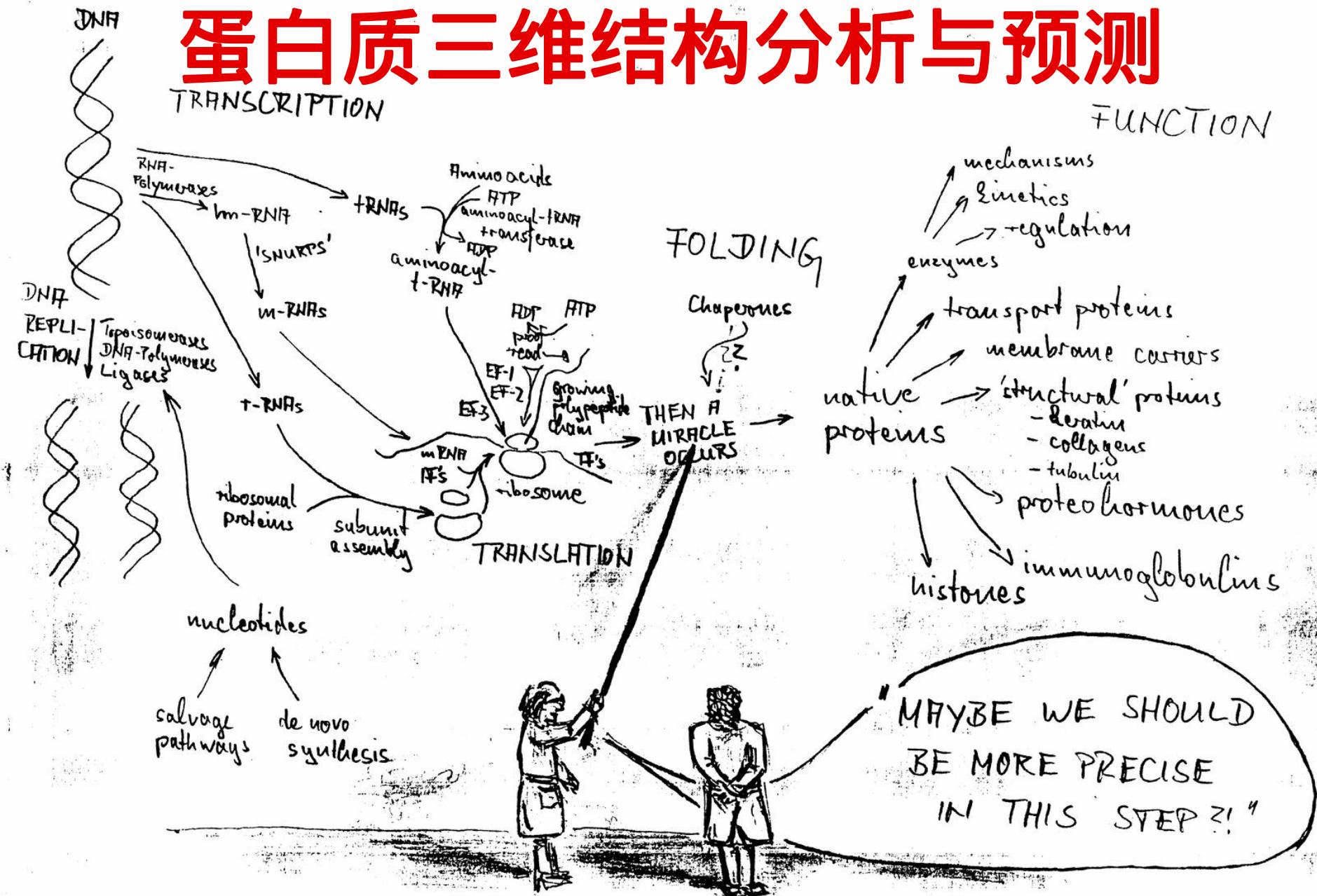


蛋白质三维结构分析与预测



吉祥如意

蛋白质结构

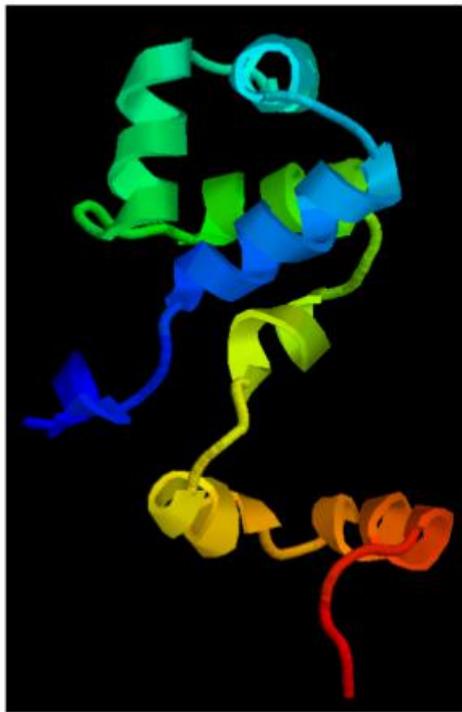
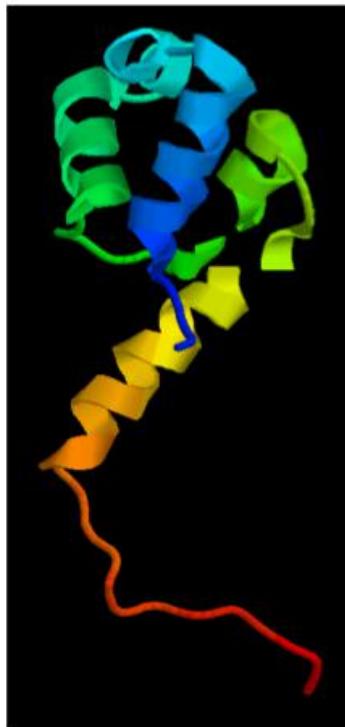
- 【 一级结构: 序列
- 【 二级结构: alpha螺旋, beta片层
- 【 三级结构: 单体
- 【 四级结构: 多聚体



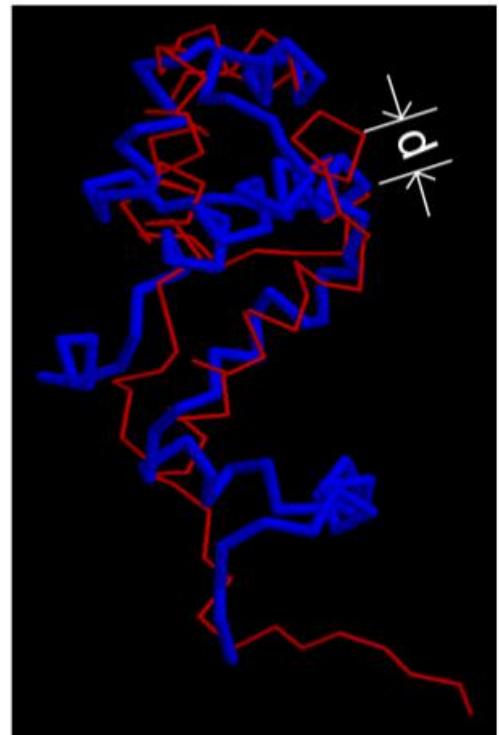
吉祥園

结构的相似性

How similar are these two structures?



RMSD
superposition



Native

Model



用RMSD衡量三维结构的差别

□ RMSD (root mean square deviation):

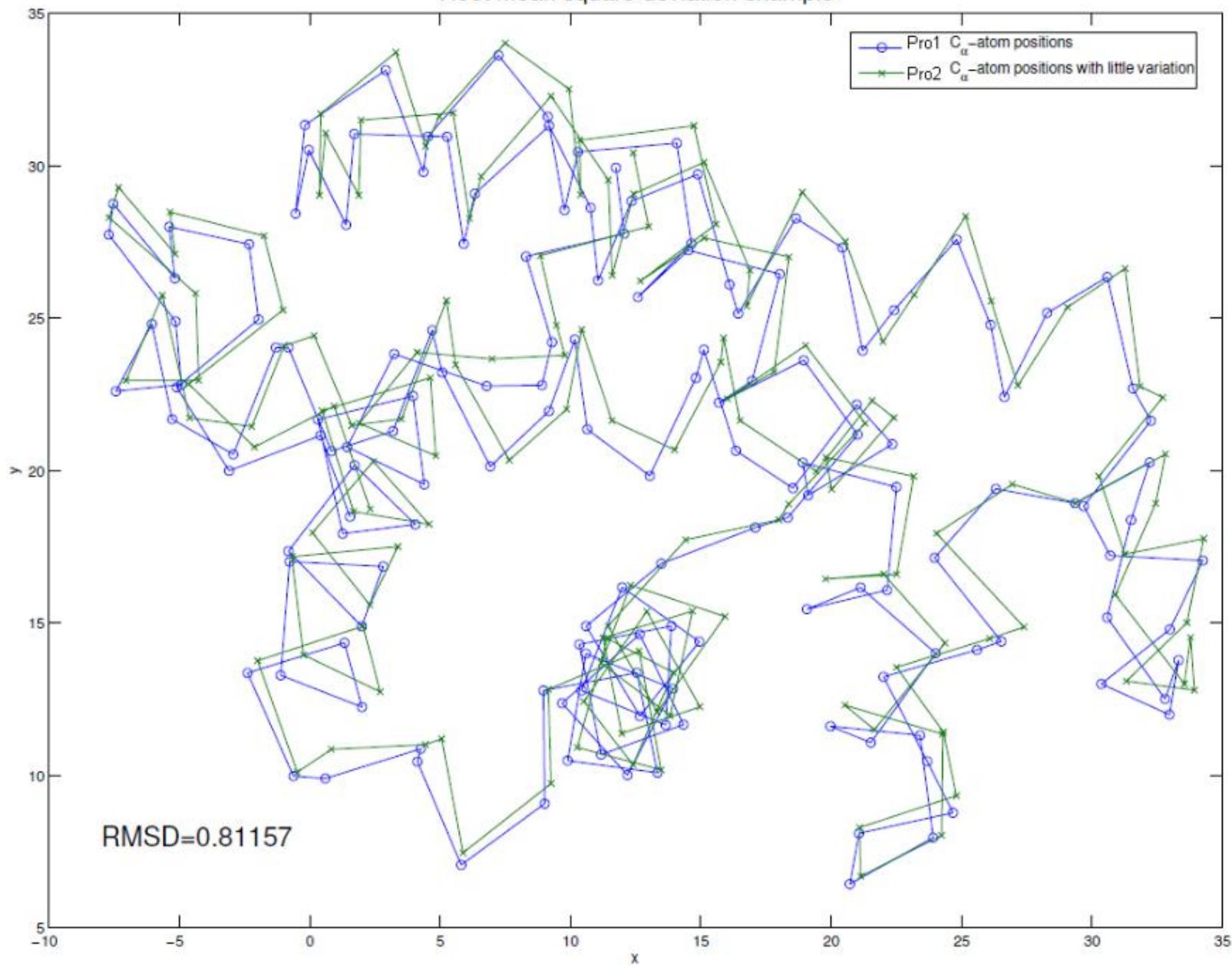
均方根偏差

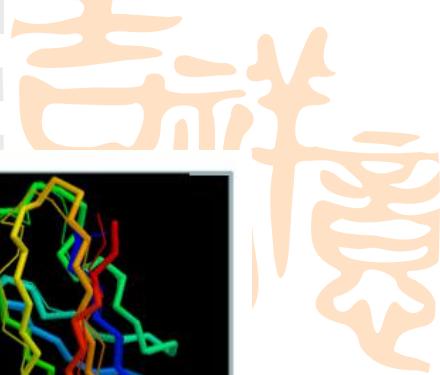
$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \|r1_i - r2_i\|^2}$$

$r1_i$ 和 $r2_i$ 分别是重叠的蛋白质1和蛋白质2对应的Ca原子的位置

RMSD的单位是 Å

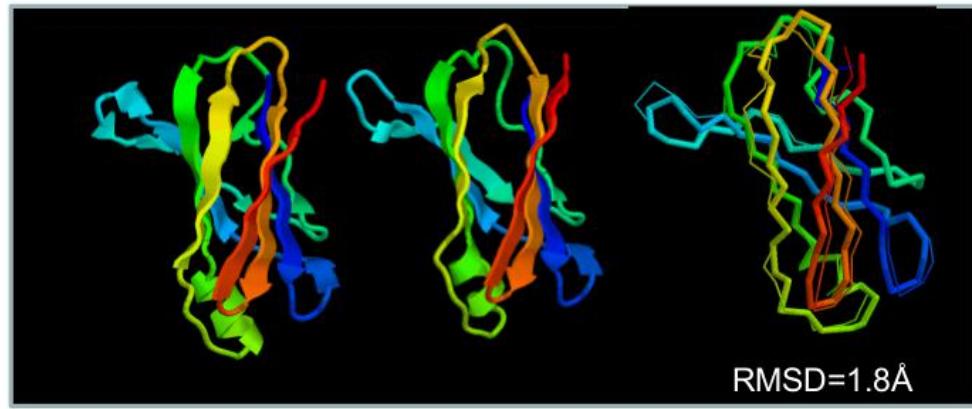
Root mean square deviation example



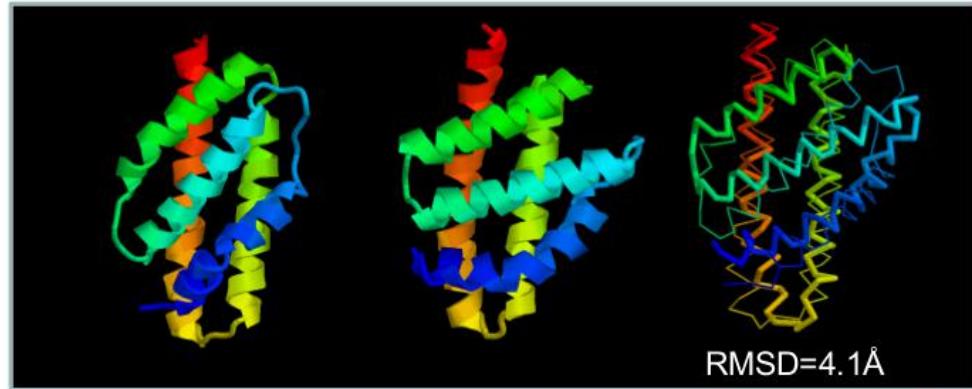


General RMSD values

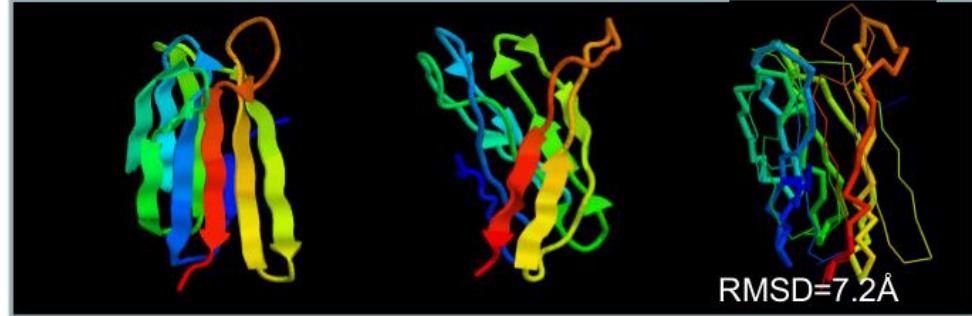
- RMSD in $[0, 2\text{\AA}]$, high resolution structures of close similarity



- RMSD in $[2\text{\AA}, 6\text{\AA}]$, similar topology, medium resolution



- RMSD $> 6.5\text{\AA}$, different topology



Determine Protein Structure

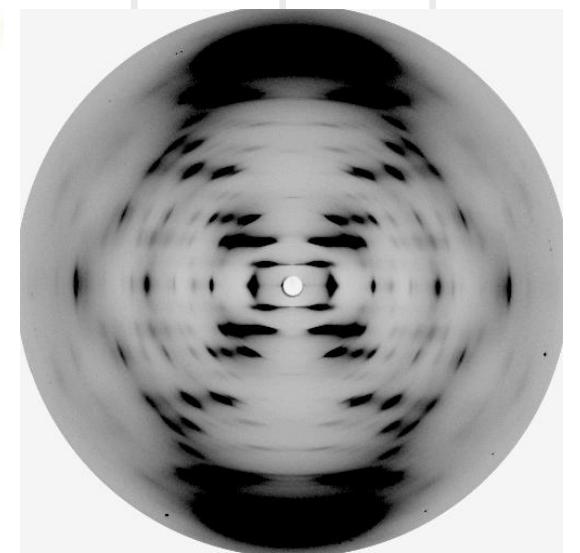
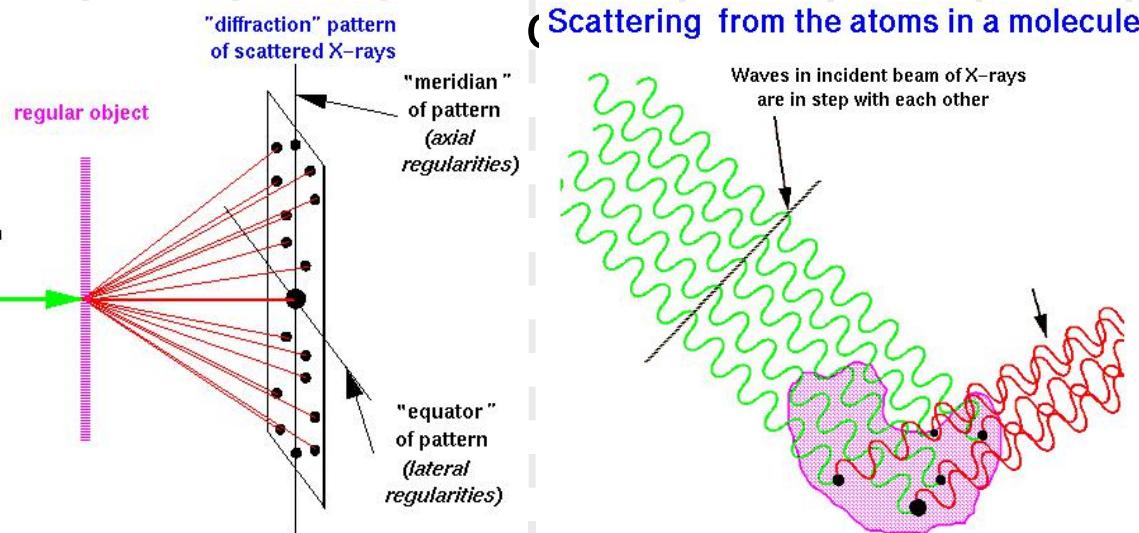


[[X-ray crystallography (gold standard)

Grow crystals, rate limiting, relies on the repeating structure of a crystalline lattice

Collect a diffraction pattern

Map to real space electron density, build and refine structural model

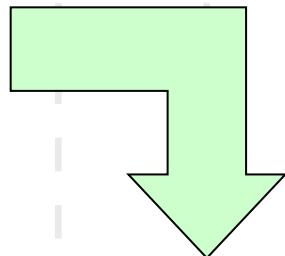
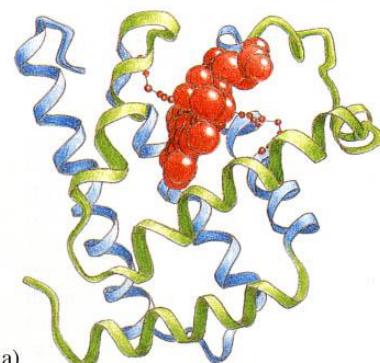
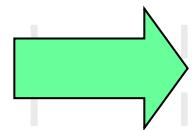
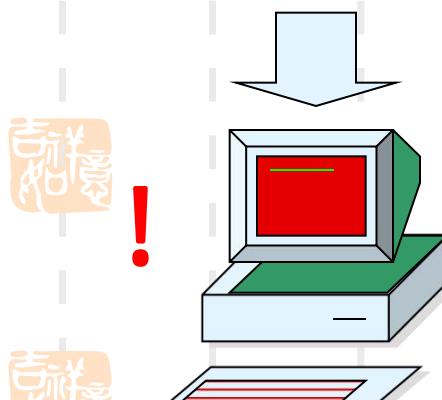


吉祥如意

蛋白质结构预测问题

序列——结构——功能

....-Gly-Ala-Glu-Phe-....



FUNCTION

吉祥如意

吉祥如意

吉祥如意

吉祥如意

吉祥如意

蛋白质三维结构预测

吉祥如意

历史悠久: 超过30年

蛋白质三维结构预测被誉为计算生物学的“圣杯”



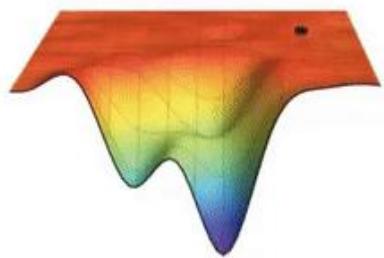
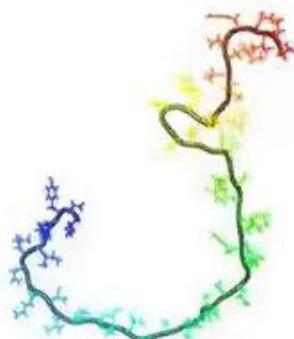
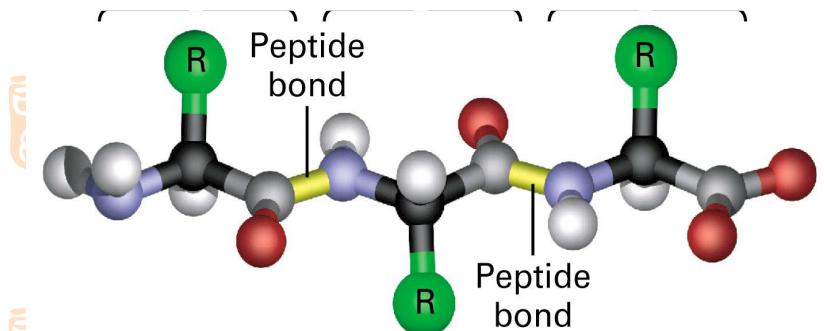
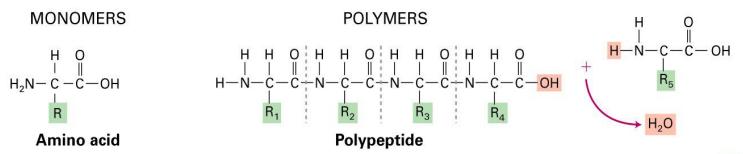
持续性的研究, 专门的比赛: **CASP** (1994-2014)



用途: 蛋白质相互作用、大分子复合物的组装、优化和完善实验所得结构 (X-ray、NMR以及电镜结构)、药物设计、功能注释、蛋白质工程等

吉祥如意

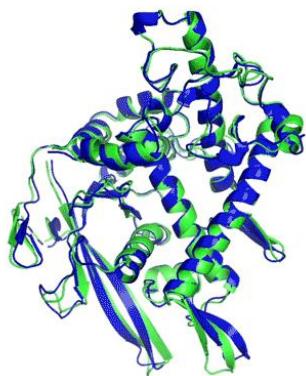
背景：蛋白质的折叠



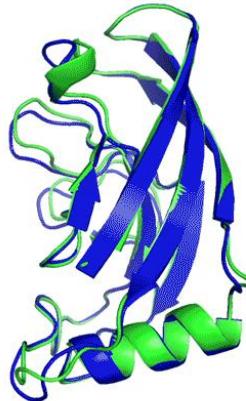


14th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction

- 152 targets
- 215 methods
- 5.18~9.22, 2020



T1037 / 6vr4
90.7 GDT
(RNA polymerase domain)



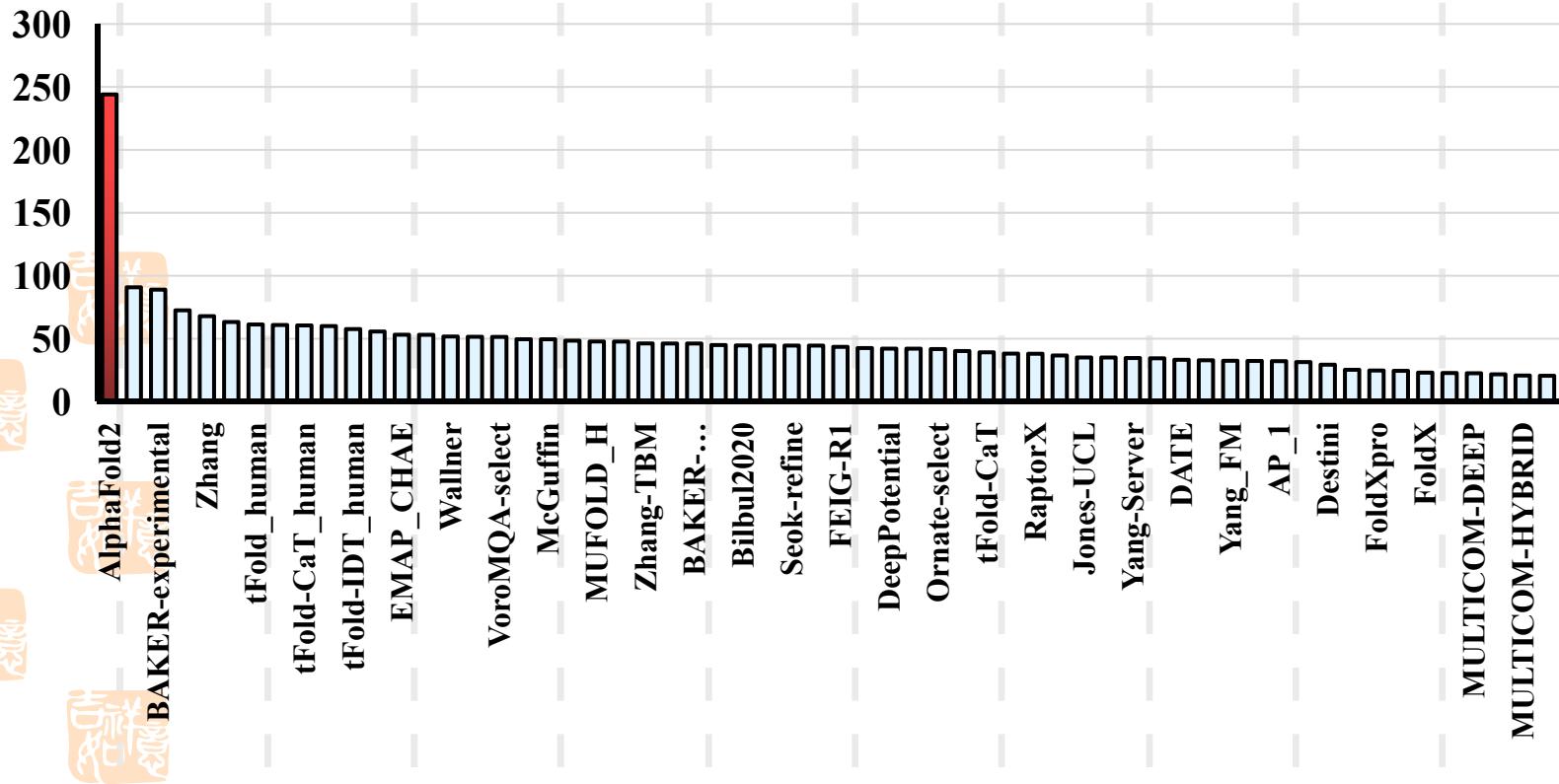
T1049 / 6y4f
93.3 GDT
(adhesin tip)

- Experimental result
- Computational prediction

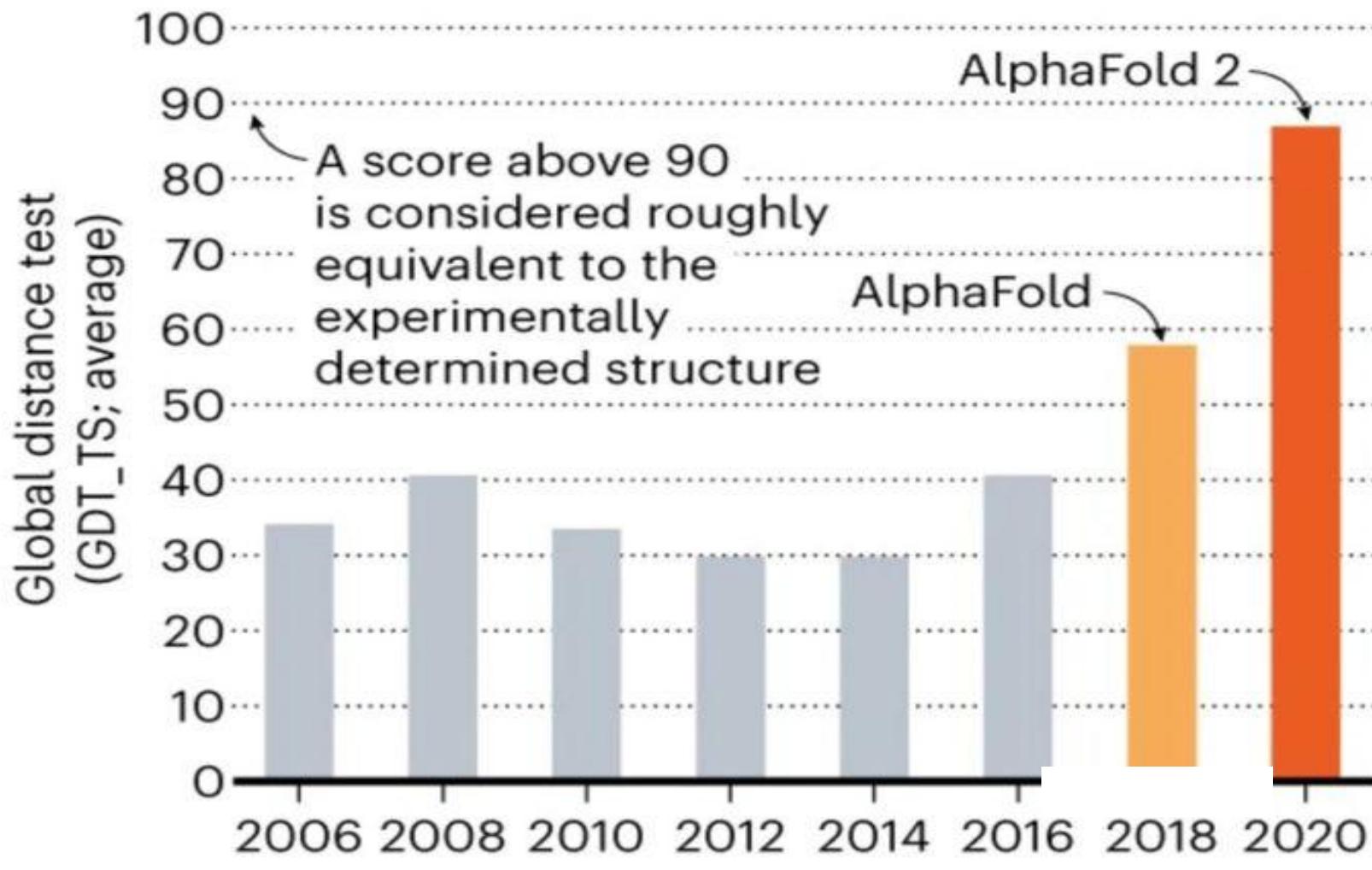


CASP14 比赛结果

Sum Z-score (>-2.0)

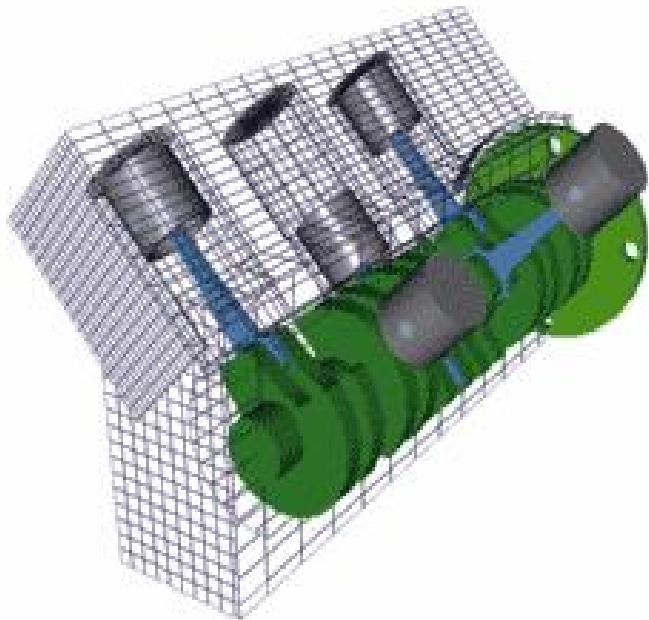


CASP7 ~ CASP14 比赛结果



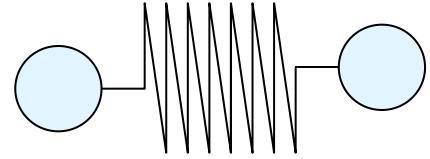
吉祥如意

怎样到达折叠好的状态？

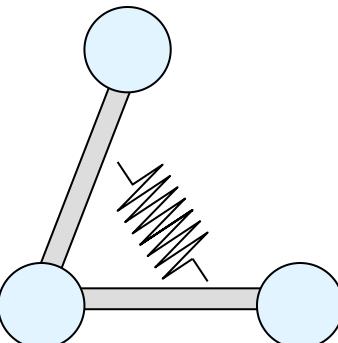


能量(打分)函数

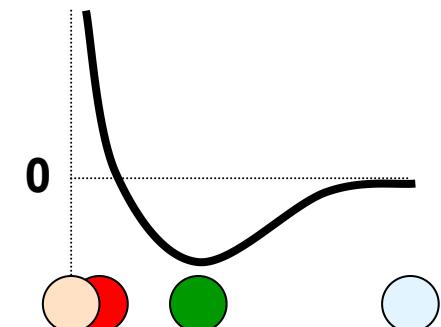
吉祥



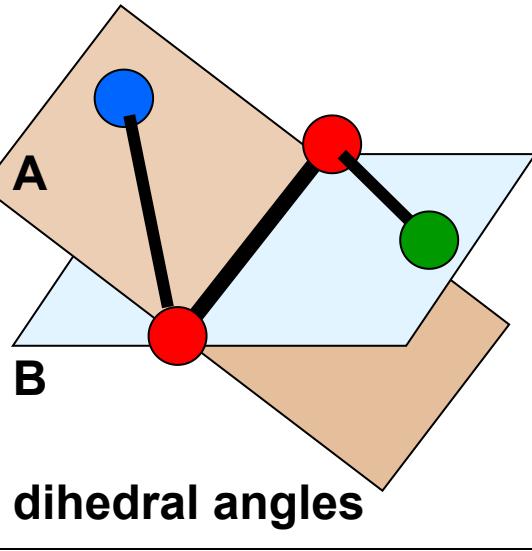
bond lengths



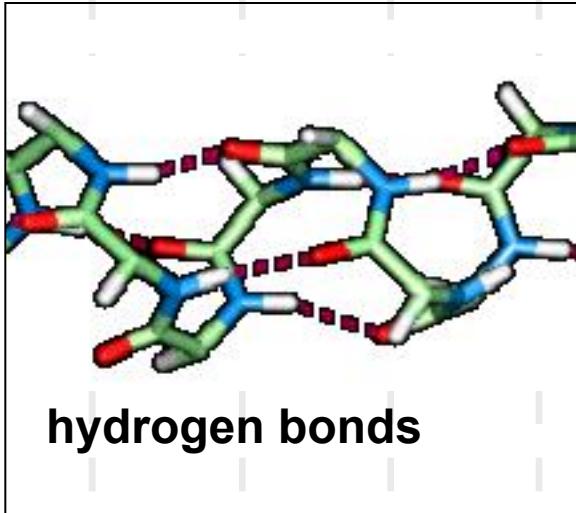
bond angles



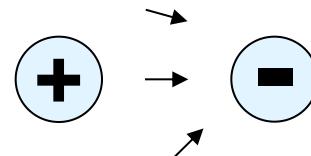
van der Waals



dihedral angles



hydrogen bonds



electrostatics

吉祥如意

搜索算法

- 【 蒙特卡罗
- 【 遗传算法
- 【 动态规划
- 【 宽/深度优先搜索
- 【 DEE
- 【 贪婪算法
-



蛋白三维结构预测的主要方法

基于模板的结构预测

(1) 同源模建

Homology modeling

(2) 折叠识别

Fold Recognition/Threading

不依赖于模板的结构预测

(3) 从头计算

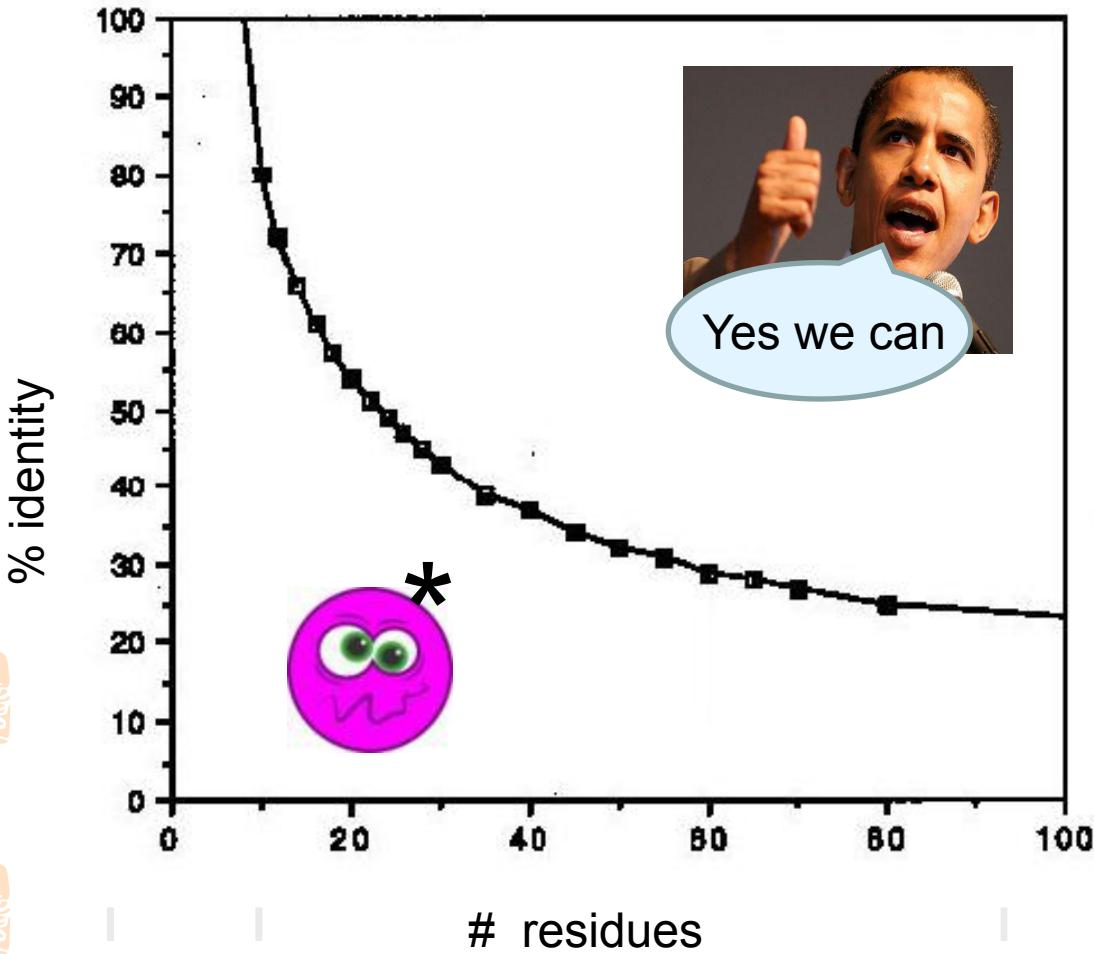
Ab initio/de novo prediction



同源模建

Homology Modeling

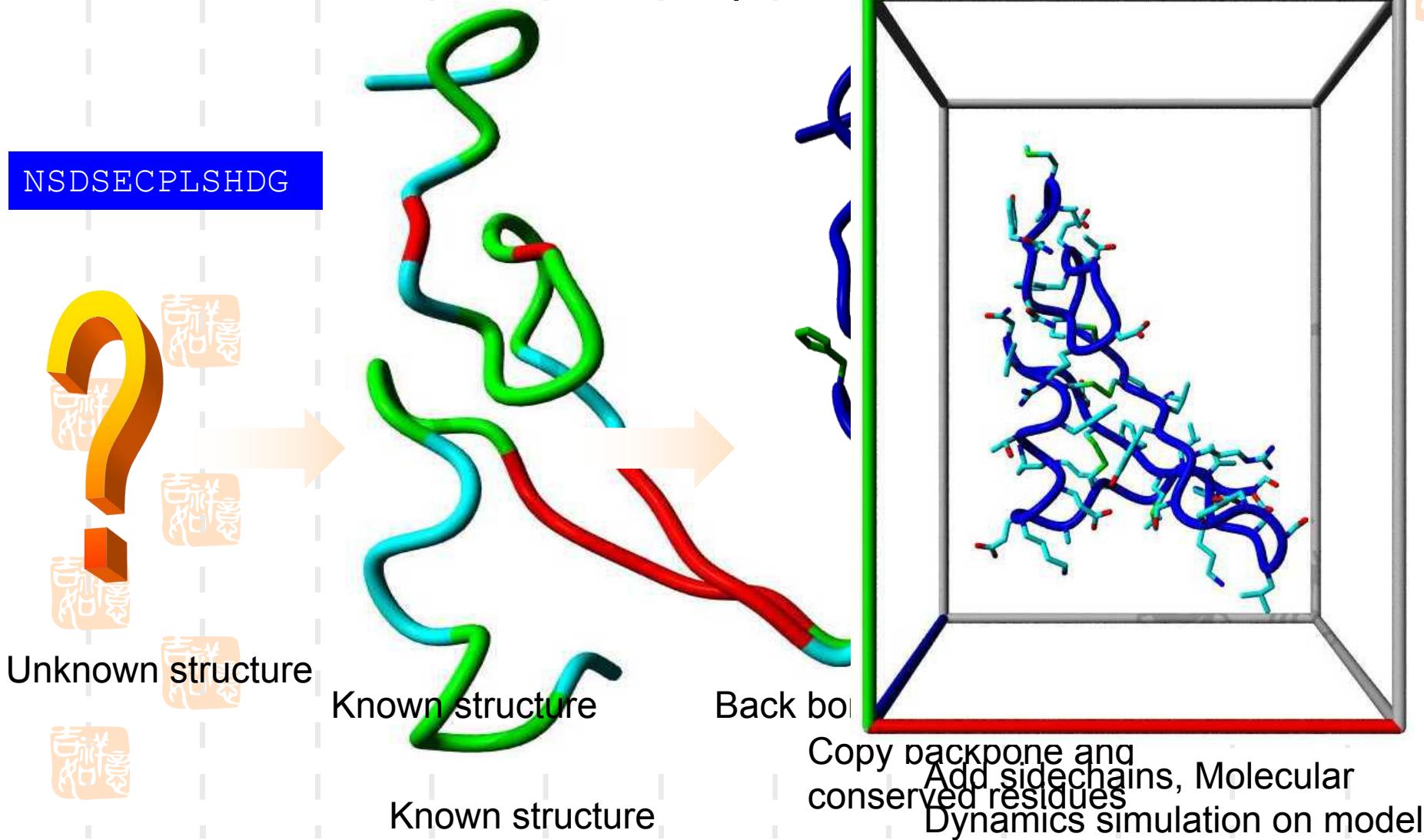
Homology modeling



* Actually, modelling is possible here, but we cannot get an alignment...

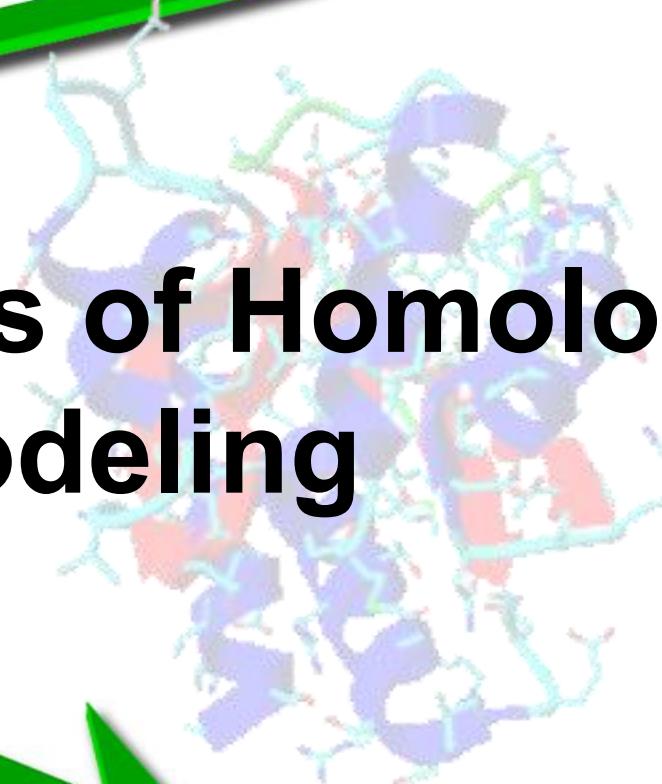
Homology modeling in short...

Prediction of structure based upon a highly similar structure



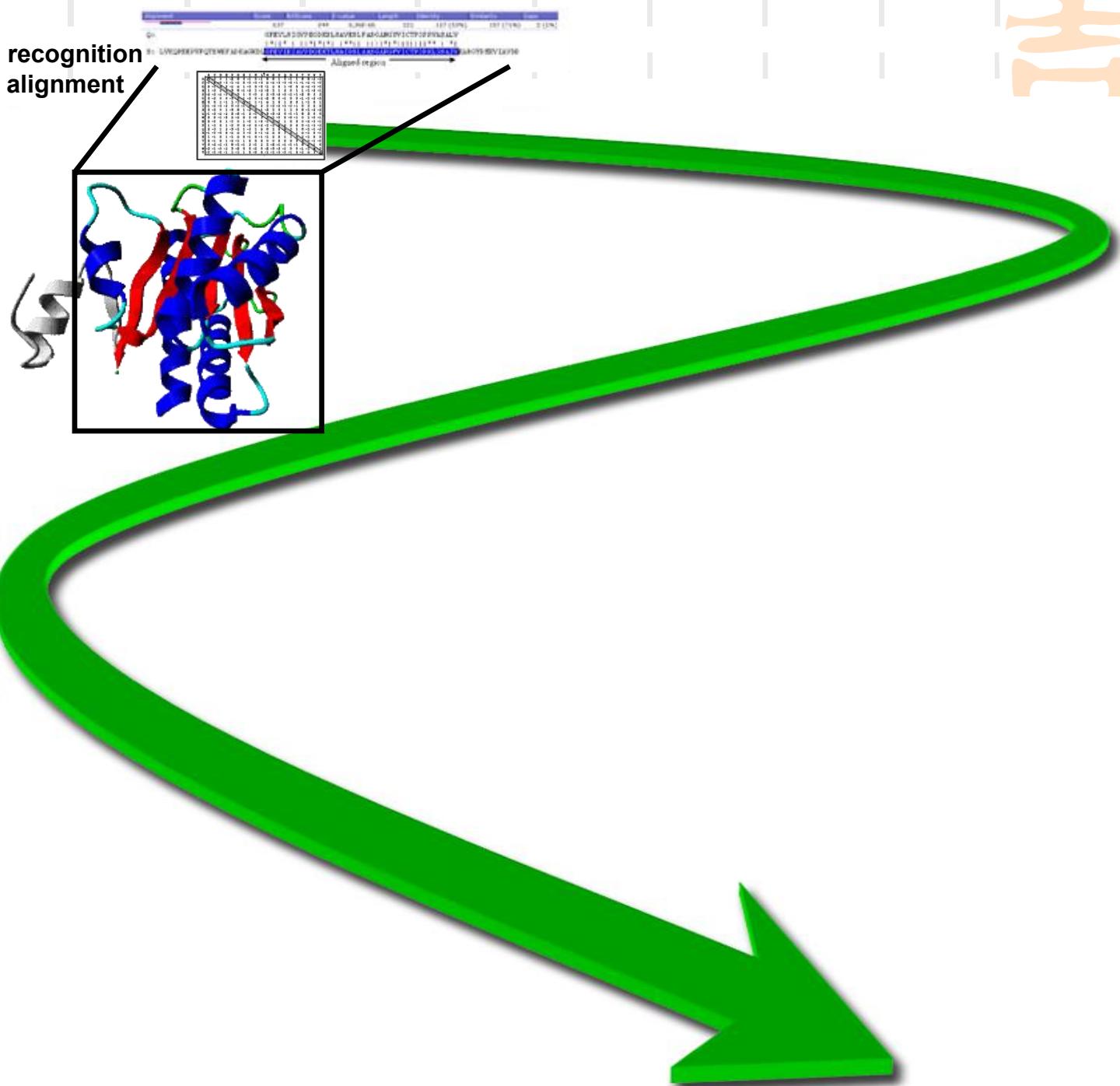


The 8 steps of Homology modeling



吉
祥

1: Template recognition
and initial alignment



1: Template recognition and initial alignment

BLAST your sequence against PDB

Best hit normally template

Alignment	DB-ID	Source	Length	Score	Identity%	Positives%	E-value
1 ✓	PDB:1NQL_A	mol:protein length:624 Epidermal Growth Factor Receptor	624	2728	100	100	0.0
2 ✓	PDB:1MOX_B	mol:protein length:501 Epidermal Growth Factor Receptor	501	2728	100	100	0.0
3 ✓	PDB:1MOX_A	mol:protein length:501 Epidermal Growth Factor Receptor	501	2728	100	100	0.0
4 ✓	PDB:1IVO_B	mol:protein length:622 Epidermal Growth Factor Receptor	622	2728	100	100	0.0

Sequence

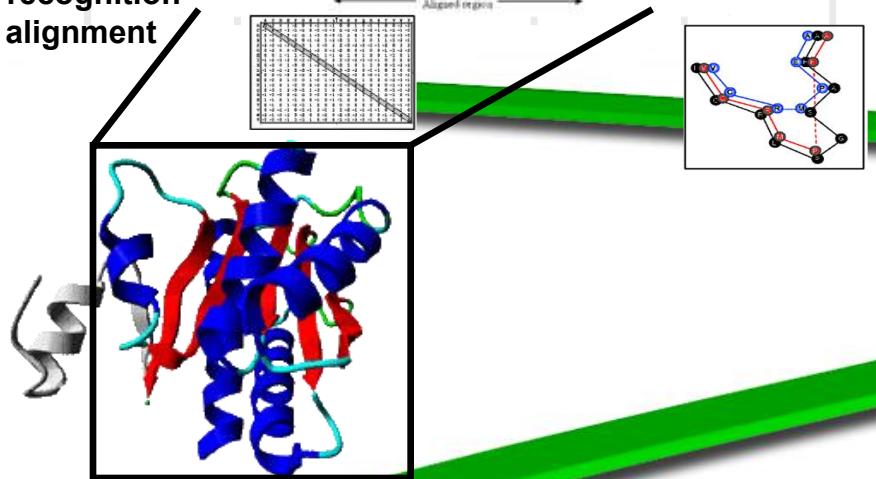
1 PDB:1NQL_A
2 PDB:1MOX_B
3 PDB:1MOX_A
4 PDB:1IVO_B
5 PDB:1IVO_A
6 PDB:1Y9_A
7 PDB:1S78_B
8 PDB:1S78_A
9 PDB:1N8Z_C
10 PDB:1M6B_B
11 PDB:1M6B_A
12 PDB:1N8Y_C
13 PDB:1IGR_A
14 PDB:1VFT_A
15 PDB:1N7D_A
16 PDB:4MT2
17 PDB:1WGT_B

1 [REDACTED]

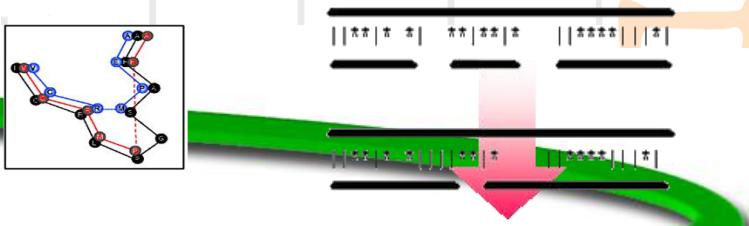
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
L E E K K V Q G T S N K L I T Q L G T F E D H F L S L Q R M F N N E V V L G N L E I T Y V Q R N Y D L S F L K T I Q E V A G Y V L I A L N T V E R I P L E N L
---- -Q V T G T D M K L R L P A S P E T H L D M L R H L Y Q G Q V V Q G N L E I T Y L P T N A S I S F L Q D I Q E V Q G Y V L I A H N Q V R Q V P L Q R L
---- -Q V T G T D M K L R L P A S P E T H L D M L R H L Y Q G Q V V Q G N L E I T Y L P T N A S I S F L Q D I Q E V Q G Y V L I A H N Q V R Q V P L Q R L
---- -Q V T G T D M K L R L P A S P E T H L D M L R H L Y Q G Q V V Q G N L E I T Y L P T N A S I S F L Q D I Q E V Q G Y V L I A H N Q V R Q V P L Q R L
---- -V P G T L N G L S V T G D A E N Q Y Q T L Y K L Y E R E V V M G N L E I V I T G H N A D I S F L Q W V R E V T G Y V L V A M N E F S T L P L P N L
---- -V P G T L N G L S V T G D A E N Q Y Q T L Y K L Y E R E V V M G N L E I V I T G H N A D I S F L Q W V R E V T G Y V L V A M N E F S T L P L P N L
---- -Q V T G T D M K L R L P A S P E T H L D M L R H L Y Q G Q V V Q G N L E I T Y V P A N A S I S F L Q D I Q E V Q G Y M L I A H N Q V R V P L Q R L
---- -Y Q Q L K R L - E N T V I E G Y I H I L L I s t S Y R F P K L T V I T E Y L - L L F R V A G L E S L g f P N L

80

1: Template recognition
and initial alignment



2: Alignment correction



2: Alignment correction



- Functional residues conserved
- Use multiple sequence alignments
- Deletions shift gaps



CPISRT**A**AS-FRCW
CPISRT**G**-SMFRCW
CPISRT**A**--TFRCW
CPISRT**A**AASHFRCW
CPISRT**G**ASIFRCW
CPISRT**A**---FRCW

Multipe sequence alignment



CPISRT**G**ASIFRCW
CPISRT**A**---FRCW

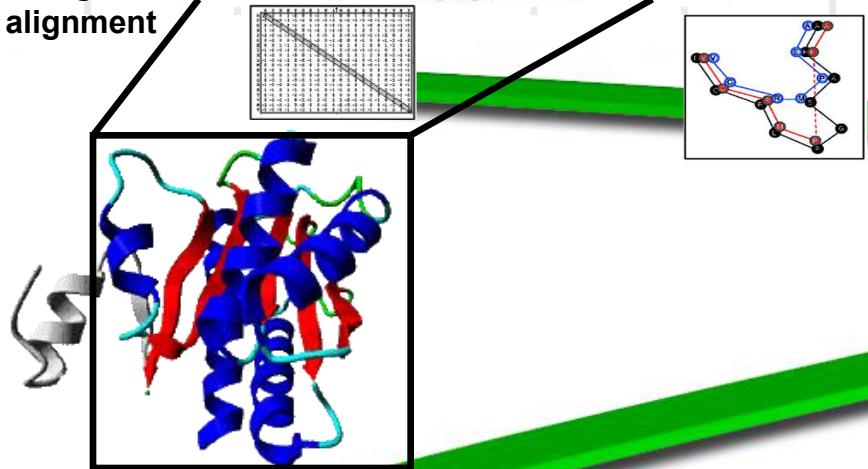
CPISRT**G**AS**I**FRCW
CPISRT---**A**FRCW

Sequence with known structure
Your sequence

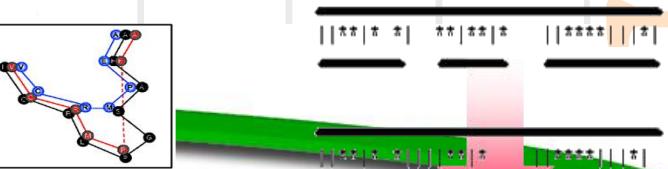
Both are possible

Correct alignment

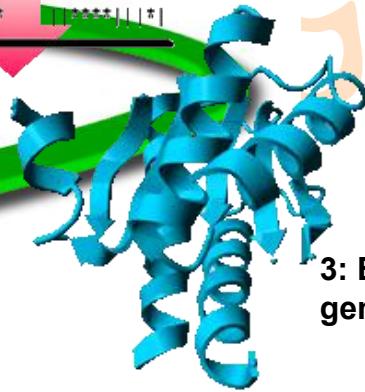
1: Template recognition
and initial alignment



2: Alignment correction

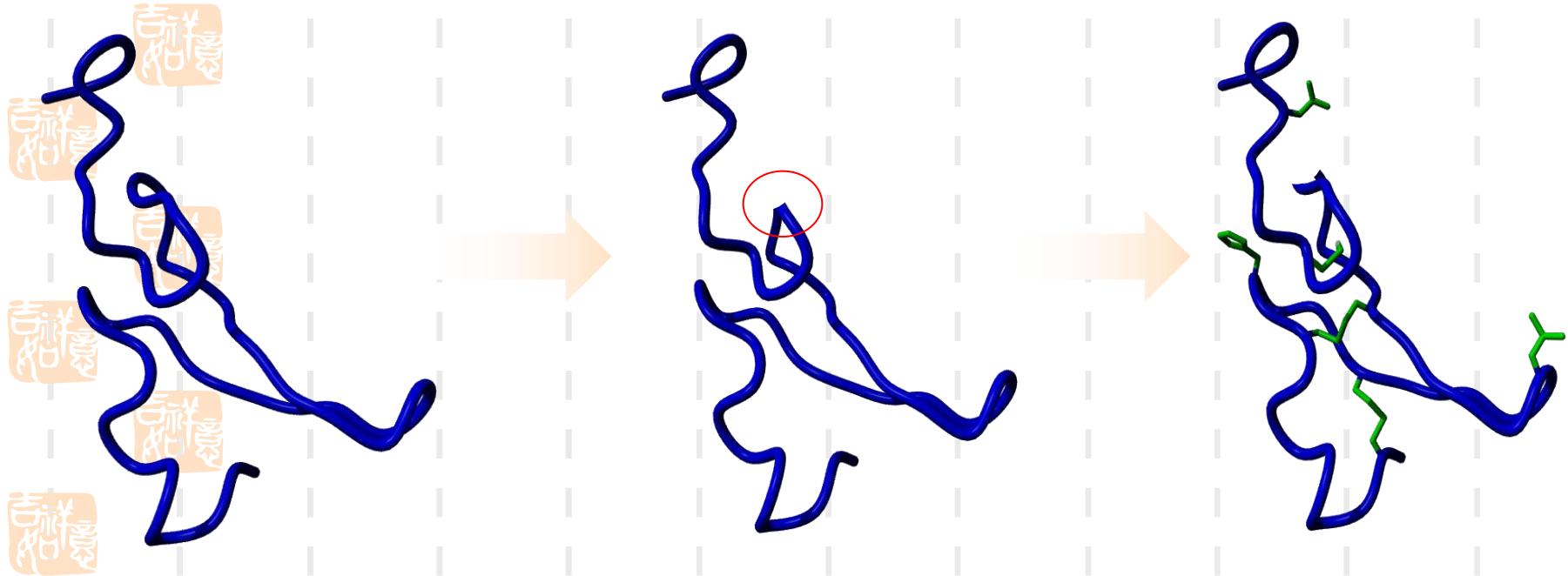


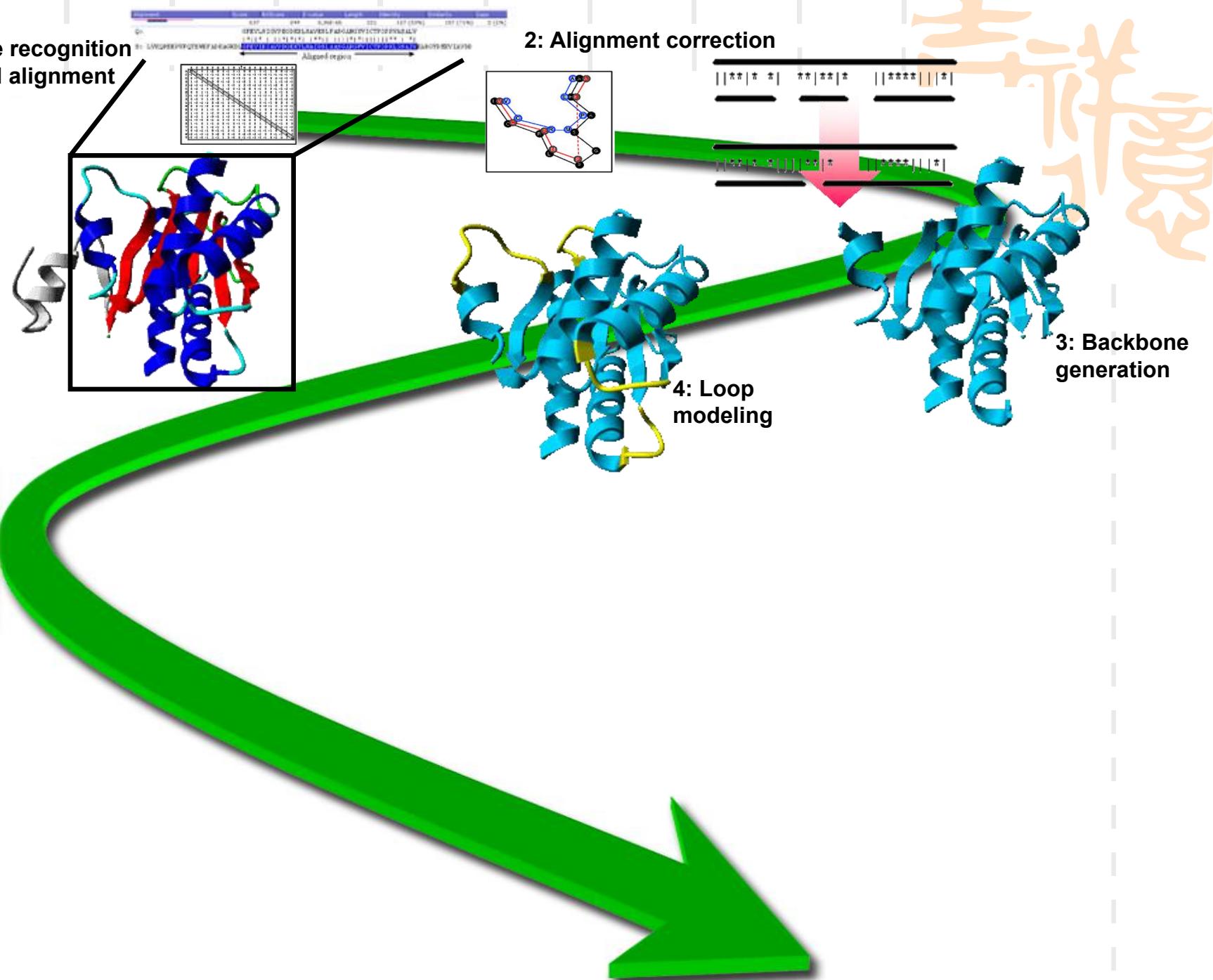
3: Backbone generation



3: Backbone generation

- Making the model.....
- Copy backbone of template to model
- Make deletions as discussed
- (Keep conserved residues)



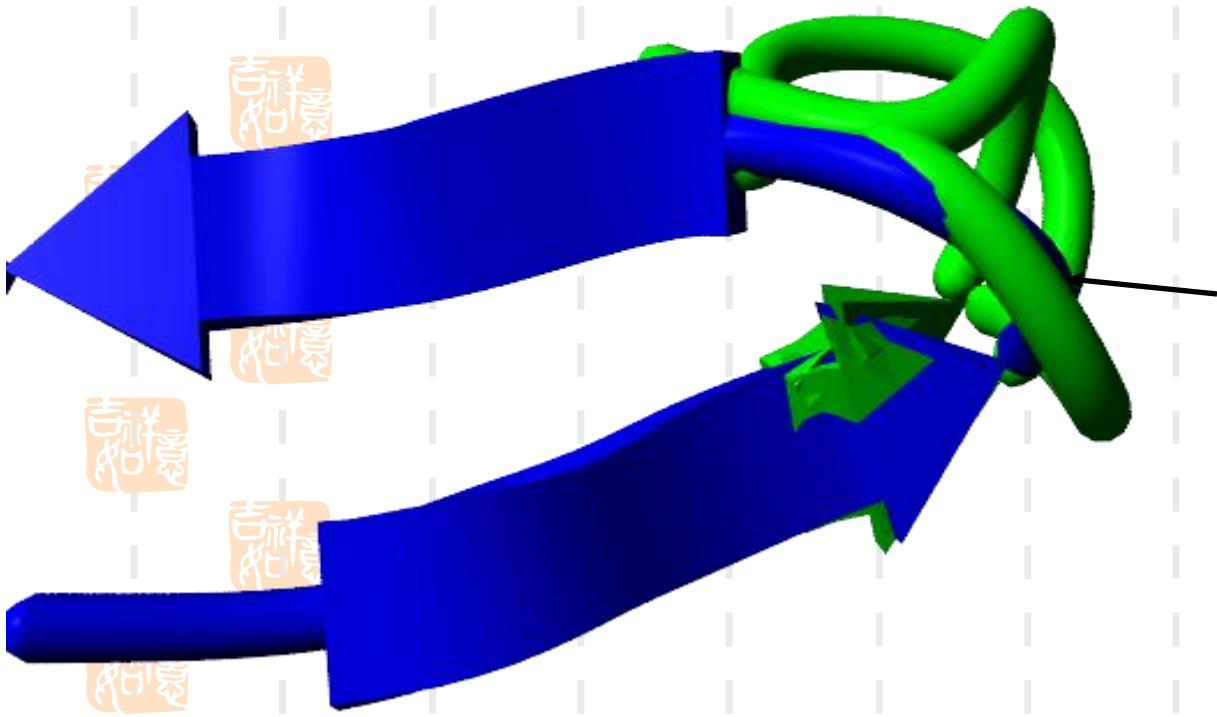


4: Loop modeling

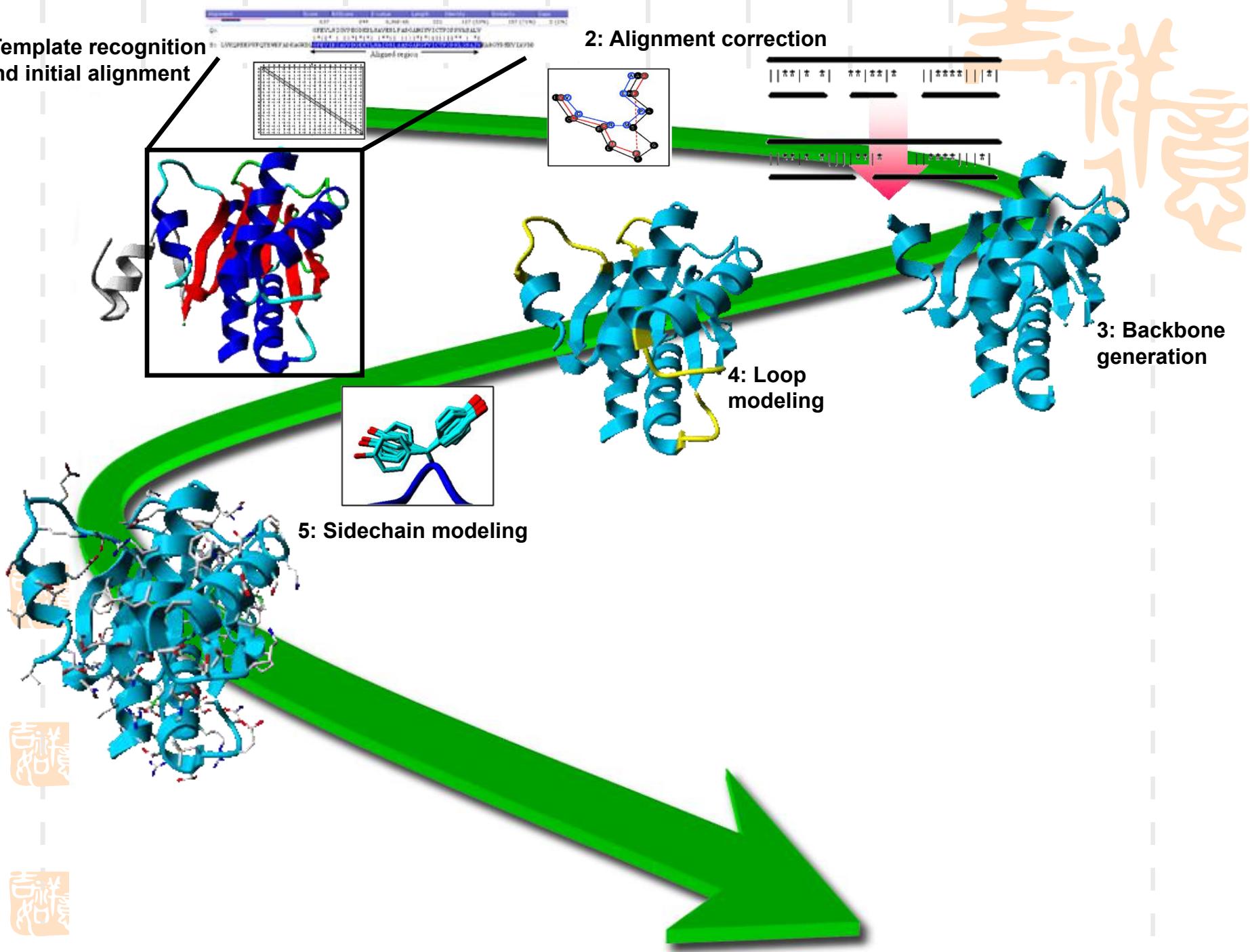
吉祥如意

Known structure **GVC MYIEA---LDKYACNC**

Your sequence **GEC FMV KDLSN P SRY LCKC**

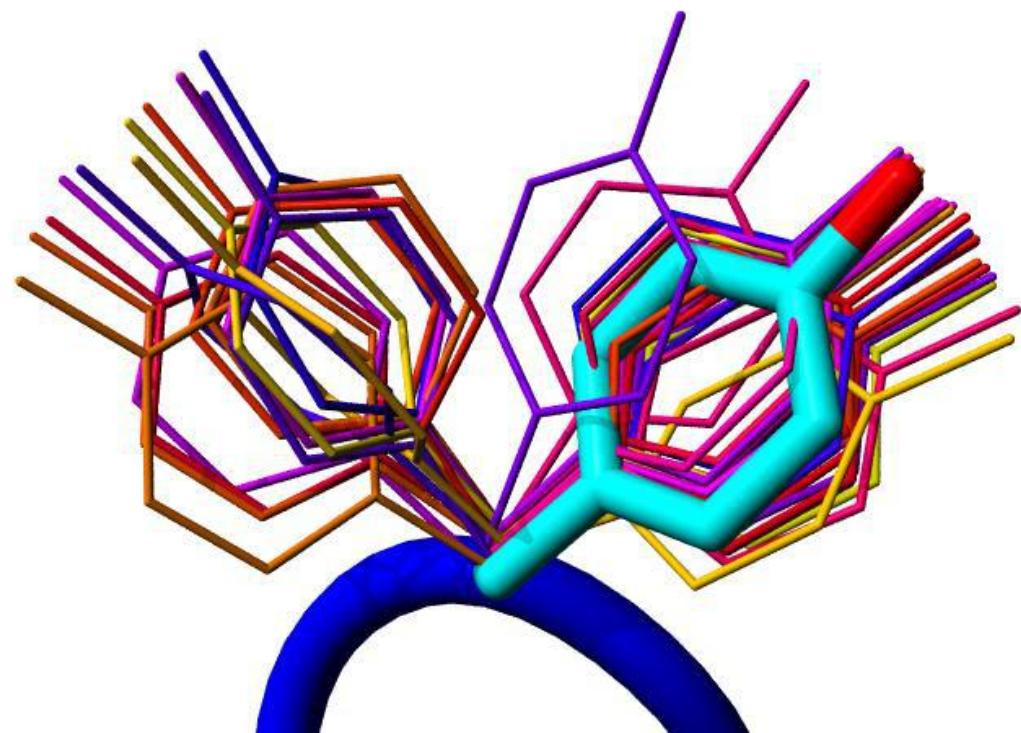


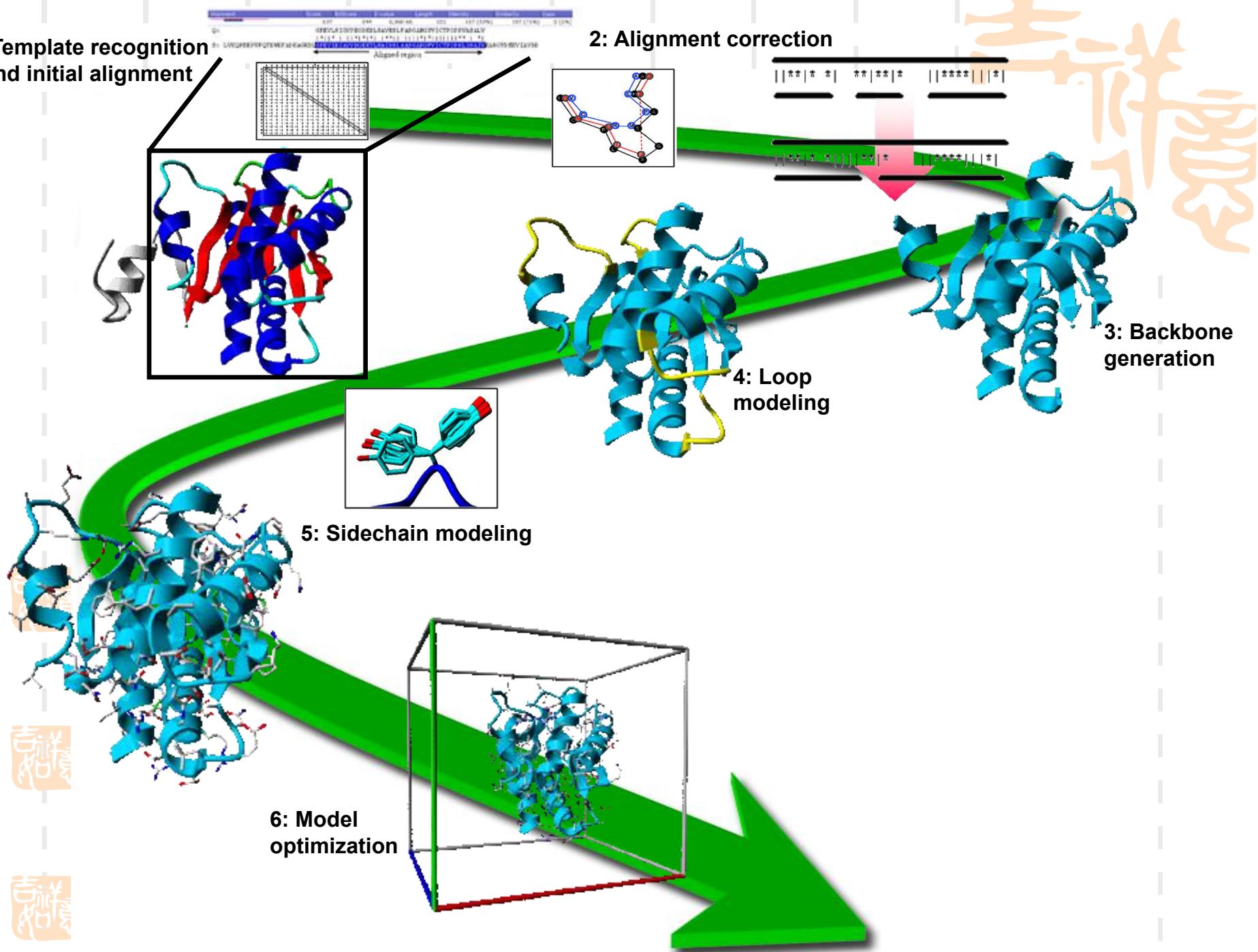
Loop library,
try different
options



5: Side-chain modeling

- Several options
- Libraries of preferred rotamers based upon backbone conformation

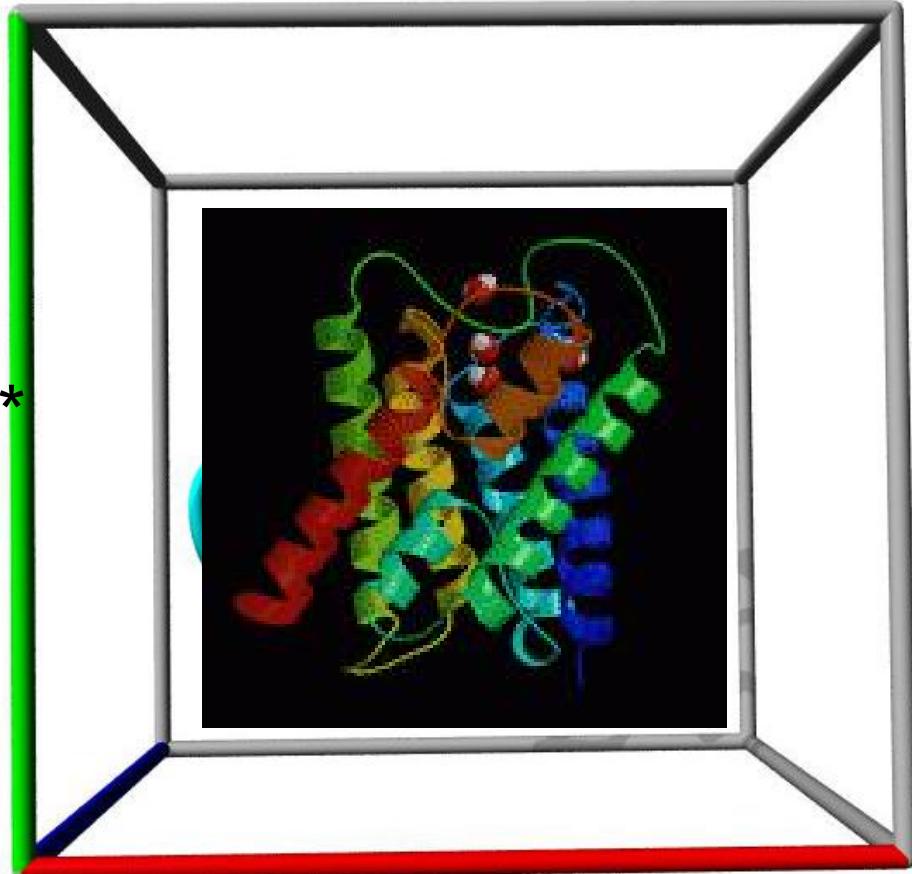




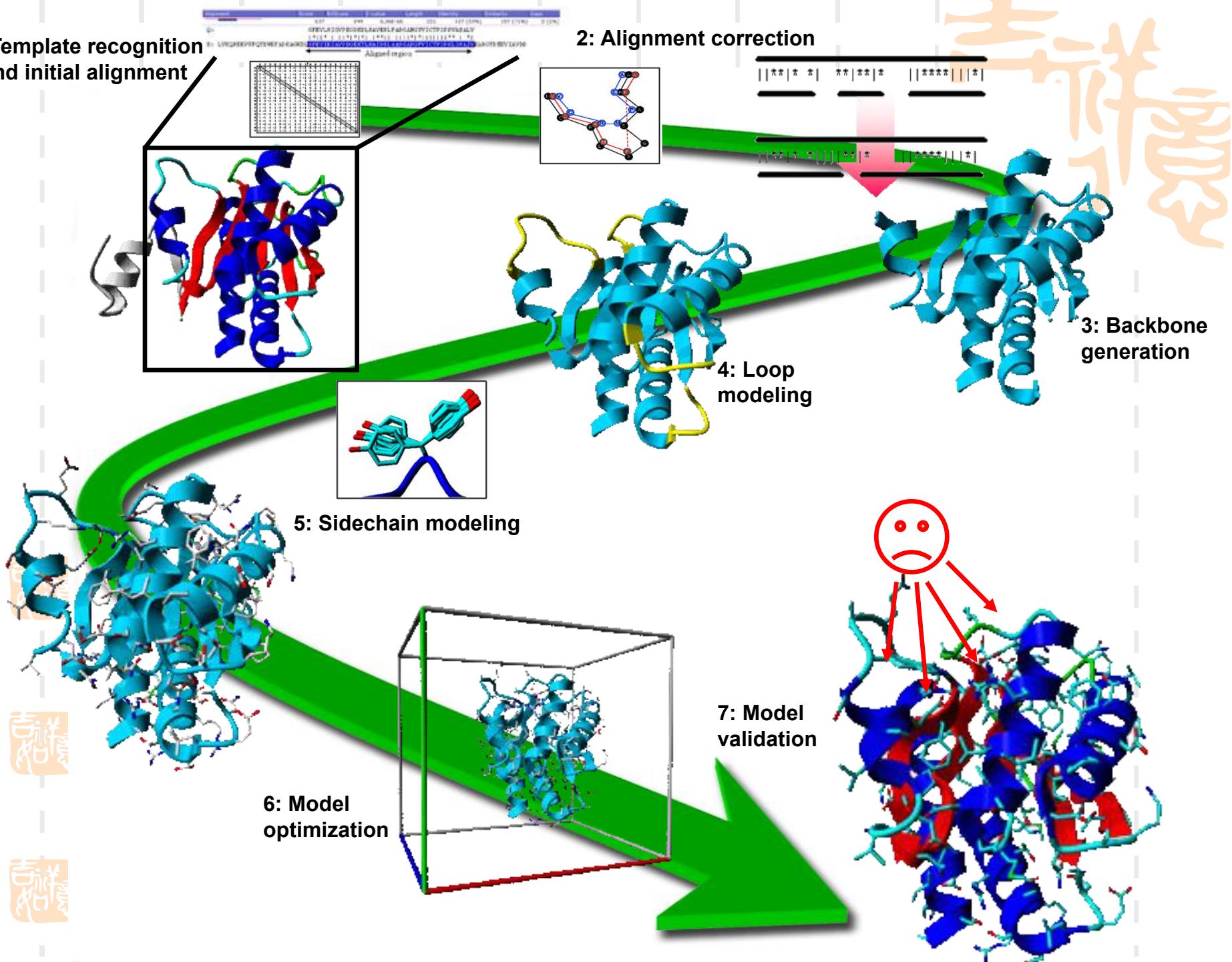


6: Model optimization

- 【 Molecular dynamics simulation
- 【 Remove big errors
- 【 Structure moves to lowest energy conformation*



* Which has nothing to do with the real coordinates...





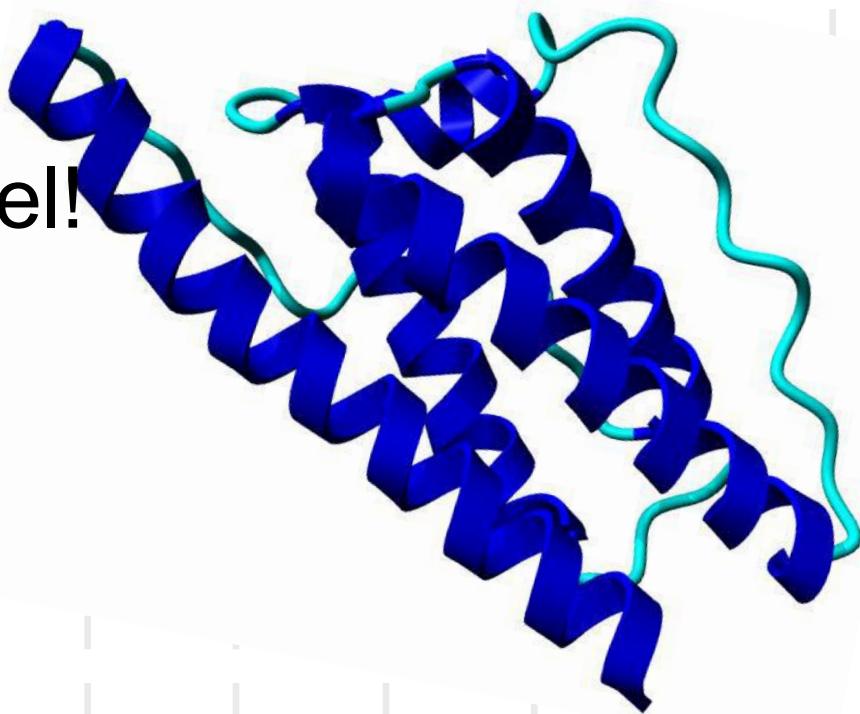
7: Model Validation

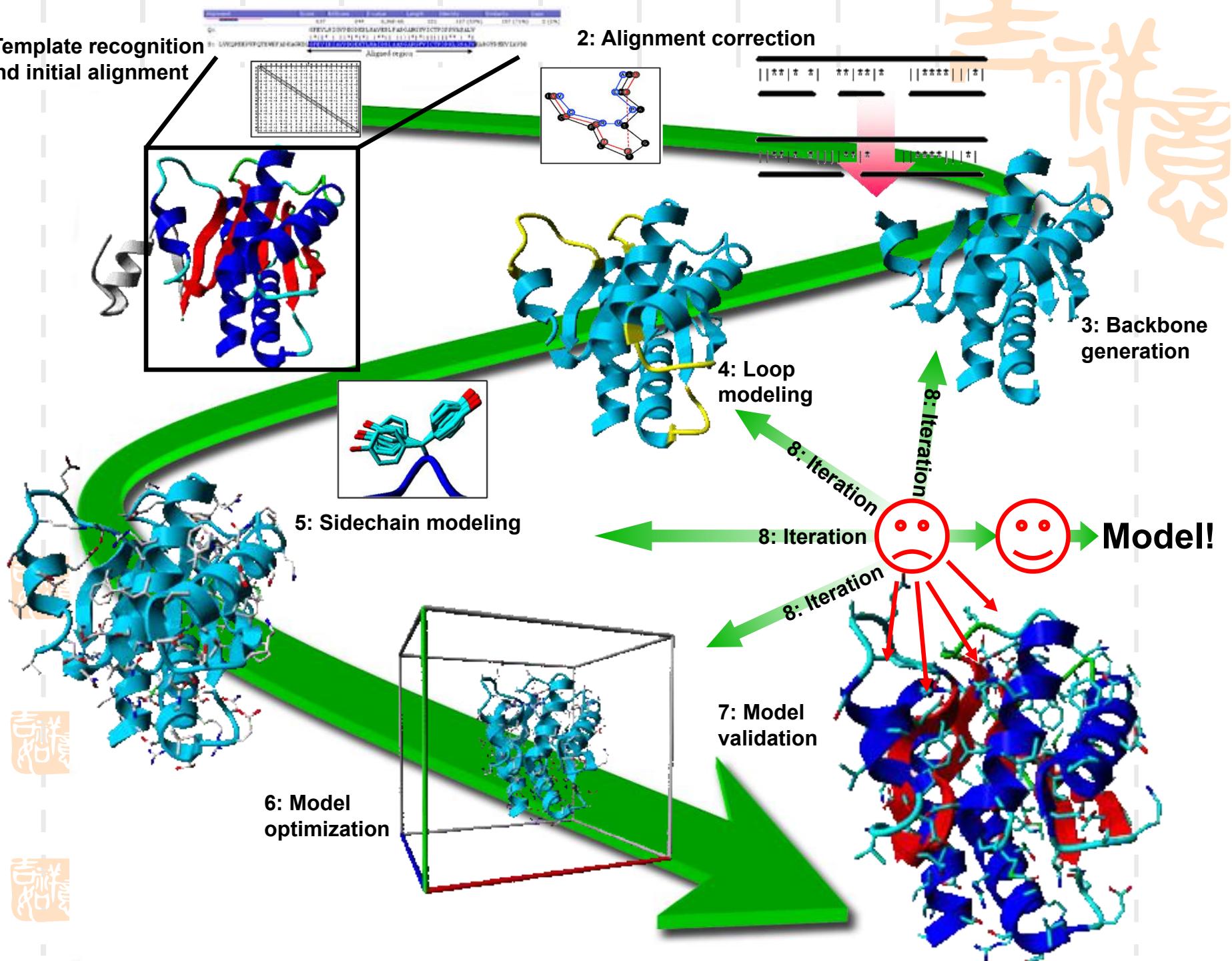
- 【 Second opinion by PDBreport /WHAT IF
- 【 Errors in active site? new alignment/template



No errors?

Model!





吉祥如意

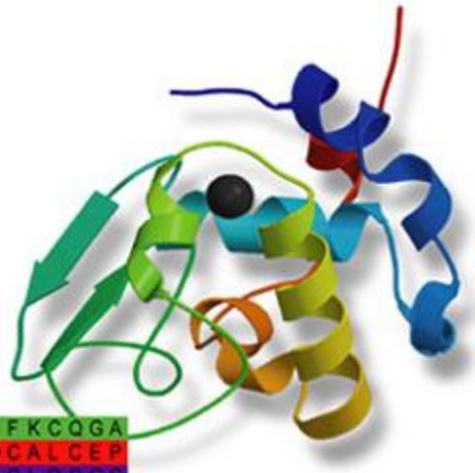
推荐软件

Modeller

Modeller

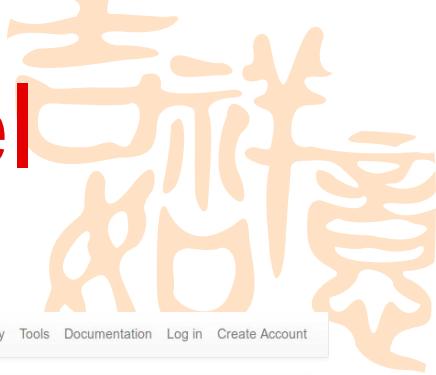
Program for Comparative Protein
Structure Modelling by Satisfaction
of Spatial Restraints

A	I	L	V	G	S	M	P	R	R	D	G	M	E	R	K	D	L	L	K	A	N	V	K	I	F	K	C	O	G	A
V	E	V	C	P	V	D	C	F	Y	E	G	P	N	F	L	V	I	H	P	D	E	C	I	D	C	A	L	C	E	P
G	A	C	K	P	E	C	P	V	N	I	I	Q	G	S	-	-	Y	A	I	D	A	D	S	C	I	D	G	S	S	
C	-	-	I	A	C	G	A	C	K	P	E	C	P	V	N	I	Q	G	S	-	-	I	Y	A	I	D	A	D	S	



推荐软件swiss model

<https://www.swissmodel.expasy.org/>



BIOZENTRUM
University of Basel
The Center for Molecular Life Sciences

Modelling Repository Tools Documentation Log in Create Account

Start a New Modelling Project

Target Sequence:
(Format must be FASTA, Clustal, plain string, or a valid UniProtKB AC)

Paste your target sequence(s) or UniProtKB AC here

+ Upload Target Sequence File... Validate

Project Title: Untitled Project

Email: Optional

Supported Inputs

Sequence(s)
Target-Template Alignment
User Template
DeepView Project

Search For Templates Build Model

By using the SWISS-MODEL server, you agree to comply with the following [terms of use](#) and to cite the corresponding [articles](#).

You are currently not logged in - to take advantage of the workspace, please [log in](#) or [create an account](#).
(There is no requirement to create an account to use any part of SWISS-MODEL, however you will gain the benefit of seeing a list of your previous modelling projects here.)



Swiss Institute of Bioinformatics Schwede Group Contact Us Terms of use

Back to the Top



预测结果

吉祥如意

BIOZENTRUM
University of Basel
The Center for Molecular Life Sciences

SWISS-MODEL

All Projects Untitled Project Created: today at 10:16

Summary Templates 41 Models 1

Model Results

Oligo-State: Monomer. Ligands: None. GMQE: 0.72.

Global Quality Estimate: QMEAN = -3.06, Cg = -2.38, All Atom = -1.64, Solvation = -1.11, Torsion = -1.96.

Local Quality Estimate: GMEAN = -3.06.

Comparison: Comparison with Non-redundant Set of PDB Structures.

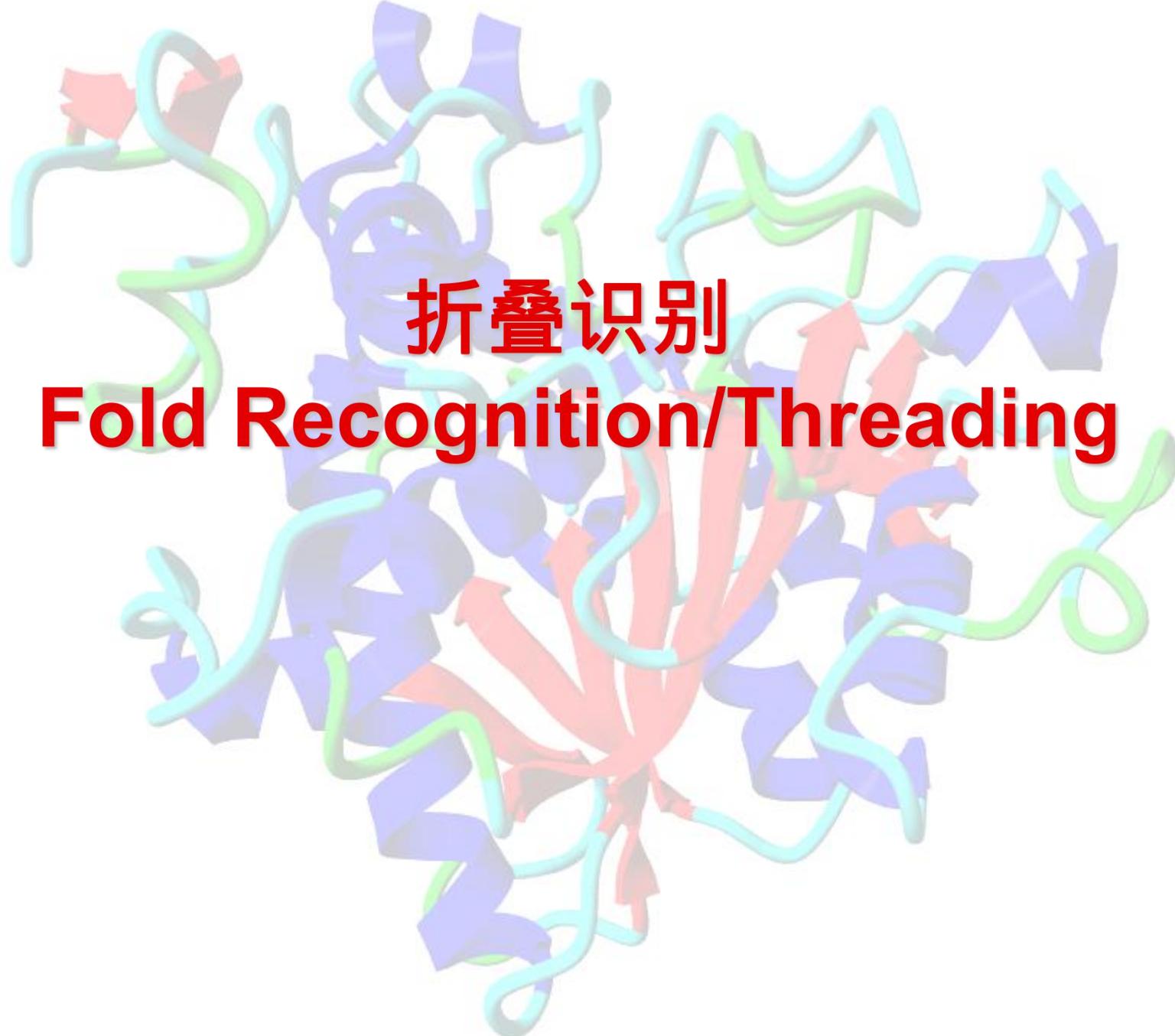
Description: Sialic acid synthase.

Model-Template Alignment:

Model_01	LGKSVVAKVVKIPIPKGTVLTDQDMLAVKAAEPMGIAAEDLCKMVGKTYTEDVEEDDSVMPENVKGYCKNKKC	69
1wov.1.A	VVAKVVKIPIPKGTVLTDQDMLAVKAAEPMGIAAEDLCKMVGKTYTEDVEEDDSVMPENVKGYCKNKKC	73

NGL Cartoon

1 69

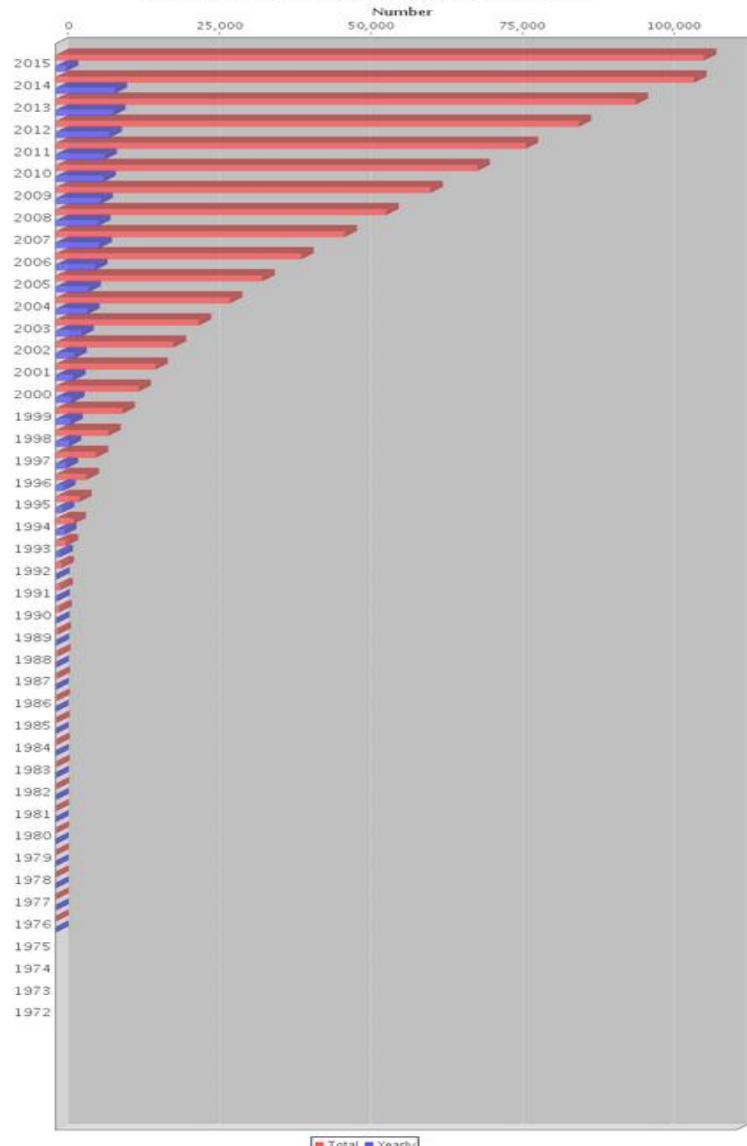


折叠识别 Fold Recognition/Threading

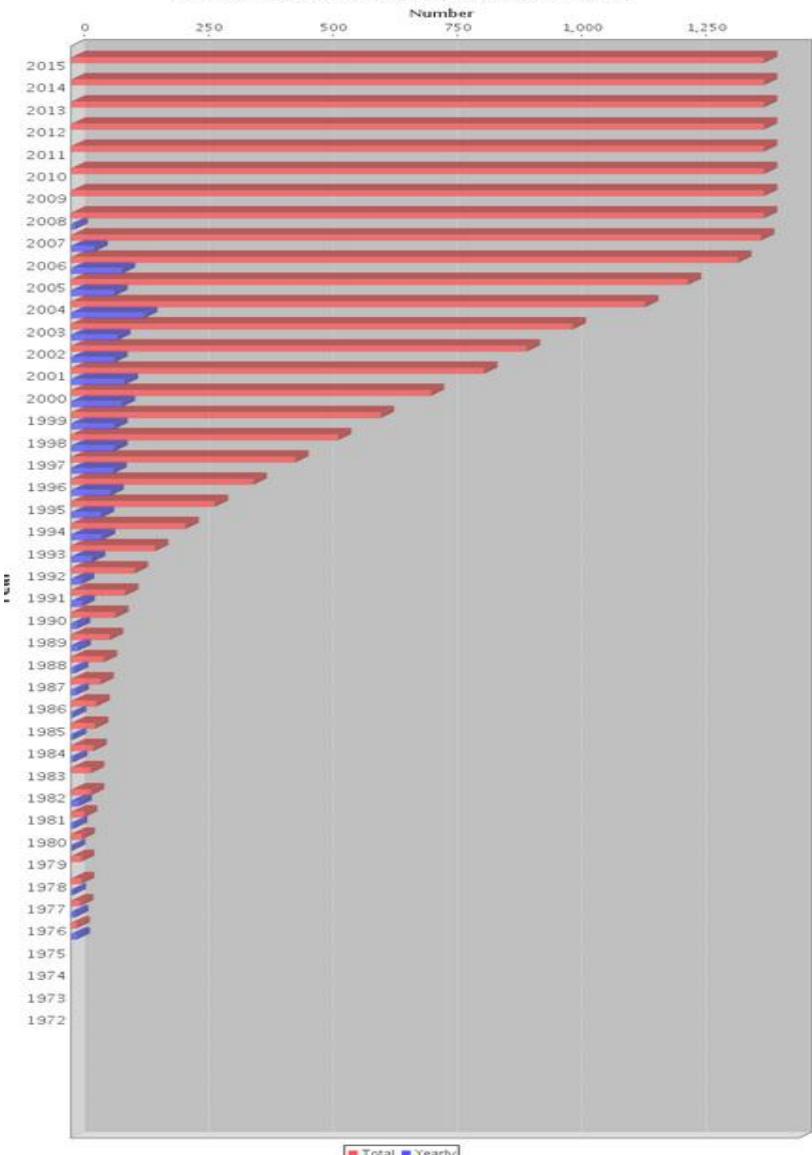
PDB统计

吉祥如意

Yearly Growth of Total Structures
number of structures can be viewed by hovering mouse over the bar

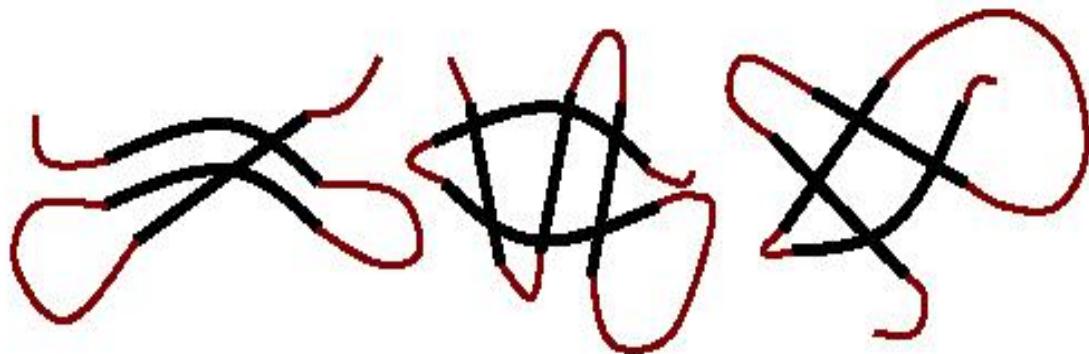
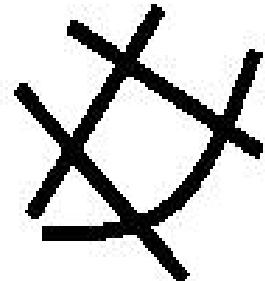
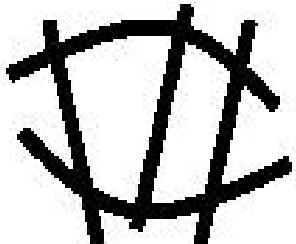
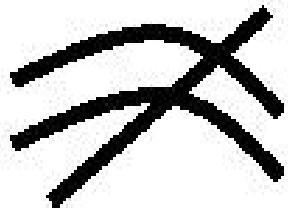


Growth Of Unique Folds Per Year
As Defined By SCOP (v1.75)
number of folds can be viewed by hovering mouse over the bar



折叠识别

叠
叠
叠
叠
叠



福

吉

福

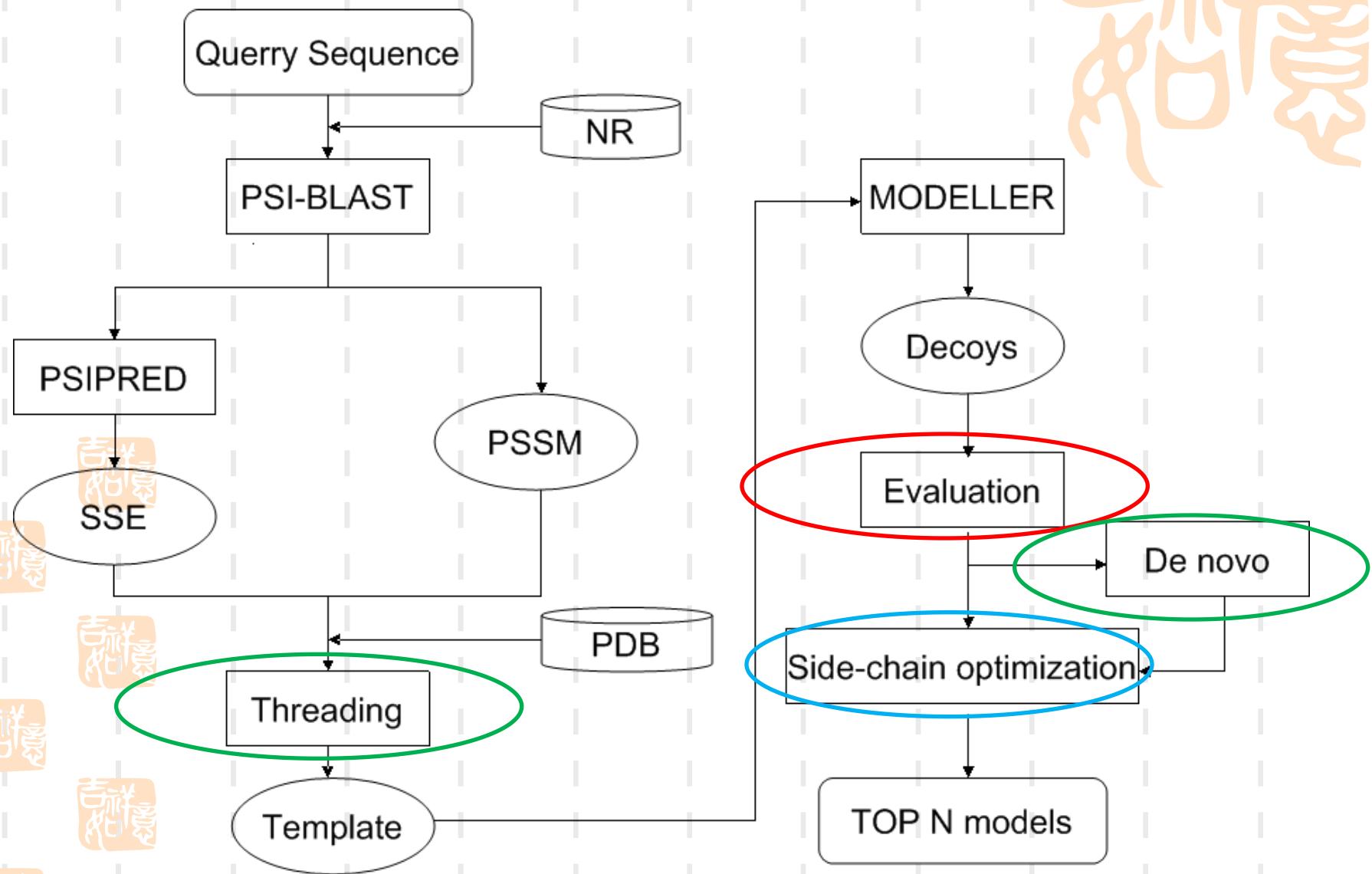
吉

福

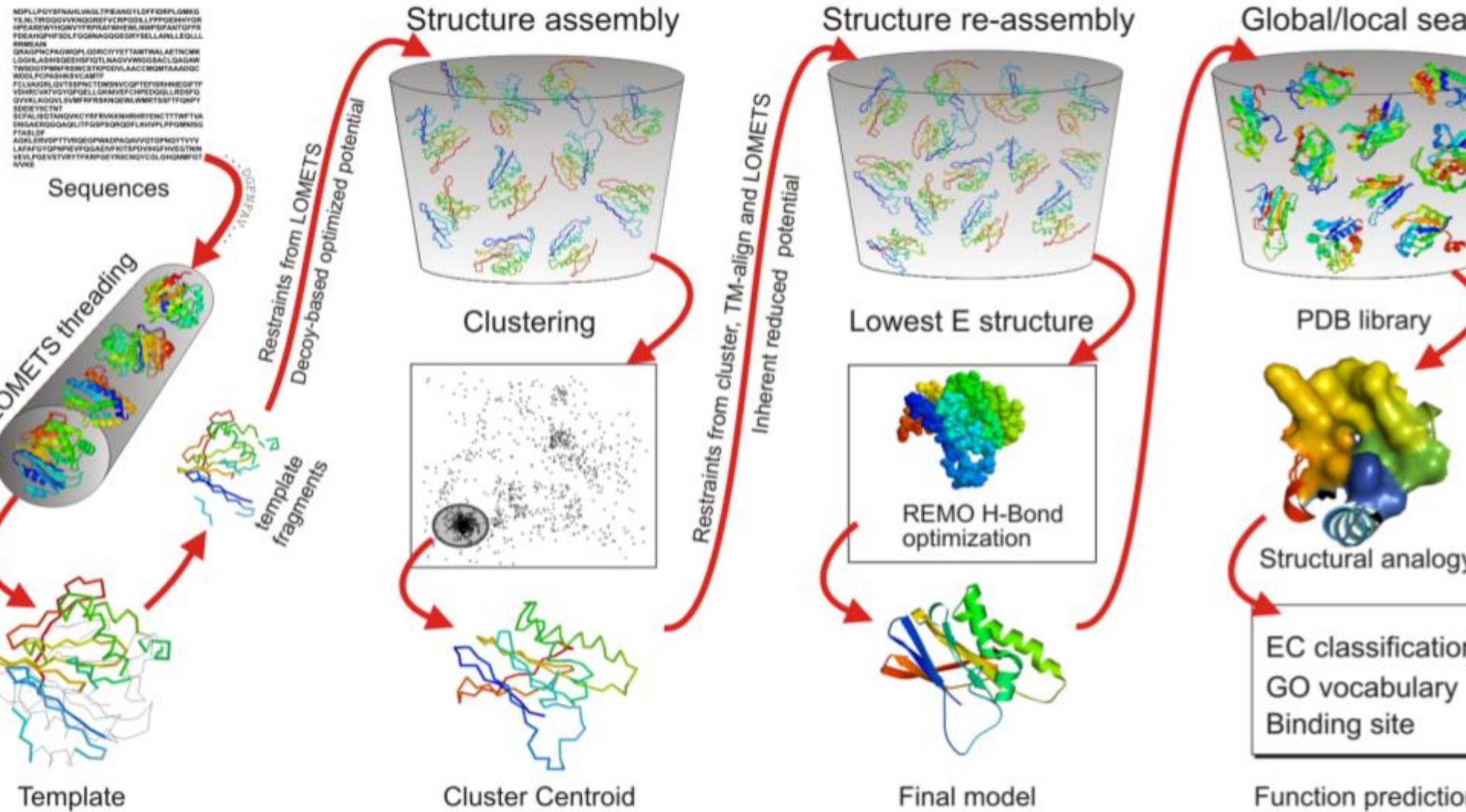
吉

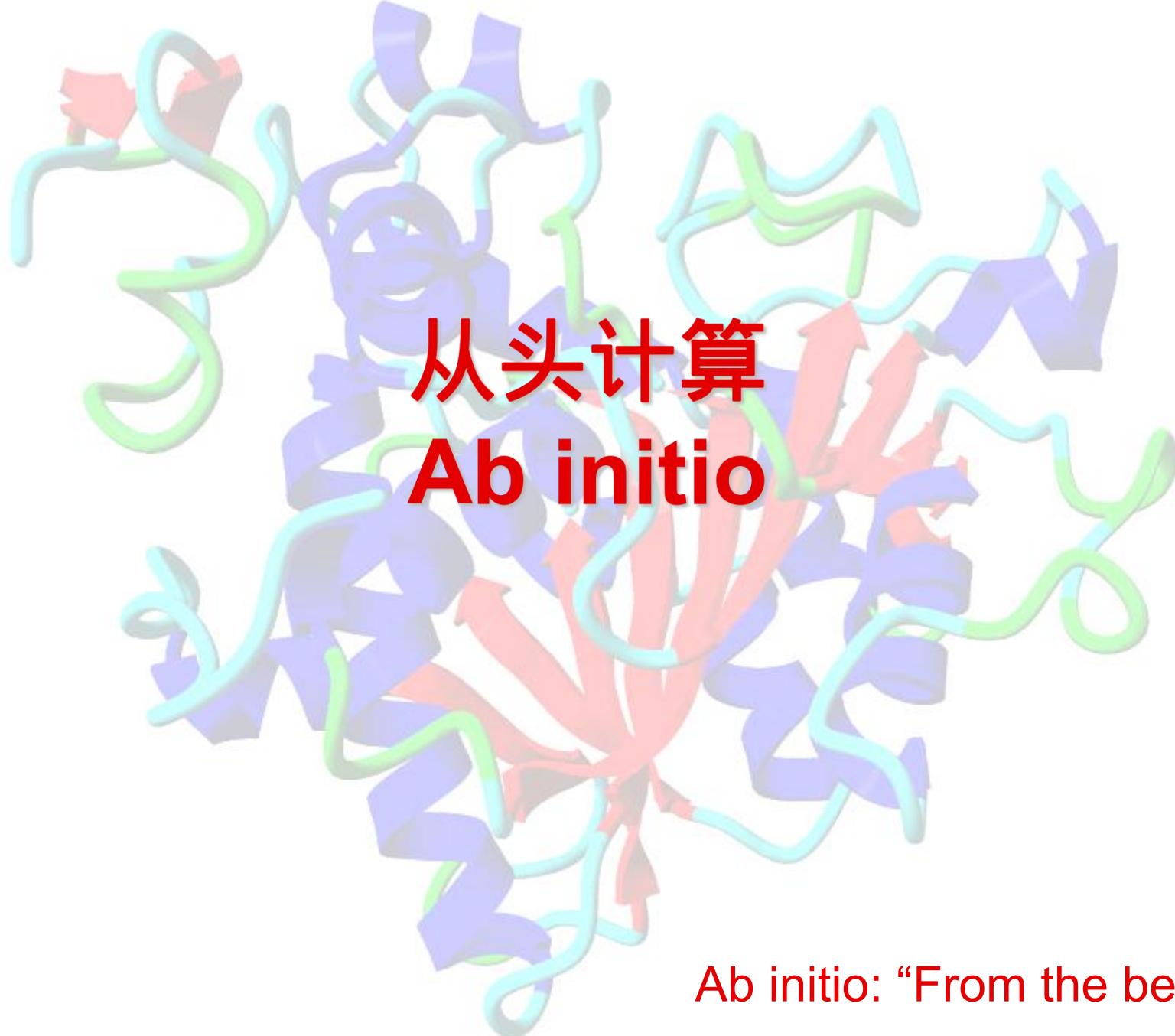


The Frame work of Jiang_Server1.0



(Iterative Threading ASSEmbley Refinement)

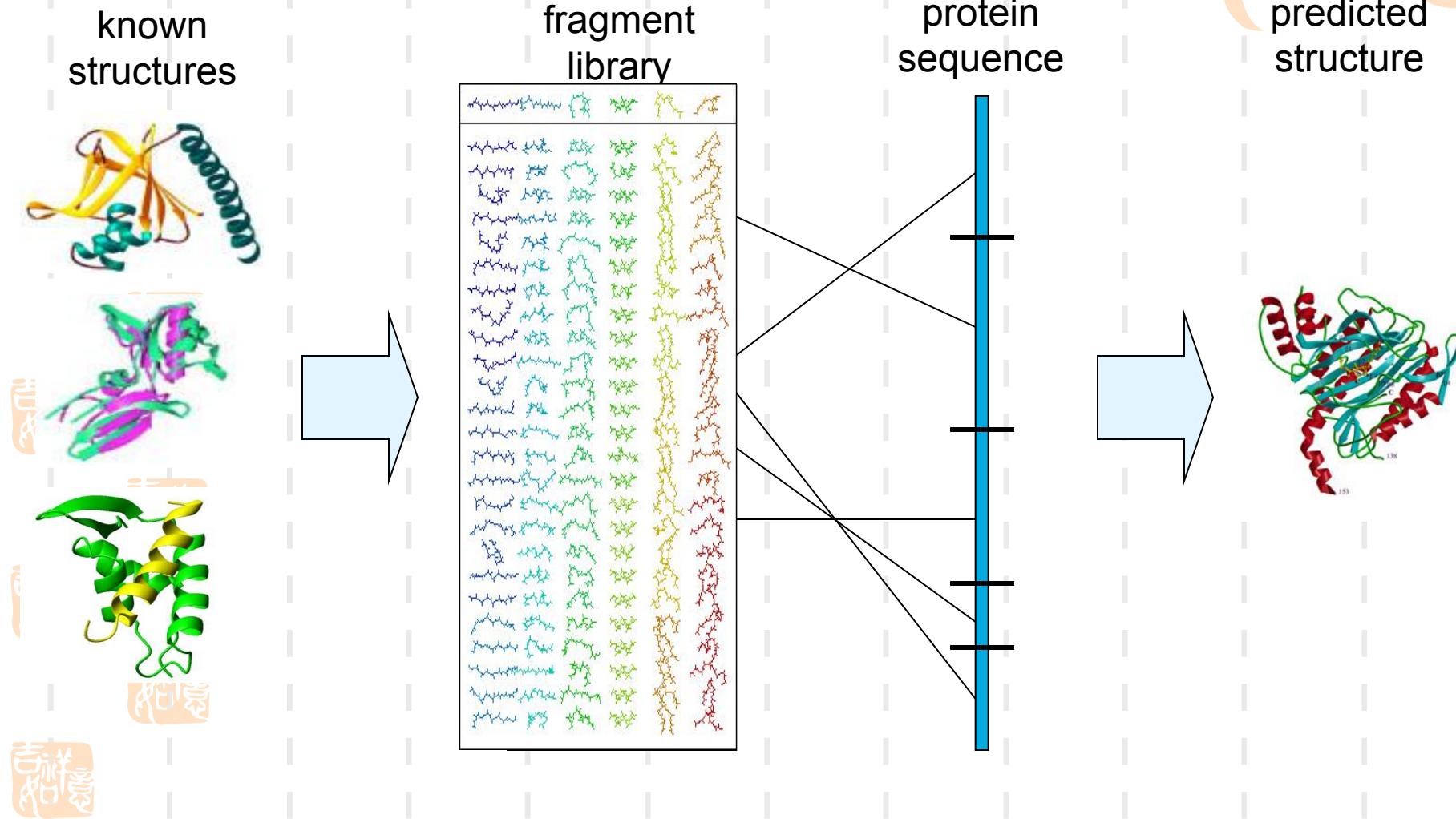




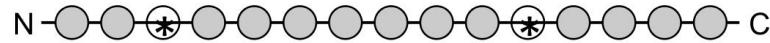
从头计算
Ab initio

Ab initio: “From the beginning”

Assembly of sub-structural units

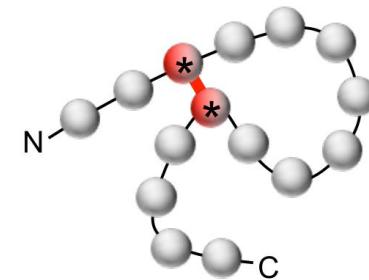
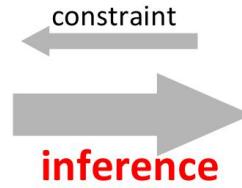


结构预测新技术：残基接触预测(Contact)



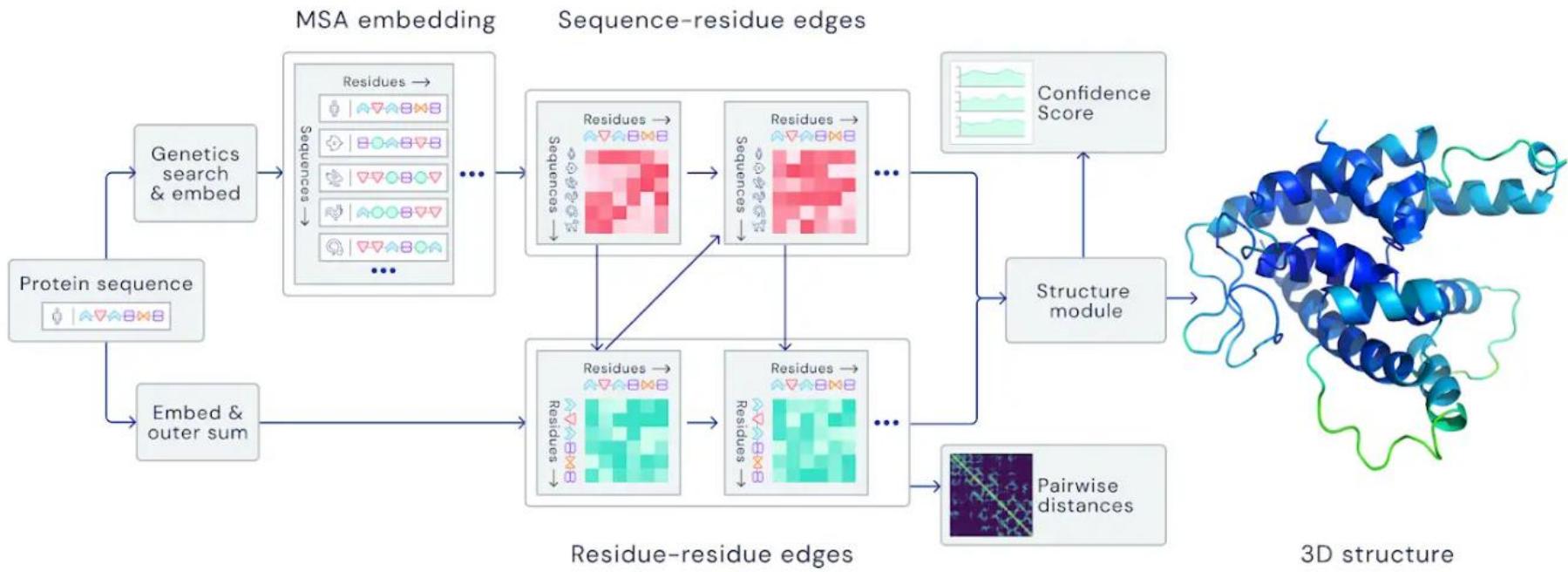
A	T	R	L	T	L	T	A	K	K	D	G	P	C	D
A	T	R	L	T	L	T	A	K	K	D	G	P	C	D
A	T	R	L	T	L	T	A	K	K	D	G	P	C	D
A	T	K	L	C	L	T	A	K	K	E	G	P	K	D
A	T	K	L	T	L	T	A	K	K	E	G	P	K	D
A	T	K	L	T	L	G	A	K	K	E	G	G	C	D
A	T	W	L	T	L	T	A	K	K	V	G	P	C	D
A	T	W	L	T	L	T	A	K	K	V	G	P	C	D

correlated



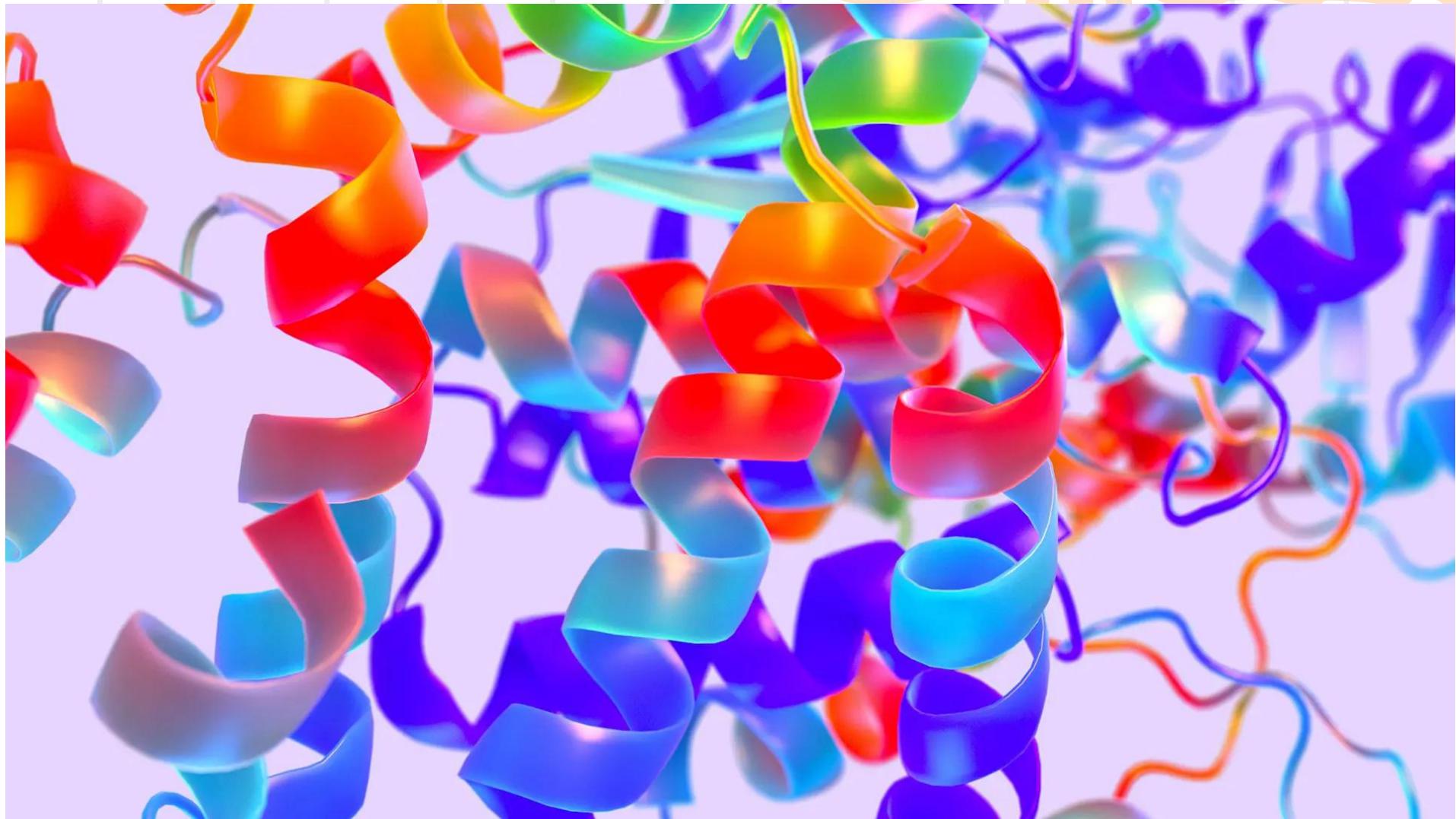
contact in 3D

结构预测新技术：AlphaFold2 算法



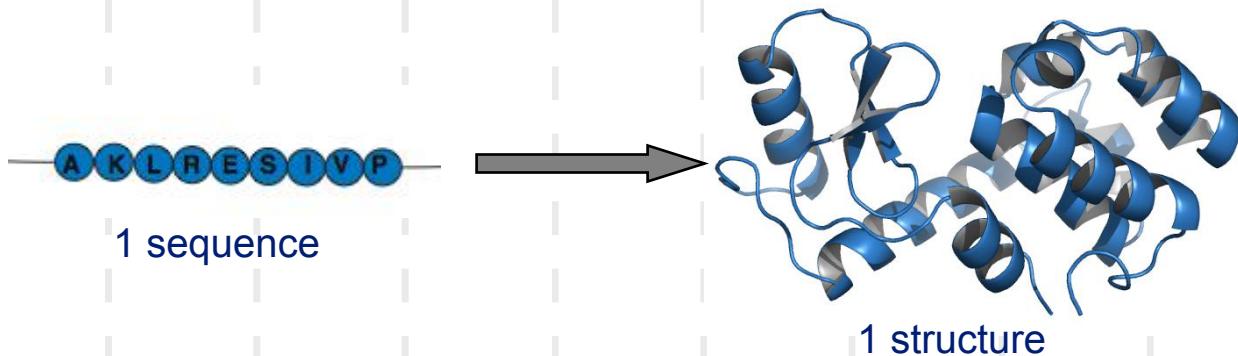
蛋白质设计

Protein Design

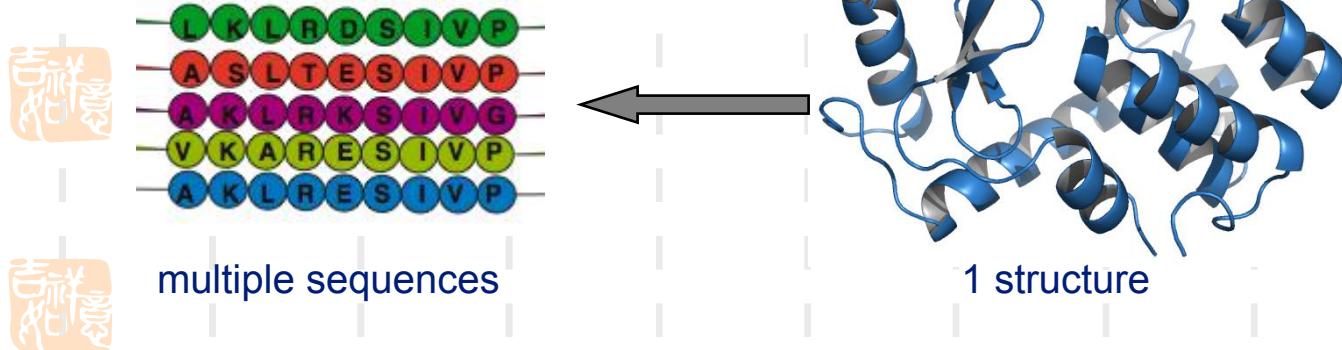


吉祥如意

蛋白质结构预测

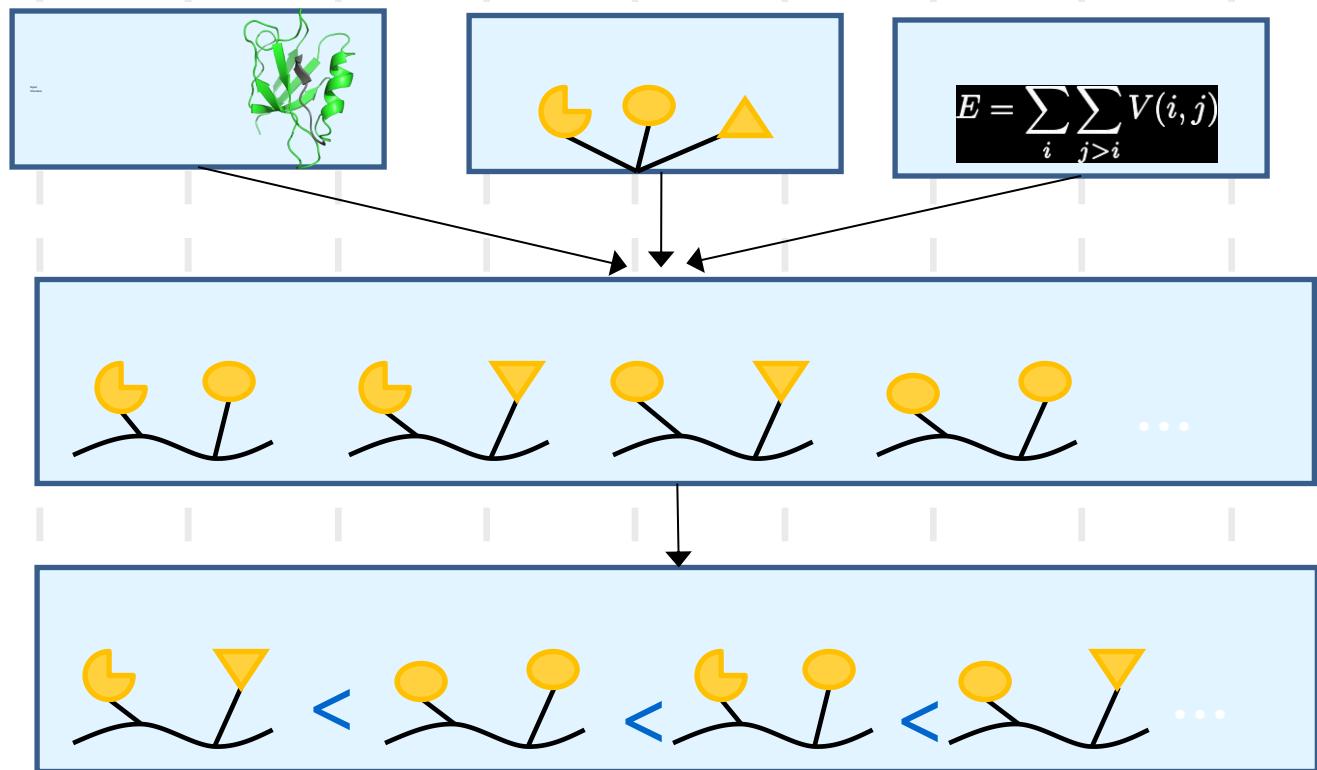
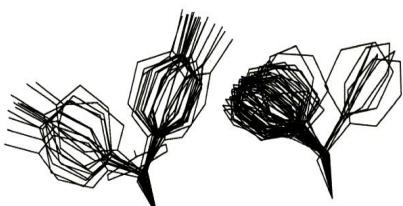
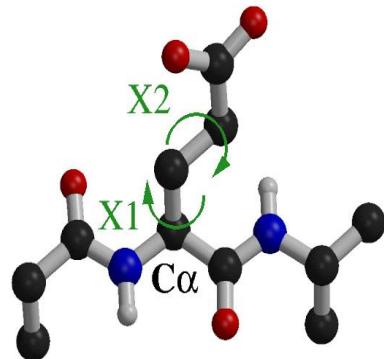


蛋白质设计



吉祥如意

蛋白质设计流程



吉祥如意

Rotamers

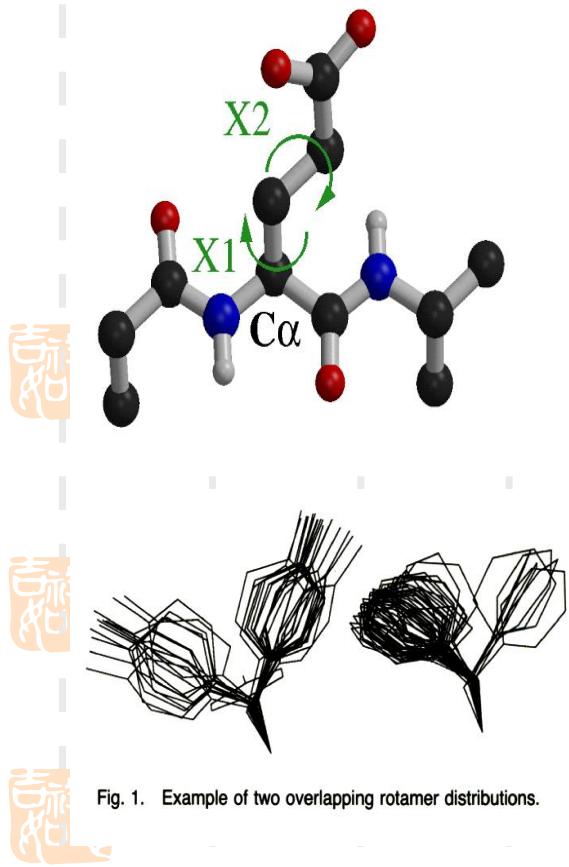
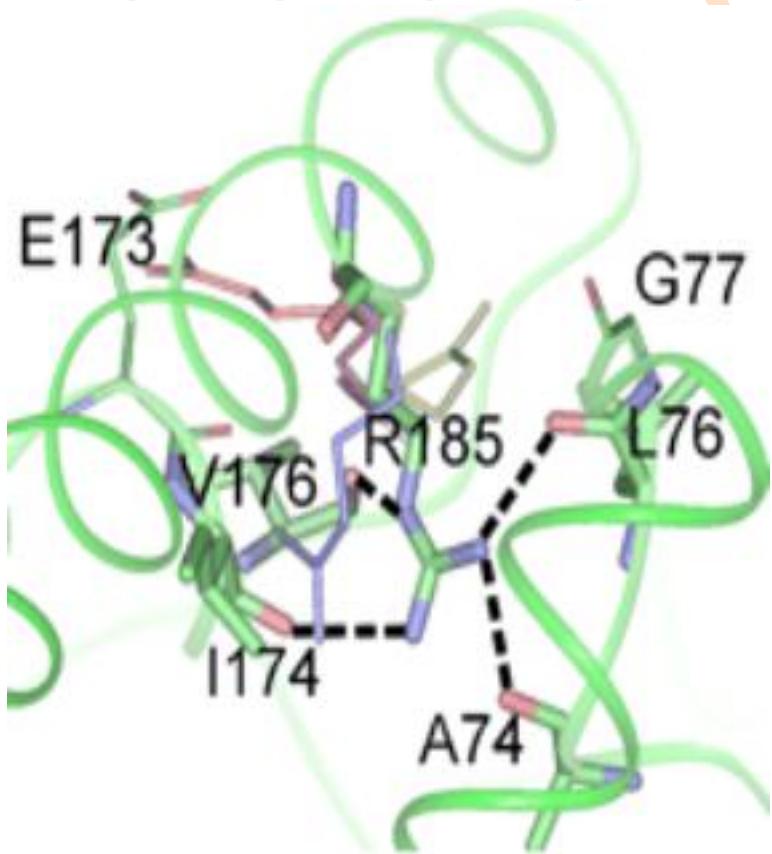


Fig. 1. Example of two overlapping rotamer distributions.



吉祥如意

搜索算法



heuristic
(MC, SCMF, GA)

3 4

6

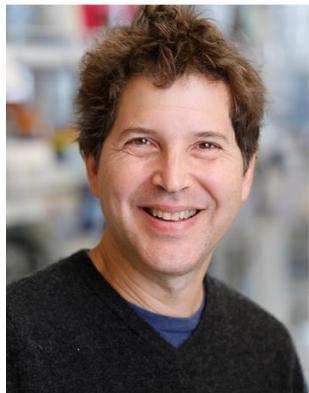
Enumeration
with gaps

provable
(A* Enumeration)

1 2 3 4 5 6 7

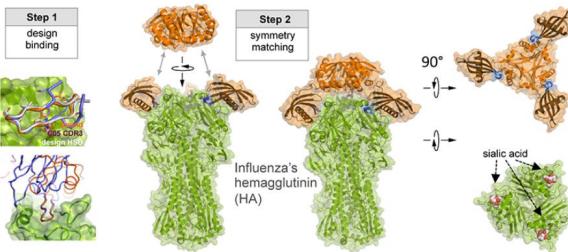
Gap free
enumeratio
n

Protein Design



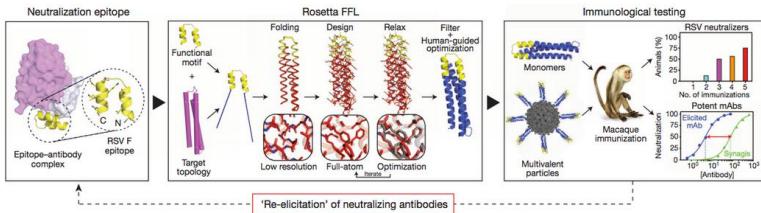
David Baker

Computational design of Flu Glue

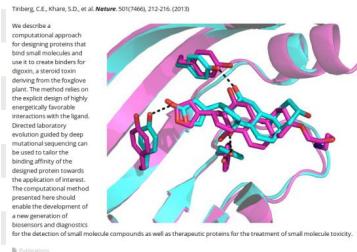


Proof of principle for epitope-focused vaccine design

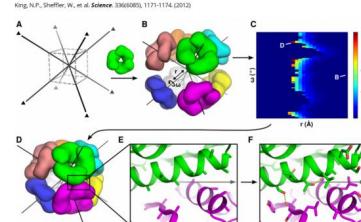
Correia, B.E. et al. *Nature* 507, 201-6 (2014)

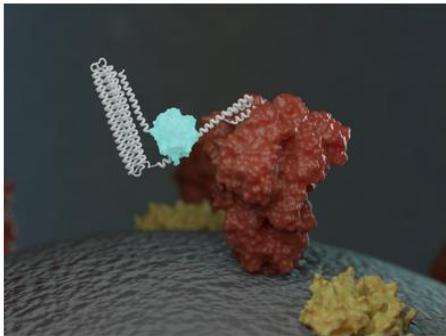


Computational design of ligand-binding proteins with high affinity and selectivity



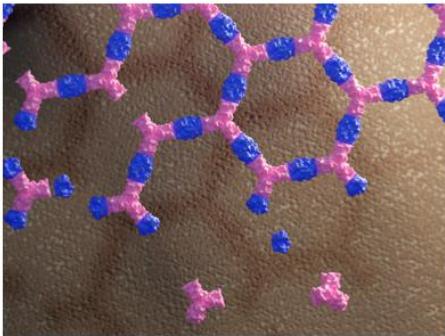
Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy





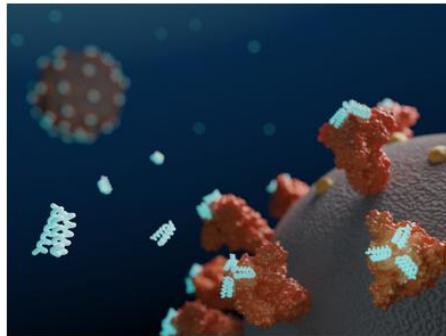
De novo design of modular protein biosensors

JANUARY 28, 2021



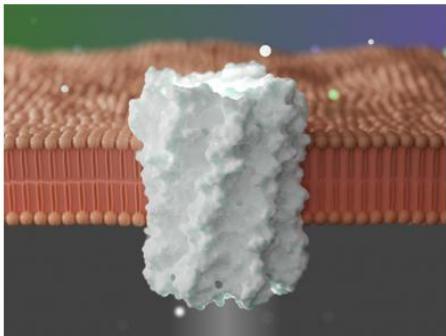
Biologically active 2D arrays

JANUARY 6, 2021



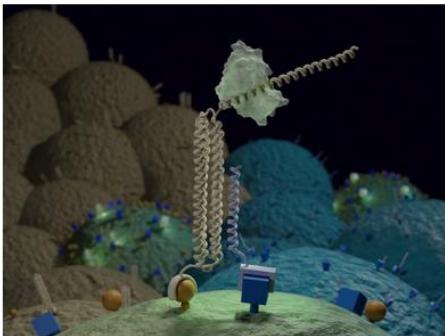
De novo minibinders target SARS-CoV-2 Spike protein

SEPTEMBER 9, 2020



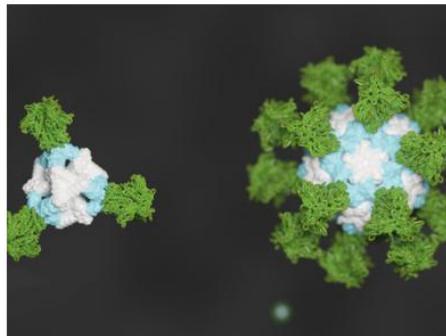
Selective ion channels designed from scratch

AUGUST 31, 2020



Introducing Co-LOCKR: designed protein logic for cell targeting

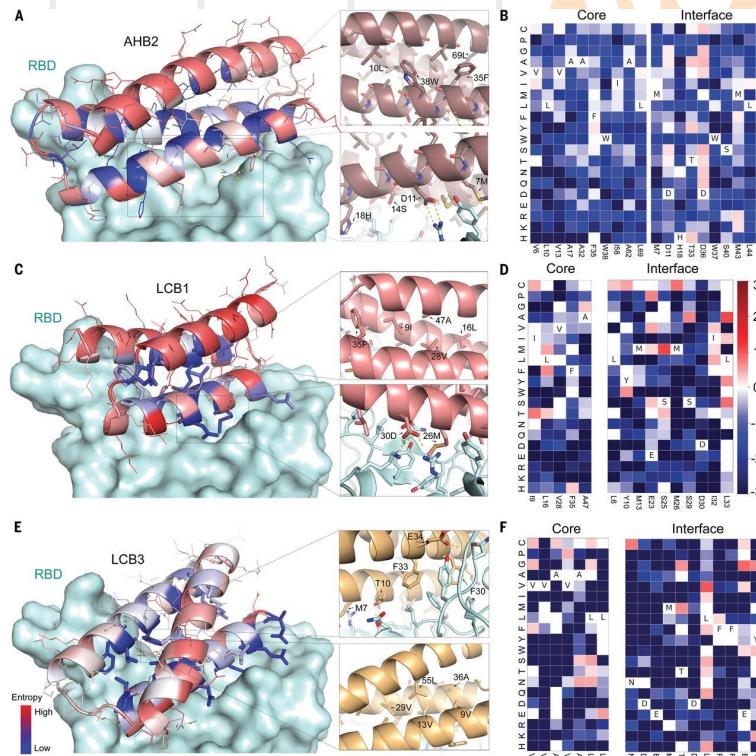
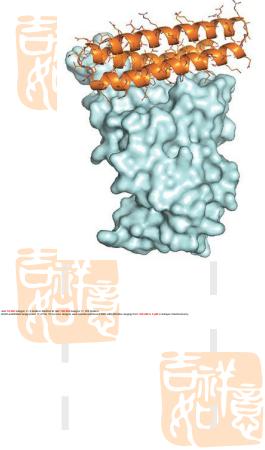
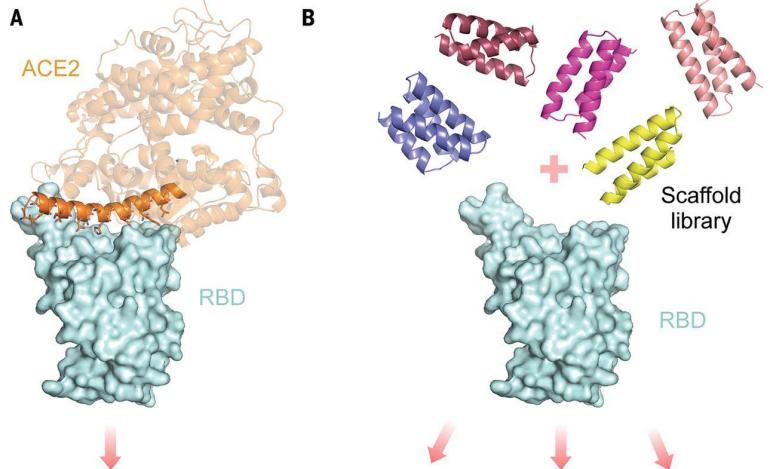
AUGUST 20, 2020

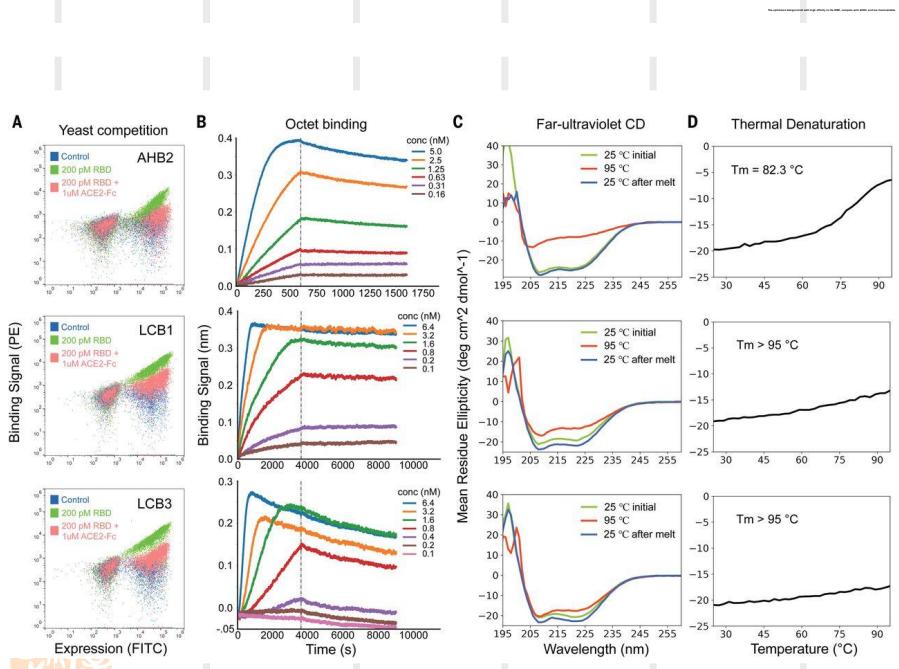
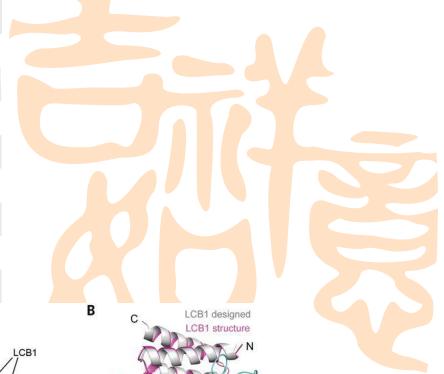


De novo nanoparticles as vaccine scaffolds

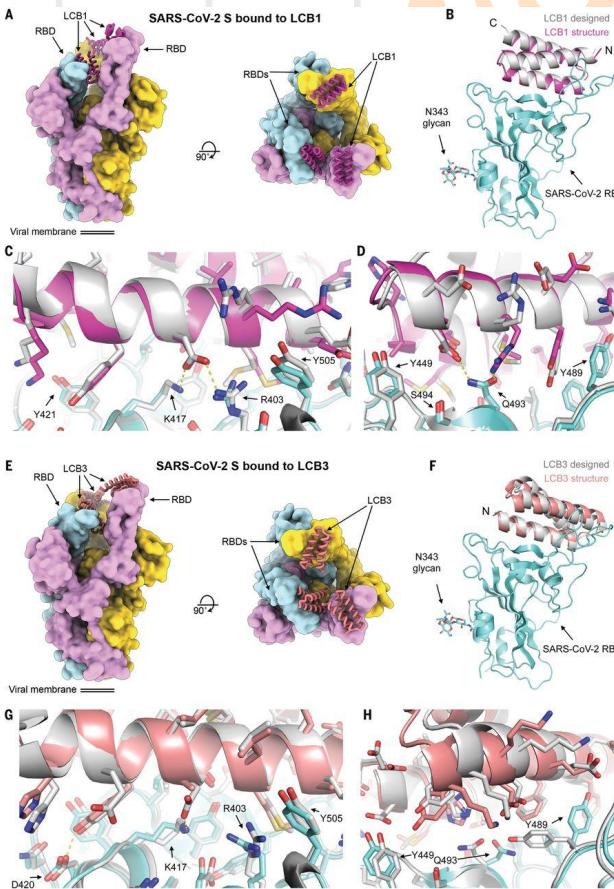
AUGUST 5, 2020







For most of these minibinders, a small number of substitutions were enriched in the FACS sorting; combinatorial libraries incorporating these substitutions were constructed for the ACE2-based design and the eight highest-affinity approach 2 designs and again screened for binding to the RBD at concentrations down to **20 pM**. Each library converged on a small number of closely related sequences; one of these was selected for each design, AHB2 or LCB1-LCB8, and found to bind the RBD with high affinity on the yeast surface in a manner competed with by ACE2.





De novo vaccine scaffolds

King Lab @KingLabIPD · 1月28日

It's a historic day for the lab: Our candidate nanoparticle #SARS-CoV2 vaccine has been injected into humans — the first in-person test of our immunogen design platform.

Thanks to SK Bioscience 🇰🇷 for leading this Phase 1 trial.

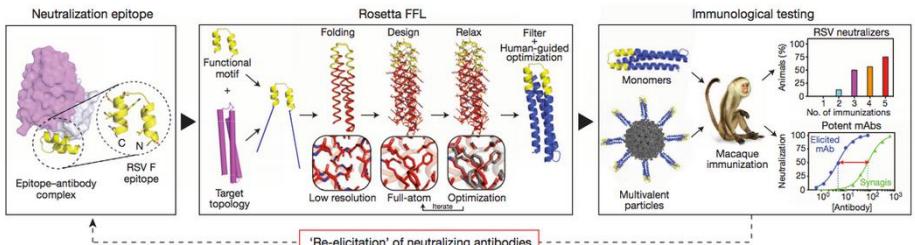
[cell.com/cell/fulltext/...](http://cell.com/cell/fulltext/)

The Veesler Lab 及 另外 3 个人

14 147 680

Proof of principle for epitope-focused vaccine design

Correia, B.E. et al. *Nature* 507, 201-6 (2014)



https://fold.it/

... 吉祥如意 吉祥如意 吉祥如意 吉祥如意 吉祥如意

foldit BETA

Solve Puzzles for Science

14:56:44 GMT

PUZZLES Blog CATEGORIES GROUPS FORUM PLAYERS WIKI FAQ RECIPES ABOUT CONTESTS CREDITS

Click to learn how you contribute to science by playing Foldit.

CARTO SCOPE CONTRIBUTE TO CITIZEN SCIENCE MAPPING!

What's New

Newsletter April 2: No Fools Here

Hey folders! Newsletter coming at ya--take a look [here!](#)
(Tue, 04/06/2021 - 01:48 | 0 comments)

Hotfixes to Main and Devprev soon!

Hey all, we're going to be pushing some hotfixes to both main and devprev here soon.

For both releases:

- * Fixed a crash on the Reaction Design intro puzzle.
- * Fixed a scoring issue affecting some design puzzles.
- * Fixed an issue where the objectives panel wouldn't re-compute bonuses when it should.

For devprev only:

- * Fixed a bug causing moves to not be counted
- * Fixed a bug causing Dojo mode to soft lock when ending the level using a recipe
- * Fixed a bug where Dojo mode would connect the first level many times

USER LOGIN

Username: *
Password: *

• Create new account
• Request new password

DOWNLOAD LINKS:

[Download Windows](#) [Download Mac](#) [Download Linux](#)

Windows (7/8/10) OSX (10.12 or later) Linux (64-bit)

Are you new to Foldit? Click here.
Are you a student? Click here.
Are you an educator? Click here.

SOCIAL MEDIA

JOIN OUR DISCORD

SEARCH

Google Search Only search fold.it