基于结构的药物设计（一）

——蛋白质的结构与预测

上机课内容

PBD数据库初探

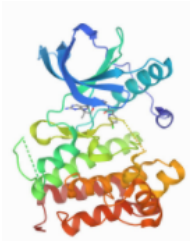
- 大分子PDB选择（选择分辨率高、有小分子配体的结构）
- 结构信息分析

同源模建构建蛋白质的三维结构

- 查找目标蛋白序列（Uniprot）
- Blast蛋白序列比对（>30%），找到相似蛋白模板（PDB库）
- 产生序列-结构对齐矩阵（align sequence to template）
- 蛋白建模
- 目标结构质量评估

PDB数据库初探

大分子PDB选择（选择分辨率高、有小分子配体的结构）



3D View

5JFX

Download File | View File

Crystal structure of TrkA in complex with PF-06273340

Jayasankar, J., Kurumbail, R., Skerratt, S., Brown, D.

(2016) J Med Chem **59**: 10084-10099

Released	2017-03-01
Method	X-RAY DIFFRACTION 1.63 Å
Organisms	Homo sapiens
Macromolecule	High affinity nerve growth factor receptor (protein)
Unique Ligands	6K4

3D View

5JFW

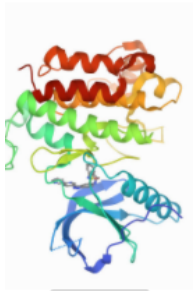
Download File | View File

Crystal structure of TrkA in complex with PF-05247452

Jayasankar, J., Kurumbail, R., Skerratt, S., Brown, D.

(2016) J Med Chem **59**: 10084-10099

Released	2017-03-01
Method	X-RAY DIFFRACTION 1.52 Å
Organisms	Homo sapiens
Macromolecule	High affinity nerve growth factor receptor (protein)
Unique Ligands	6K2



3D View

6J5L

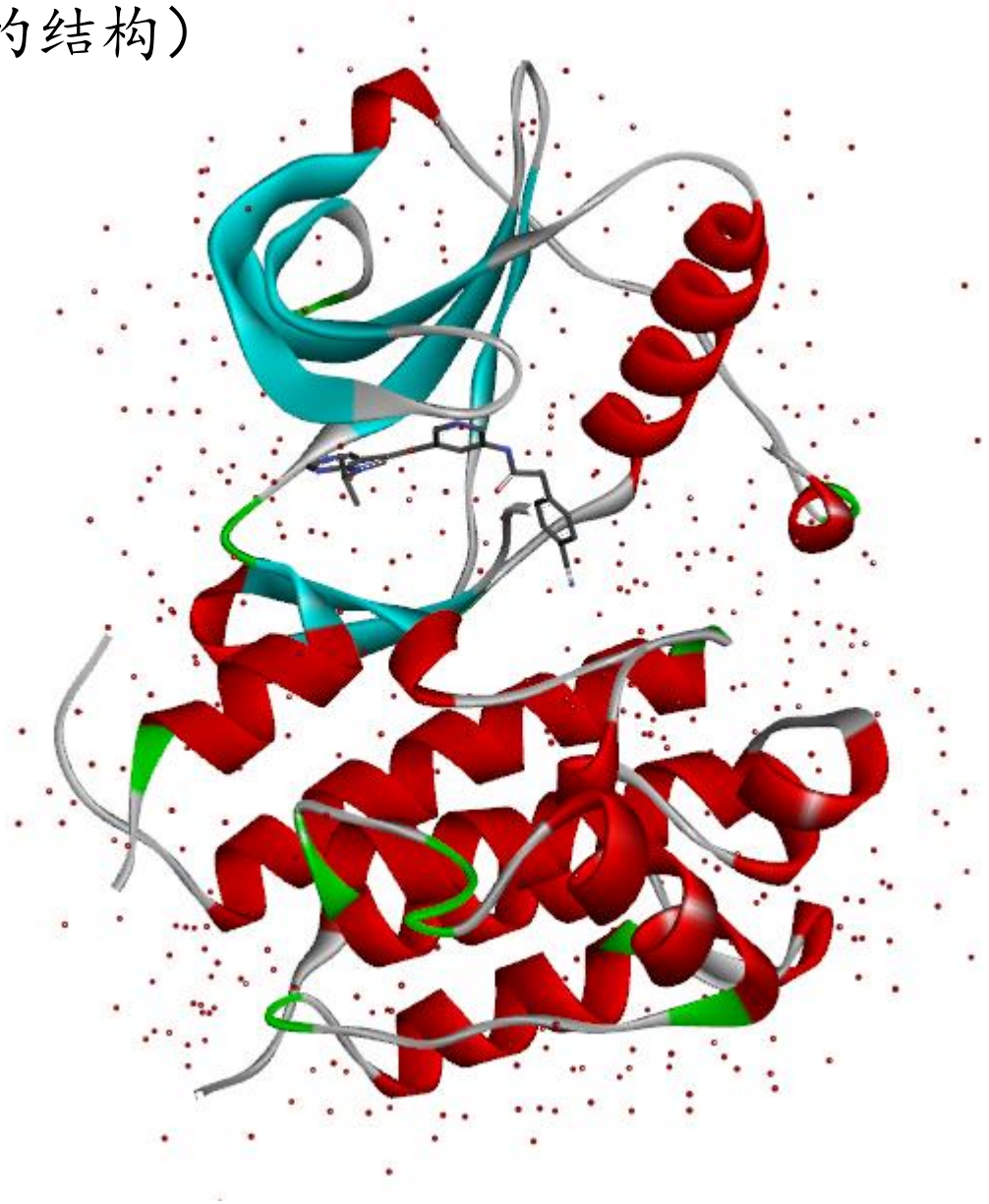
Download File | View File

Crystal structure of Trk-A in complex with the Pan-Trk Kinase Inhibitor, compound 10e

Kensuke, A., Kazutaka, I.

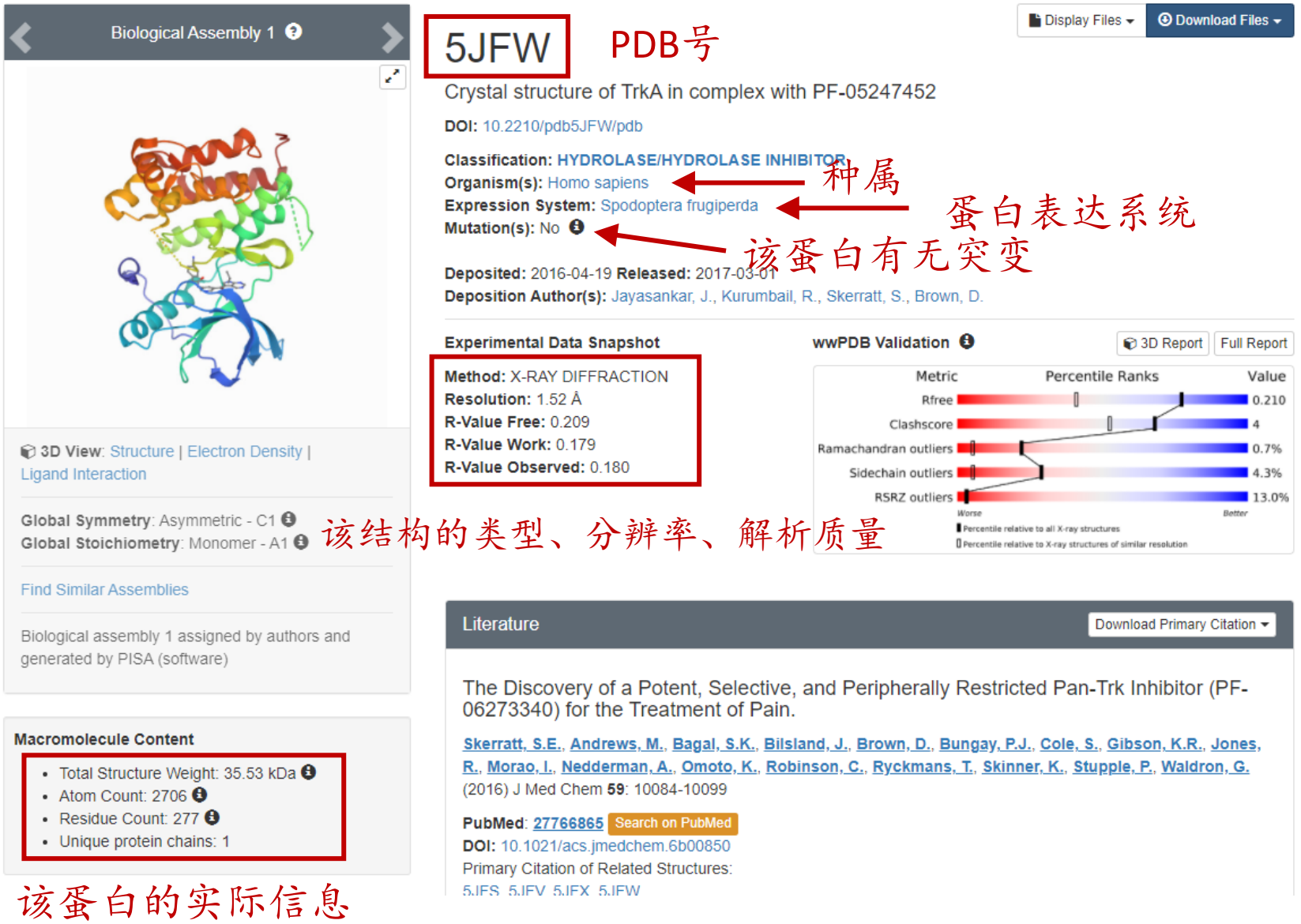
(2019) Bioorg Med Chem Lett **29**: 2320-2326

Released	2019-07-17
Method	X-RAY DIFFRACTION 2.3 Å
Organisms	Homo sapiens
Macromolecule	High affinity nerve growth factor receptor (protein)
Unique Ligands	B9C



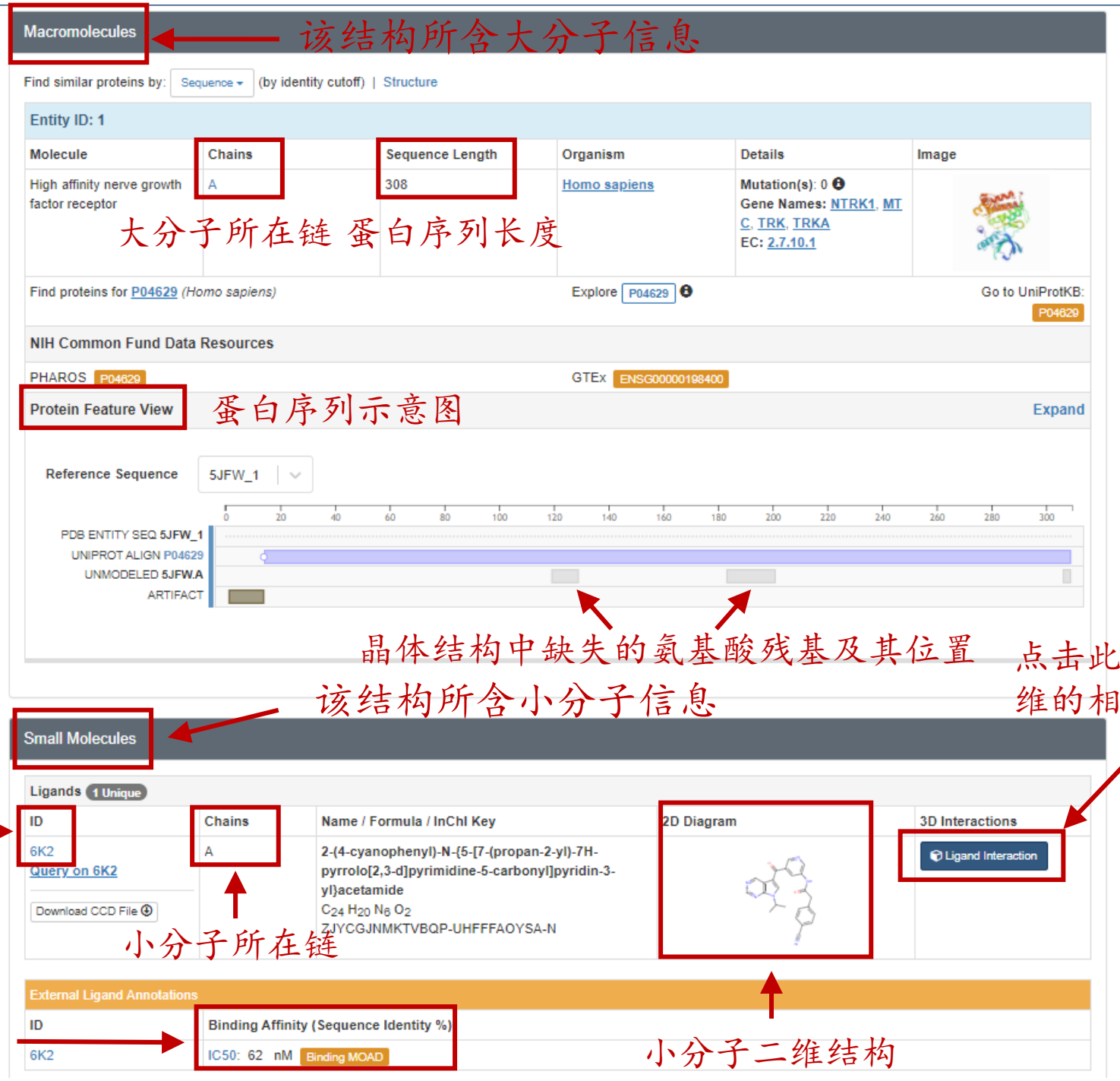
PDB数据库初探

结构信息分析



PDB数据库初探

结构信息分析



PDB数据库初探

结构信息分析

该结构的晶体学数据

Experimental Data & Validation

Experimental Data

Method: X-RAY DIFFRACTION
Resolution: 1.52 Å
R-Value Free: 0.209
R-Value Work: 0.179
R-Value Observed: 0.180
Space Group: P 6_4

Unit Cell:

Length (Å)	Angle (°)
a = 124.332	α = 90
b = 124.332	β = 90
c = 46.194	γ = 120

Software Package:

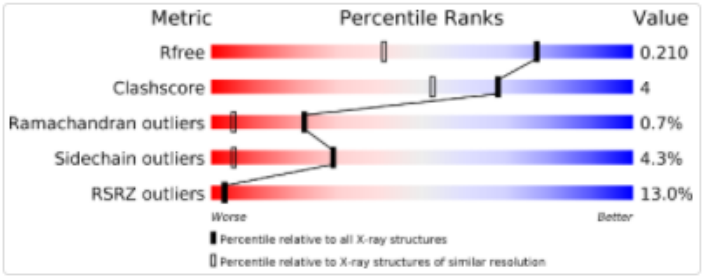
Software Name	Purpose
BUSTER	refinement
XDS	data reduction
SCALA	data scaling
PHASER	phasing

晶体解析软件

View more in-depth experimental data

Structure Validation

View Full Validation Report



该结构的数据上传、修改历史记录

Entry History

Deposition Data

Deposited Date: 2016-04-19
Released Date: 2017-03-01
Deposition Author(s): Jayasankar, J., Kurumbail, R., Skerratt, S., Brown, D.

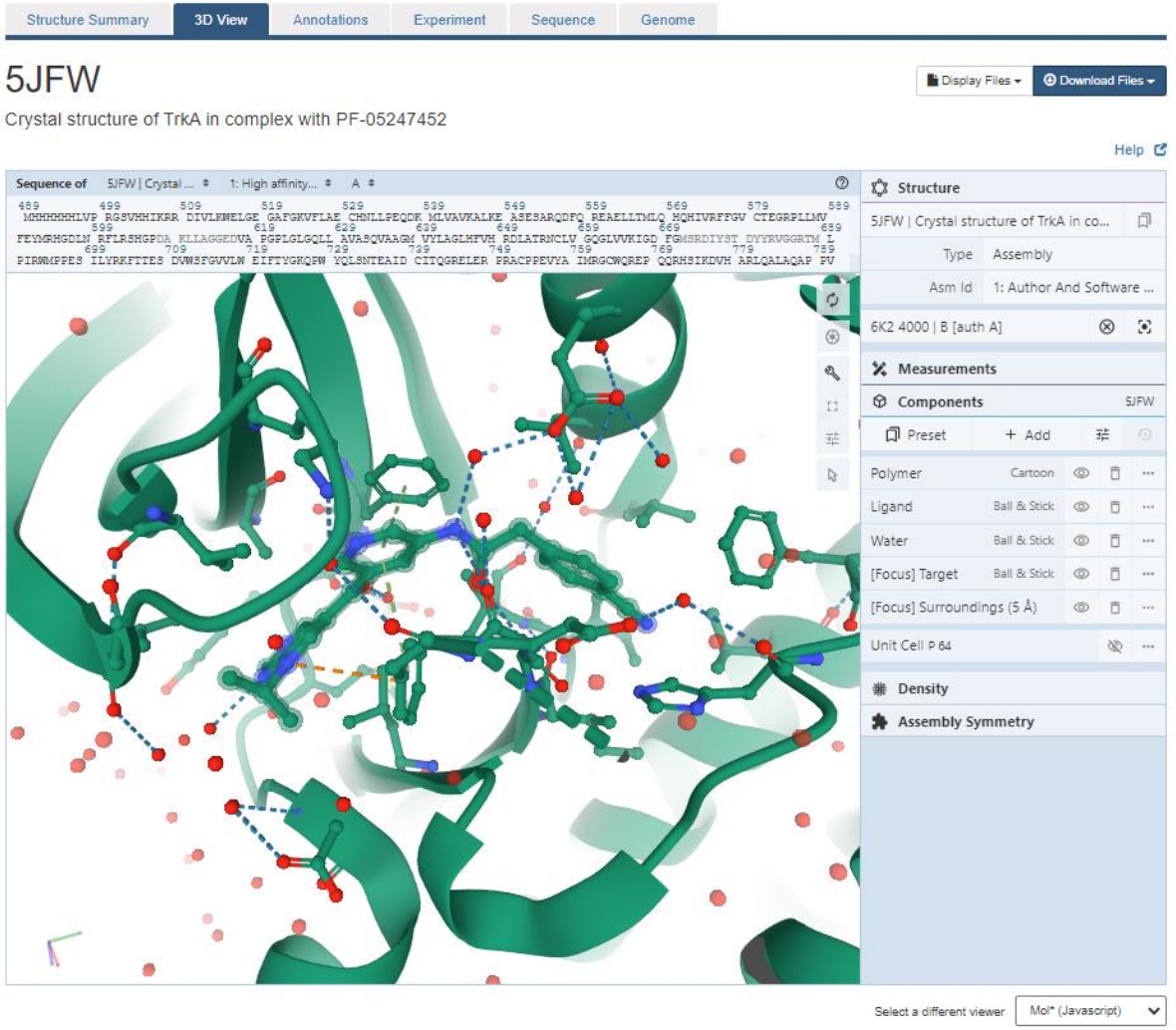
Revision History

- **Version 1.0: 2017-03-01**
  Type: Initial release

PDB数据库初探

结构信息分析

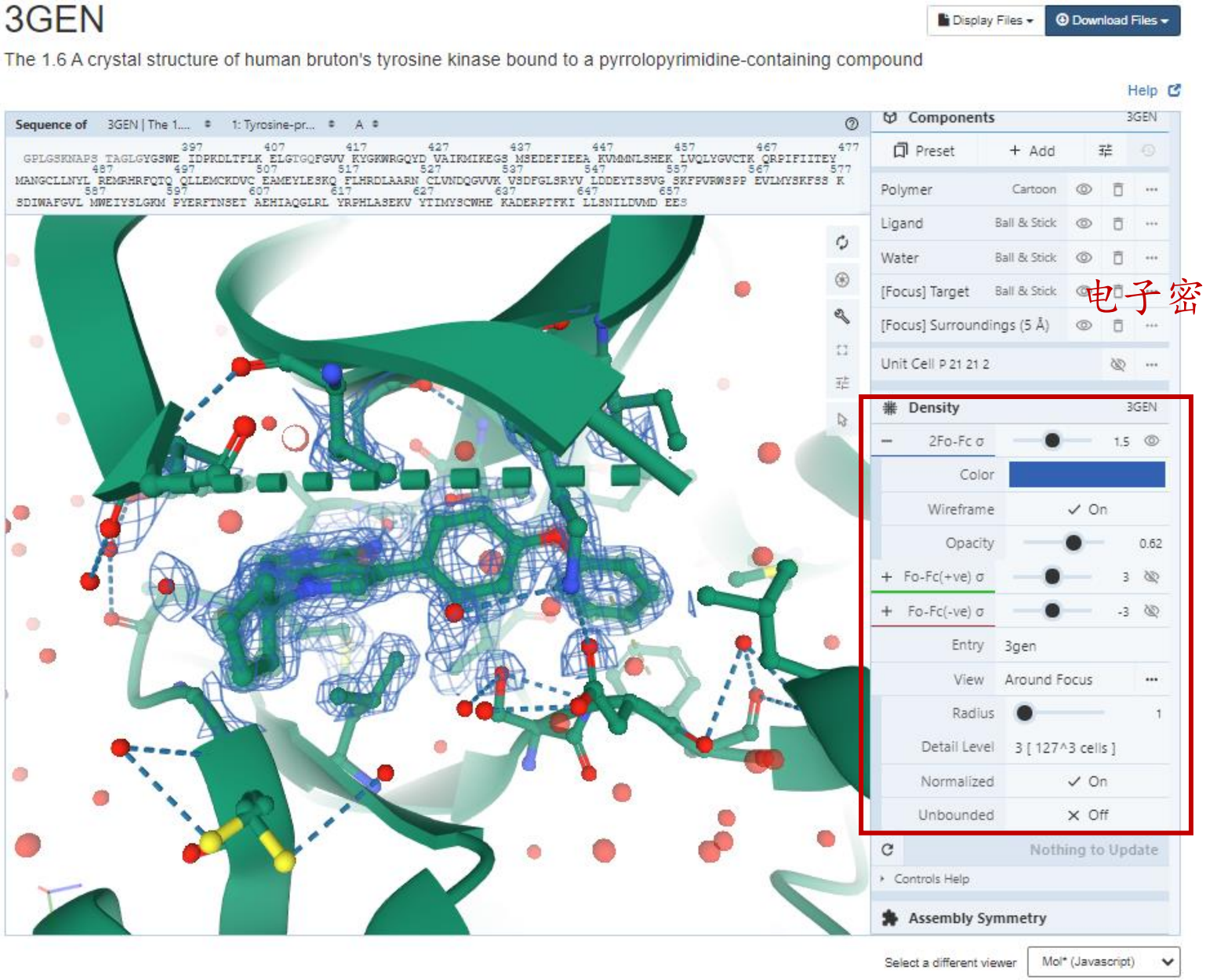
点击Ligand Interaction后，可预览配体-受体相互作用图，单击配体可现实/隐藏配体周围氨基酸残基。



PDB数据库初探

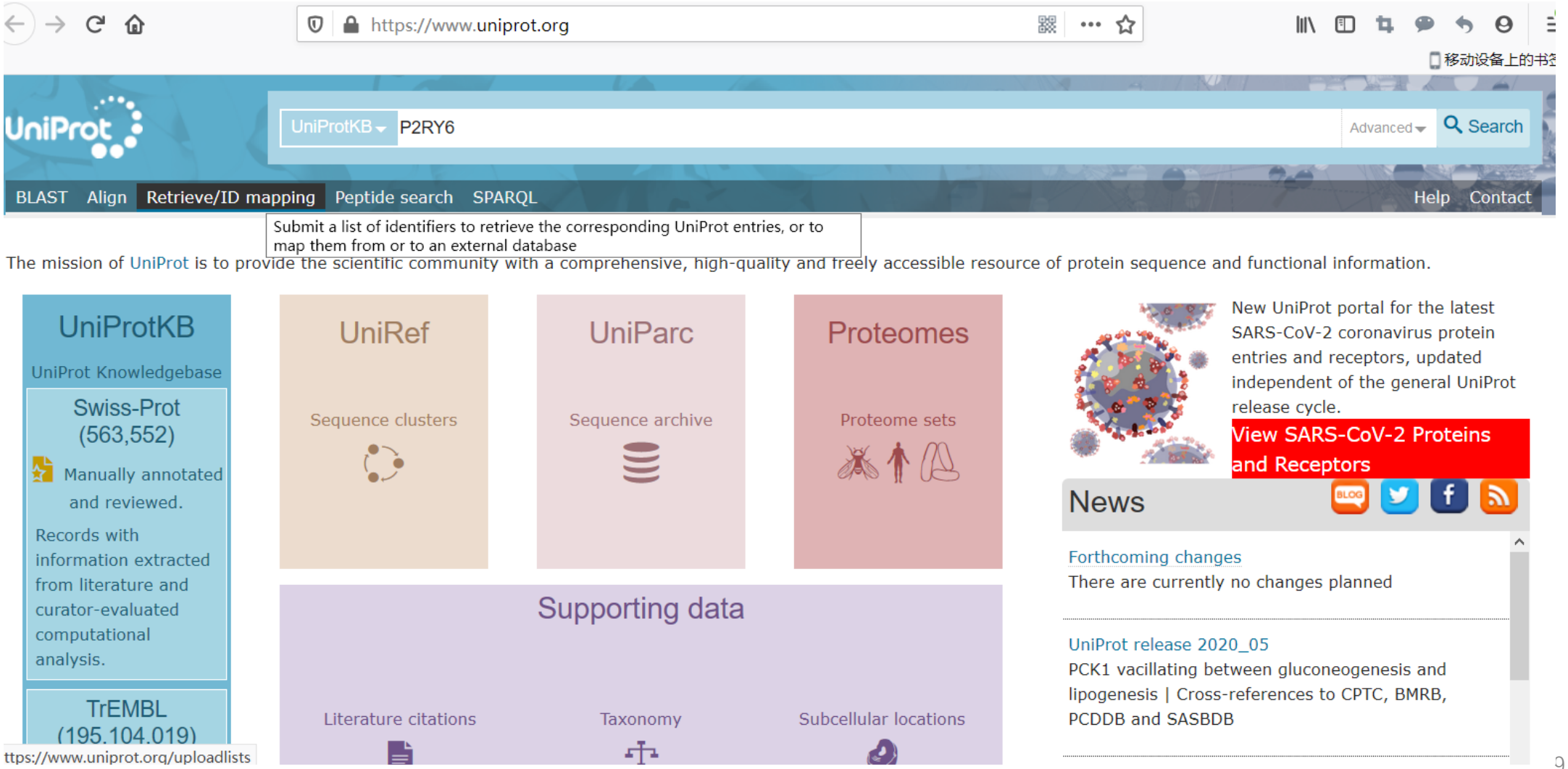
结构信息分析

开启蛋白晶体的电子密度预览模式



同源模建基本流程：（1）查找目标蛋白序列（P2RY6）

https://www.uniprot.org/



同源模建基本流程：（1）查找目标蛋白序列（P2RY6）

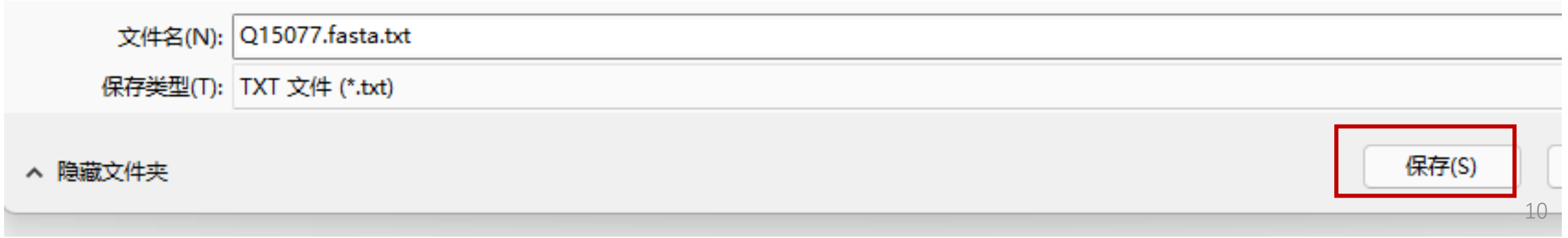
Q15077-1 [UniParc] FASTA Add to basket

« Hide

```
        10         20         30         40         50
MEWDNGTGQA LGLPPTTCVY RENFKQLLLP PVYSAVLAAG LPLNICVITQ
        60         70         80         90        100
ICTSRRALTR TAVYTLNLAL ADLLYACSLP LLIYNYAQGD HWPFGDFACR
       110        120        130        140        150
LVRFLFYANL HGSILFLTCI SFQRYLGICH PLAPWHKRGG RRAAWLVCVA
       160        170        180        190        200
VWLAVTTQCL PTAIFAATGI QRNRTVCYDL SPPALATHYM PYGMALTVIG
       210        220        230        240        250
FLLPFAALLA CYCLLACRLC RQDGPAEPVA QERRGKAARM AVVVAAAFAI
       260        270        280        290        300
SFLPFHITKT AYLAVRSTPG VPCTVLEAFA AAYKGTRPFA SANSVLDPIL
       310        320
FYFTQKKFRR RPHELLQKLT AKWQRQGR
```

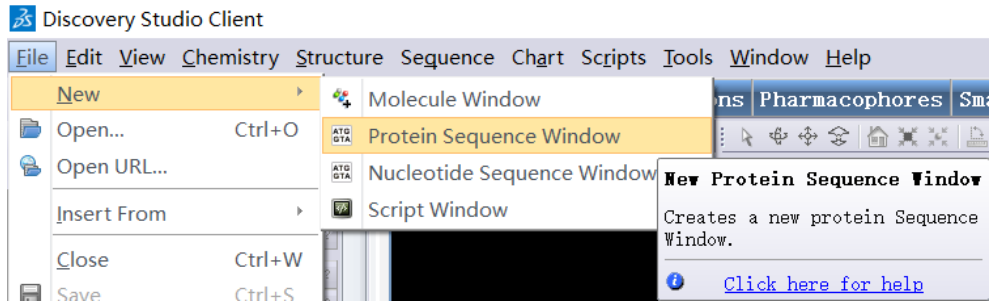
Length: 328
Mass (Da): 36,429
Last modified: November 1, 1996 - v1
Checksum: AAD6C55A43818107

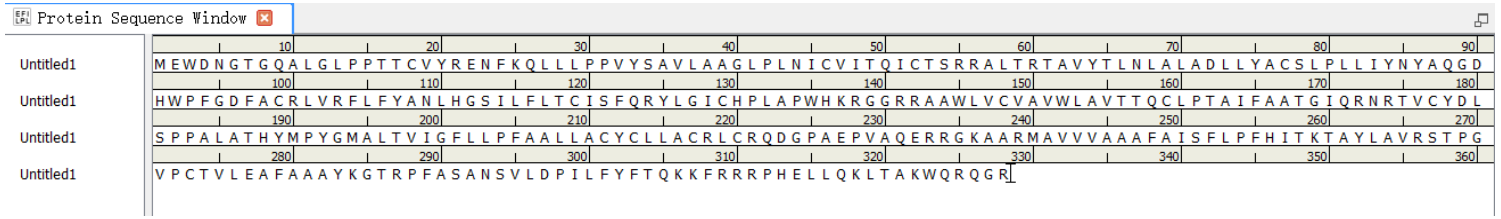
BLAST GO



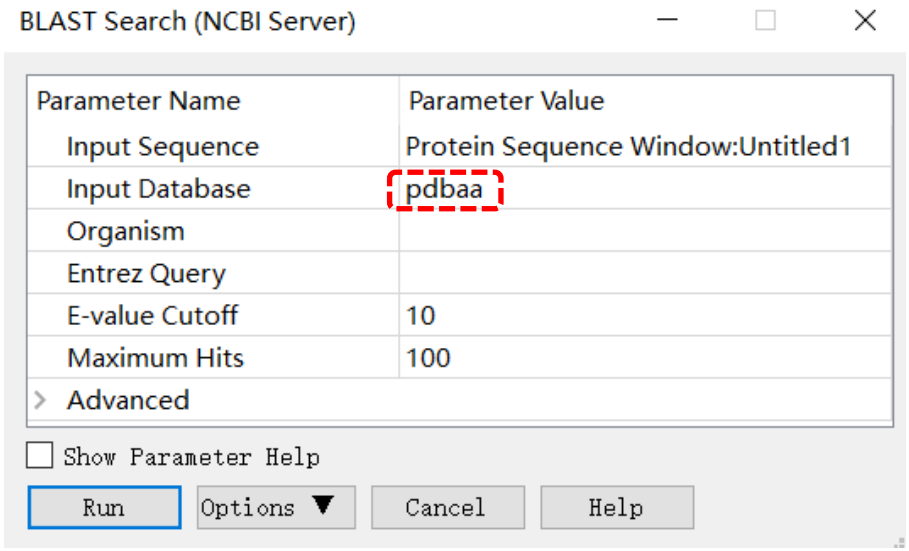
同源模建基本流程：（2）Blast蛋白序列比对

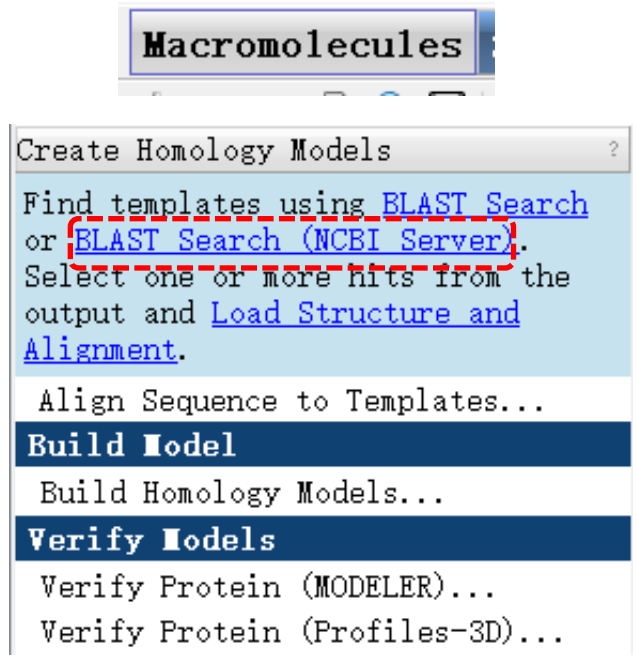
- 创建蛋白序列窗口，填入目标蛋白序列



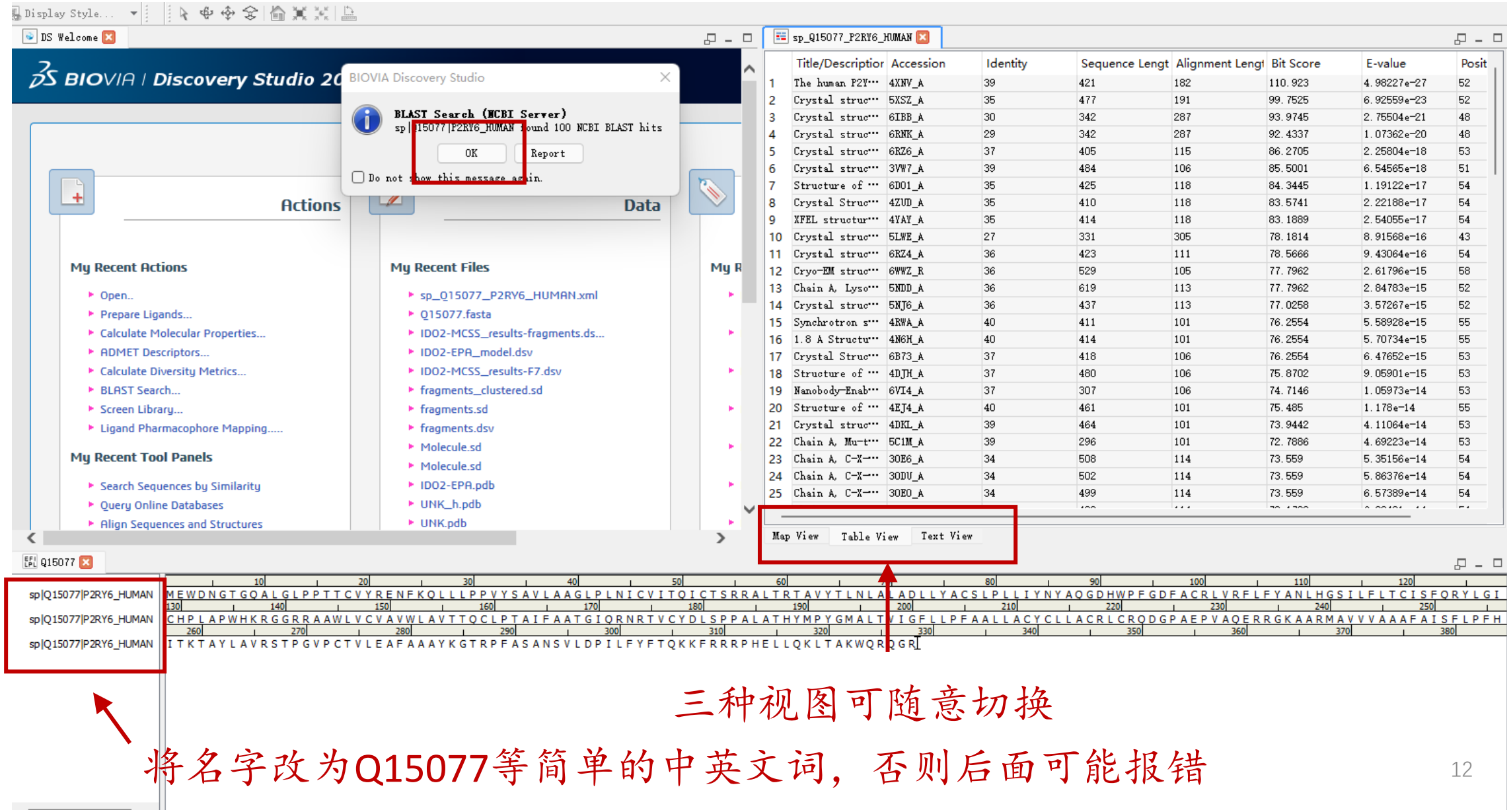


- 使用蛋白晶体数据库找寻相似模板蛋白





同源模建基本流程：（2）Blast蛋白序列比对



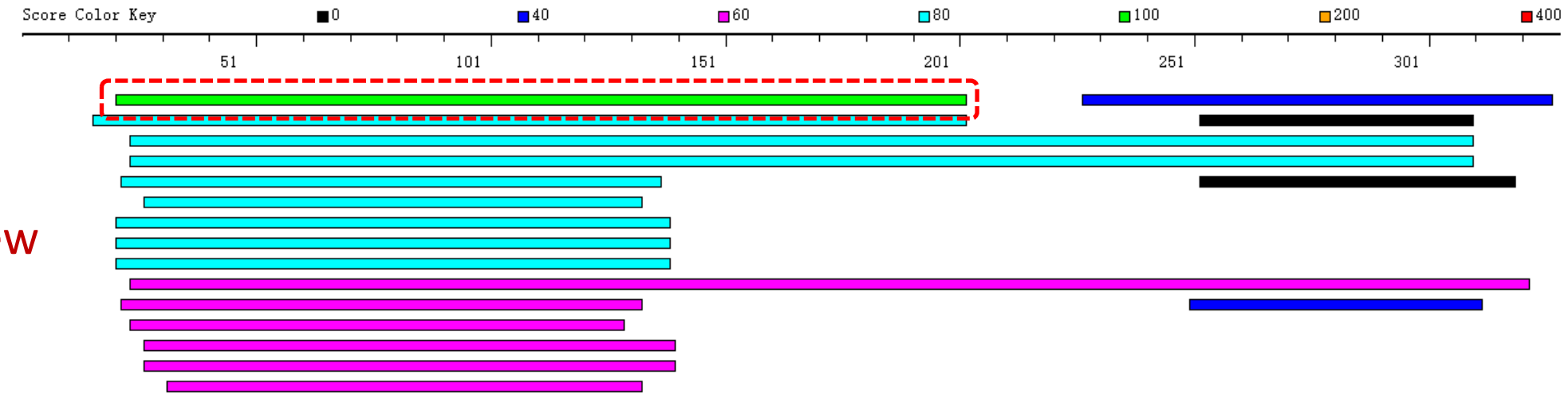
	Title/Description	Accession	Identity	Sequence Length	Alignment Length	Bit Score	E-value	Position
1	The human P2Y…	4XNV_A	39	421	182	110.923	4.98227e-27	52
2	Crystal struc…	5XSZ_A	35	477	191	99.7525	6.92559e-23	52
3	Crystal struc…	6IBB_A	30	342	287	93.9745	2.75504e-21	48
4	Crystal struc…	6RNK_A	29	342	287	92.4337	1.07362e-20	48
5	Crystal struc…	6RZ6_A	37	405	115	86.2705	2.25804e-18	53
6	Crystal struc…	3VW7_A	39	484	106	85.5001	6.54565e-18	51
7	Structure of …	6DO1_A	35	425	118	84.3445	1.19122e-17	54
8	Crystal Struc…	4ZUD_A	35	410	118	83.5741	2.22188e-17	54
9	XFEL structur…	4YAY_A	35	414	118	83.1889	2.54055e-17	54
10	Crystal struc…	5LWE_A	27	331	305	78.1814	8.91568e-16	43
11	Crystal struc…	6RZ4_A	36	423	111	78.5666	9.43064e-16	54
12	Cryo-EM struc…	6WWZ_R	36	529	105	77.7962	2.61796e-15	58
13	Chain A, Lyso…	5NDD_A	36	619	113	77.7962	2.84783e-15	52
14	Crystal struc…	5NJ6_A	36	437	113	77.0258	3.57267e-15	52
15	Synchrotron s…	4RWA_A	40	411	101	76.2554	5.58928e-15	55
16	1.8 A Structu…	4N6H_A	40	414	101	76.2554	5.70734e-15	55
17	Crystal Struc…	6B73_A	37	418	106	76.2554	6.47652e-15	53
18	Structure of …	4DJH_A	37	480	106	75.8702	9.05901e-15	53
19	Nanobody-Enab…	6VI4_A	37	307	106	74.7146	1.05973e-14	53
20	Structure of …	4EJ4_A	40	461	101	75.485	1.178e-14	55
21	Crystal struc…	4DKL_A	39	464	101	73.9442	4.11064e-14	53
22	Chain A, Mu-t…	5C1M_A	39	296	101	72.7886	4.69223e-14	53
23	Chain A, C-X-…	3OE6_A	34	508	114	73.559	5.35156e-14	54
24	Chain A, C-X-…	3ODU_A	34	502	114	73.559	5.86376e-14	54
25	Chain A, C-X-…	3OEO_A	34	499	114	73.559	6.57389e-14	54

三种视图可随意切换

将名字改为Q15077等简单的中英文词，否则后面可能报错

同源模建基本流程：选择合适蛋白模板

Map View



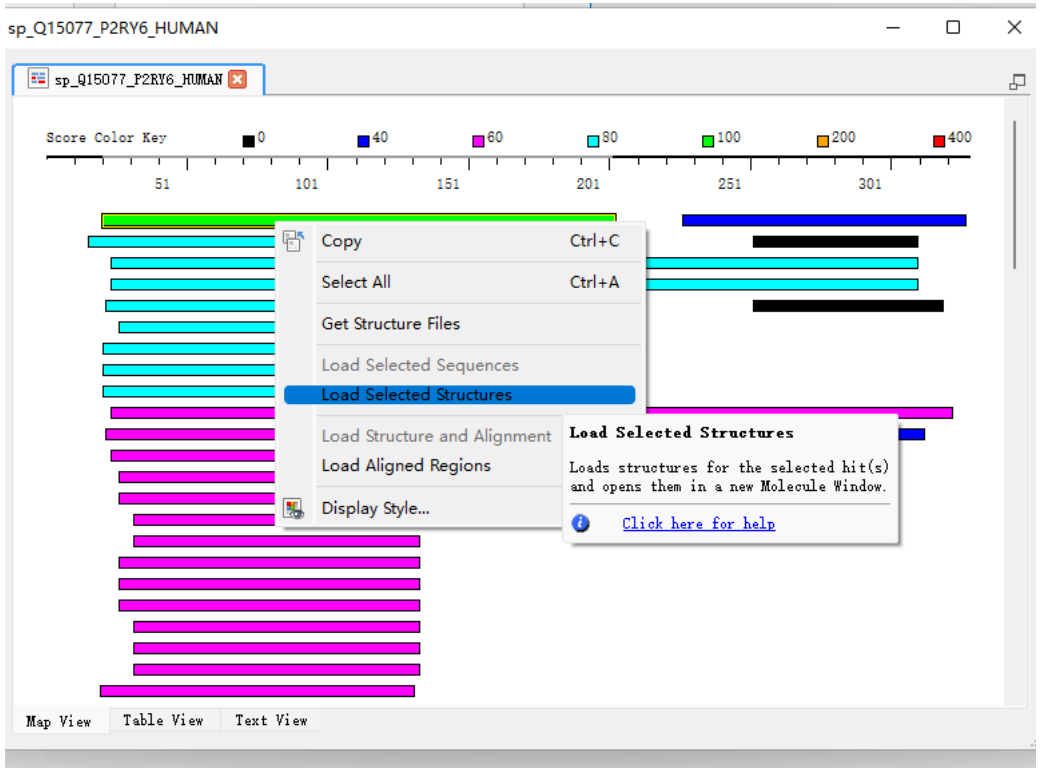
	Title/Description	Accession	Identity	Sequence Length	Alignment Length	Bit Score	E-value	Positive
1	Chain A, P2Y purino…	4XNV_A	39	421	182	110.923	4.42019e-27	52
2	Chain A, Lysophosph…	5XSZ_A	35	477	191	99.7525	6.14428e-23	52
3	Chain A, Succinate …	6IBB_A	30	342	287	93.9745	2.44423e-21	48
4	Chain A, Succinate …	6RNK_A	29	342	287	92.4337	9.52501e-21	48
5	Chain A, Cysteinyl …	6RZ6_A	37	405	115	86.2705	2.0033e-18	53
6	Chain A, Proteinase…	3VW7_A	39	484	106	85.5001	5.8072e-18	51
7	Chain A, Type-1 ang…	6DO1_A	35	425	118	84.3445	1.05683e-17	54
8	Chain A, Chimera pr…	4ZUD_A	35	410	118	83.5741	1.97122e-17	54
9	Chain A, Soluble cy…	4YAY_A	35	414	118	83.1889	2.25394e-17	54
10	Chain A, C-C chemok…	5LWE_A	27	331	305	78.1814	7.90985e-16	43
11	Chain A, Cysteinyl …	6RZ4_A	36	423	111	78.5666	8.36671e-16	54
12	Chain R, C-C chemok…	6WWZ_R	36	529	105	77.7962	2.32261e-15	58
13	Chain A, Lysozyme,P…	5NDD_A	36	619	113	77.7962	2.52655e-15	52
14	Chain A, Proteinase…	5NJ6_A	36	437	113	77.0258	3.16962e-15	52
15	Chain A, Soluble cy…	4RWA_A	40	411	101	76.2554	4.95872e-15	55
16	Chain A, Soluble cy…	4N6H_A	40	414	101	76.2554	5.06346e-15	55
17	Chain A, Soluble cy…	6B73_A	37	418	106	76.2554	5.74586e-15	53
18	Chain A, Kappa-type…	4DJH_A	37	480	106	75.8702	8.03701e-15	53
19	Chain A, Kappa opio…	6VI4_A	37	307	106	74.7146	9.40175e-15	53

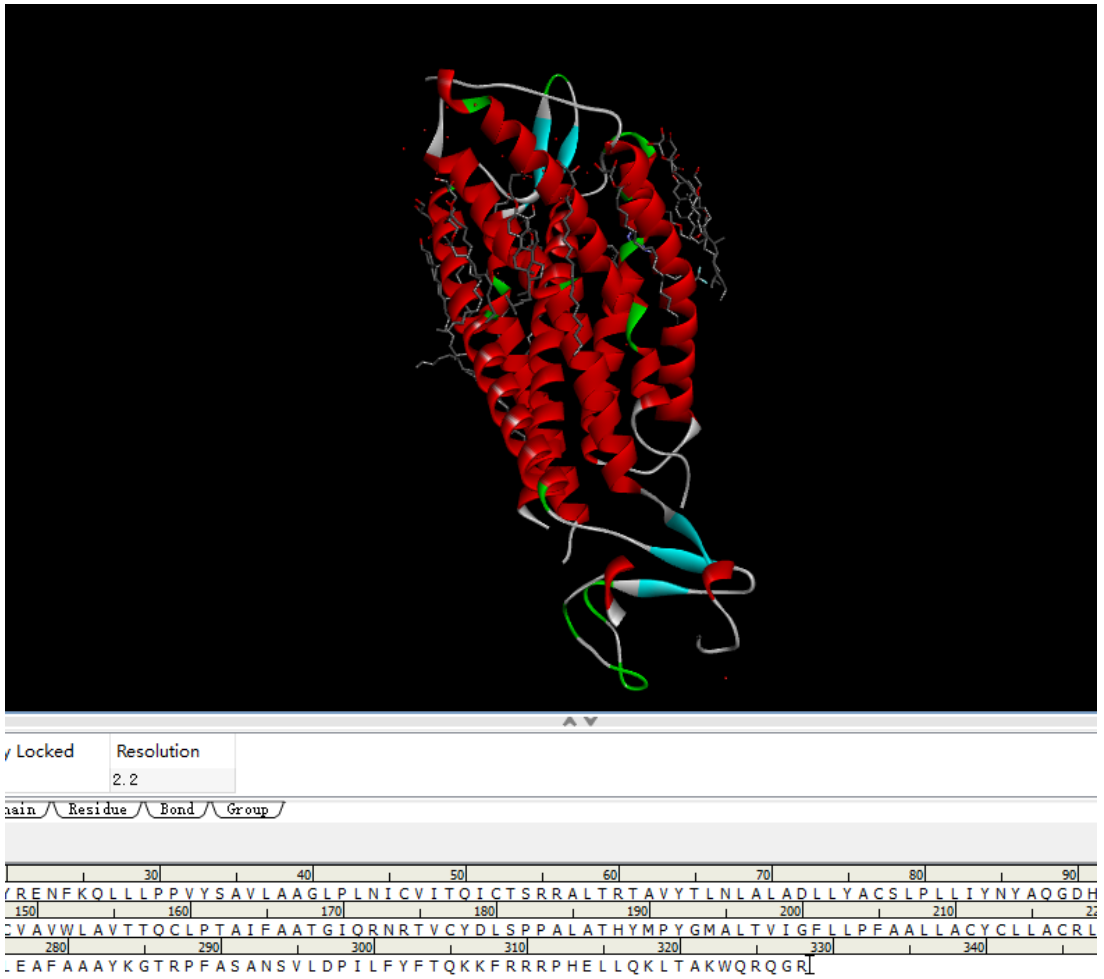
- 选择合适蛋白模板并载人结构

Table View

同源模建基本流程：选择合适蛋白模板

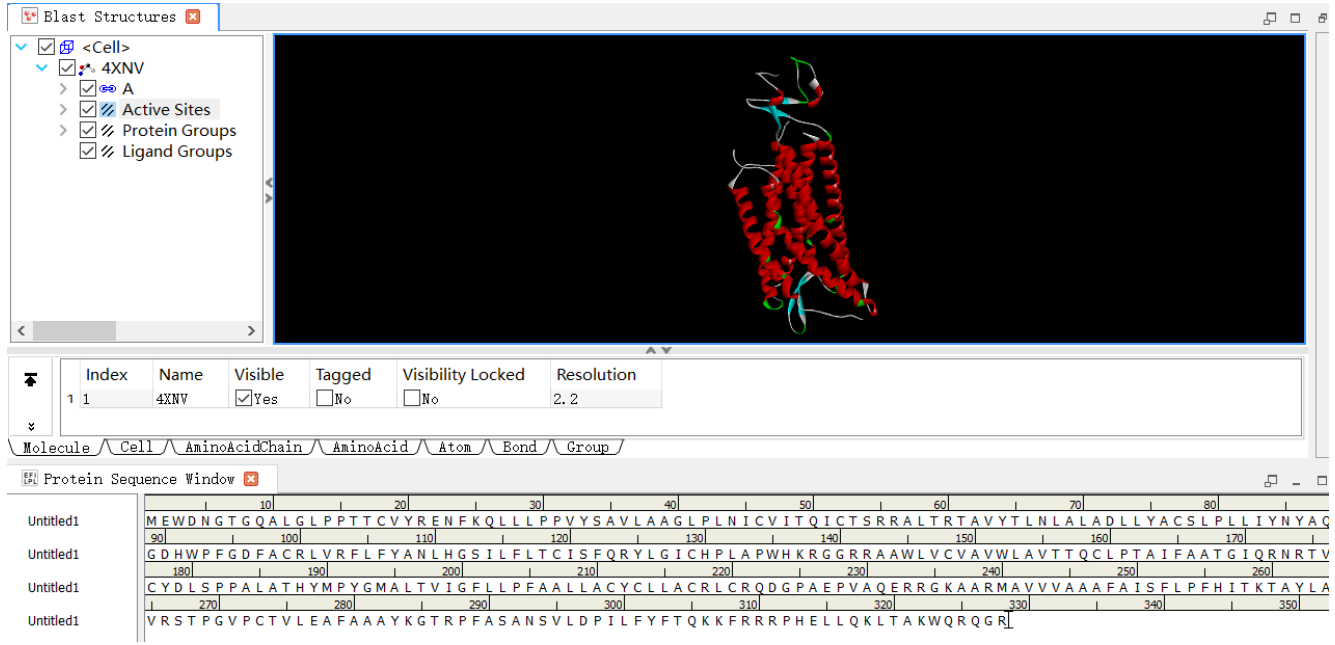
Map View

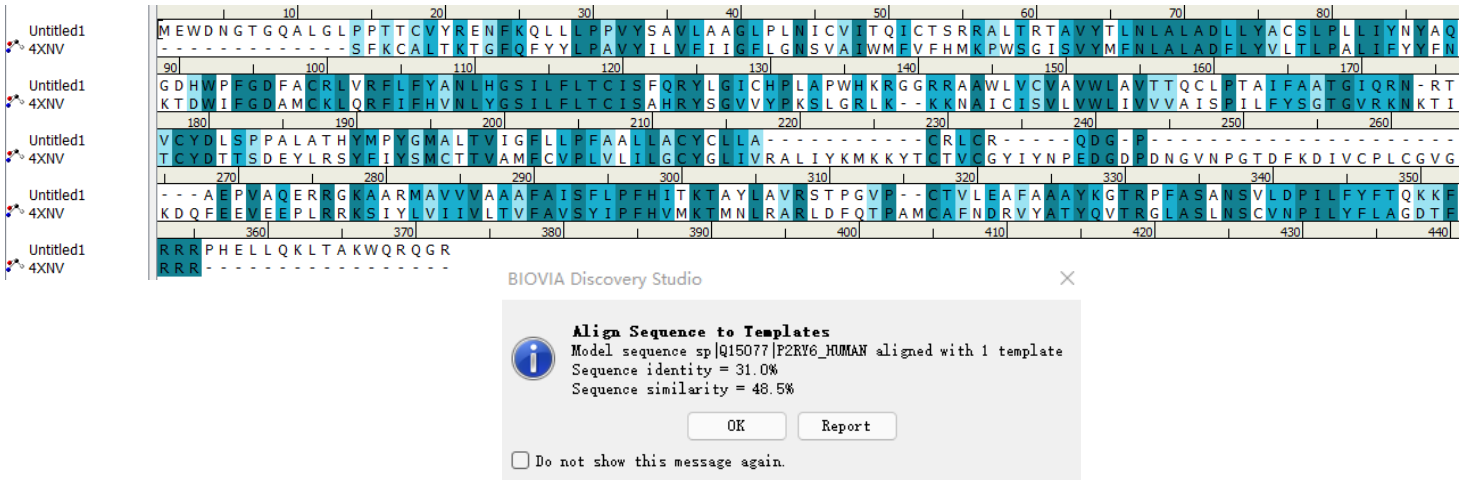


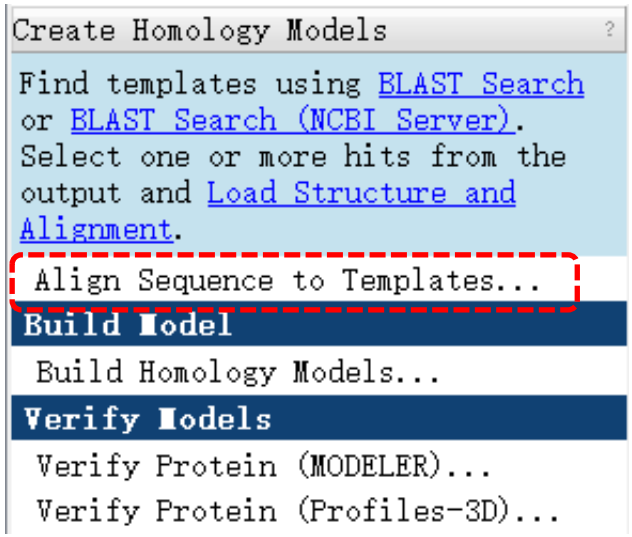


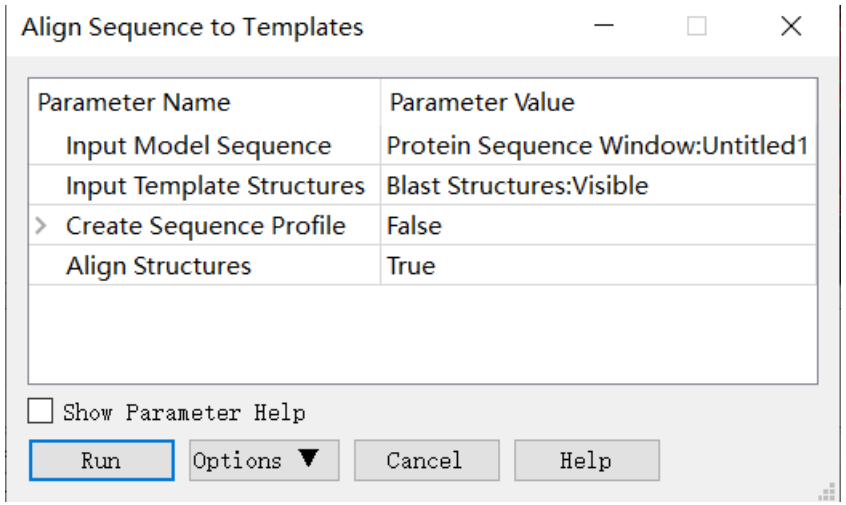
- 载入模板蛋白

同源模建基本流程：（3）创建序列-结构对齐矩阵

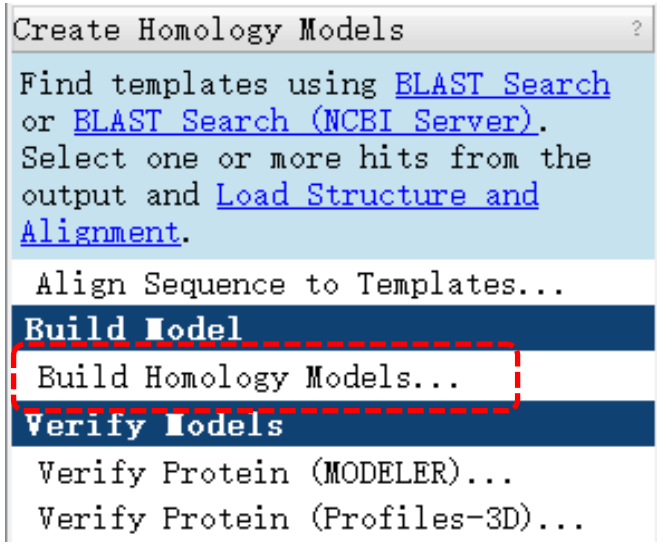


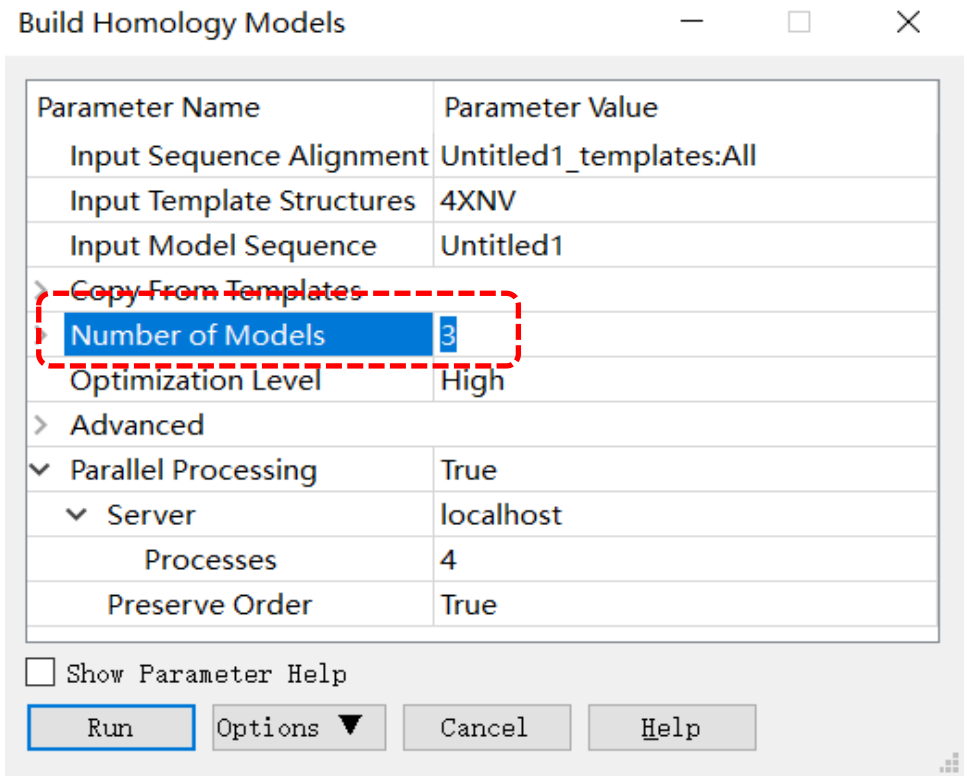


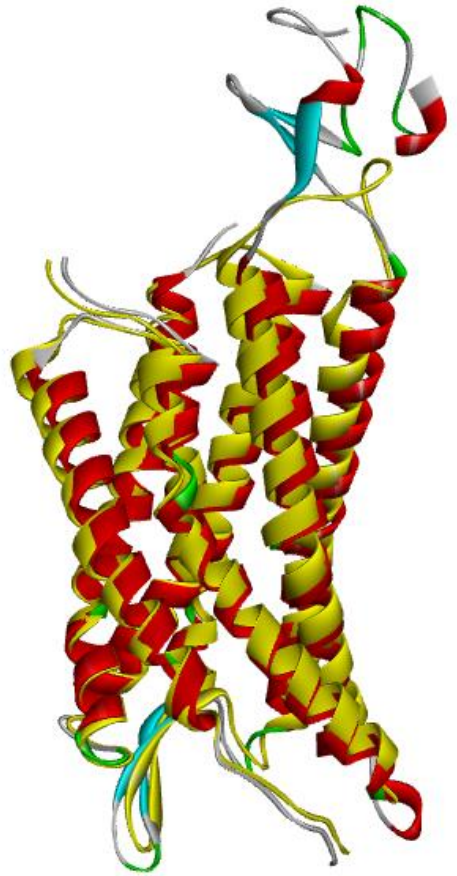




同源模建基本流程：（4）蛋白建模

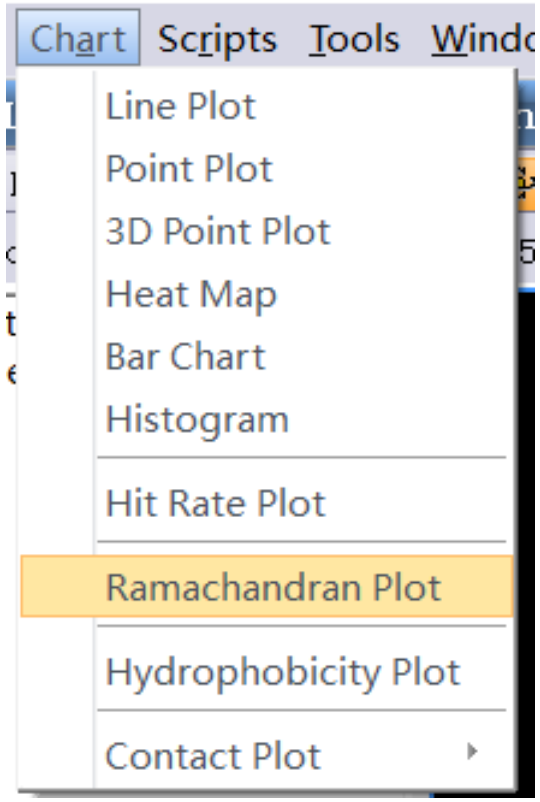


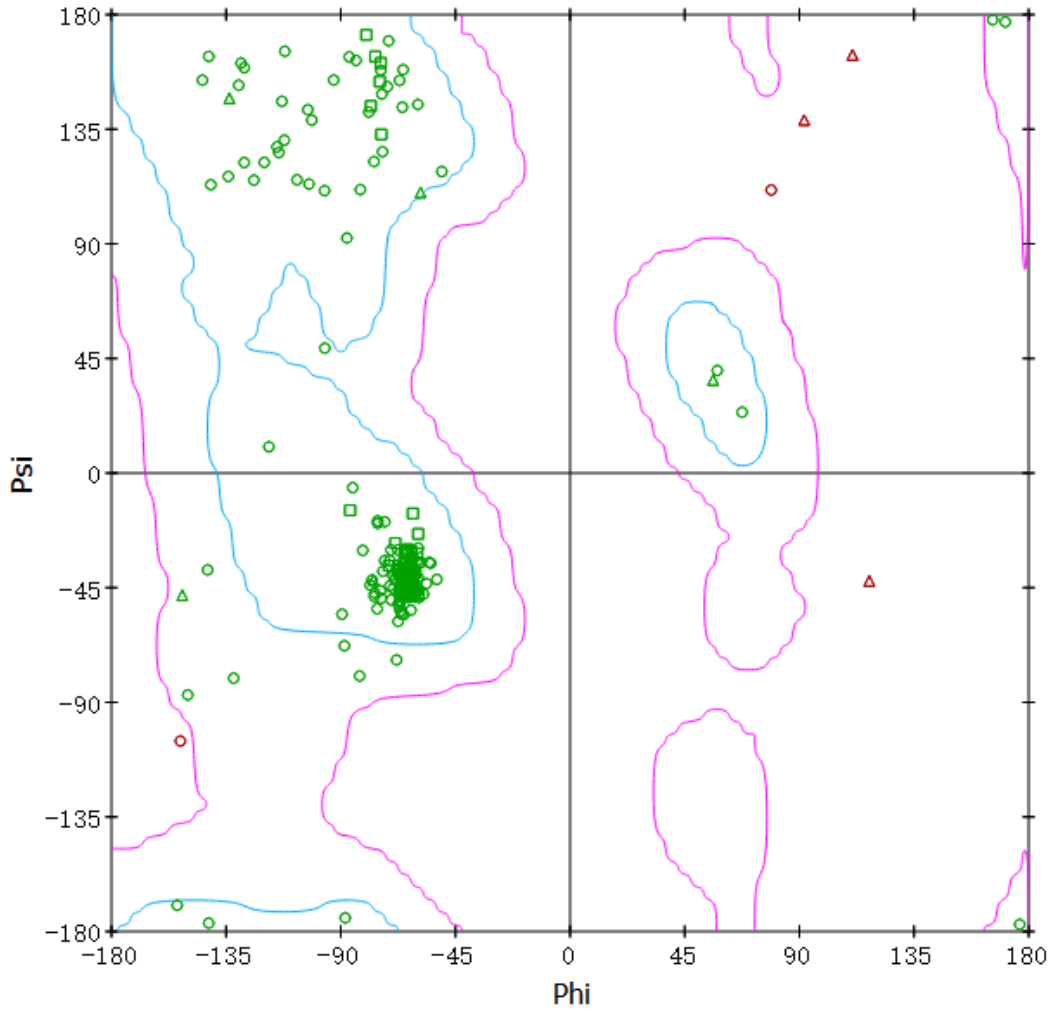




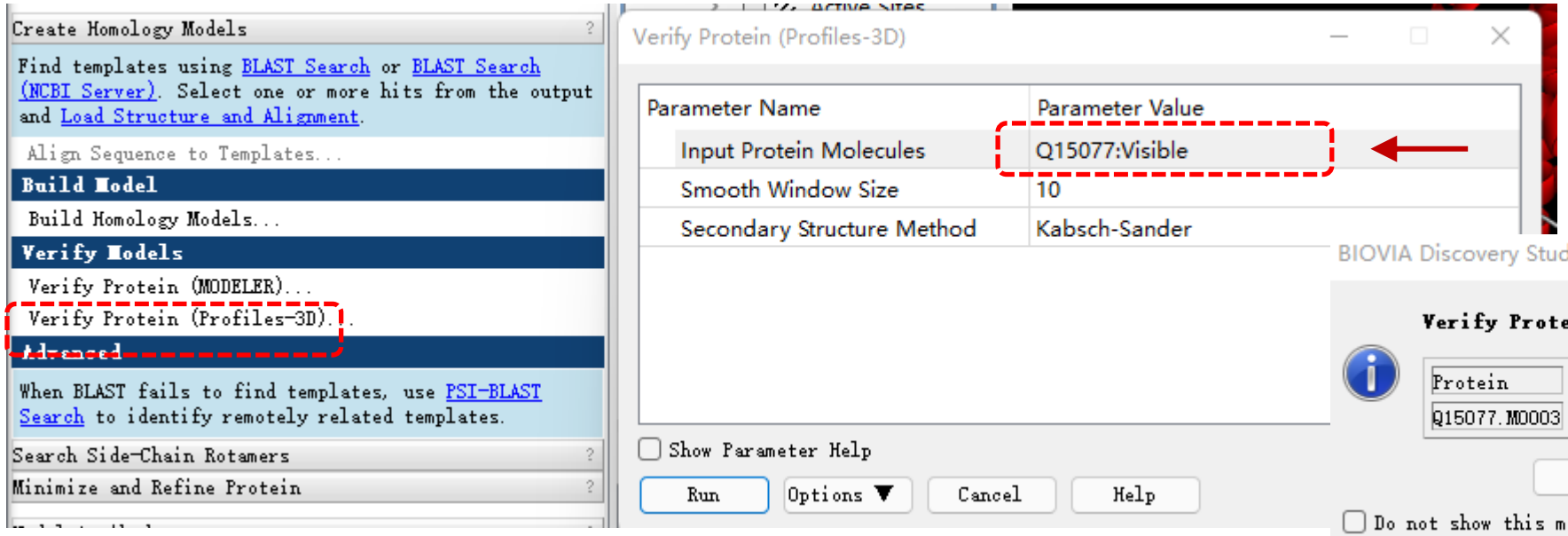
PDF Total Energy	PDF Physical Energy	DOPE Score	RMSD	NOverlap
1,219.29	742.159	-36,820.3	3.744	296
1,238.35	754.255	-36,606.4	1.447	298
1,412.02	783.256	-36,696.1	1.16	298

同源模建基本流程：（5）模建蛋白质量评估——拉氏图

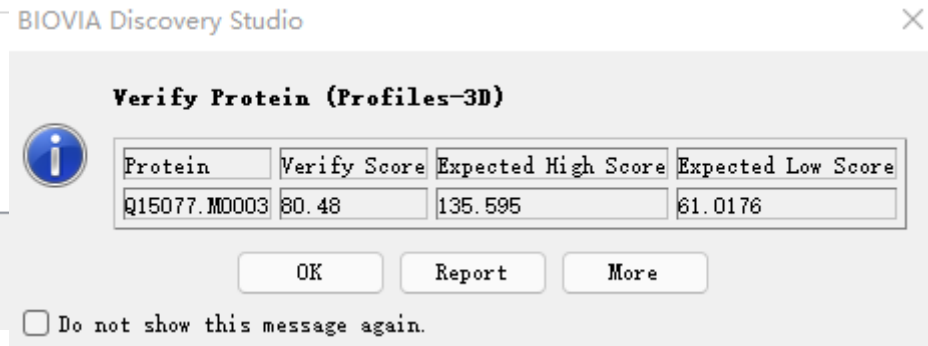




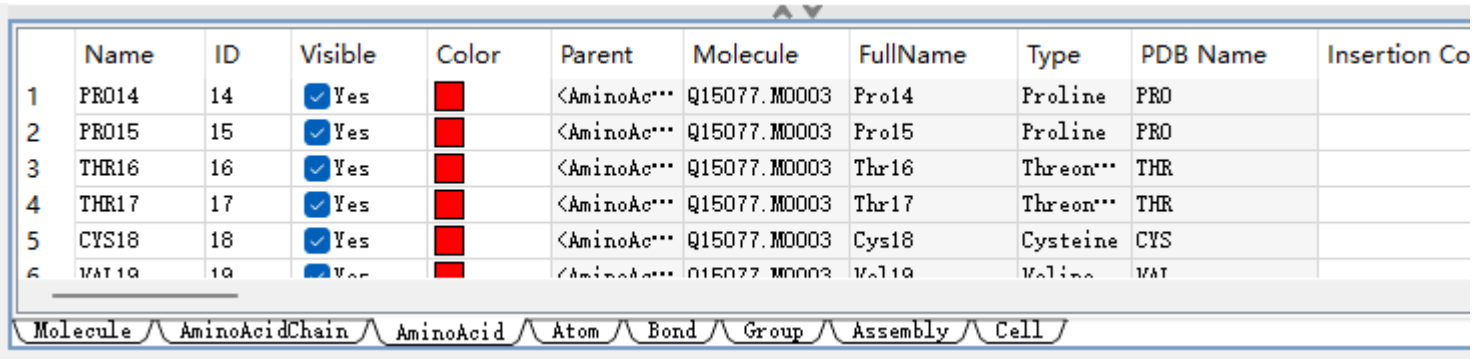
同源模建基本流程：（5）模建蛋白质量评估——Profile-3D



选中需要计算的结构

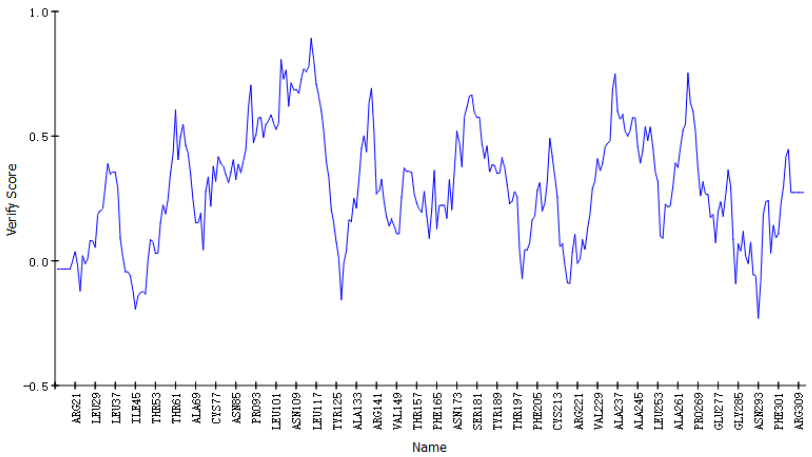


- 以残基打分作图，选择Aminoacid选项卡，Chart-Line Plot，X轴设为name，Y轴设为verify score



	Name	ID	Visible	Color	Parent	Molecule	FullName	Type	PDB Name	Insertion Co
1	PRO14	14	Yes		<AminoAc…	Q15077.M0003	Pro14	Proline	PRO	
2	PRO15	15	Yes		<AminoAc…	Q15077.M0003	Pro15	Proline	PRO	
3	THR16	16	Yes		<AminoAc…	Q15077.M0003	Thr16	Threon…	THR	
4	THR17	17	Yes		<AminoAc…	Q15077.M0003	Thr17	Threon…	THR	
5	CYS18	18	Yes		<AminoAc…	Q15077.M0003	Cys18	Cysteine	CYS	

Molecule | AminoAcidChain | AminoAcid | Atom | Bond | Group | Assembly | Cell



任务：

（1）从PDB数据库获取感兴趣蛋白结构并挑选优质的蛋白晶体结构（如EGFR、c-MET、VEGFR等）

（2）从Uniprot数据库获取问询蛋白序列（P2RY6）

（3）通过BLAST序列比对获得PDB库中相似度最高模板蛋白

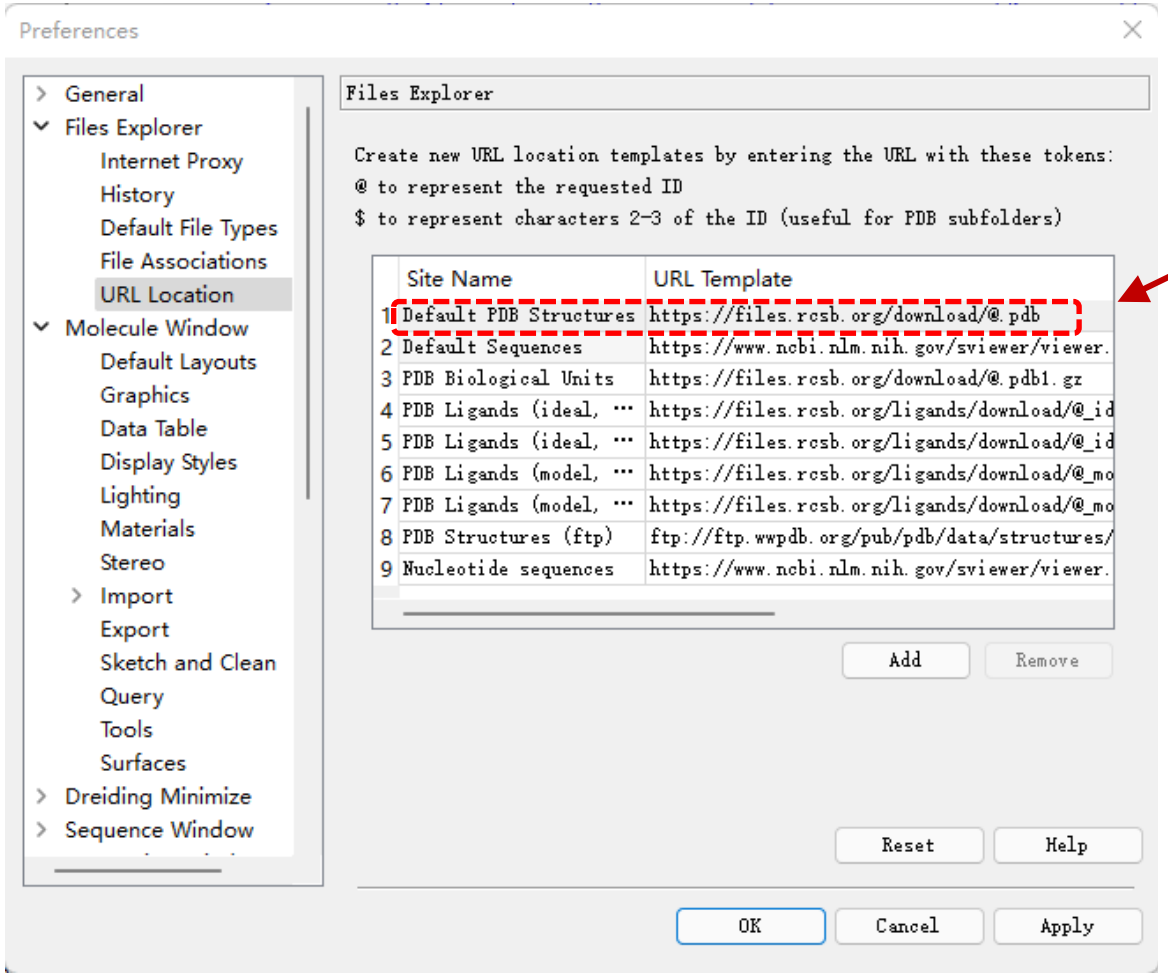
（4）序列对齐，同源模建

（5）结果分析（使用拉氏图和Profile-3D验证模型的合理性）

注意事项：

操作前，需要确保DS软件中，PDB数据库的地址是正确的。

EDIT—Preferences—Files Explorer



https://files.rcsb.org/download/@.pdb

实验报告：

（1）实验目的

模建未知跨膜蛋白P2RY6三维结构

（2）操作流程

Unprot蛋白序列搜索、BLAST蛋白序列比对，确定模板蛋白、序列-模板对齐、蛋白建模

（3）结果与讨论

观测蛋白结构，用拉氏图和Profile-3D分析蛋白结构合理性