

SPDBV 使用介绍

2020 年春季 BBT 课程第四组

组长：刘晨露 执笔：毛澍霖

组员：李擎宇 田鹭

1. SPDBV 介绍

Swiss-PdbViewer 是一个用户友好的应用程序，它允许同时分析多种蛋白质。可以叠加蛋白质以推断结构比对并比较其活性位点或任何其他相关部分。通过直观的图形和菜单界面，研究者可以轻松获得氨基酸突变，氢键，原子之间的角度和距离。

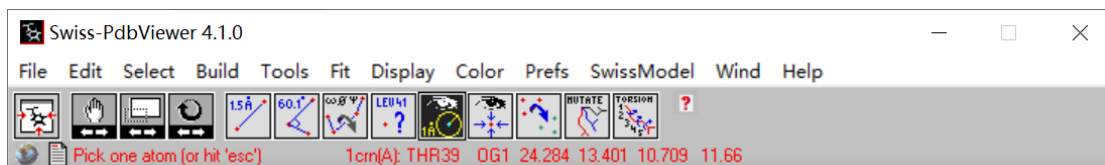
1994 年，Nicolas Guex 开发了 Swiss-PdbViewer。Swiss-PdbViewer 最初是为瑞士生物信息学研究所（SIB）给巴塞尔的结构生物信息学小组开发的自动同源性建模服务器 SWISS-MODEL 数据库服务。经过了 30 余年的发展，Swiss-PdbViewer 已经成为了一款成熟的、功能强大而广受欢迎的蛋白质结构分析工具，受到全世界的结构生物学家的欢迎。

Swiss-PdbViewer 还可以读取电子密度图，并提供各种工具来建立电子密度图。此外，Swiss-PdbViewer 还集成了各种建模工具，并且可以对残基突变进行预测。

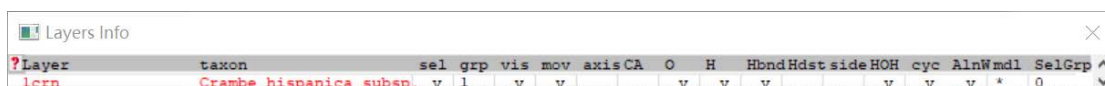
2. SPDBV 界面介绍

a) 主要窗口

- i. 最上方是“工具栏”窗口，它包含了常用的定心和测量工具。

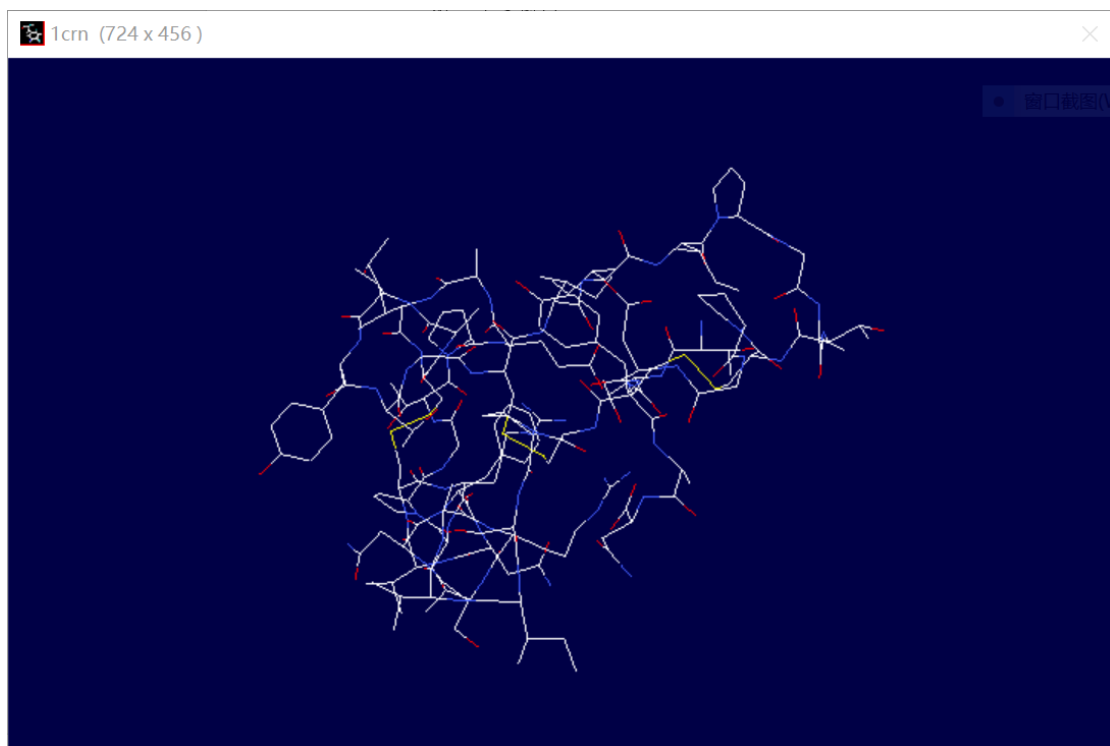


- ii. 下一个窗口显示加载到工作区中的每个图层（结构）的信息，并允许更改其各个显示选项。

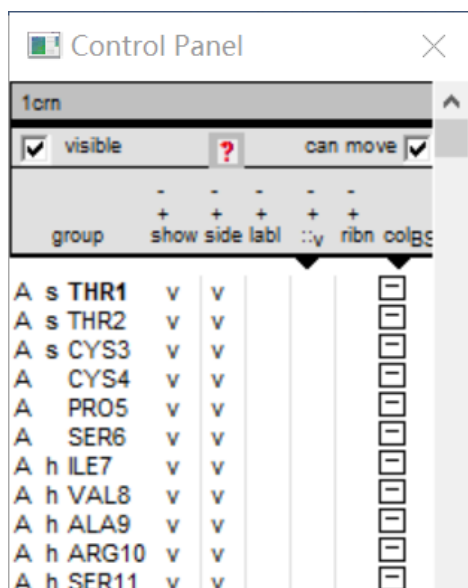


Layers Info	
sel	分子是否被选中
vis	控制蛋白的可见性
mov	控制这个蛋白质是否可以移动
axis	给该图形加一个坐标轴，该坐标轴有三个方向：X,Y,Z
CA	用来控制是否只显示蛋白质的 α -碳原子
O	用来控制是否显示骨架的氧原子
H	用来控制是否显示氢原子
Hbnd	用来控制是否显示氢键
Hdst	显示氢键间的距离
cyc	控制不同 layer 之间的轮换显示
AlnW	控制给出的序列是否会出现现在比对窗口
mdl	允许去选择模板来构建初级模型或提交给 SwissModel 进行同源模建。标记“*”的 layer 会被作为参考 layer，这一 layer 会被作为同源模建时的主要模板

- iii. 中间有一个主窗口，可以在其中实时操作分子，测量距离和角度，进行突变并比较结构。



- iv. 在屏幕右侧，有一个控制面板，可以通过它来选择和操作各个组的属性。



第一行是控制当前活动分子的可见性和可移动性，以及针对 Control Panel 寻求帮助。

第二行的功能从左至右依次是：显示氨基酸残基(show)，显示侧链(side)，标注残基(lable)，显示分子表面()，以条带形式显示螺旋和折叠(ribn)，并且可以有选择的对各种氨基酸残基进行颜色修饰(color)。当点击“+”时，就会针对选中的氨基酸显示相应的侧链、标签、电子云等，点击“-”时，这些相应的侧链、标签、电子云消失。

b) 工具栏中按钮介绍



SPDBV 共有 13 个控制按钮，从左至右依次为：

- i. 将显示的蛋白质结构模型在窗口中居中



- ii. 保持蛋白质结构模型不旋转的情况下在显示窗口中移动模型








- iii. 放大或者是缩小结构模型





- iv. 旋转该蛋白质结构模型



- v. 测量两个原子间的距离

- vi. 测量三个相邻原子的角度

- vii. 测量相邻四个原子的二面角

- viii. 显示鼠标点击的氨基酸残基信息，并且标出该氨基酸残基在蛋白质多肽中的位置

- ix. 是以某个氨基酸残基为中心，显示在一定范围内的所有原子

- x. 将所选定的氨基酸残基移至中心

- xi. 将一个分子叠加到另一个分子上

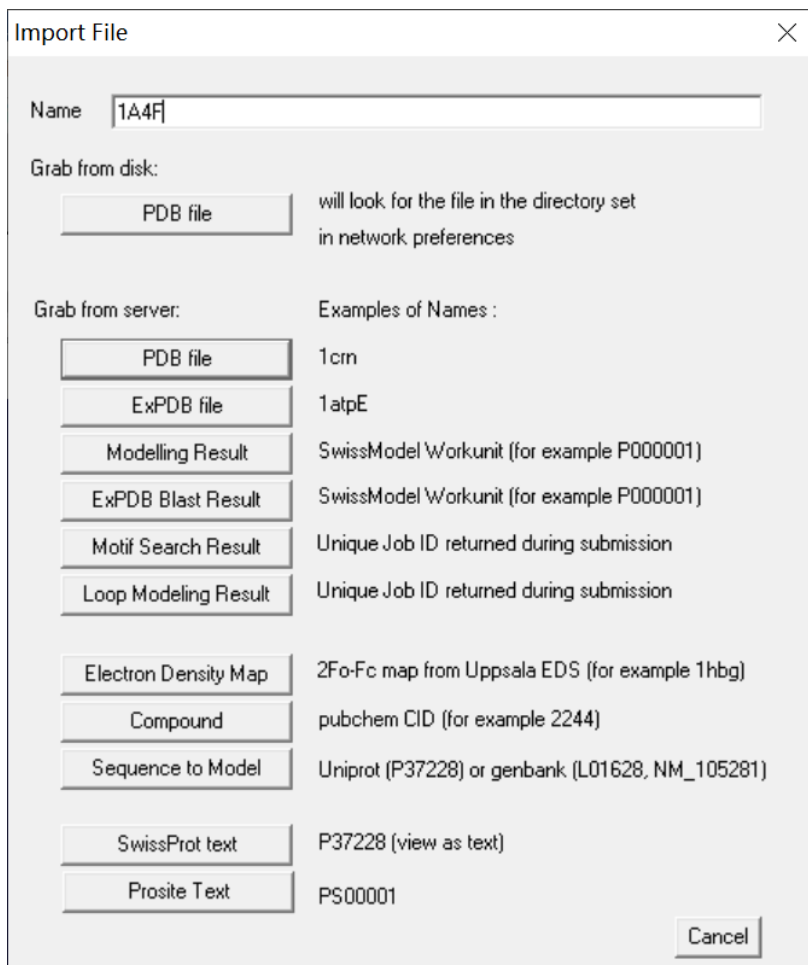
- xii. 对某一点进行定点突变并预测

- xiii. 对所选的氨基酸残基的某一链进行旋转扭曲


3. SPDBV 常用功能使用方法

a) 打开数据文件

- i. 直接在菜单选择“file→open PDB file”
- ii. 打开 PdbViewer 之后再直接把 pdb 文件拖进去

- iii. 在菜单 file 中选择 Import, 出现对话框, 在 Name 栏中键入 PDB ID, 如“1A4F”, 点击“Grab from SERVER”下面的“PDB File”, 模型就会联网下载后显示 (此处支持使用 ExPDB ID, SwissModel 中的编号, pubchem CID, Uniprot 和 GenBank 中的序列号等)



b) 保存数据文件

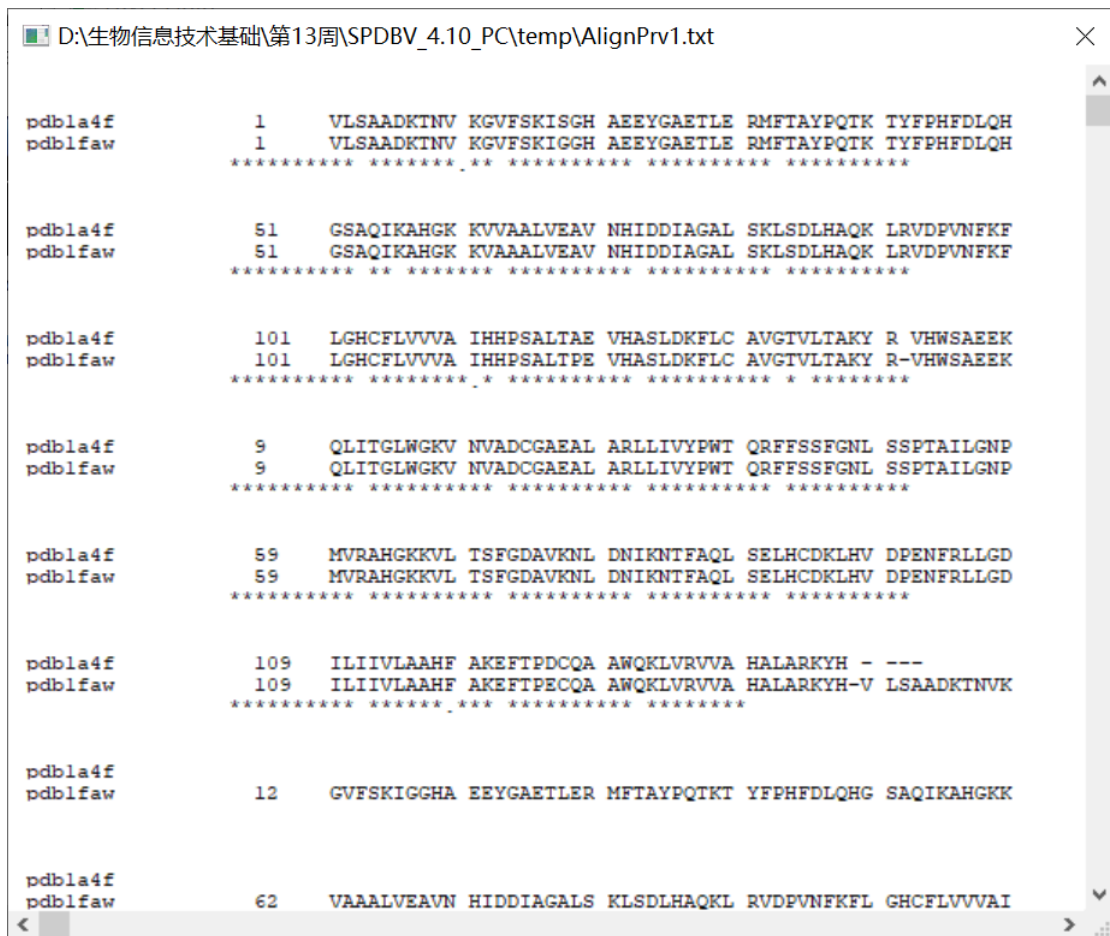
- i. 可以通过“file→Save→ Current Layer(Ctrl+S)”保存在显示窗口的蛋白分子的.pdb格式文件
- ii. Project(shift+ctrl+S)可以保存所有同时打开显示的蛋白质分子
- iii. Save selected residues of Current Layer 可以保存被选择的氨基酸残基
- iv. Sequence (FASTA) 按键可以保存蛋白序列的.fasta格式文件

c) 序列比对功能

Alignment 可以进行序列的比对, 将斑头雁的血红蛋白 (1a4f) 与灰雁的血红蛋白 (1faw) 文件打开, 应用 Alignment 进行序列比对。



点击红圈内图标可以以文本形式显示比对的结果。



```
D:\生物信息技术基础\第13周\SPDBV_4.10_PC\temp\AlignPrv1.txt
pdbla4f      1      VLSAADKTNV KGVFSKISGH AEEYGAETLE RMFTAYPQTK TYFPFDLQH
pdblfaw      1      VLSAADKTNV KGVFSKIGGH AEEYGAETLE RMFTAYPQTK TYFPFDLQH
*****_******

pdbla4f     51      GSAQIKAHGK KVAALVEAV NHIDDIAGAL SKLSDLHAQK LRVDPVNFKF
pdblfaw     51      GSAQIKAHGK KVAALVEAV NHIDDIAGAL SKLSDLHAQK LRVDPVNFKF
*****_******

pdbla4f    101      LGHCFLVVVA IHHSALTAE VHASLDKFLC AVGTVLTAKY R-VHWSAEEK
pdblfaw    101      LGHCFLVVVA IHHSALTPE VHASLDKFLC AVGTVLTAKY R-VHWSAEEK
*****_******

pdbla4f      9      QLITGLWGKV NVADCGAEAL ARLLIVYPWT QRFFSSFGNL SSPTAILGNE
pdblfaw      9      QLITGLWGKV NVADCGAEAL ARLLIVYPWT QRFFSSFGNL SSPTAILGNE
*****_******

pdbla4f     59      MVRAHGKKVL TSFGDAVKNL DNIKNTFAQL SELHCDKLHV DPENFRLLGD
pdblfaw     59      MVRAHGKKVL TSFGDAVKNL DNIKNTFAQL SELHCDKLHV DPENFRLLGD
*****_******

pdbla4f    109      ILIIVLAAHF AKEFTPCQA AWQKLVVVVA HALARKYH - ---
pdblfaw    109      ILIIVLAAHF AKEFTPECQA AWQKLVVVVA HALARKYH-V LSAADKTNVK
*****_******

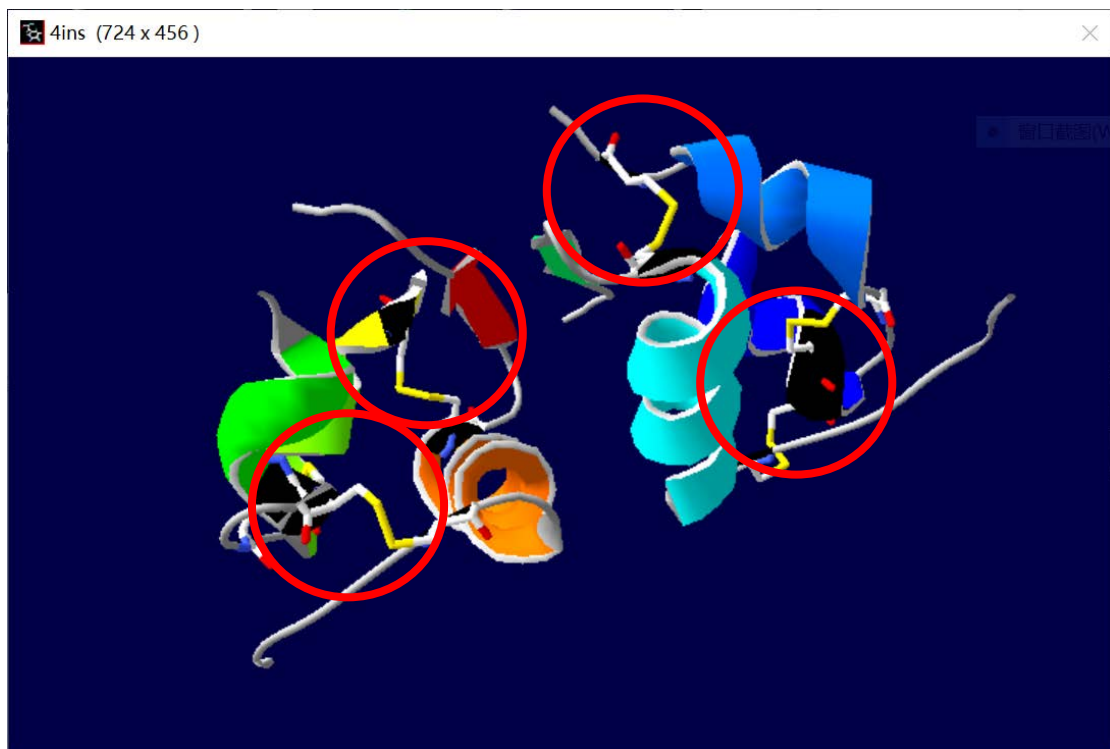
pdbla4f     12      GVFSKIGGHA EEEYGAETLER MFTAYPQTKT YFPFDLQHG SAQIKAHGKK

pdbla4f     62      VAAALVEAVN HIDDIAGALS KLSDLHAQKL RVDPNFKFL GHCFVVVAI
```

d) 显示功能

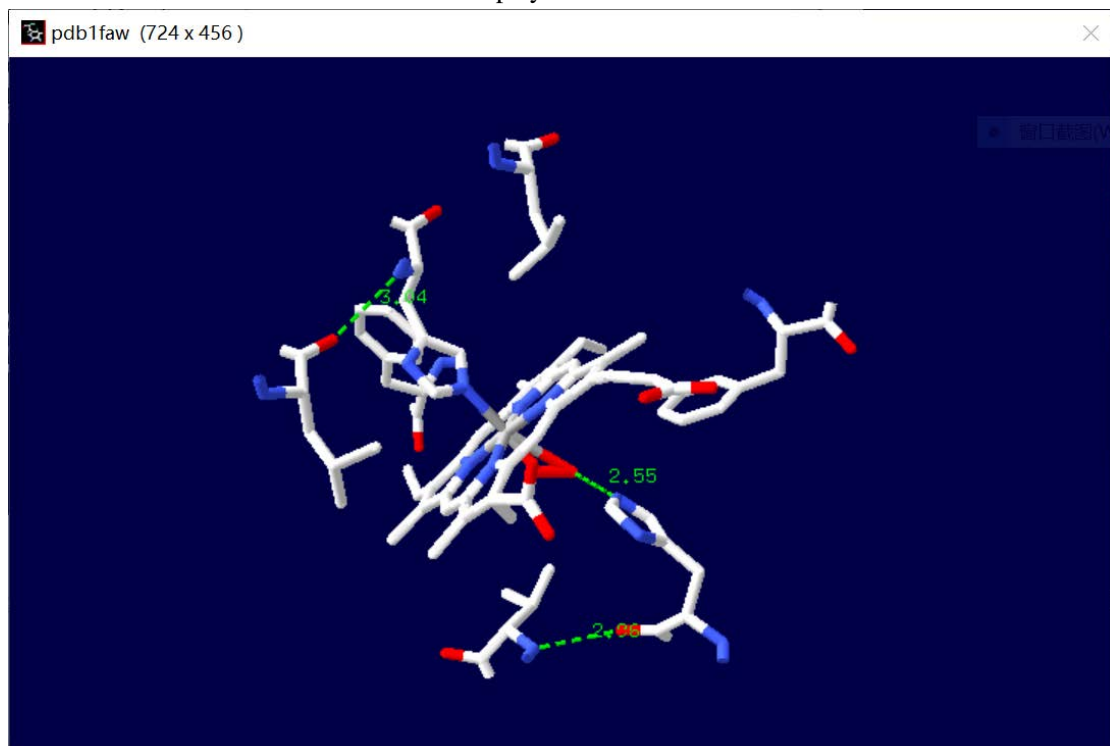
- i. 点击 select, 根据需要点击“group kind”/“group properties”/“secondary structure”, 选择具体需要显示的功能。

这里以显示胰岛素中的二硫键为例：可以看到红圈标注的是显示的二硫键



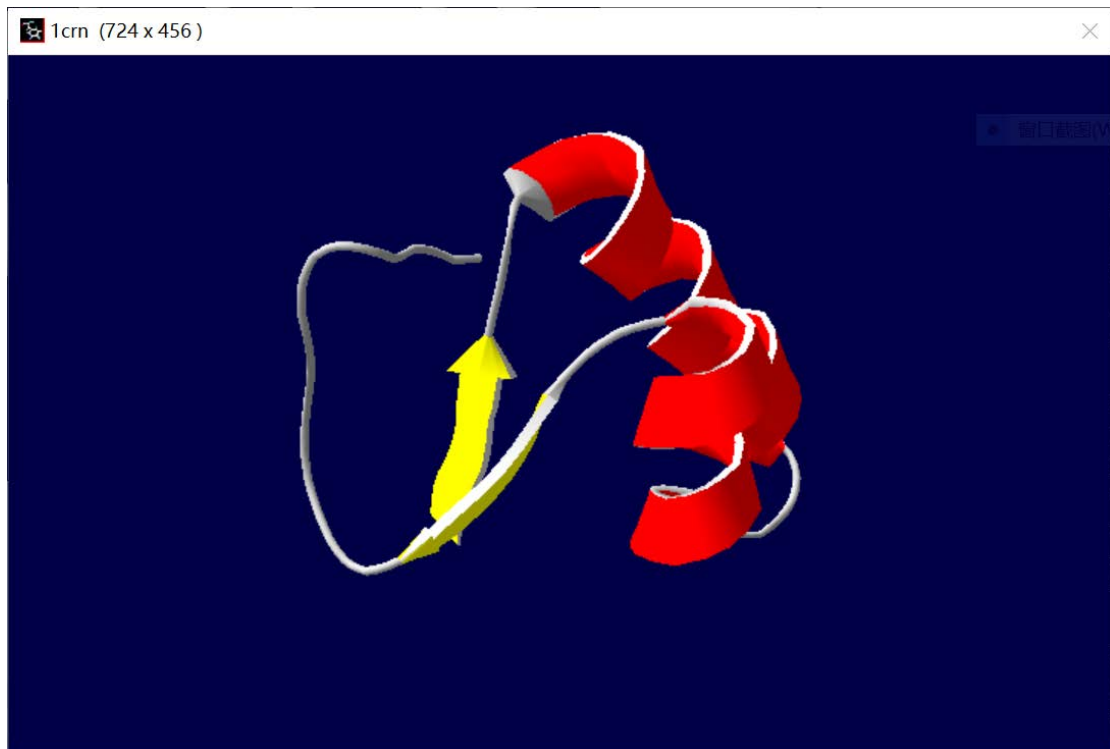
ii. 显示氢键

选择Tool->Compute H-Bonds, 再选择Display->show H-bonds, 如果想进一步查看氢键的长度, 可以选择Display->show H-bonds distances.

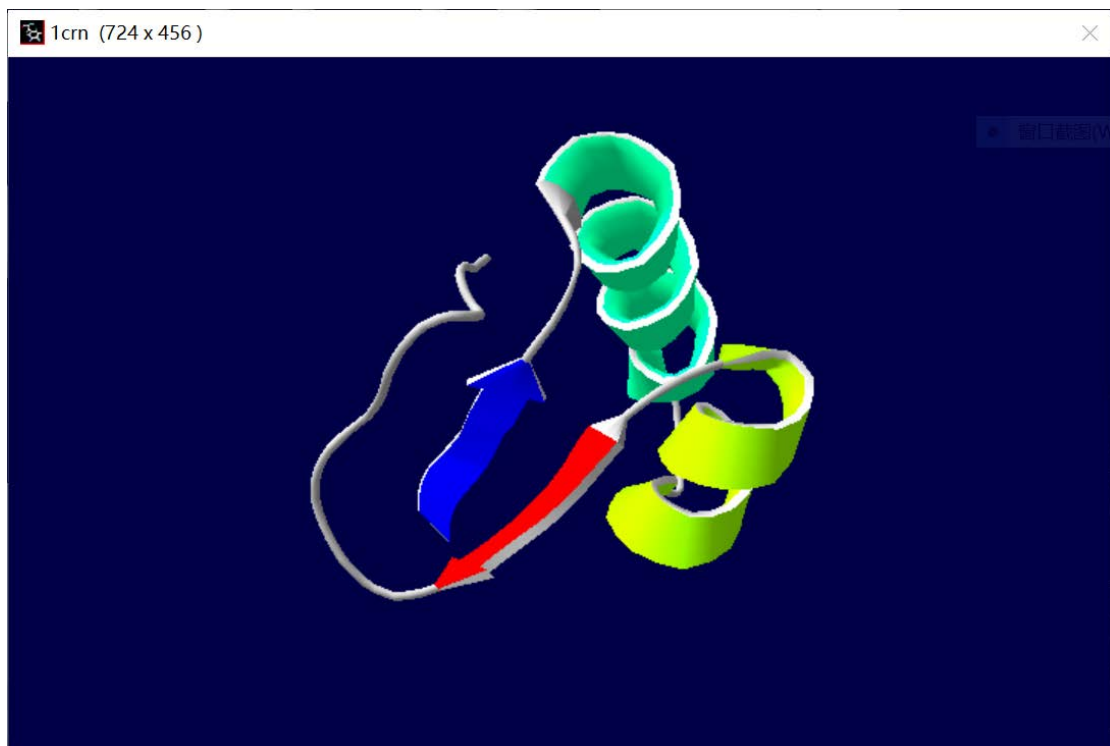


e) 对蛋白结构上色

- i. 对二级结构上色
 - 1. Color->Secondary Structure, 将 alpha-helix/beta-sheet/coil 等二级结构使用不同颜色显示

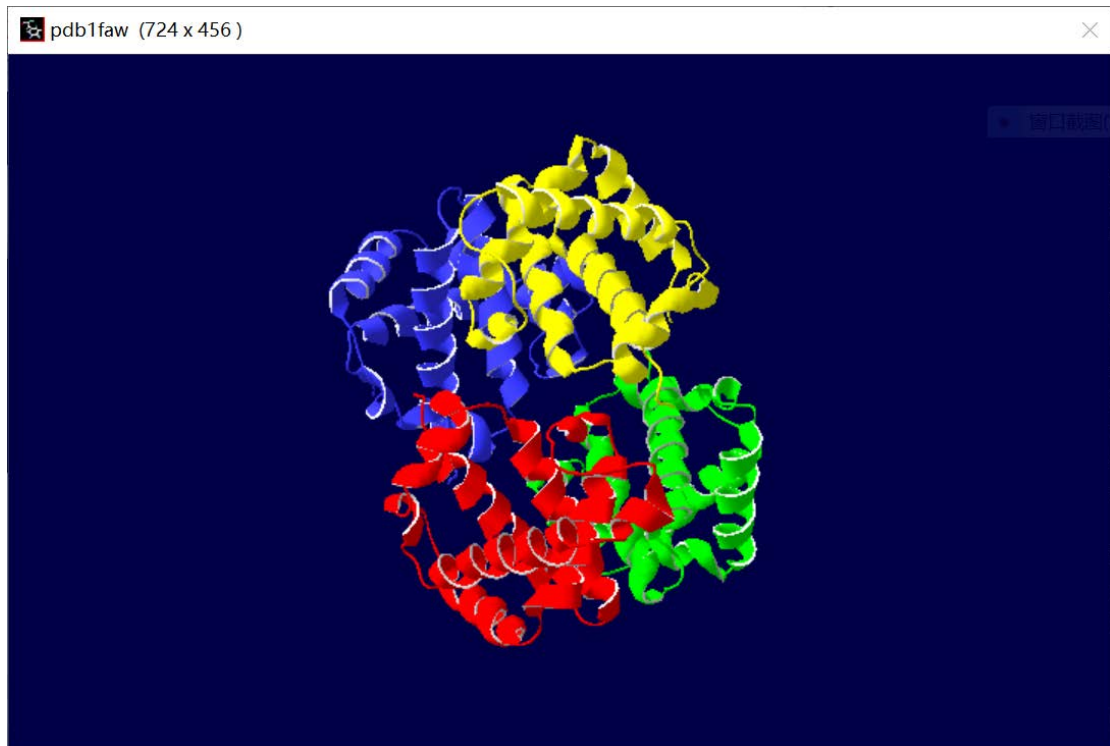


- 2. Color->Secondary Structure Succession, 将整个序列的每个二级结构用不同的颜色显示出来,方便使用者更清楚的看到从氨基端到羧基端二级结构间的顺序



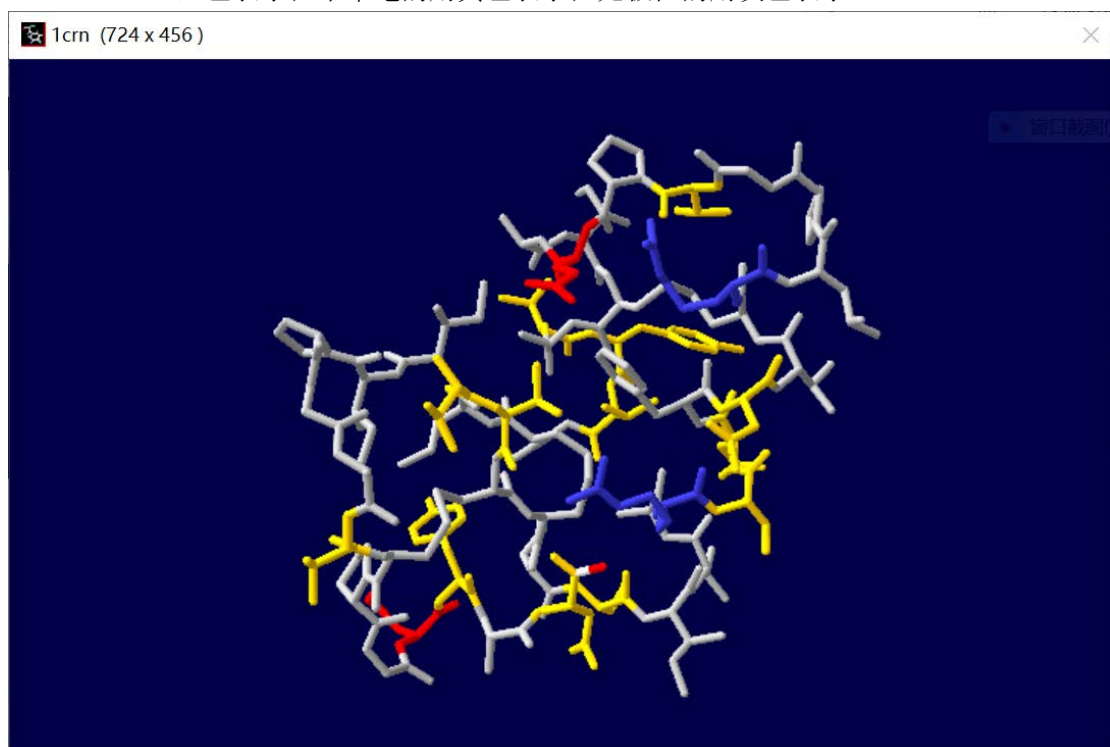
ii. 对不同链上色

Color->by Chain, 用于区分含有多条链的结构, 不同的链用不同的颜色表示出来 (例子中显示了血红蛋白的四条链)



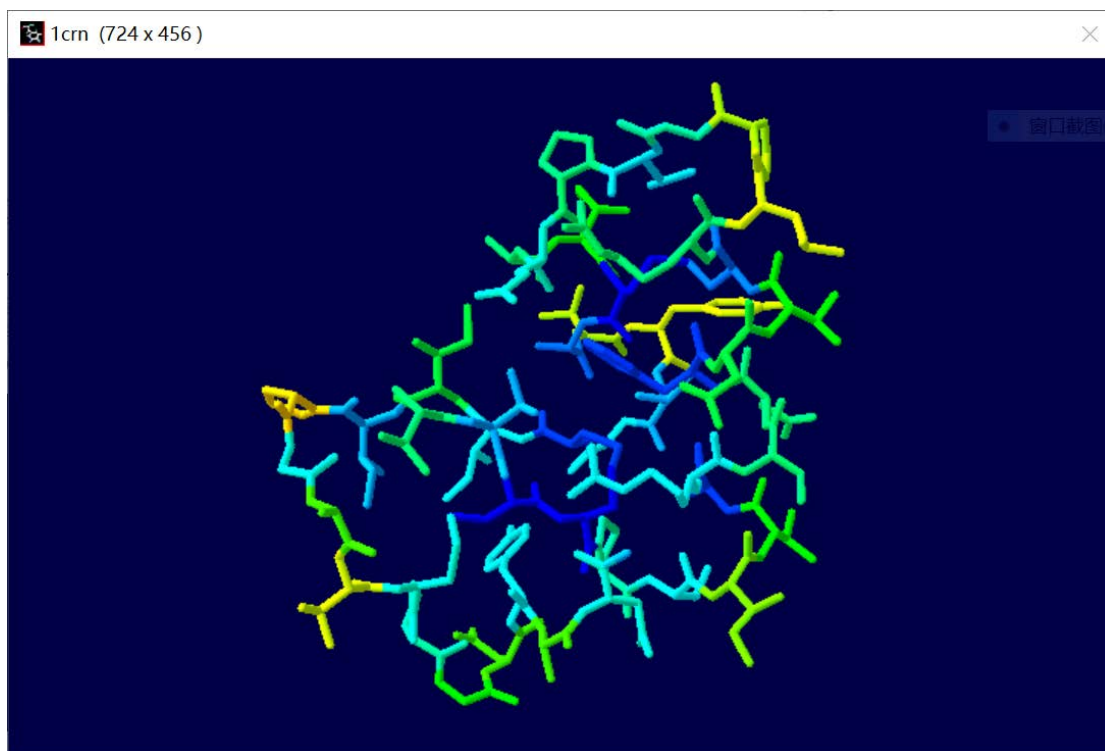
iii. 按照氨基酸类型对整个一级结构上色

Color->by Type, 按照氨基酸类型对结构模型染色, 带正电/带负电/不带电/无极性的氨基酸分别用不同颜色标出 (例图中带正电的用蓝色表示; 带负电用红色表示; 不带电的用黄色表示; 无极性的用灰色表示)



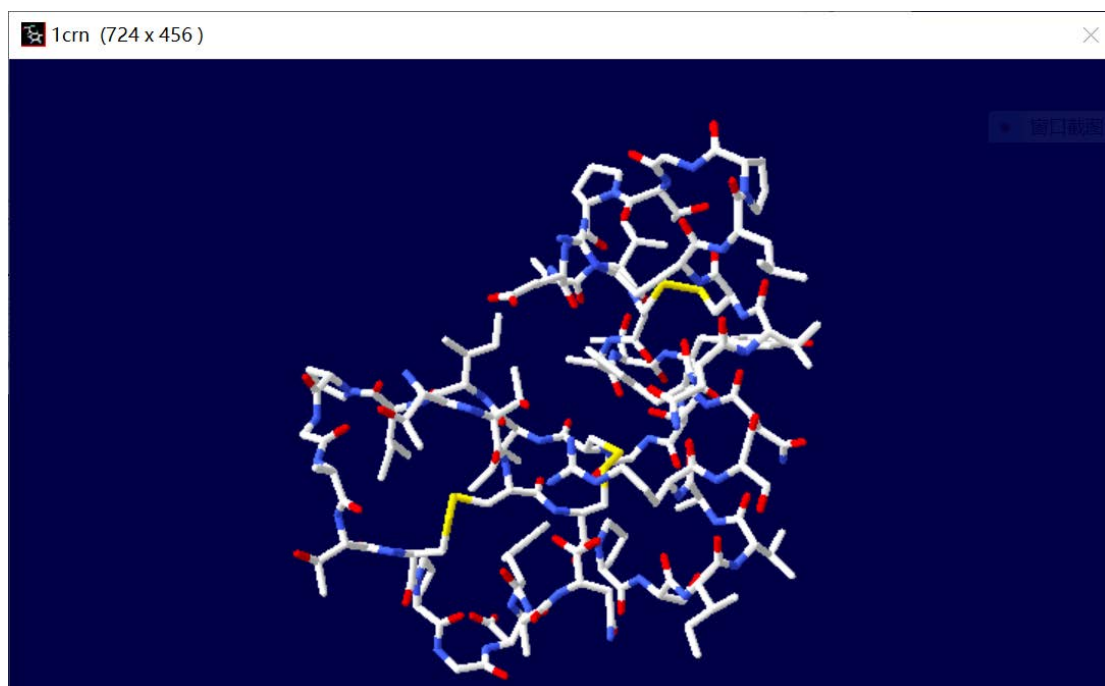
iv. 按照氨基酸溶剂可及性上色

Color->by Accessibility, 在结构中每个氨基酸残基与周围溶剂接触程度的多少决定了残基的颜色。与溶剂接触最少的是蓝色, 完全露在分子表面的是红色, 接触介于二者之间的, 用蓝色和红色中间的颜色表示, 如蓝绿色和洋红色



v. 按照不同元素进行上色 (默认)

Color ->by CPK, 将所有残基集团的颜色恢复到默认状态: C 原子用白色表示; O 原子用红色表示; N 原子用蓝色表示; S 原子用黄色表示




f) 测量功能

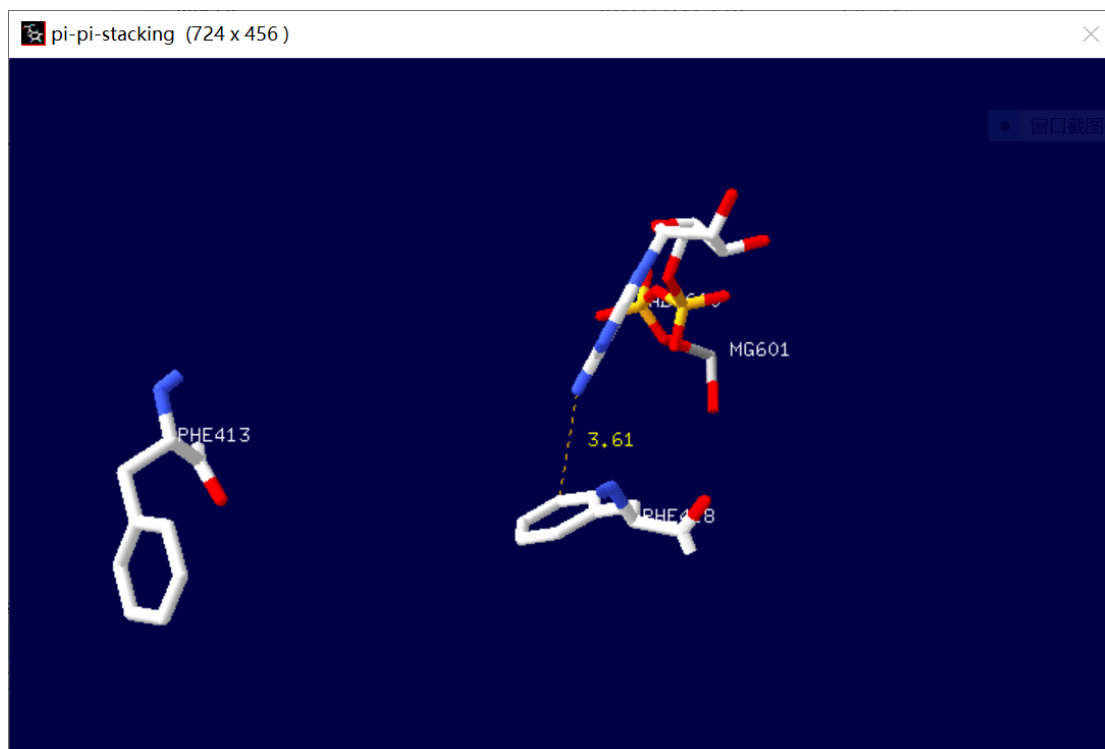
软件可以测量 2 点之间的距离、相邻 3 个原子的角度以及相邻 4 个原子的二面角，并同时在图上标注出来。

i. 测量两点之间距离




点击图标 ，选取任意两点，即可测得这两点间的距离

(例子中测量的的是 ATP synthase 中 ADP 与 Phe418 之间的距离，显示为 3.61Å)

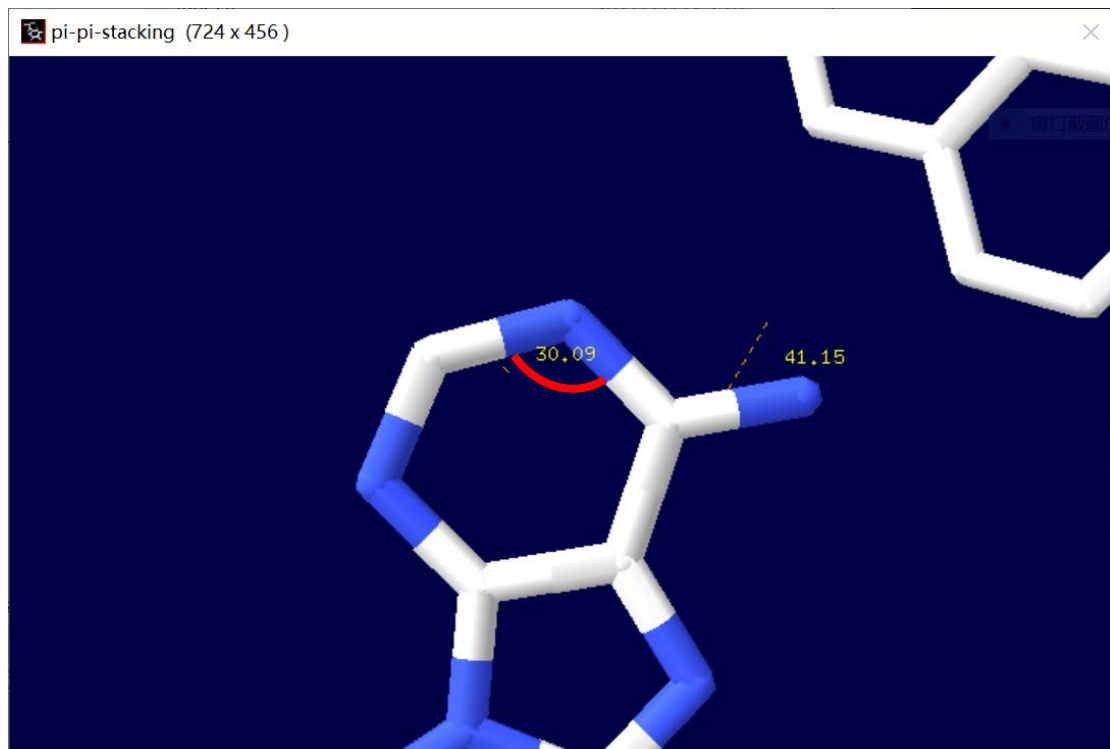


ii. 测量三个原子夹角



点击图标 ，如想测得三个相邻原子的夹角，先点击角的顶点，再点击余下两点，即可测得夹角值

(例子中测量了ADP嘌呤环中三个原子的夹角(红线所示)，测量值为30.09°)



g) 绘制 Ramachandran Plot

i. Ramachandran Plot 简介

Ramachandran Plot(拉氏构象图), 通过它可以了解氨基酸之间的 ϕ 和 ψ 角的信息。理论上 ϕ 和 ψ 角这两个角度可以自由旋转, 可实际上在某些特定的角度有空间上的障碍, Ramachandran Plot 主要目的是显示哪些角度的组合是空间上所允许的。这些在空间上允许出现的区域以蓝色与黄色等高线表示。若将已知的蛋白质结构中的角度绘在图上, 大部分 **alpha-helix** 的点都出现在第三象限的黄色区域中; 而 **beta-sheet** 的数据集中在第二象限的黄色区域中。

拉氏图中黄线封闭的区域是“允许区”, 这个区域内的任何成对二面角所规定的构象都是立体化学所允许的。此时, 构象能量最低, 构象稳定。

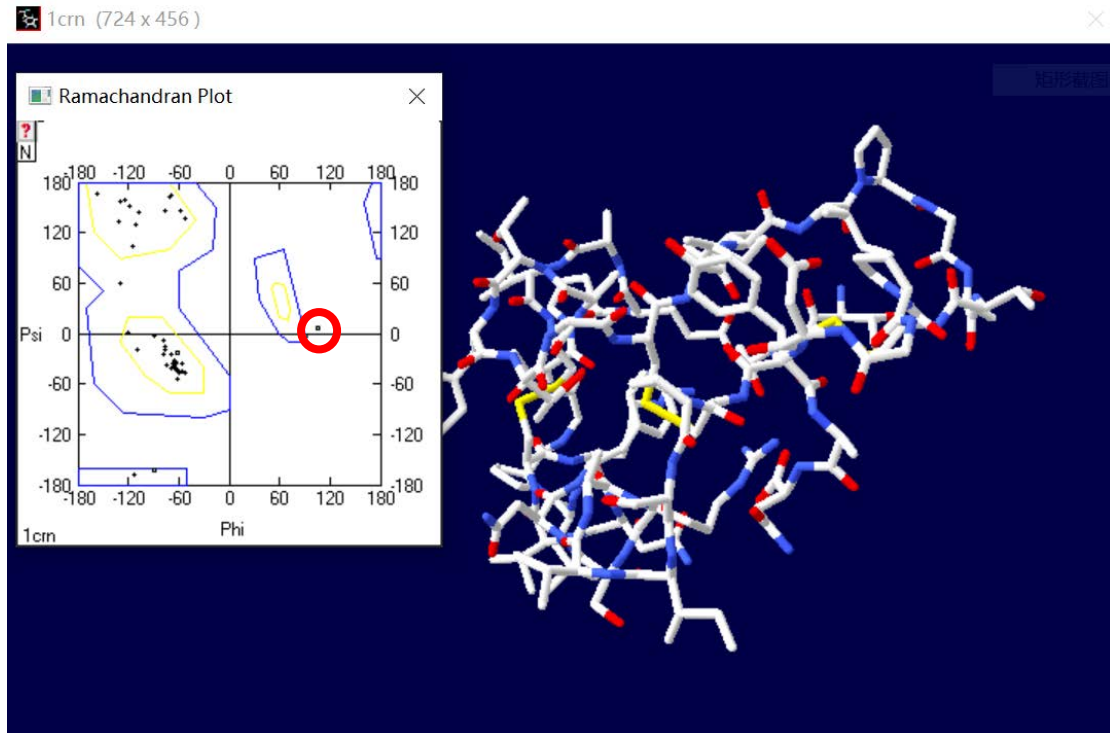
黄线外蓝线内的其他区域为“不完全允许区”。这个区域内的任何二面角所规定的主链构象虽是立体化学可允许的, 但不够稳定。

蓝线外的区域是“不允许区”。该区域内的任何二面角所规定的主链构象都是立体化学所不允许的, 因为在此构象中非共价键合原子间的距离小于极限距离, 斥力很大, 构象能量很高, 因此这种构象极不稳定, 不能存在。

ii. 绘制方法

在 **contral panel** 中选中所需要的氨基酸, 点击 **wind->Ramachandran Plot**

可以发现在 **SPDBV** 所给的例子 **1crn** 中也有一些区域在 Ramachandran Plot 的不允许区 (红圈标注), 这可能是测定结构/提取蛋白样品的问题。



h) 突变预测

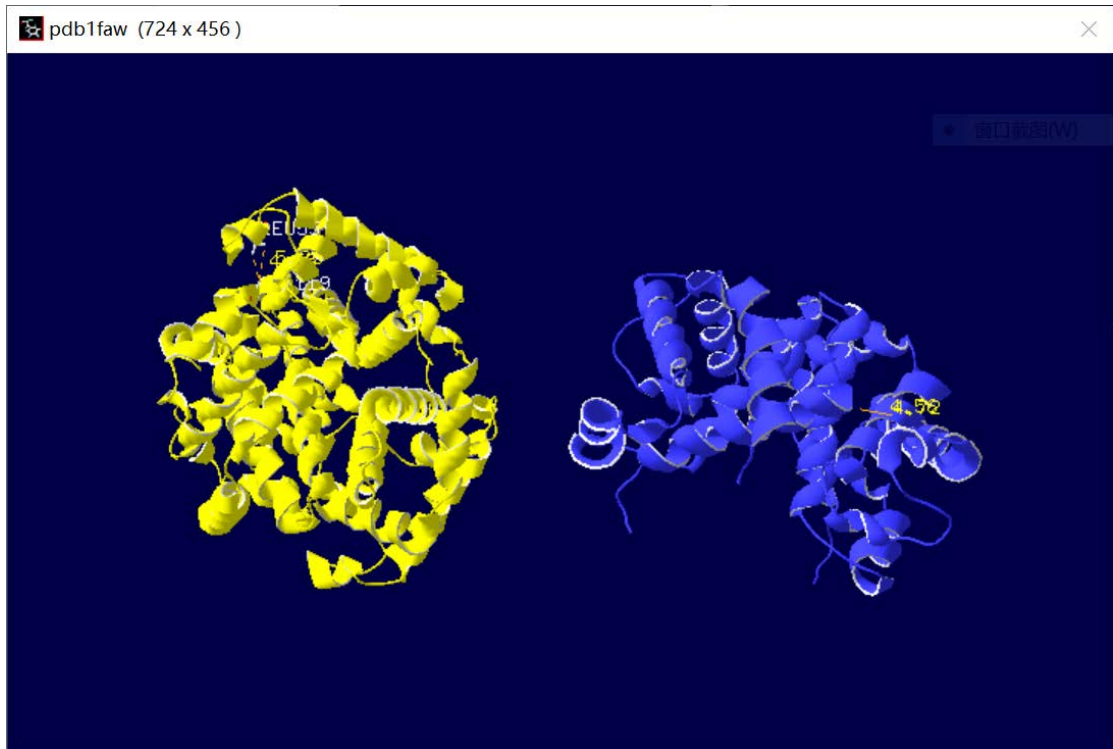
i. 内容

在上课的过程中,我们了解了灰雁与斑头雁血红蛋白序列差异导致的结构差异。1983年,佩鲁茨(Max Perutz)发现斑头雁和灰雁的血红蛋白氨基酸序列仅有4个位点差异,而其中 α -亚基的119位可能对整个结构产生了较大的影响。斑头雁 α -亚基该位点为丙氨酸(A119Ala),而灰雁该位点为脯氨酸(A119Pro)。

佩鲁茨(Max Perutz)认为斑头雁 α -亚基119位的丙氨酸侧链仅有一个甲基,与 β -亚基55位亮氨酸侧链距离较远,有利于构象变化;而灰雁该位点侧链脯氨酸有3个甲基,与 β -亚基55位亮氨酸侧链距离较近,不利于构象变化。这两种鸟类血红蛋白序列结构的差异,可能影响了两者结合氧气的能力。20世纪90年代,北京大学生物系蛋白质结构功能研究组用X射线晶体衍射的方法,分别测定了斑头雁和灰雁血红蛋白的结构,并进行了比较分析,证实了这种猜测。

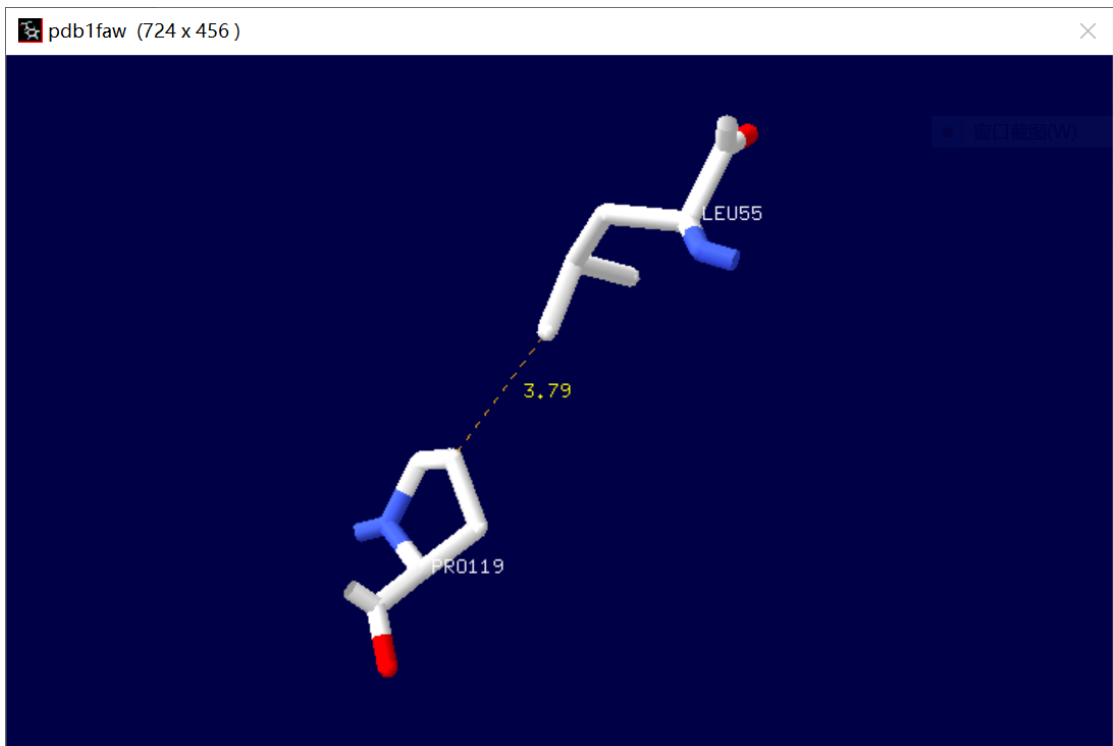
今天,我们试图使用SPDBV软件将灰雁的A119Pro突变为斑头雁中的Ala,观察预测结果与实测结果的差异。

(下面图中所示的是两种蛋白的三维结构)

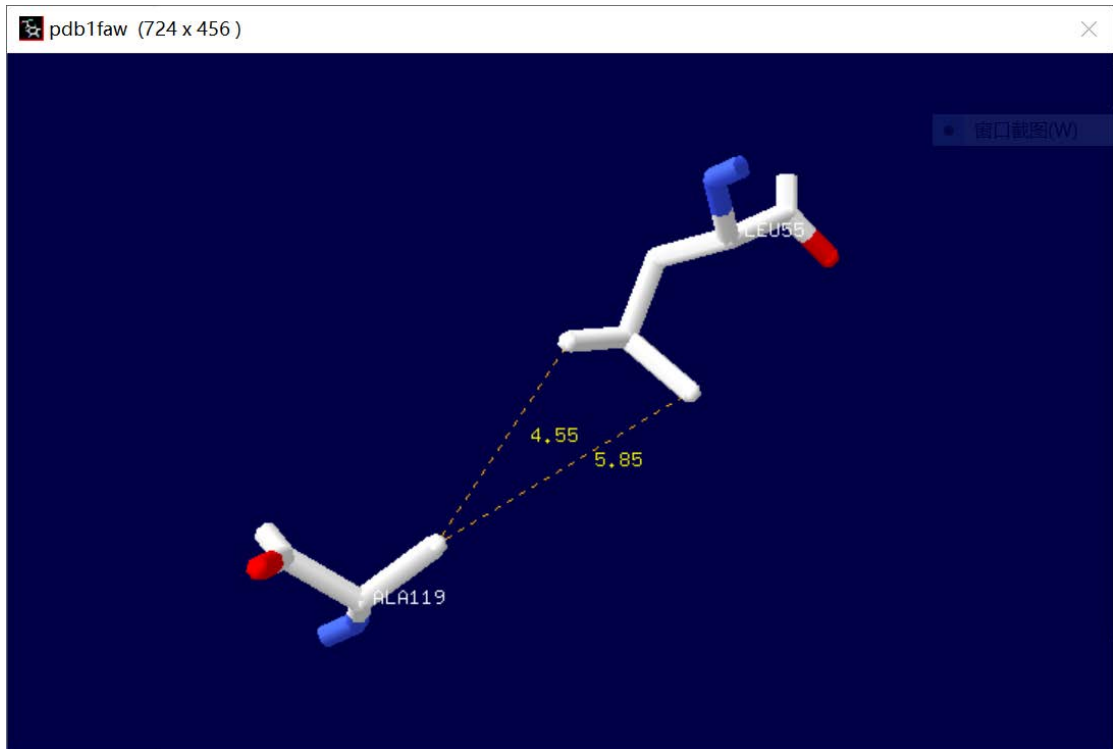


ii. 结果

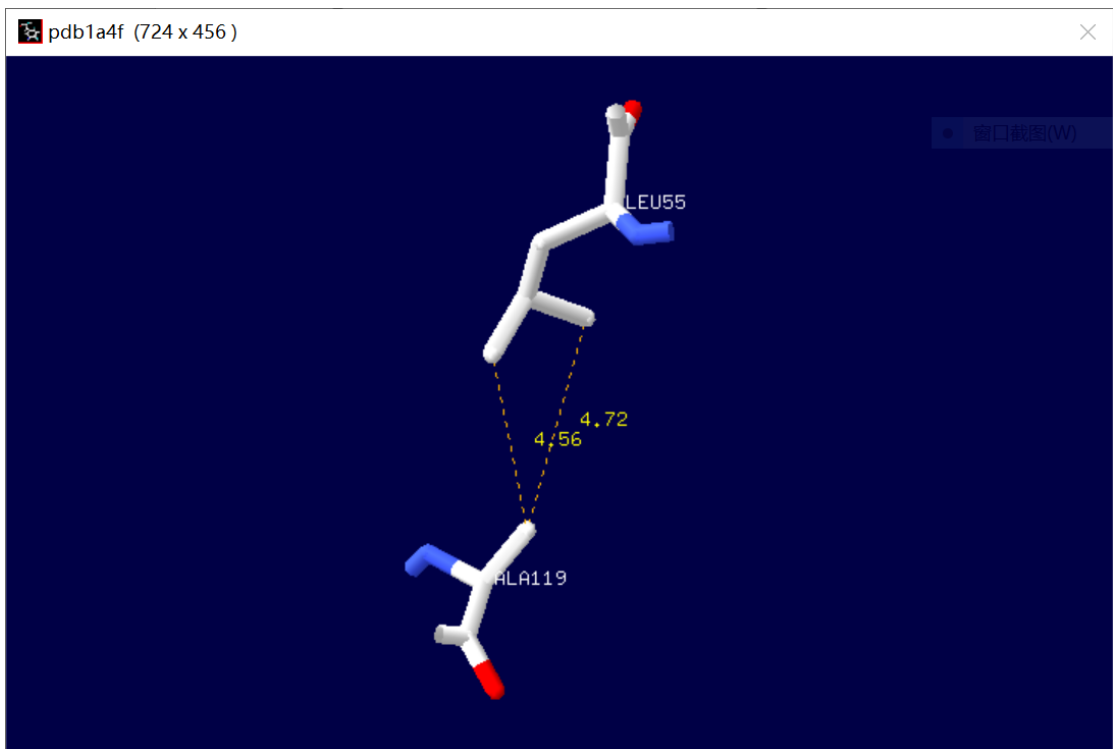
1. 未突变的灰雁血红蛋白

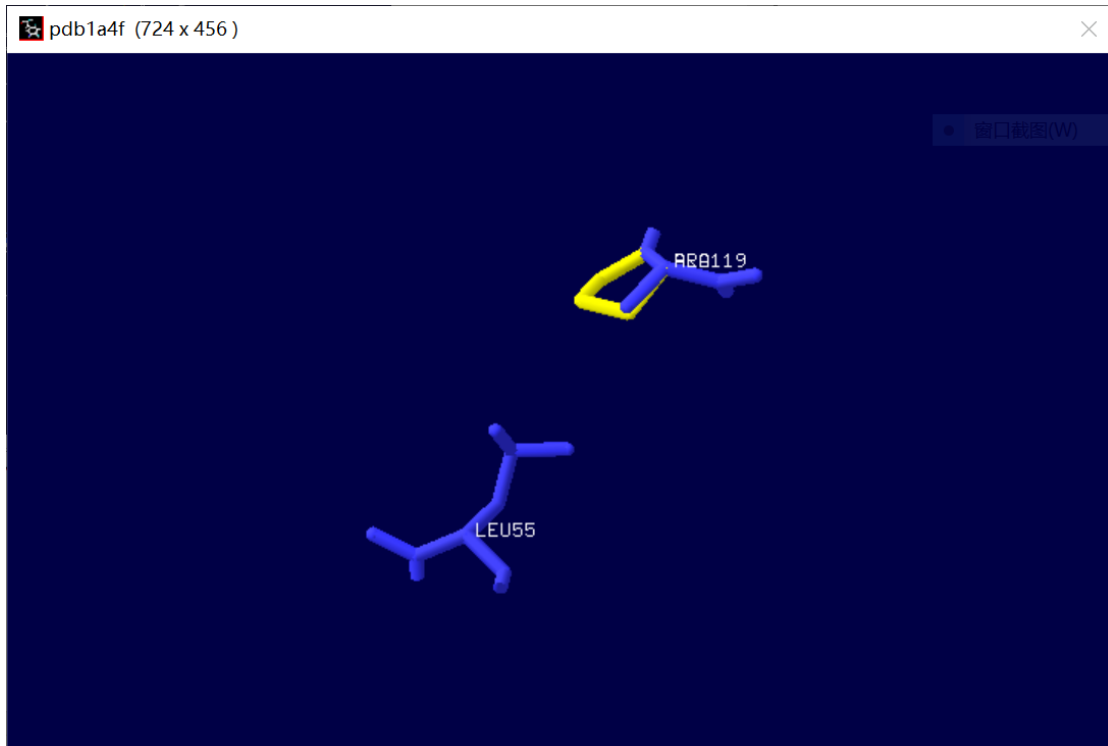


2. 突变后的灰雁血红蛋白



3. 斑头雁血红蛋白与对比图





iii. 分析

总的来说，预测的整体方向是对的，可以观察到突变之后两个残基之间的距离确实观察到变大，但是，预测得到的距离的数值与实测值还有差异，这说明该预测模型还存在提高空间。

另一方面，笔者尝试反复进行突变和回复突变（将 Ala 变为 Pro 再变回 Ala，不断进行），观察到随着迭代次数的增多，残基的位置与原先的位置之间的差异越来越大，这也同样表明了这个突变预测功能尚存缺陷。

iv. 突变预测的改进

1. Deep View 通过对 AMBER, CHARMM, GROMOS 程序递交能量最小化请求，进行模型能量最小化，使得突变后的结构得到最稳定构象，来消除不合理的构象。

2. 结果

能量最小化之后的结果，依然是预测方向正确，但是预测值还有误差。

```

D:\生物信息技术基础\第13期\SPDBV_4.10_PC\temp\energy.E22
/ Computations were done in vacuo with the GROMOS96 43BI parameters set, without reaction field.
/ For more information about GROMOS96, refer to: W.F. van Gunsteren et al. (1996) in Biomolecular
/ simulation: the GROMOS96 manual and user guide, Vdf Hochschulverlag ETH (http://acc.ethz.ch/gromos).
/ When using those results, please mention that energy computations were done with the GROMOS96
/ implementation of Swiss-PdbViewer.
-----
/ residue      bonds      angles     torsion    improper   nonBonded  electrostatic  constraint // TOTAL
-----
HHT  A  1      0.000      6.183      7.540      0.000      0.00      16.41      0.0000 // E= 30.136
VAL  A  1      2.489      0.961      1.203      0.068      3.30      114.23      0.0000 // E= 122.755
LEU  A  2      1.056      6.159      18.656      0.050      -6.96      14.75      0.0000 // E= 33.820
SER  A  3      1.040      1.143      1.207      0.992      -7.09      -39.09      0.0000 // E= -41.799
ALA  A  4      1.124      0.244      0.498      0.405      -8.97      8.29      0.0000 // E= 1.592
ALA  A  5      0.793      0.468      1.899      0.621      -10.42      3.22      0.0000 // E= -3.495
ASP  A  6      0.496      1.960      2.596      0.291      -30.50      -10.96      0.0000 // E= -36.152
LYSH  A  7      2.153      1.130      0.420      0.560      -43.54      -6.80      0.0000 // E= -46.086
THR  A  8      1.964      0.625      1.969      0.892      -7.92      -6.06      0.0000 // E= -9.126
ASN  A  9      2.654      1.950      4.109      0.629      -33.27      -168.80      0.0000 // E= -193.942
VAL  A  10     0.889      0.914      1.829      0.557      -26.36      -4.79      0.0000 // E= -26.996
LYSH  A  11     1.797      1.515      7.408      0.909      -32.39      34.03      0.0000 // E= 13.269
GLY  A  12     0.922      1.689      1.867      1.810      -11.67      26.87      0.0000 // E= 20.484
VAL  A  13     4.177      1.467      0.979      0.330      -11.26      -9.79      0.0000 // E= -14.441
PHE  A  14     2.962      12.652      12.564      3.205      -55.68      -5.76      0.0000 // E= -30.050
SER  A  15     1.203      1.330      2.422      0.005      -15.42      0.24      0.0000 // E= -10.223
LYSH  A  16     3.871      2.470      7.072      1.531      -3.29      7.44      0.0000 // E= 19.090
IIE  A  17     1.911      1.608      8.732      0.597      31.40      0.0000 // E= 46.559
GLY  A  18     1.119      3.522      2.916      0.241      -5.27      59.38      0.0000 // E= 61.908
GLY  A  19     0.854      4.519      0.689      0.275      -7.96      30.42      0.0000 // E= 28.902
HISA  A  20     13.481      9.494      3.500      0.417      -22.14      -21.51      0.0000 // E= -16.758
ALA  A  21     0.866      0.474      3.625      0.402      -16.23      -11.66      0.0000 // E= -18.504
GLU  A  22     2.263      2.726      4.616      0.383      -17.43      -0.41      0.0000 // E= -7.954
GLU  A  23     0.681      2.604      3.217      0.475      -26.77      -4.68      0.0000 // E= -24.468

```