

Как создать белок с желаемыми свойствами?

Структурная Биоинформатика

Головин А.В.¹

¹МГУ им М.В. Ломоносова, Факультет Биоинженерии и Биоинформатики

Москва, 2015

Содержание

Введение

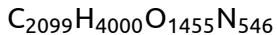
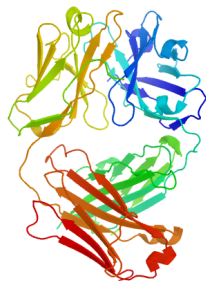
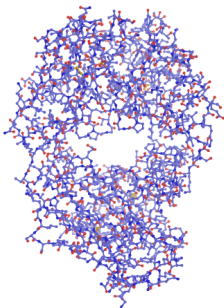
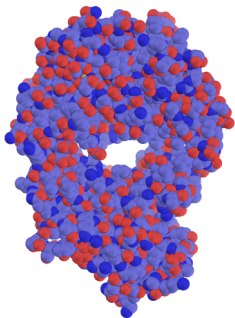
Самосборка белка

Folding@home

FoldIt Game



Белок



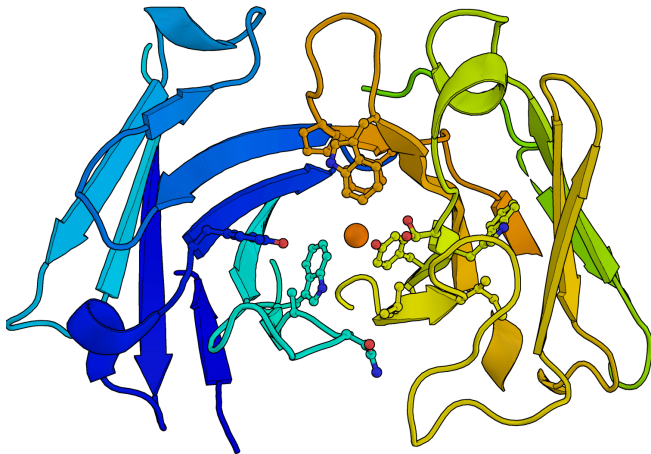
Основная причина сложности это способность узнавать и действовать



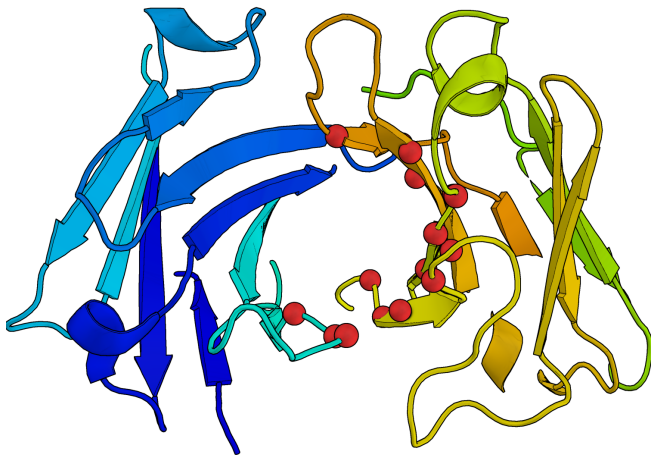
Реакция



Активный центр



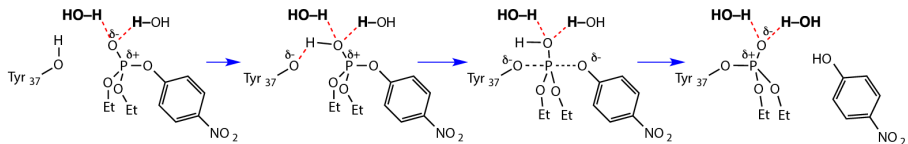
Место для мутации



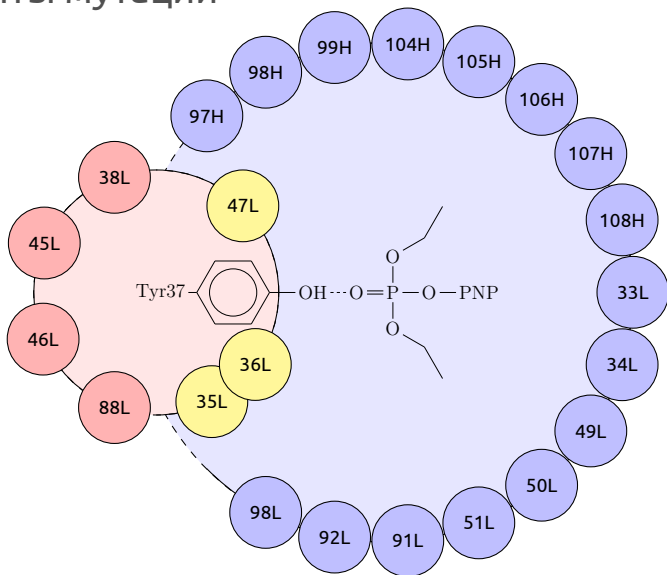
17 позиций, 20^{17} возможных мутантов



Реальная реакция



Варианты мутаций

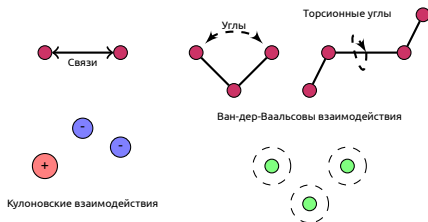


20 позиций, $2 \cdot 10^6$ возможных мутантов



Простое уравнение силового поля (СП)

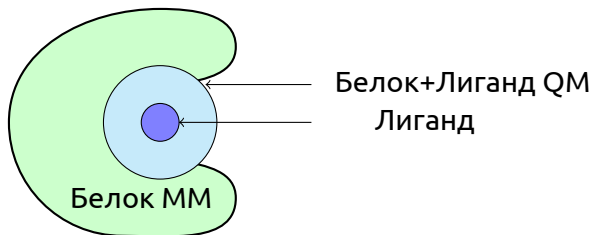
$$\begin{aligned}
 U = & \sum_{bonds} \frac{k_i}{2} (l_i - l_0)^2 + \sum_{angles} \frac{k_i}{2} (\phi_i - \phi_0)^2 + \sum_{torsions} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) + \\
 & + \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right)
 \end{aligned}$$



В результате построения моделей и их анализа только 375 мутантов показали эффективное взаимодействие с лигандом.

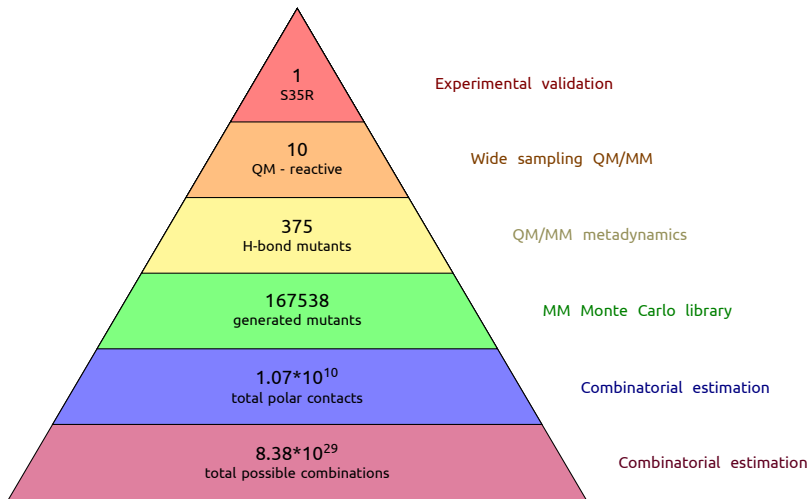
Гибридное QM/MM моделирование

- Основная идея: разделить большую систему на квантовую и молекулярную части.
- Электростатическое окружение из MM части чувствуется QM частью.
- MM часть принимает силы из QM части и соответственно адаптируется.



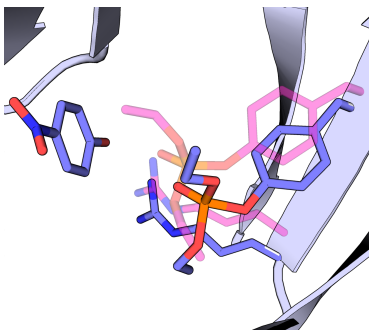
Моделирование реакции для 375м мутантов позволило выделить 10 лучших



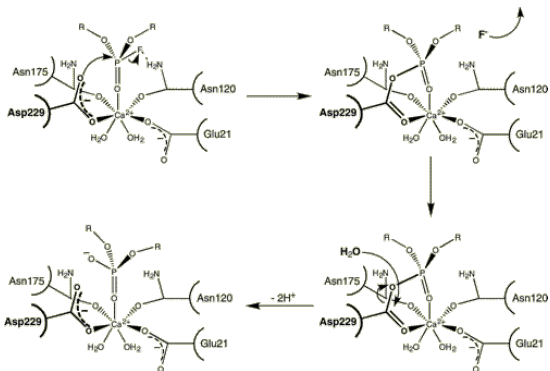
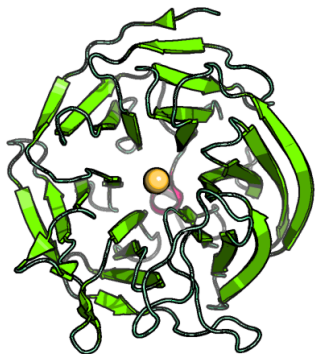


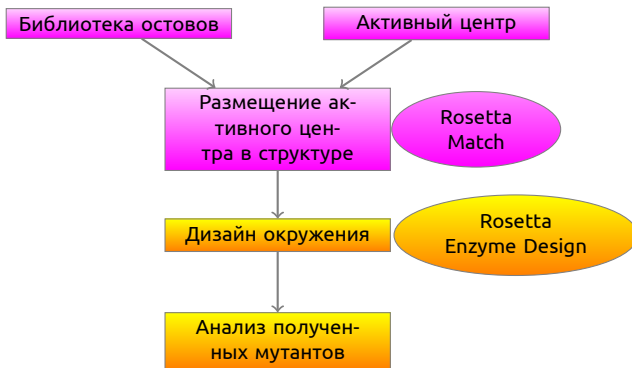
Итог

- 3-4 кандидата на экспериментальную проверку
- все кандидаты показали активность, два лучших ускорение в 200 раз
- получен кристалл и решена структура одного из лидеров.
- предсказание совпало с экспериментом

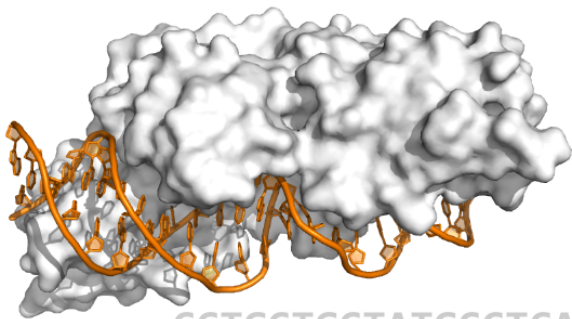


Перенос активного центра









CCTGCTGCTATGCCTCATCTTC

CCTGCTGCTATGCCTCATCTTC	HBV
AATGCTCCTATACGACGTTTAG	I-Ltrl
TGGTGGACTTCTCTCAATTTTC	
CTGAGGAGGTTTCTCTGTAAAG	I-Anil
TGCACTTCGCTTCACCTCTGCA	
TTGAGGAGGTTTCTCTGTTAAT	I-HjeMI
GCACTTCGCTTCACCTCTGCAA	
CTGACTCTCTTAAGGTAGCCAA	I-Vdi14I

I-Ltrl (*Leptographium truncatum*)

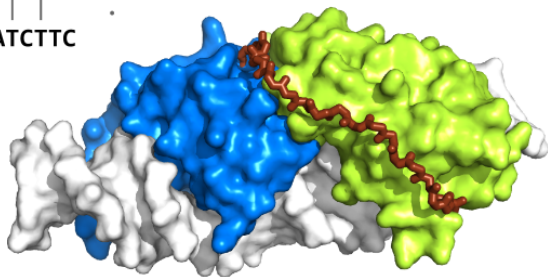
I-Ltrl

AATGCTCCTATACGACGTTTAG

||||| ||||| | | |

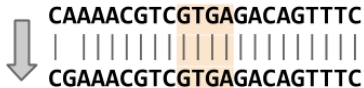
CCTGCTGCTATGCCTCATCTTC

I-Ltrl `HBV`



Проверка протокола

I-Crel



I-Crel `G`



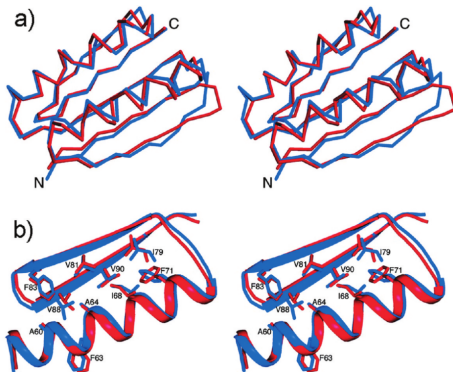
контроль



эксперимент

Дизайн de novo

- Надо менять не только конформацию в ходе моделирования но и последовательность.
- Была построена модель белка, укладки которого не было в PDB



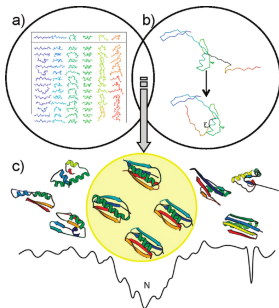
Другие способы дизайна

- Редизайн известных белков для увеличения устойчивости структуры
- Изменение места контакта с другими биополимерами полимерами (нуклеаза-ДНК)
- Дизайн ферментов, оптимизация структуры и последовательности белка для улучшения заданного переходного состояния.

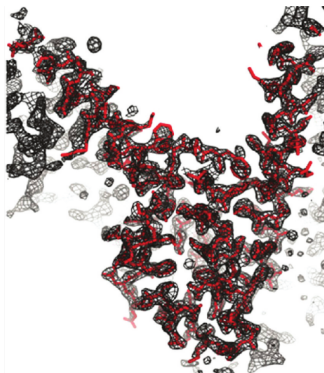
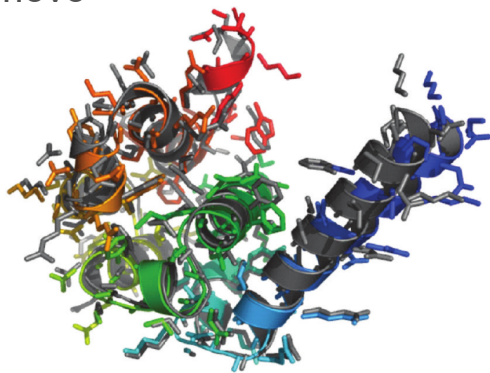


De Novo Folding Simulation

- В линейную конформацию вставляют фрагменты и быстро сканируют пространство укладывая остов белка. 9-мер 30000 раз затем 10000 раз 3-мер. Таким образом собирают 25000-50000 моделей.
- Выбранные модели оптимизируют на полноатомном уровне.



Использование экспериментальных данных в de novo

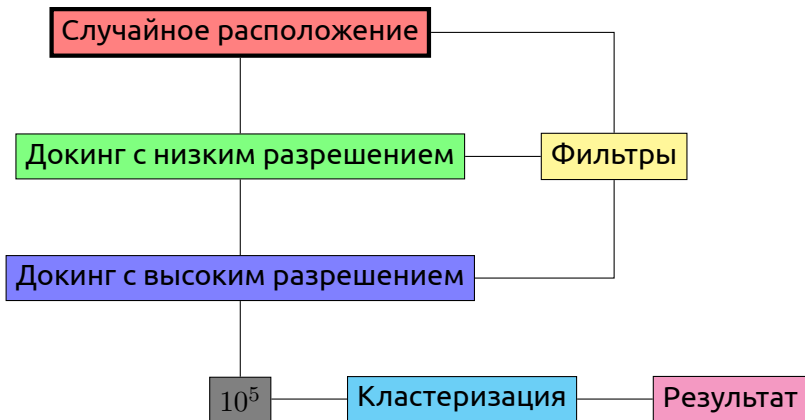


DOI: 10.1021/bi902153g

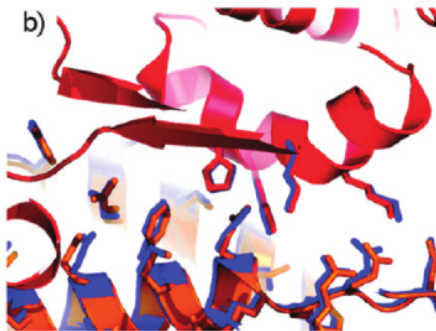
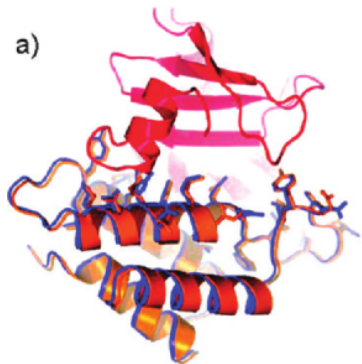
- Слева ЭПР данные + Rosetta, серым отмечена структура по данным PCA
- Справа: решение проблемы фазы с помощью моделей Rosetta и автоматического алгоритма молекулярного



Белок-Белок докинг: RosettaDock



Rosettadock, результат



МО-

дель окрашена
СИНИМ

DOI: 10.1021/bi902153g



Использование Rosetta

AbinitioRelax AnchoredDesign
antibody_graft BuildPeptide
docking_protocol enzyme_design
FlexPepDocking FloppyTail
hbs_design homodimer_design
ligand_dock loopmodel
loops_from_density

membrane_abinitio2 relax
remodel rna_denovo rna_design
rosettaDNA rosetta_scripts
rotamer_recovery SymDock
UnfoldedStateEnergyCalculator
zinc2_homodimer_design

Подробнее:

<https://www.rosettacommons.org/docs/latest/Application-Documentation.html>



Разработка Rosetta, PyRosetta

- Язык программирования в Rosetta C++
- К основным функциям и протоколам есть Python bindings

```
In [1]: import sys
import os
import operator
from itertools import *

sys.path.append('/home/anur/pyrosetta')

from rosetta import *
from toolbox import *
from rosetta.utility import vector1_bool
init()
#from prody import *
#from pylab import *
import numpy as np

from numpy.random import randint

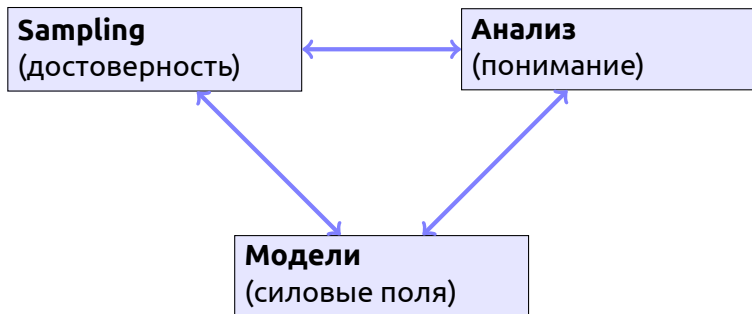
from joblib import Parallel, delayed
```

let use pymol for visual anlysis and debugging

```
In [2]: pymover = PyMOL_Mover()
```



Вызовы молекулярного моделирования



Моделирование самосборки белка

- Как иметь дело с большими временами?
Самые "быстрые" белки собираются за 10-100 μs ,
моделирование должно быть на порядок длинее.
- На сколько хороши наши силовые поля?
Сможем ли мы получать нативное состояние белка без
знания структуры.
Будут ли предсказанные структуры иметь правильные
параметры: ΔG , K_f .
- Сможем ли мы изучать сборку моделированием?
Механизмы, Теория ...



МД в моделировании самосборки белка

- Традиционный подход: несколько длинных траекторий
- Альтернатива: методы стохастического кинетического сканирования.

Фолдинг это стохастический процесс с экспоненциальной кинетикой, т.е. количество молекул, которые собрались:

$$f(t) = M[1 - \exp(-kt)]$$

для малых времён:

$$f(t) \sim Mkt \quad M \sim 10,000 \text{procs}, k \sim 1/10,000 \text{ns}, t \sim 20 \text{ns/proc}$$

ожидается, что можно увидеть 20 раз сборку белка.

- Это эффективно.
- Эргодично



Folding@home

English

Folding@home DISTRIBUTED COMPUTING

Home Learn Stats About Us Search

Help unlock the mysteries of disease.

READY.

- What is protein folding?
- Why does it matter?
- How can I help?

SET.

Download Folding@home

64bit Linux
 .deb + .rpm

- Other platforms
- Older versions

FOLD.

- Start folding now!
- Get help.
- Earn points, join a team.

You can help scientists studying Alzheimer's, Huntington's, and many cancers by simply running a piece of software on your computer or **game console**.

Join others around the world to form the world's largest distributed supercomputer.

Make a gift now!

[Click here to make a gift.](#)



Folding@home

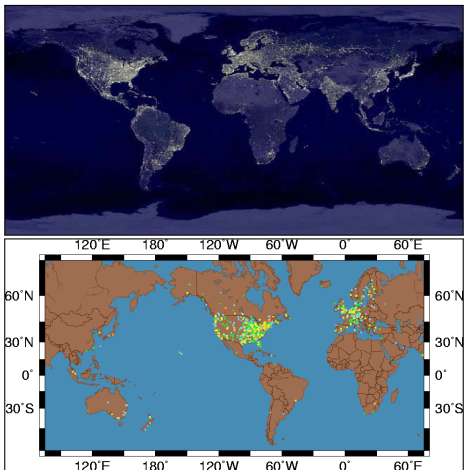


Illustration from Stefan Larson



Результаты

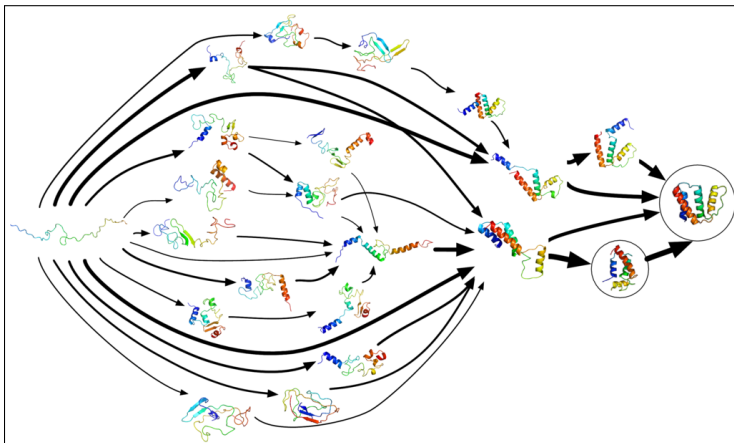


Illustration from Wikipedia

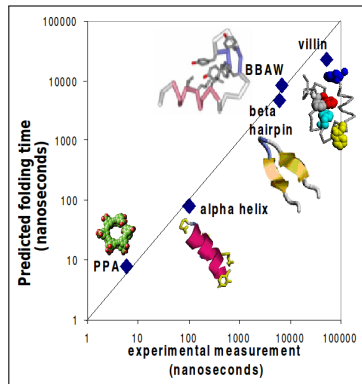


Результаты и ответы

Ответы на наши вопросы:

- Выборка достаточна.
- Силовые поля работают
Сборка происходит правильно
и с нужной скоростью
- Механизм
Вероятно, что самосборка
это не универсальный
механизм и индивидуален
для каждого белка

Illustration from Stefan Larson



FoldIt Game

Основная идея:

- Использование **crowdsourcing** для поиска оптимальной геометрии.
- Пользователь может изменять торсионные углы и добиваться оптимальной энергии.
- В ходе игры для прохождения уровня пользователь должен достичь нужного уровня энергии для структуры
- Создатели программы собираются утилизировать способы придуманные людьми для разработки новых алгоритмов.



FoldIt

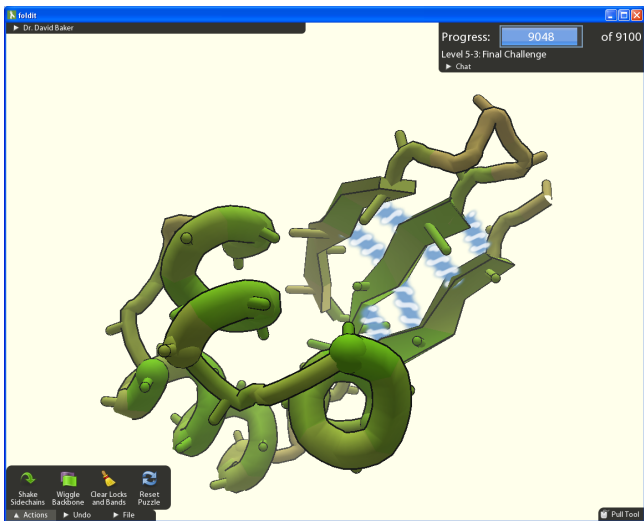


Illustration from Wikipedia