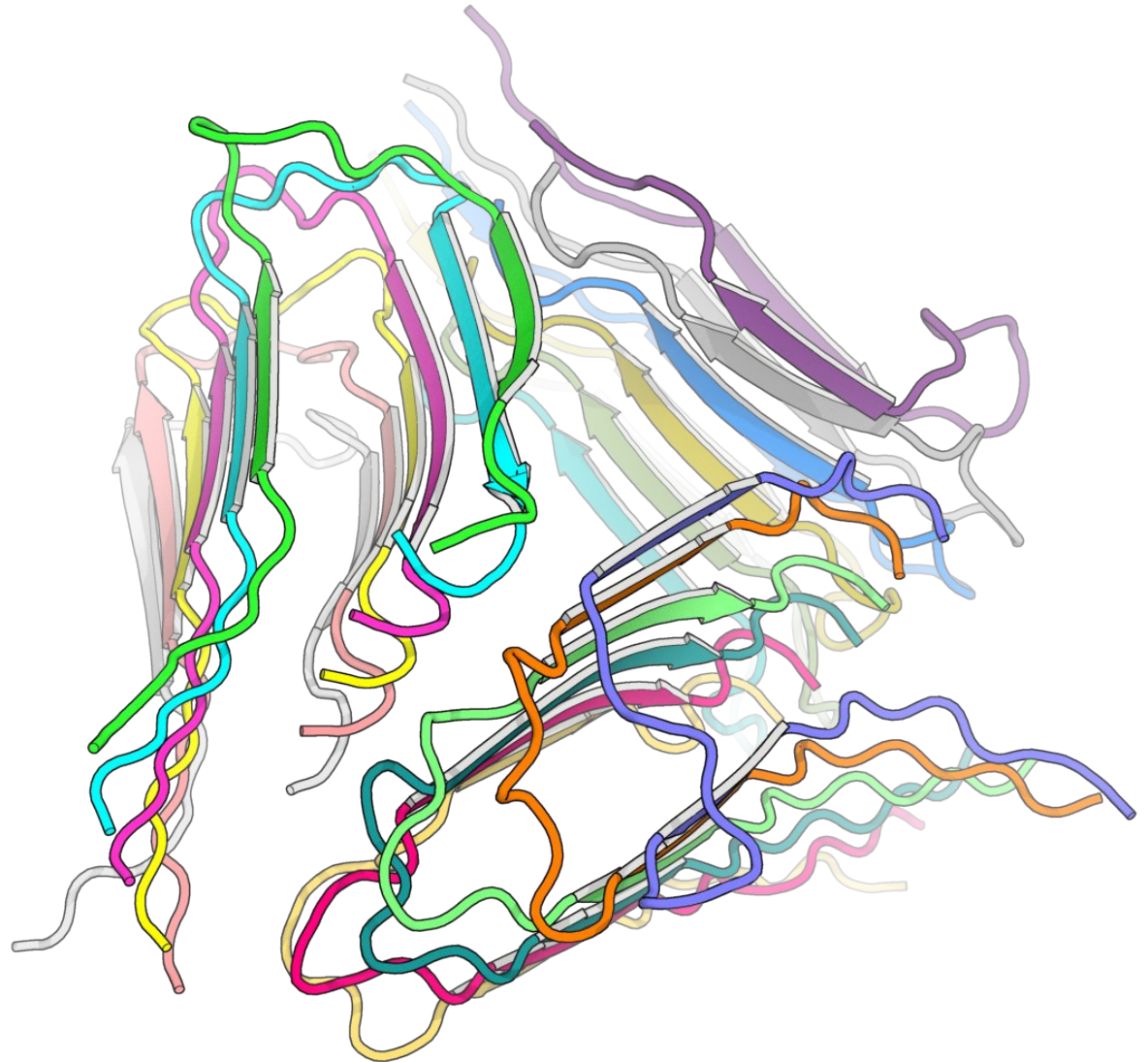
Одиночный бета-амилоидный пептид

# Амилоиды



2-fold бета-амилоидная фибрилла

# Амилоиды



3-fold бета-амилоидная  
фибрилла

# Предсказание укладки белка

CASP8 target 512-D1  
all models  
(3dsm)

