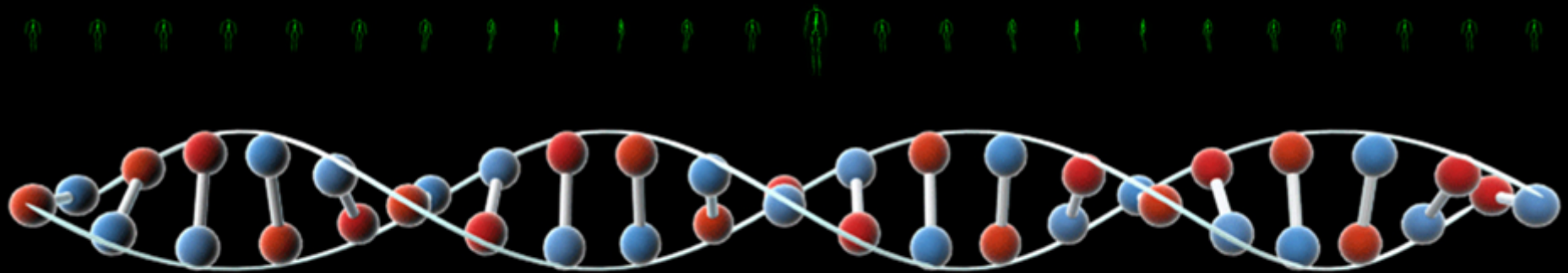


# STRUCTURE :TOOLS & PREDICTION



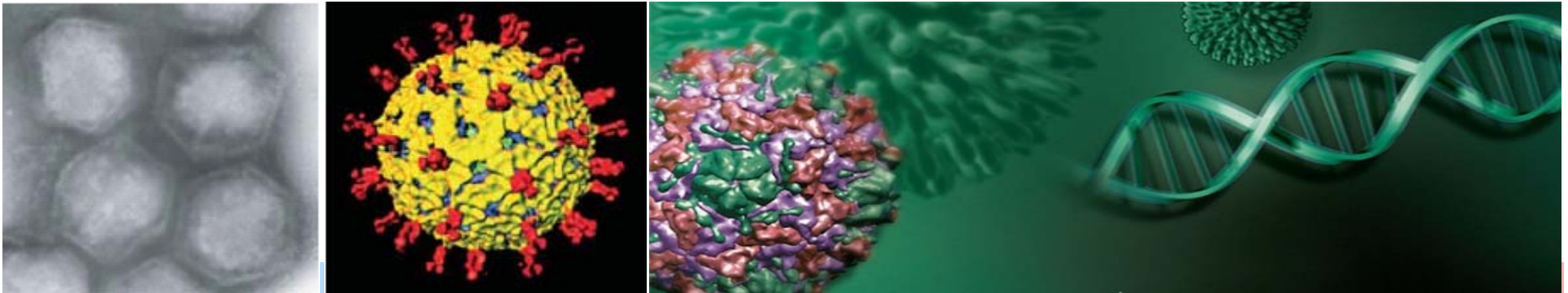
*Zhu Hongwei, Cao Zhongzan, Ma Wenge, Zheng Aijuan*

## 蛋白质三维结构的可视化工具

- ❖ Swiss PDBViewer
- ❖ RasMol
- ❖ Cn3D

## 蛋白质三维结构预测

- ❖ 蛋白质结构解析(X-Ray,NMR)
- ❖ homology modeling:SWISS-MODEL
- ❖ Threading
- ❖ 结构预测相关程序及数据库



## ➤ Swiss PDB Viewer

- ❖ 特点：界面友好、可同时分析几个PDB文件、可叠加起来分析结构类似性…
- ❖ 可与Swiss-Model服务器连接，从软件直接连到Swiss-Model服务器进行理论蛋白立体结构构建。



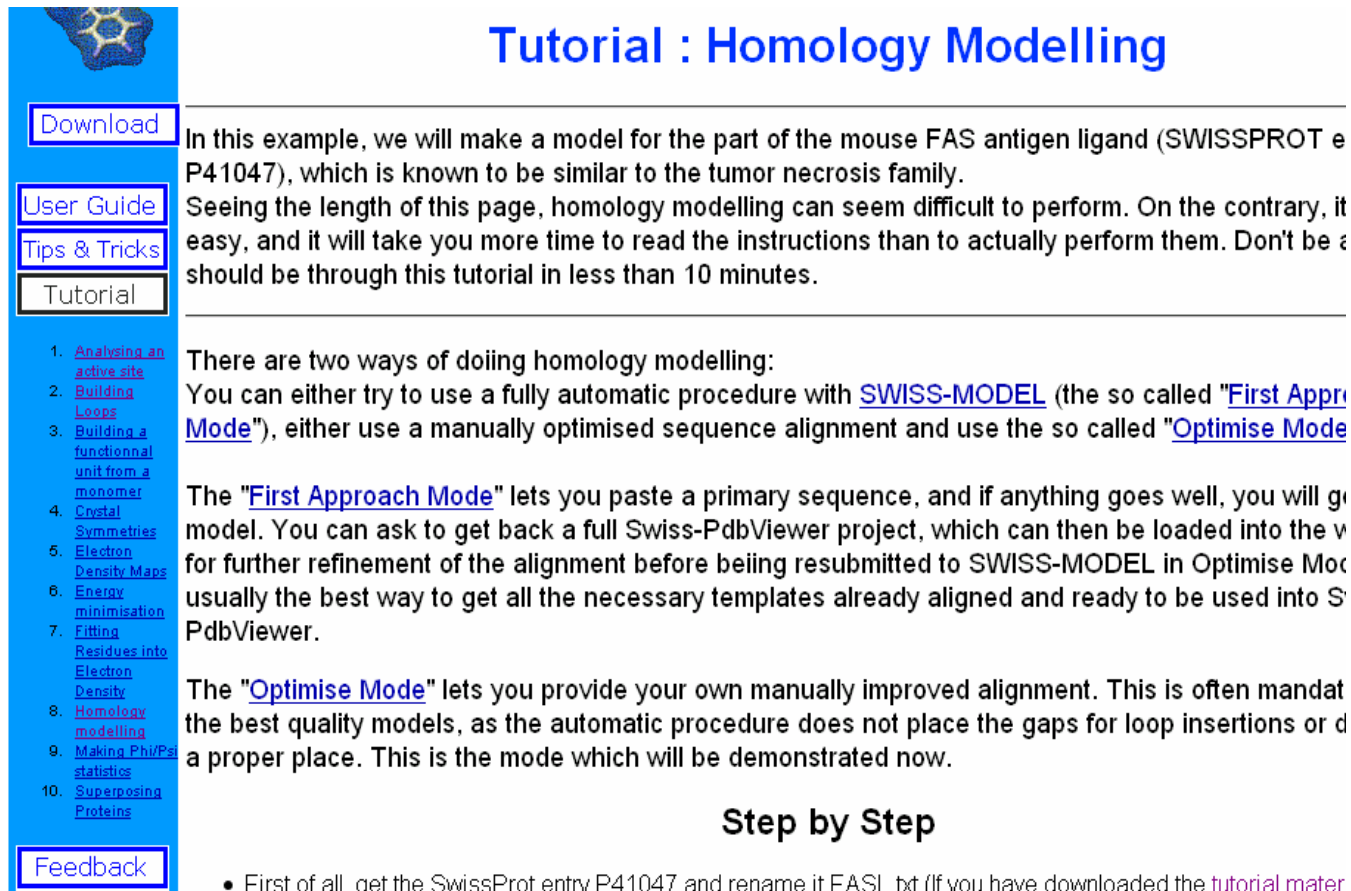
## ❖ Swiss PDB Viewer的主要功能

- ❖ 1、分析蛋白特别是具有催化作用的酶的活化位点、抗原决定簇位点、蛋白的金属结合位点及蛋白间的相互作用位点；
- ❖ 2、创建一个loop；
- ❖ 3、通过应用**non-crystallographic symmetries**工具从只含有一个蛋白单体的**pdb**文件构建一个对称的聚合体；
- ❖ 4、构建晶胞图形；
- ❖ 5、构建电子密度图形；

- ❖ 6、分析一个人为突变或认为重建一个loop后可进行最小能量优化；
- ❖ 7、将肽段或残基与电子云图谱进行匹配；
- ❖ 8、可以通过Swiss PDB Viewer的“SwissModel”进行蛋白质三维结构建模；
- ❖ 9、查看Phi/Psi数据；
- ❖ 10、超级移动（即类似于fit的叠合）。

## ❖ tips & tricks

❖ OpenGL或POV-Ray程序也可以加工这些蛋白图形，使其更加美观。



**Tutorial : Homology Modelling**

[Download](#)

[User Guide](#)

[Tips & Tricks](#)

[Tutorial](#)

[1. Analysing an active site](#)

[2. Building Loops](#)

[3. Building a functional unit from a monomer](#)

[4. Crystal Symmetries](#)

[5. Electron Density Maps](#)

[6. Energy minimisation](#)

[7. Fitting Residues into Electron Density](#)

[8. Homology modelling](#)

[9. Making Phi/Psi statistics](#)

[10. Superposing Proteins](#)

[Feedback](#)

In this example, we will make a model for the part of the mouse FAS antigen ligand (SWISSPROT e P41047), which is known to be similar to the tumor necrosis family.

Seeing the length of this page, homology modelling can seem difficult to perform. On the contrary, it is easy, and it will take you more time to read the instructions than to actually perform them. Don't be afraid, you should be through this tutorial in less than 10 minutes.

There are two ways of doing homology modelling:

You can either try to use a fully automatic procedure with [SWISS-MODEL](#) (the so called "[First Approach Mode](#)"), either use a manually optimised sequence alignment and use the so called "[Optimise Mode](#)".

The "[First Approach Mode](#)" lets you paste a primary sequence, and if anything goes well, you will get a model. You can ask to get back a full Swiss-PdbViewer project, which can then be loaded into the viewer for further refinement of the alignment before being resubmitted to SWISS-MODEL in Optimise Mode. This is usually the best way to get all the necessary templates already aligned and ready to be used into Swiss-PdbViewer.

The "[Optimise Mode](#)" lets you provide your own manually improved alignment. This is often mandatory to get the best quality models, as the automatic procedure does not place the gaps for loop insertions or deletions in a proper place. This is the mode which will be demonstrated now.

**Step by Step**

- First of all, get the SwissProt entry P41047 and rename it FASL.txt (If you have downloaded the [tutorial material](#)

## ❖ RasMol

- ❖ RasMol读取PDB格式文件，显示生物大分子三维结构图像的软件；
- ❖ 系统的要求很低，应用广，可由Unix、Windows及Macintosh 平台支持运行。
- ❖ RasMol 最大的特点是界面简单，基本操作简单，运行非常迅速。

- ❖ 在Unix系统下，可以直接用命令的方式，如 RasMol 1a4h.pdb
- ❖ `help <主题>`，`help <主题>`，`<亚主题>`来获得帮助，键入会得到一份完整的命令清单。
- ❖ 打开RasMol后，会出现两个窗口，一个是图形窗口，另外一个 一个是命令行窗口。



- ❖ Cn3D 的含义为：“See in 3-D”，是一个生物分子的三维结构、序列以及序列比对结果的可视化工具。
- ❖ Cn3D读取MMDB数据库来源的ASN.1格式的结构信息。
- ❖ MMDB从蛋白质数据库中获取数据信息，然后分析每一个PDB文件，进行验证和纠错，以更的格式存储信息。

- ❖ 特定结构的查找:
- ❖ Entrez 直接获得
- ❖ 根据 Entrez 临近序列
- ❖ 通过 BLAST 搜索
- ❖ 通过已知的 PDB 标识

- ❖ Cn3D中的结构比对
- ❖ Cn3D 可处理单独结构，也可在显示多个蛋白的结构比较。
- ❖ NCBI创建并且维护了一个比对数据库-VAST
- ❖ VAST为每一对相关蛋白做两件事：
  - ❖ 计算保守区的最佳三维重叠；
  - ❖ 构建基于空间结构相关的序列比对

## ❖ 其它可视化软件

❖ CHIME 2.6 IE与NetScape浏览器插件

❖ 安装后可以直接用浏览器观看PDB格式的文件；

❖ ICMLite三维分子浏览工具

❖ 是MolSoft公司的软件ICM的简化版，免费推出，功能强大。可读取PDB格式及其他数种格式的分子文件。

- ❖ POV-Ray3.5 是一个高质量、完全免费创造三维图像的非常有名的工具。
- ❖ 3D-mol viewer, 带动windows系统下生物软件革新的vector NTIsuite中的一员。
- ❖ Molmol 主要用来进行各种分子文件的转换, 支持30种主要分子格式。
- ❖ MAGE是能读取分子结构信息文件的非标准格式。

## ❖ 蛋白质结构解析方法:

### ❖ X-Ray和NMR (nuclear magnetic resonance)

#### ❖ X晶体衍射:

❖ 首先要得到蛋白质的晶体。大肠杆菌表达重组蛋白;

❖ 提纯之后摸索结晶条件;

❖ 拿到晶体之后, 将晶体进行x射线衍射, 收集衍射图谱;

❖ 通过计算, 能得到蛋白质的原子结构。

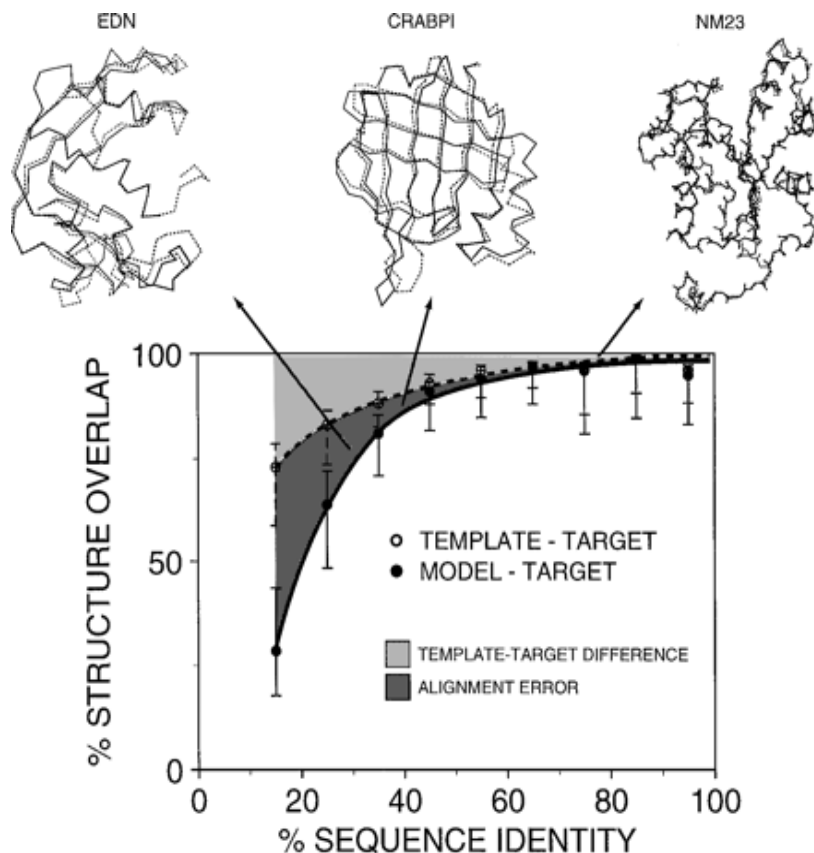
- ❖ 优点：速度快，通常只要拿到晶体，就能得到结构；
- ❖ 不受大小限制，无论是多大的蛋白，或者复合体，无论是蛋白质还是RNA、DNA，还是结合了什么小分子，只要能够结晶就能够得到其原子结构。
- ❖ x射线方法解析蛋白的关键是摸索蛋白结晶的条件。

- ❖ 首先通过基因工程的方法，表达出目的蛋白；
- ❖ 提纯之后，摸索一下蛋白稳定的条件，如蛋白没有聚合，折叠良好，装入核磁管中，放入程控核磁谱仪中，电磁波激发蛋白中的原子；
- ❖ 收集受激发的原子所放出的“能量”，收集处理数据、谱图处理、电脑计算从而得到蛋白的原子结构；
- ❖ 优点：在液体中得到结构，动态的结构；
- ❖ 发表的NMR结构都是多个结构的集合。

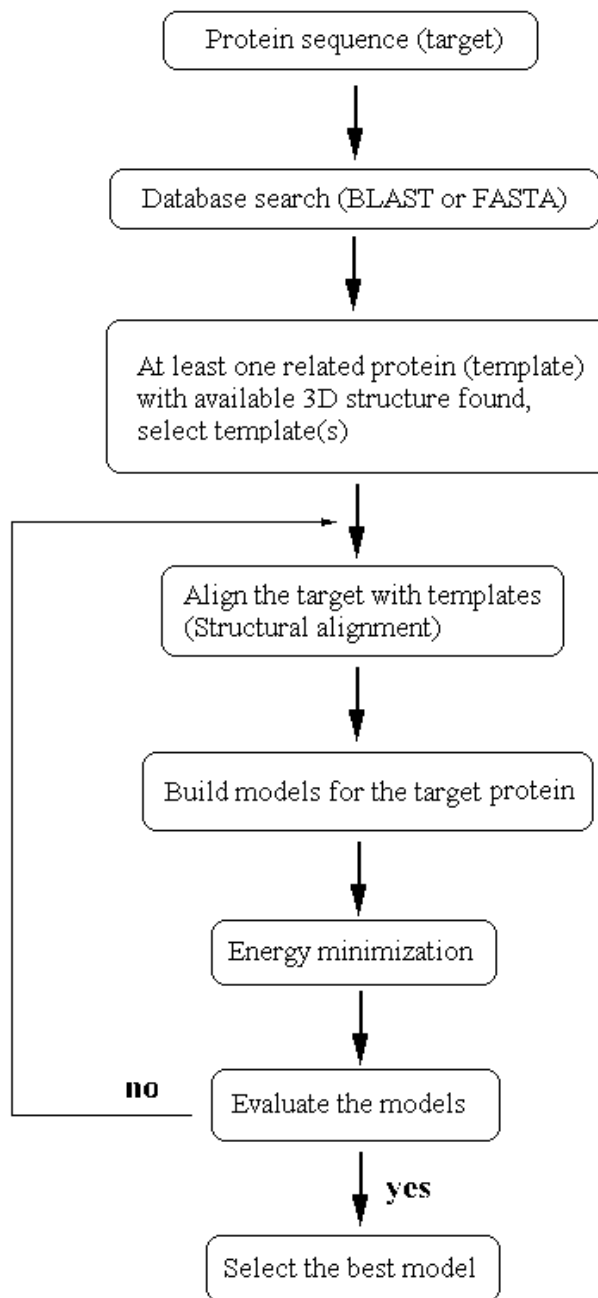


- ❖ 结构预测是最复杂和最困难的预测技术。
- ❖ 序列差异与三维构象
- ❖ 最常见的是同源建模（homology modeling）和穿针引线法（threading）。
- ❖ PSI-BLAST 方法也可以把查询序列分配到合适的蛋白质折叠家族。
- ❖ .....

- ❖ 根据同源蛋白质三级结构的保留性超过蛋白质序列的理论（ $\text{identity} > 30\%$ ）。
- ❖ 序列一致性越高，三维结构其准确性越高。



# 同源建模



数据库搜寻

选择模版

多重序列排列

骨架的建构，环状  
结构仿真及侧

链仿真能量最小化

结构合理性评估

- ❖ SWISS-MODEL与ExPASy网站是紧密相联系。
- ❖ 目前有30000个蛋白质三维结构模型。
- ❖ SWISS-MODEL由Manuel Peitsch 于1993年发起成立。
- ❖ 服务公众。

- ❖ **SWISS- MODEL 的工作模式:**
- ❖ 简捷模式(first approach mode):
- ❖ 通过web 界面提交氨基酸序列或者Swiss-Prot编号，服务器自动建模。
- ❖ 若相似度大于25%，建模程序便自动开始运行。
- ❖ 此模式只能进行大于25个残基的单链蛋白质三维结构预测。

- ❖ 比对界面( **alignment interface**):
- ❖ 用户可以通过此界面上上传以FASTA、MSF、CLUSTALW、PFAM和SELEX 格式的多重序列比对结果。
- ❖ 序列中必须至少包含一条目标序列和一条来自于Expasy Protein Database(ExPdb)的模板序列。

- ❖ 项目模式(**project mode**):
- ❖ 用户可将经过手工优化的请求提交给服务器。
- ❖ 可使用DeepView来建立一个项目文件, 它包含模板结构、目标序列和模板序列的比对结果。
- ❖ 蛋白质较验工具
- ❖ WhatCheck: <http://www.cmbi.kun.nl/gv/whatcheck>
- ❖ 平均势能的web界面<http://swissmodel.expasy.org/anolea>
- ❖ 或者使用某些商业软件包提供的检验算法来计算。

- ❖ AGADIR 是一种基于螺旋/卷曲转化理论。
- ❖ 可以在残基水平上准确预测单体肽螺旋行为的算法。
- ❖ 可以预测肽链的平均螺旋含量、 $\alpha$  碳和  $\alpha$  氢原子的构象、耦合常数等参数。
- ❖ 此算法对短肽链，即三级互作不明显时，预测准确很高。
- ❖ 利用AGADIR的预测数据，可以参考其对肽链螺旋，及蛋白结构进行适当修饰。



# 其它结构预测相关程序及数据库

LOGO

- ❖ CPHmodels神经网络同源模建方法: <http://www.cbs.dtu.dk/services/CPHmodels/>
- ❖ threading法3D-PSSM : <http://www.sbg.bio.ic.ac.uk/~3dpssm/>
- ❖ 从头预测法HMMSTR/Rosetta:  
<http://www.bioinfo.rpi.edu/~bystrc/hmmstr/server.php>
- ❖ PROSITE蛋白质功能位点: <http://kr.expasy.org/prosite/>
- ❖ FSSP已知空间结构的蛋白质家族: <http://www.ebi.ac.uk/dali/fssp/fssp.html>
- ❖ SCOP蛋白质分类数据库: <http://scop.mrc-lmb.cam.ac.uk/scop/>
- ❖ CATH蛋白质分类数据库: <http://www.biochem.ucl.ac.uk/bsm/cath/>
- ❖ Pfam蛋白质家族和结构域: <http://pfam.wustl.edu/>
- ❖ Bryant-Lawrence: <ftp://ncbi.nlm.nih.gov/pub/pkb>
- ❖ DALI: <http://www.embl-heidelberg.de/dali/dali.html>
- ❖ FSSP: <http://www.embl-heidelberg.de/dali/fssp/fssp.html>
- ❖ TOPITS: [http://www.embl-heidelberg.de/predictprotein/phd\\_help.html](http://www.embl-heidelberg.de/predictprotein/phd_help.html)

## ❖ 三级结构分析:

❖ STRAP: <http://www.charite.de/bioinf/strap/>

❖ iMolTalk: <http://i.moltalk.org/>

