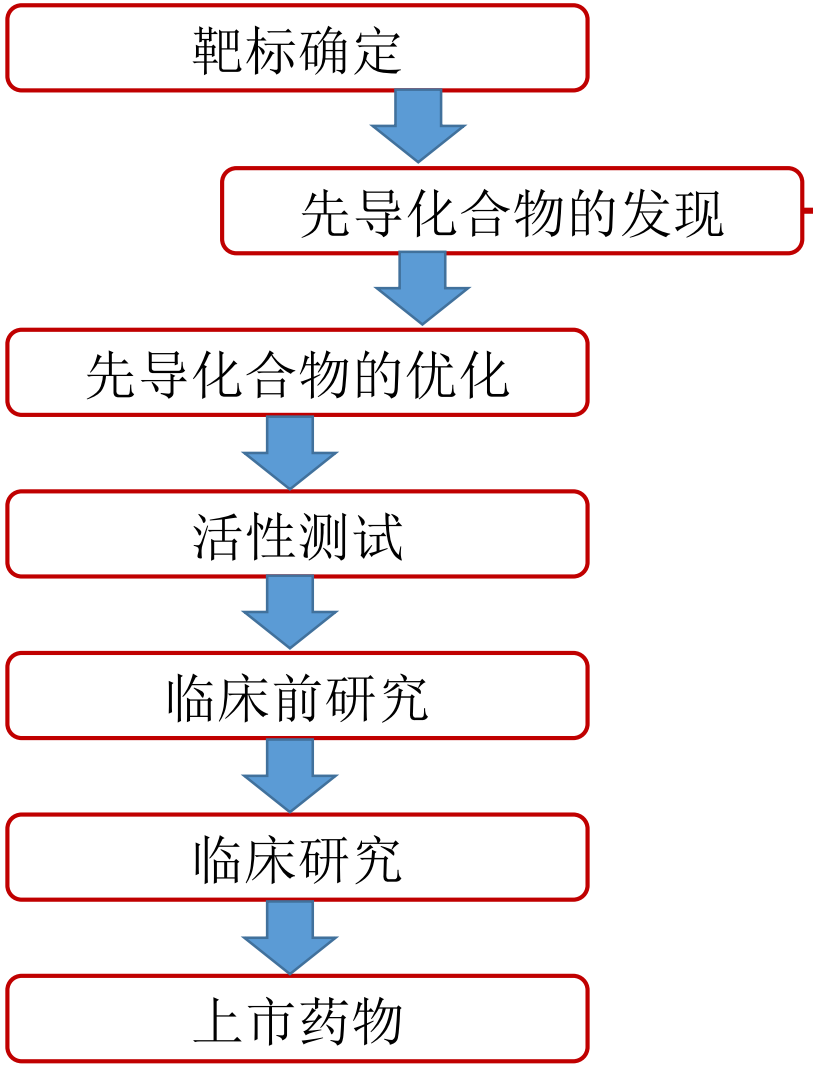
药物设计

G04小组
G07小组

药物设计的一般过程

虚拟筛选逐渐成为发现先导化合物的重要手段



高通量实验筛选

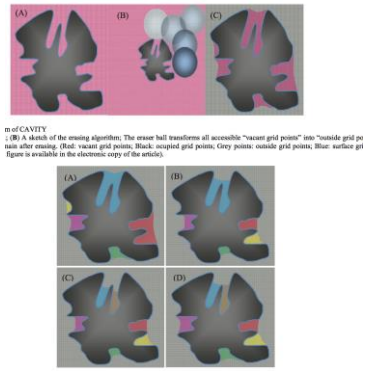
- 大量体外实验测试（高成本）
- 需要大量化合物
- 命中率极低

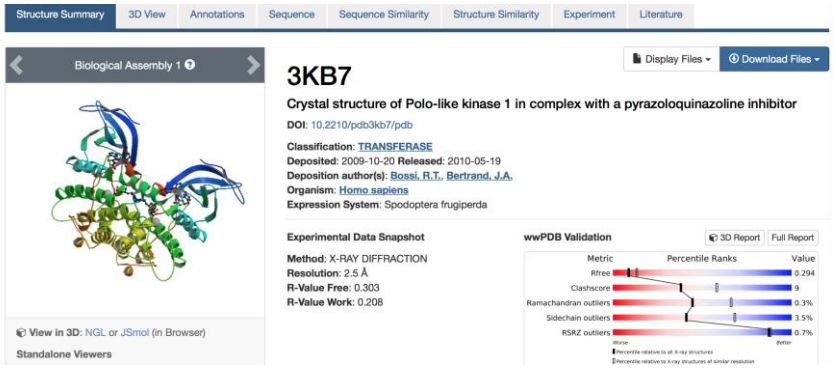
计算机辅助的虚拟筛选

- 计算机模拟分析（低成本）
- 需要有限化合物
- 命中率相对较高

虚拟筛选的一般流程

RCSB PDB PROTEIN DATA BANK
An Information Portal to
130599 Biological
Macromolecular Structures





$$\text{Cavity Score} = \frac{\text{Volume} - \text{Adjust volume}}{\text{Surface area} - \text{Adjust surface area}}$$

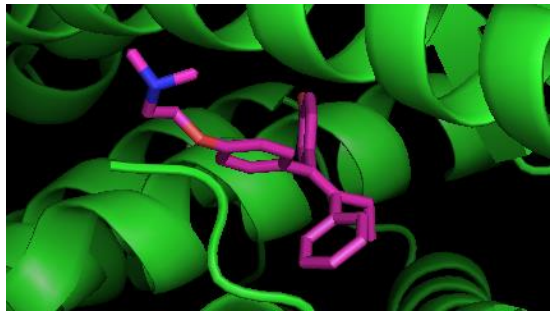
D U D • E
A Database of Useful Decoys: Enhanced

ZINC15
Welcome to ZINC, a free database of commercially-available compounds for virtual screening. ZINC contains over 100 million purchasable compounds in ready-to-dock, 3D formats.

specs
www.specs.net

AutoDock

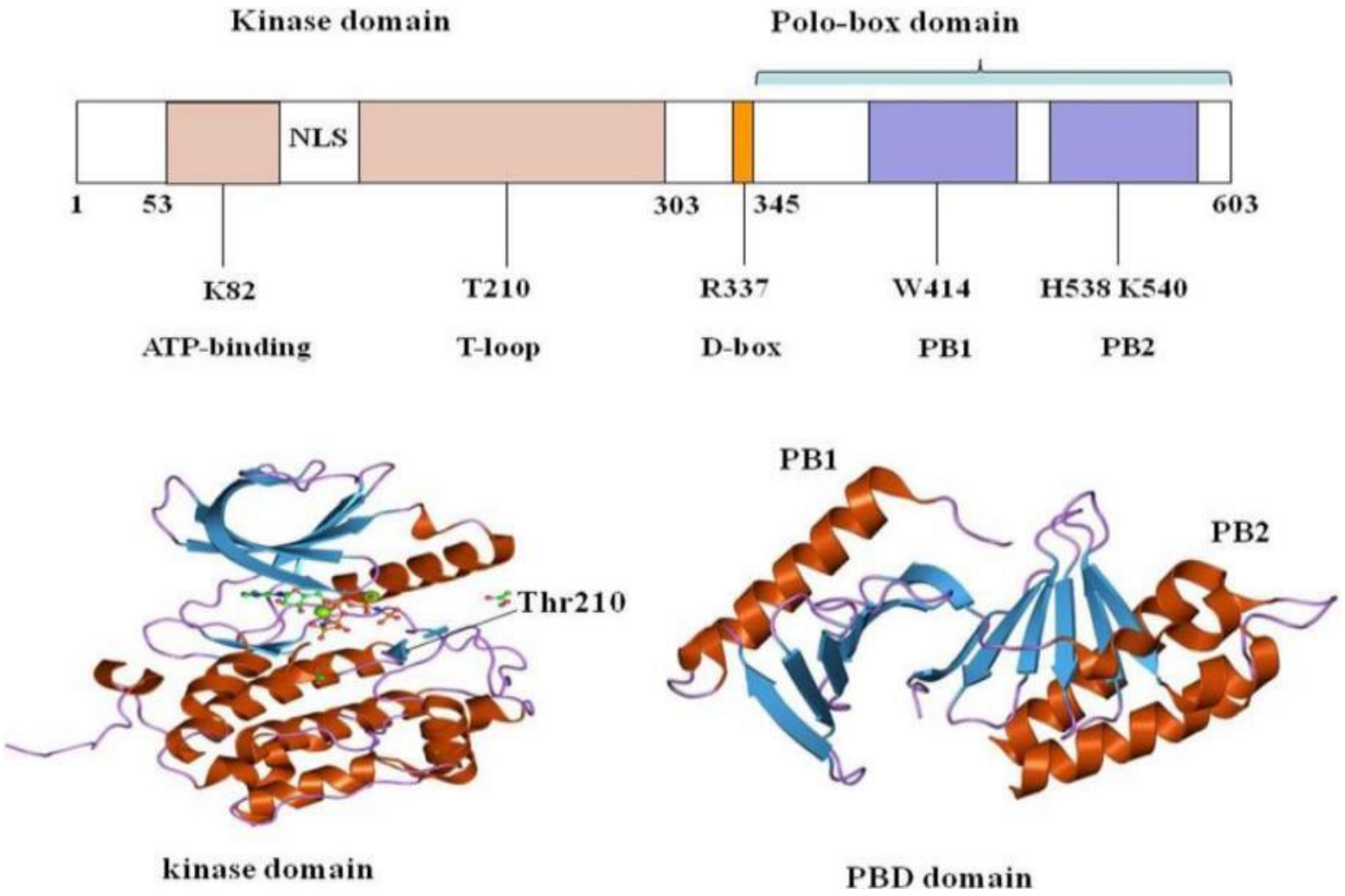
SCHRÖDINGER



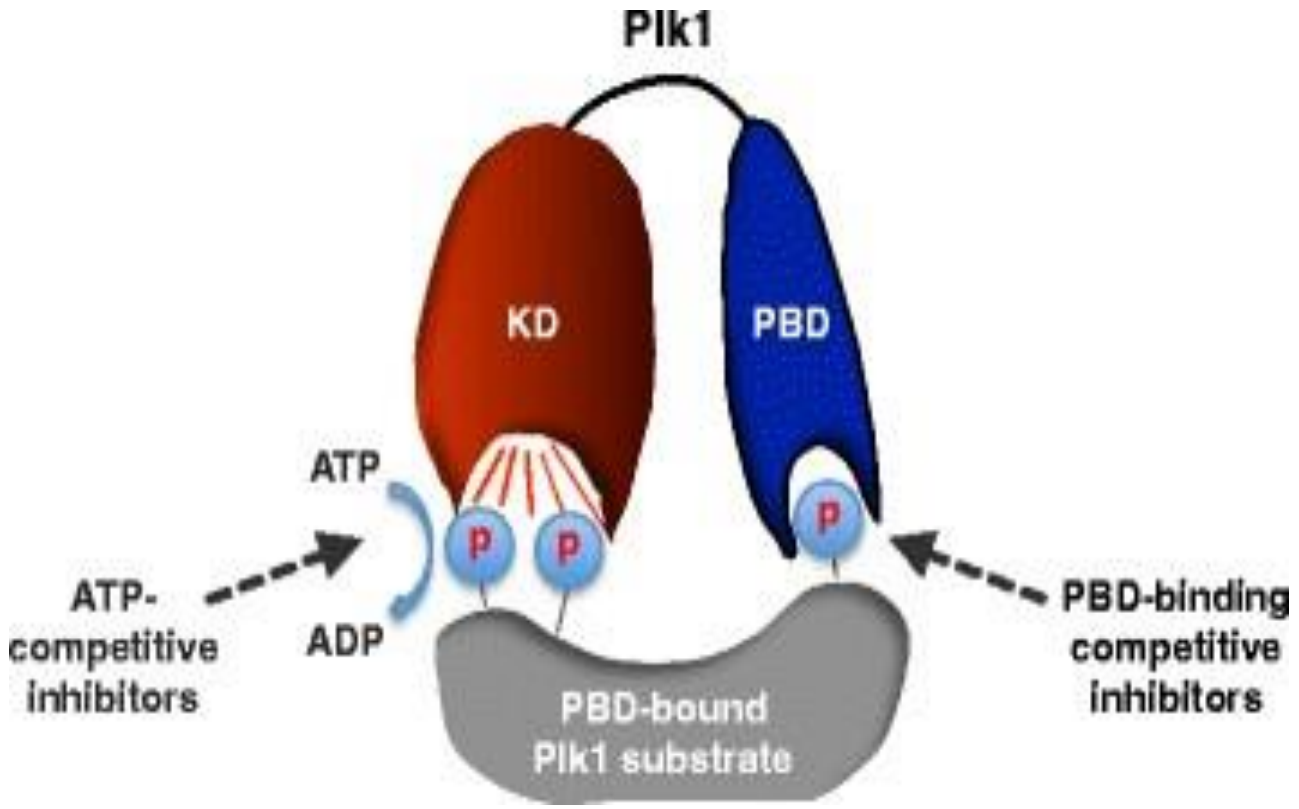
虚拟筛选的一般流程

[以PLK1（polo-like kinase 1）激酶口袋抑制剂的虚拟筛选为例]

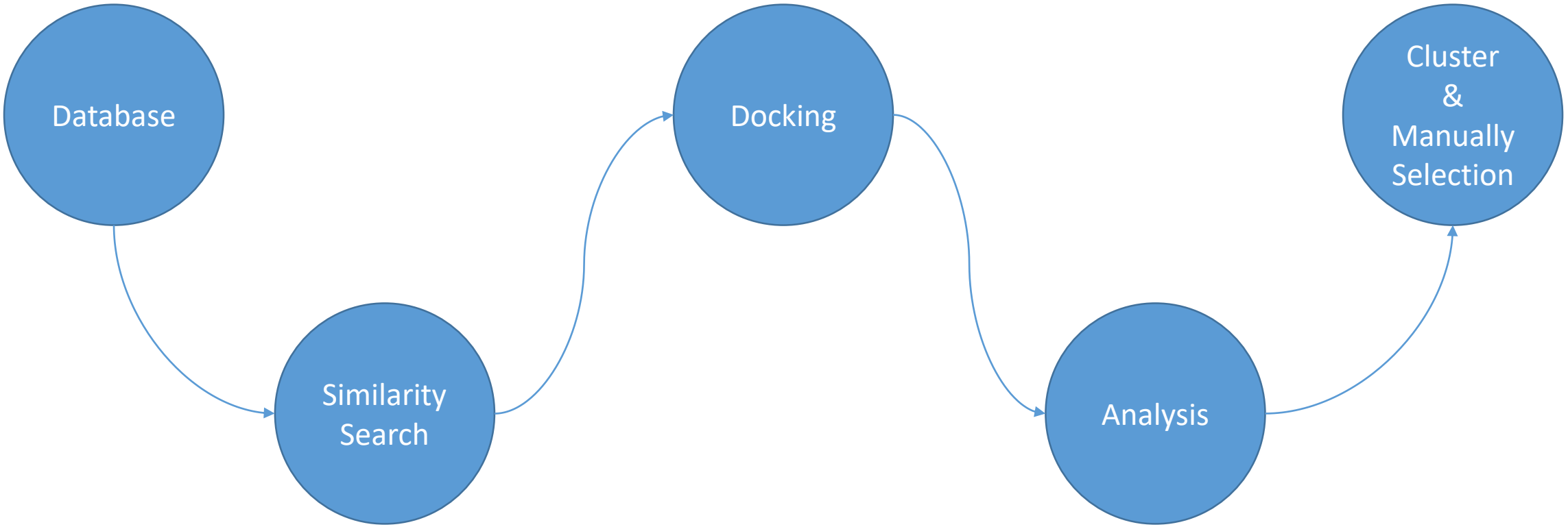
- Structure-Based
  - Docking
  - *De Novo* Dessign
- Ligand-Based
  - QSAR
  - Similarity-Based
  - Pharmacophore-Based



Liu, Z., Sun, Q., & Wang, X. (2017). PLK1, A Potential Target for Cancer Therapy. *Translational Oncology*, 10(1), 22-32.



Park, J. E., Kim, T. S., Meng, L., Bang, J. K., Kim, B. Y., & Lee, K. S. (2015). Putting a bit into the polo-box domain of polo-like kinase 1. *Journal of analytical science and technology*, 6(1), 27.


Database
Similarity Search
Docking
Analysis
Cluster & Manually Selection

Rre Docking

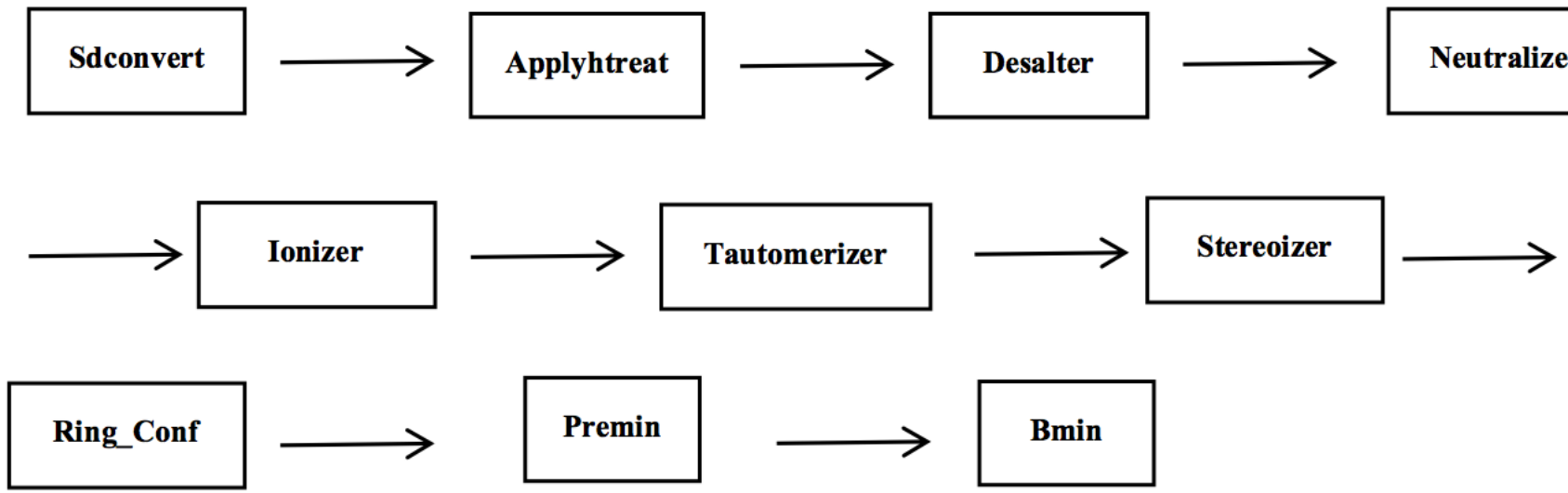
虚拟筛选的一般流程

小分子的前处理

OpenEye：FILTER　　对库中成药性与可计算性不佳的分子进行提前筛除

OpenEye：OMEGA　　快速地生成类药分子的构象空间

薛定谔LigPrep



虚拟筛选的一般流程

蛋白的选择方法

你的蛋白晶体适合对接吗？

1、蛋白来源？

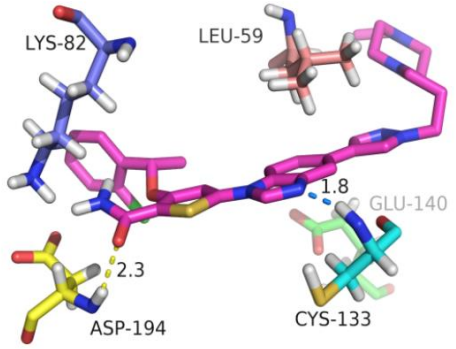
2、蛋白是否有和小分子结合的？

3、蛋白是否有突变？

4、体系蛋白分辨率？

虚拟筛选的一般流程

蛋白的选择方法



1、看PDB中蛋白晶体的分辨率，一般小于3A会比较好

2、对比蛋白的FASTA序列，有可能会有解错的（比如：PDB编号3KB7的蛋白晶体结构解析有误，查询它的fasta文件显示末端残基序列为XFSIA，而其他PLK1正确序列为RFXIA）

3、做第一次对接：

看最低自由结合能（Estimated Free Energy of Binding），此值在9kcal/mol上下，对应K_i为nM水平。

看结合口袋的疏水亲水区域，判断小分子构象（比如PLK1结合口袋有个疏水区域，所以亲水的哌嗪基朝向蛋白的外面更合适。）

与原配体氢键相互作用（个数，对应残基）

蛋白和原配体重对接的RMSD，此值在2-3A之下说明重对接比较成功，**原配体重对接**：RMSD < 1埃。

虚拟筛选的一般流程

验证库的选择

		真实值		总数
		p	*n*	
预测输出	*p′*	真阳性 (TP)	伪阳性 (FP)	P'
	n′	伪阴性 (FN)	真阴性 (TN)	N'
总数		P	N	

Active

Decoy（诱饵化合物）：与活性化合物的化学性质应当相似

1、相似的分子量
2、相似的LogP
3、相似的柔性键数
4、相似的氢键受体和供体
5、不同的分子拓扑结构

Similarity Search

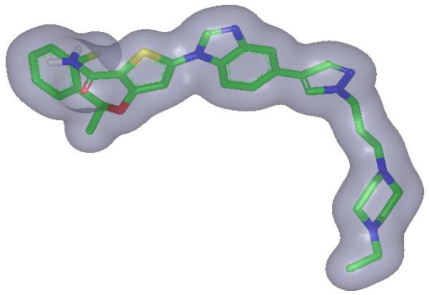
虚拟筛选的一般流程

三维相似性搜索

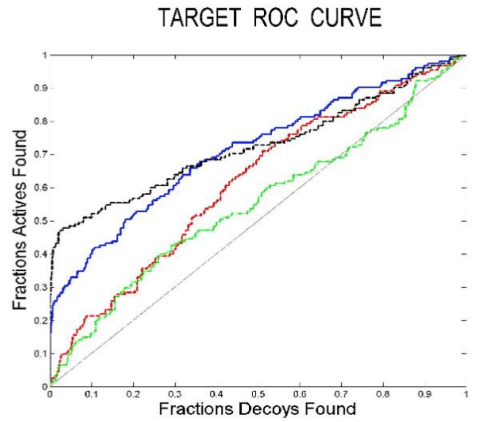
骨架跃迁（Scaffold Hopping）：是指从已知的活性分子结构出发，通过传统的类似物设计方法或者计算化学的方法在已知的数据库中寻找与苗头化合物（Hit）完全不同的拓扑骨架，又称为先导化合物跃迁。骨架跃迁主要分为四类：环系替换、开环闭环、拟肽伪肽和基于化合物**三维拓扑结构的骨架跃迁（三维相似性搜索）**方法。

前三类：二维　　一般成功率较高，但是因为模式的限定不容易实现明显的骨架突破

第四类：三维　　基于化合物拓扑结构的骨架跃迁，可以迅速地提供比较合理的新骨架化合物，实现原有骨架的局限实现专利突破。



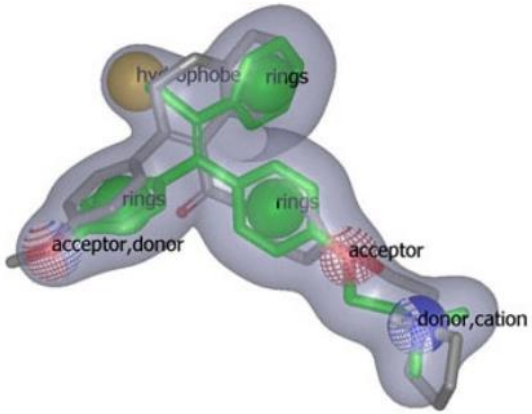
Query：问询式



ROC Curve

虚拟筛选的一般流程

三维相似性搜索



形状三维相似性搜索　ROCS
静电三维相似性搜索　EON

非Shape Only打分函数：

Tanimoto-Combo

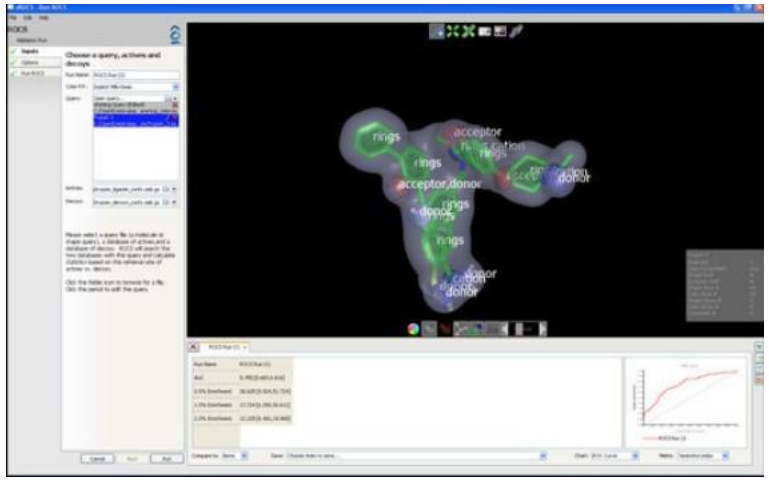
Color-Tanimoto

Shape-Tanimoto

Shape Only打分函数：

Shape-Tanimoto

保留Database前百分之2



Parameter：

AUC

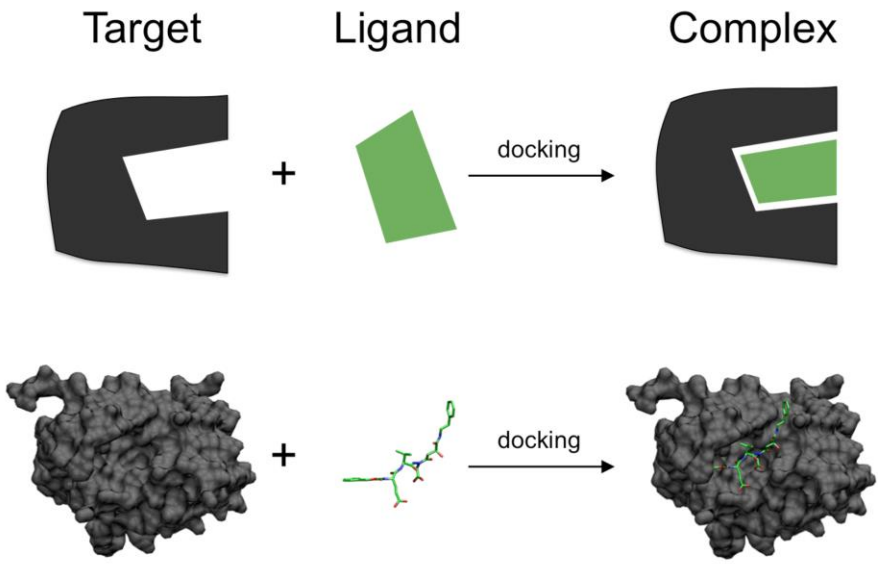
Enrichment Factor

ROC曲线

Docking

虚拟筛选的一般流程

分子对接



Lock-and key→Induced-fit

Applications:

Hit identification(virtual screening)

Lead optimization

Bioremediation

虚拟筛选的一般流程

分子对接

List of Protein-ligand Docking Software

Program	Search Algorithm	Scoring Functions	Country of Origin	Year Published
DOCK	Shape-Complementarity	Force Field	USA	1988
AutoDock	Simulated/Lamarckian Genetic Algorithms	Empirical	USA	1990
AutoDock Vina	Genetic Algorithms	Knowledge-Based	USA	2010
GOLD	Genetic Algorithms	Empirical	UK	1995
Glide	Shape-Complementarity	Empirical	USA	2004

虚拟筛选的一般流程

分子对接

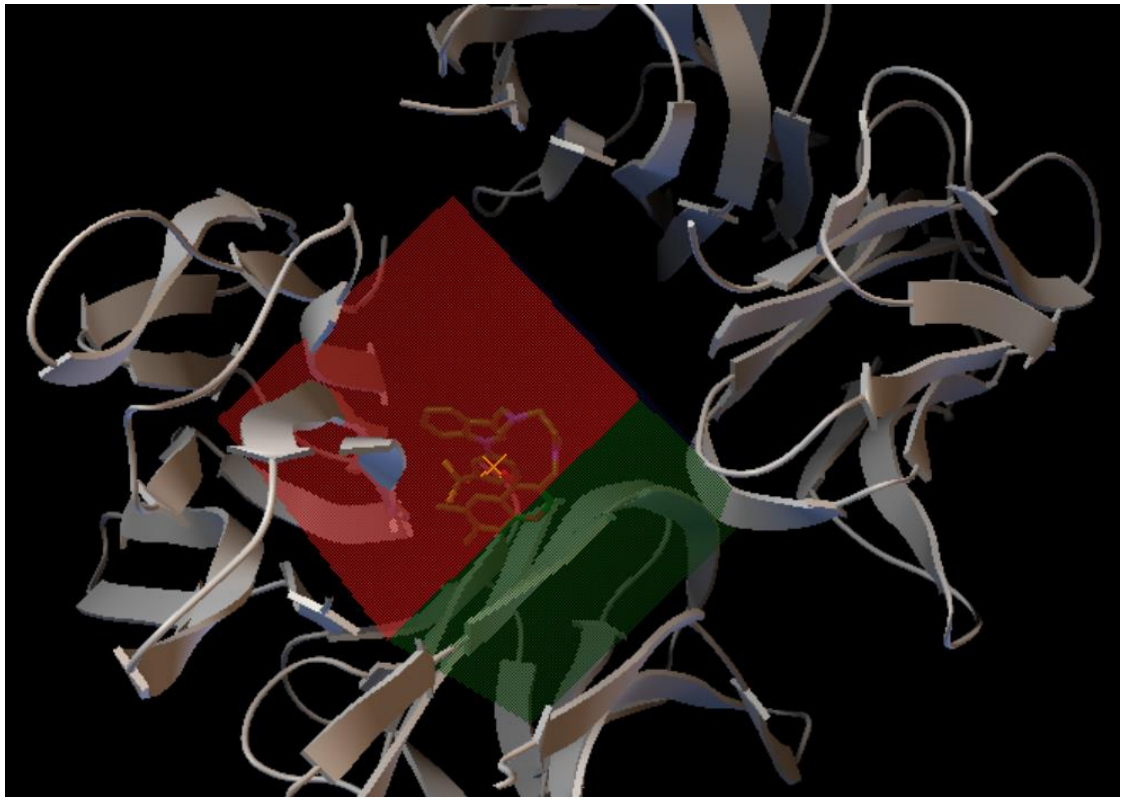
1、蛋白和小分子的前处理（这部分可以在AutoDock Tools 和薛定谔做）

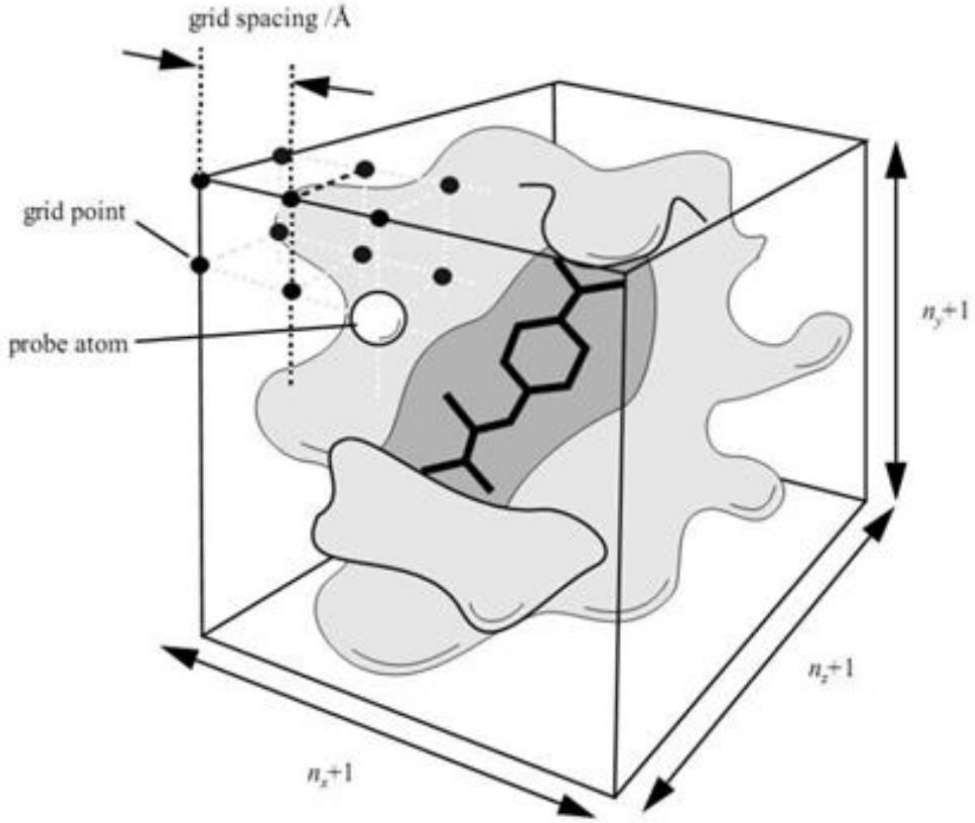
2、Autodock和Vina对接

3、会影响对接成功与否的地方？

虚拟筛选的一般流程

分子对接　**AutoDock**





虚拟筛选的一般流程

分子对接　**AutoDock**

```
Number of distinct conformational clusters found = 7,  out of 100 runs,
Using an rmsd-tolerance of 2.0 A


CLUSTERING HISTOGRAM
___________________

_____________________________________________________________________________
     |             |     |             |     |                                 
Clus | Lowest      | Run | Mean        | Num | Histogram                       
-ter | Binding     |     | Binding     | in  |                                 
Rank | Energy      |     | Energy      | Clus|    5    10   15   20   25   30   35
_____|_____________|_____|_____________|_____|____:____|____:____|____:____|____:___
   1 |      -9.82 |  12 |      -8.73 |  85 |#################################################################################
   2 |      -7.94 |  41 |      -7.89 |   2 |##
   3 |      -7.42 |   5 |      -7.20 |   6 |######
   4 |      -7.37 |  18 |      -7.37 |   1 |#
   5 |      -6.93 |  72 |      -6.93 |   1 |#
   6 |      -6.92 |  90 |      -6.77 |   4 |####
   7 |      -6.67 |  69 |      -6.67 |   1 |#
_____|_____________|_____|_____________|_____|_____________________________________


Number of multi-member conformational clusters found = 4, out of 100 runs.
```

虚拟筛选的一般流程

分子对接 **AutoDock Vina**

The usage summary can be obtained with "vina --help":

```
Input:
  --receptor arg        rigid part of the receptor (PDBQT)
  --flex arg            flexible side chains, if any (PDBQT)
  --ligand arg          ligand (PDBQT)

Search space (required):
  --center_x arg        X coordinate of the center
  --center_y arg        Y coordinate of the center
  --center_z arg        Z coordinate of the center
  --size_x arg          size in the X dimension (Angstroms)
  --size_y arg          size in the Y dimension (Angstroms)
  --size_z arg          size in the Z dimension (Angstroms)
```

… …

虚拟筛选的一般流程

分子对接 **AutoDock Vina**

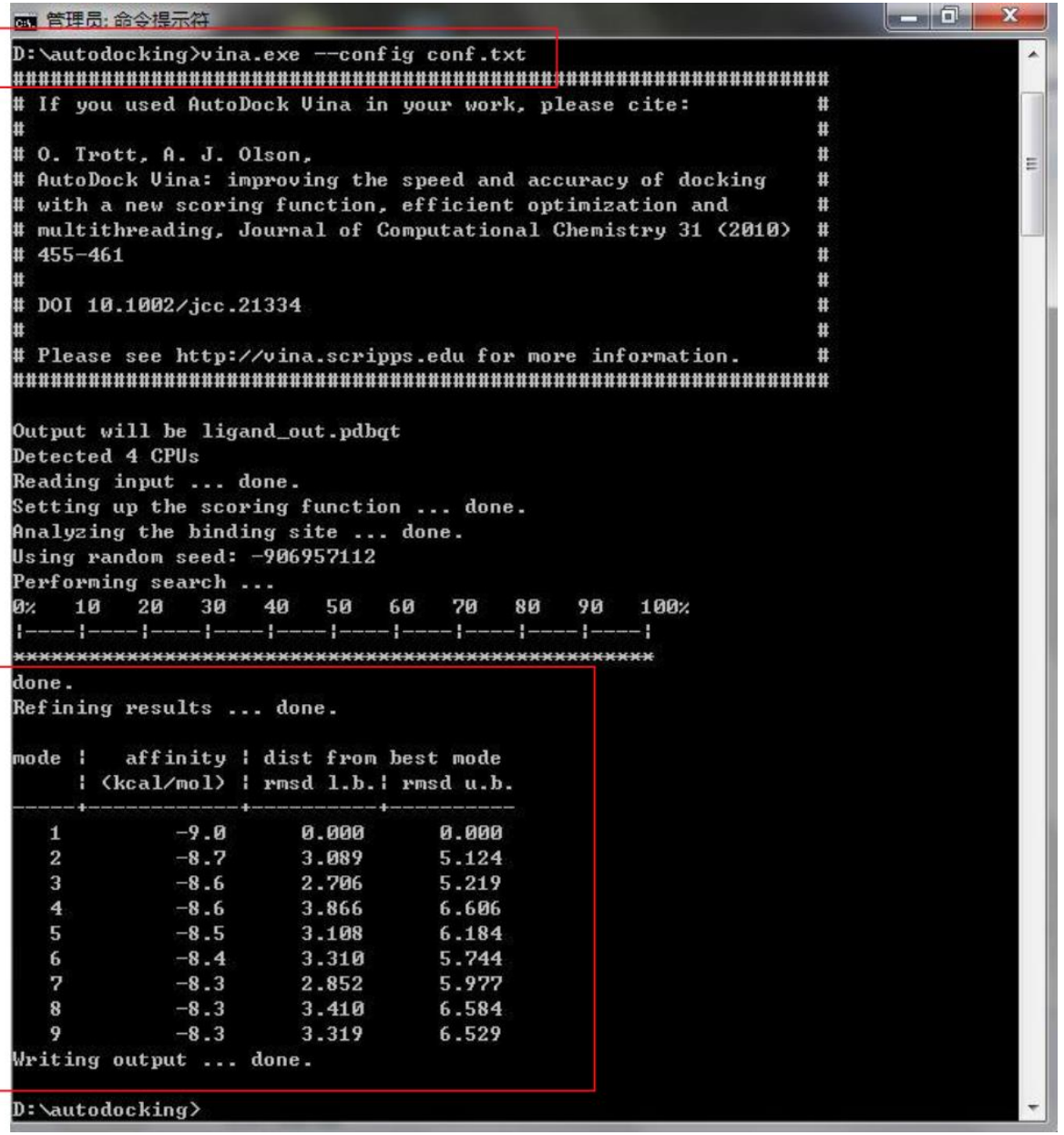
Configuration file

For convenience, some command line options can be placed into a configuration file.
For example:

```
receptor = hsg1/rigid.pdbqt
ligand = ligand.pdbqt

center_x =  2
center_y =  6
center_z = -7

size_x = 25
size_y = 25
size_z = 25

energy_range = 4
```

虚拟筛选的一般流程

分子对接　**AutoDock Vina**

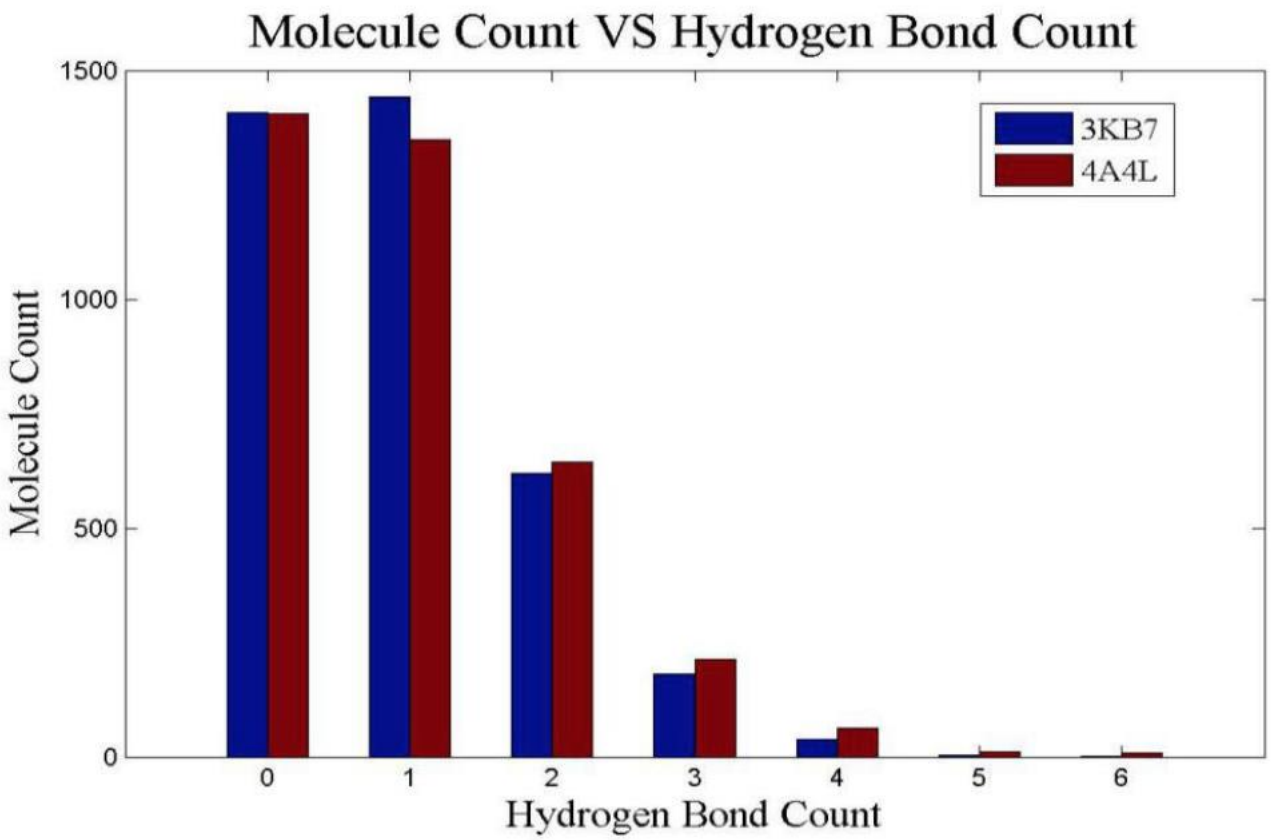


Analysis

虚拟筛选的一般流程

Analysis

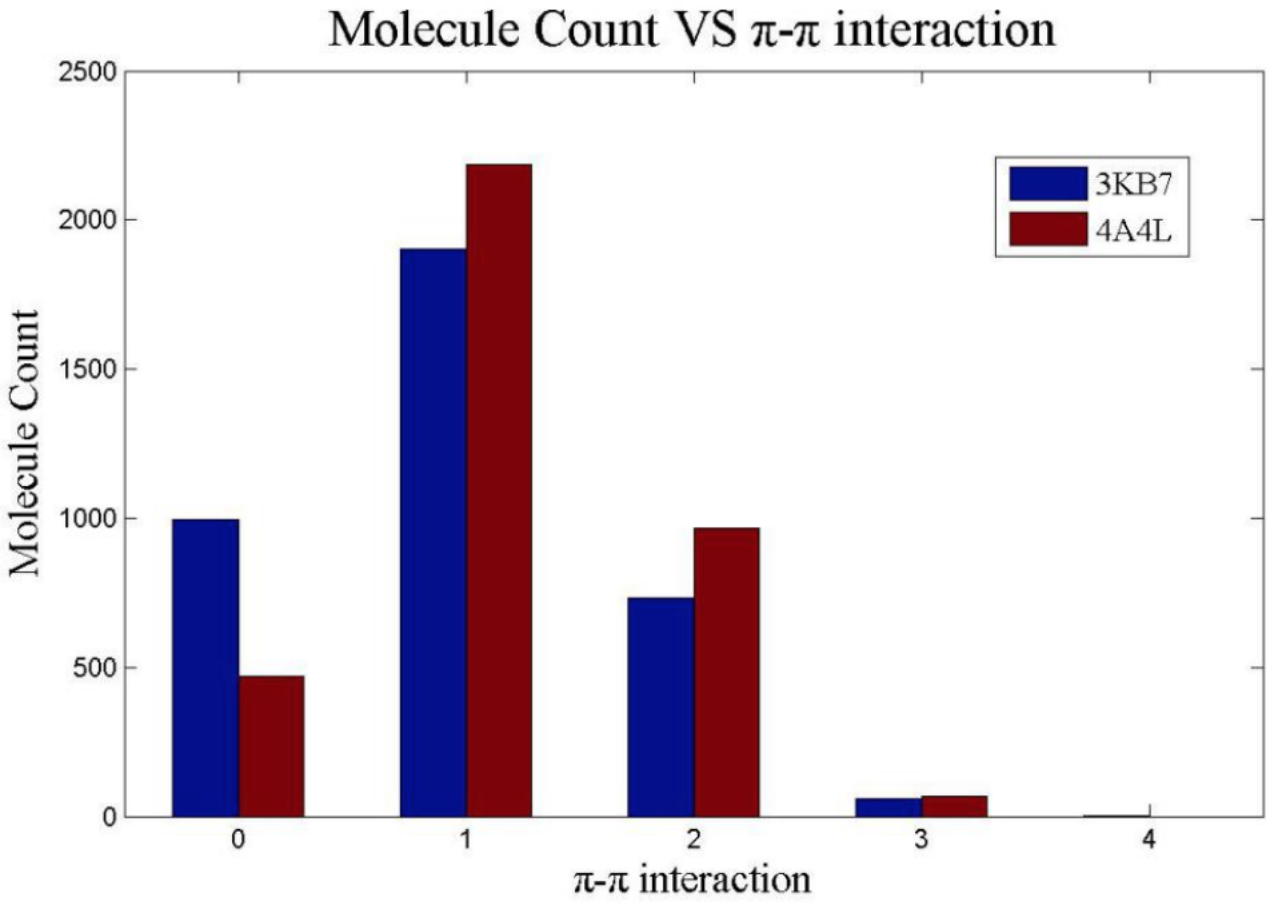
氢键相互作用



虚拟筛选的一般流程

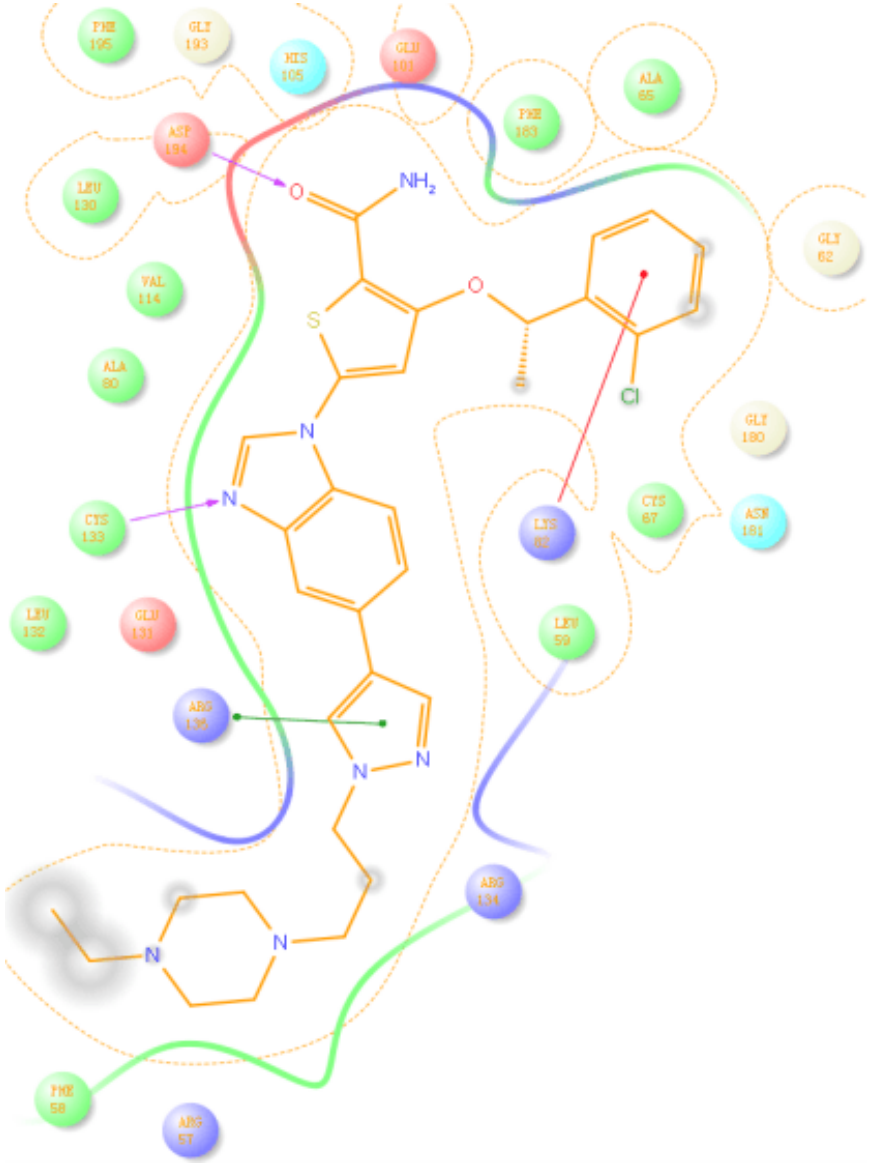
Analysis

π-π相互作用



虚拟筛选的一般流程

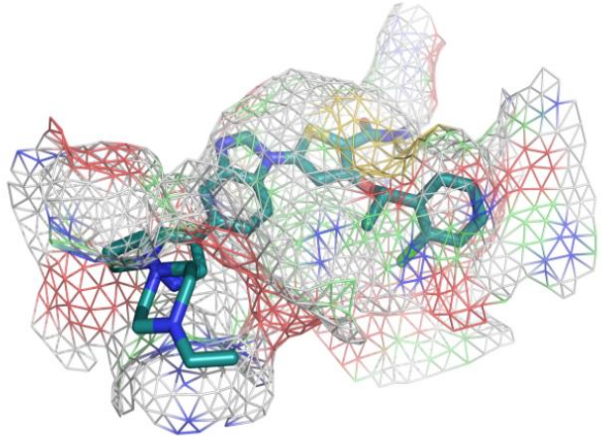
Analysis



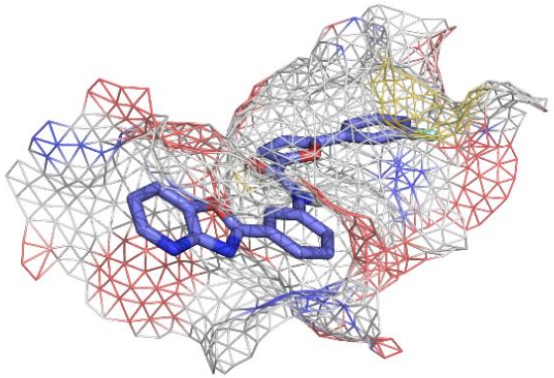
虚拟筛选的一般流程

Analysis

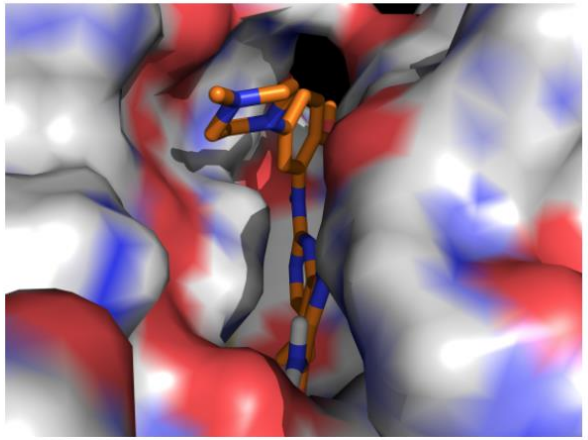
对重要分子的最佳结合构象分析



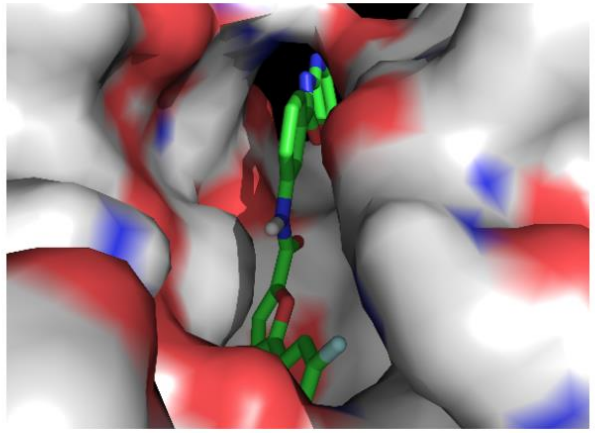
A



B



C

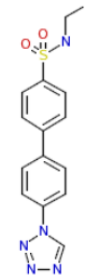
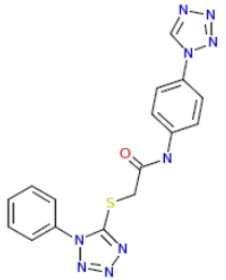
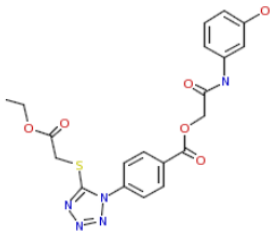
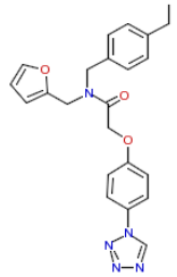
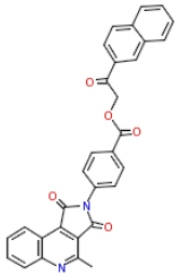
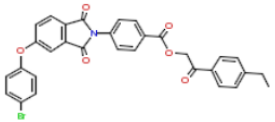
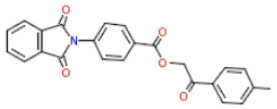
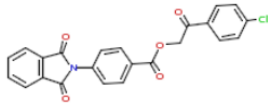
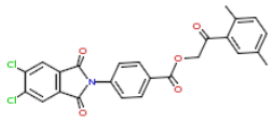
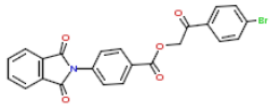
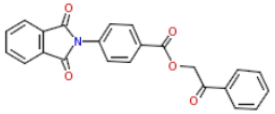
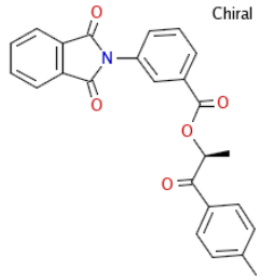
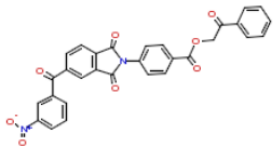
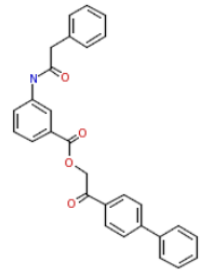


D

Cluster

虚拟筛选的一般流程

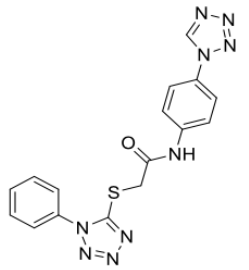
Cluster

Cluster: 1				
				
F:\senior\fenzi\specs-top3692_000596_01log.txt.pdbqt	F:\senior\fenzi\specs-top3692_001146_1001log.txt.pdbqt	F:\senior\fenzi\specs-top3692_003436_3001log.txt.pdbqt	F:\senior\fenzi\specs-top3692_001932_1001log.txt.pdbqt	
Cluster: 120				
				
F:\senior\fenzi\specs-top3692_000574_01log.txt.pdbqt	F:\senior\fenzi\specs-top3692_000366_01log.txt.pdbqt	F:\senior\fenzi\specs-top3692_001731_1001log.txt.pdbqt	F:\senior\fenzi\specs-top3692_001878_1001log.txt.pdbqt	F:\senior\fenzi\specs-top3692_002888_2001log.txt.pdbqt
		 Chiral		
F:\senior\fenzi\specs-top3692_002092_2001log.txt.pdbqt	F:\senior\fenzi\specs-top3692_001288_1001log.txt.pdbqt	F:\senior\fenzi\specs-top3692_002635_2001log.txt.pdbqt	F:\senior\fenzi\specs-top3692_000191_01log.txt.pdbqt	F:\senior\fenzi\specs-top3692_000180_01log.txt.pdbqt

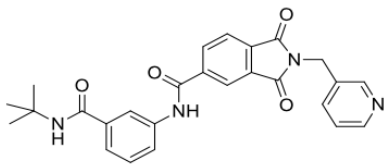
Manually Selection

虚拟筛选的一般流程

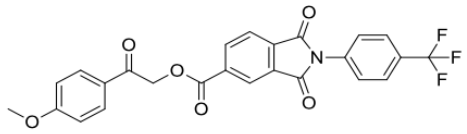
Manually Selection



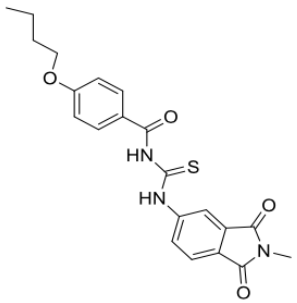
AN-648/37428008



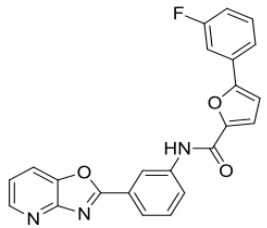
AN-979/41713531



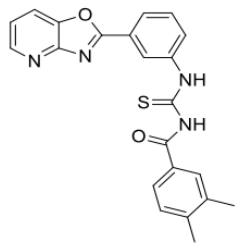
AK-918/41779618



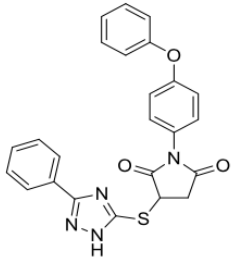
AN-465/40934528



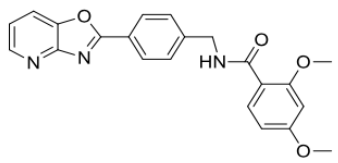
AP-970/43374400



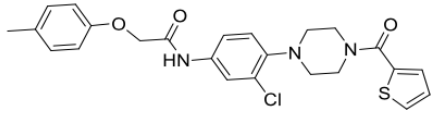
AG-690/15429814



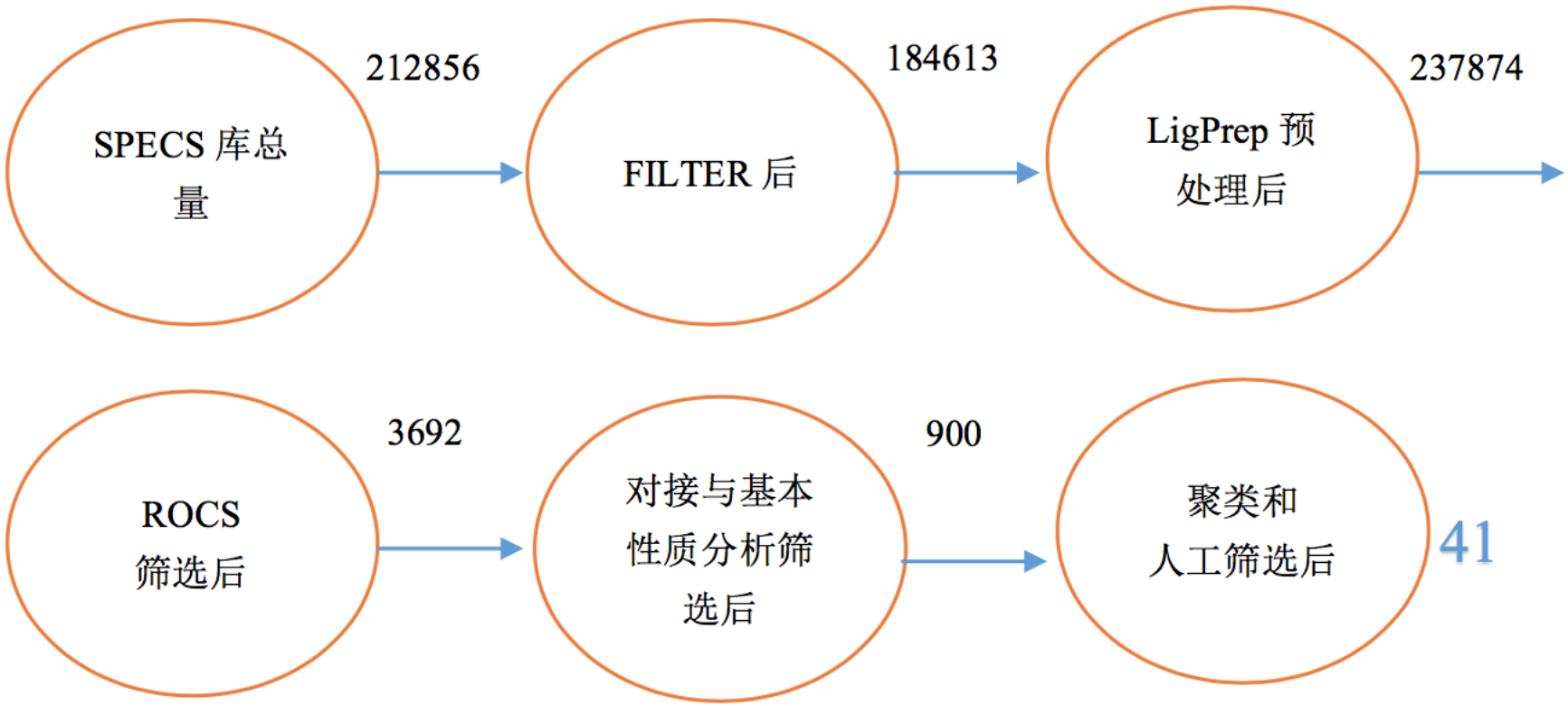
AJ-292/41686368



AP-970/43482873



AP-970/42777298


SPECS 库总量
212856
FILTER 后
184613
LigPrep 预处理后
237874
ROCS 筛选后
3692
对接与基本性质分析筛选后
900
聚类和人工筛选后
41

PyMOL实例应用——小分子配体与蛋白互相作用

G04+G07

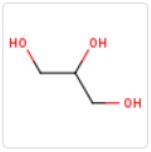
2018/01/15

小分子配体与蛋白互相作用

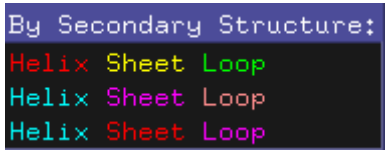
通过PDB查询蛋白后，在页面下方会有关于小分子的信息

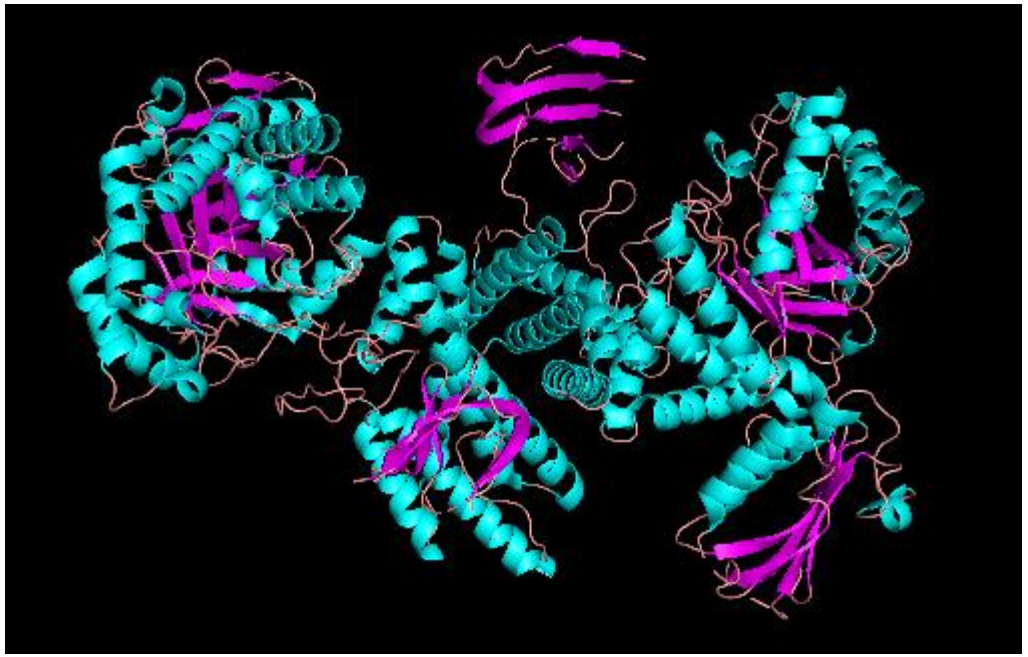
Small Molecules

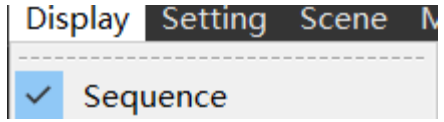
Ligands 4 Unique

ID	Chains	Name / Formula / InChI Key	2D Diagram & Interactions	3D Interactions
5FN Query on 5FN Download SDF File Download CCD File	A, B	**(3AS,5S,6S,7R,7AR)-5-FLUORO-5-(HYDROXYMETHYL)- 2-METHYL-5,6,7,7A-TETRAHYDRO-3AH-PYRANO[3,2- D][1,3]OXAZOLE-6,7-DIOL** $C_8 H_{12} F N O_5$ GCSIDVFZDNBNLE-QQGCVABSSA-N		Ligand Explorer NGL Binding Pocket (JSmol) Electron Density (JSmol)
GOL Query on GOL Download SDF File Download CCD File	A, B	GLYCEROL **2**. *(Synonym)* $C_3 H_8 O_3$ PEDCQBHIVMGVHV-UHFFFAOYSA-N		Ligand Explorer NGL Binding Pocket (JSmol) Electron Density (JSmol)
CA Query on CA Download SDF File Download CCD File	A, B	**CALCIUM ION** Ca BHPQYMZQTOCNFJ-UHFFFAOYSA-N	Ca^{2+}	Ligand Explorer NGL Binding Pocket (JSmol)
NA Query on NA Download SDF File Download CCD File	B	**SODIUM ION** Na FKNQFGJONOIPTF-UHFFFAOYSA-N	Na^{+}	Ligand Explorer NGL Binding Pocket (JSmol)

点击all中的H，选择everything，隐藏所有
点击2wzi中的S，选择cartoon，以cartoon形式显示蛋白质
点击2wzi中的C，选择by ss，以二级结构分配颜色，从下图选择一种

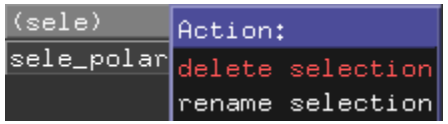






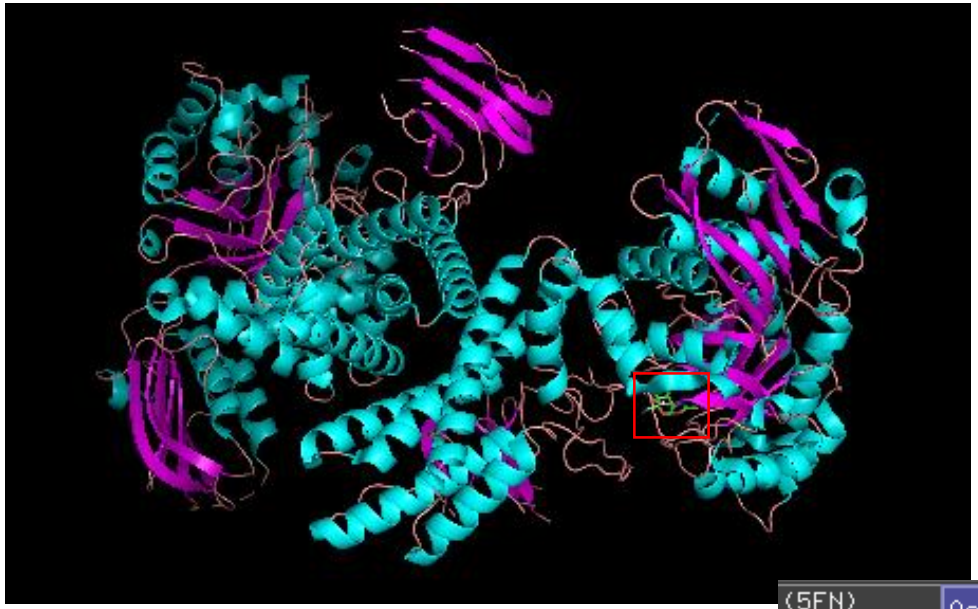
窗口里面会出现氨基酸序列，找到5FN这个小分子配体

出现sele后对其进行改名称，点击A后出现

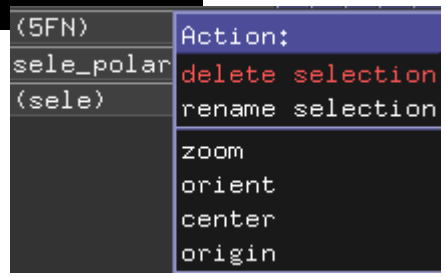


点击rename selection

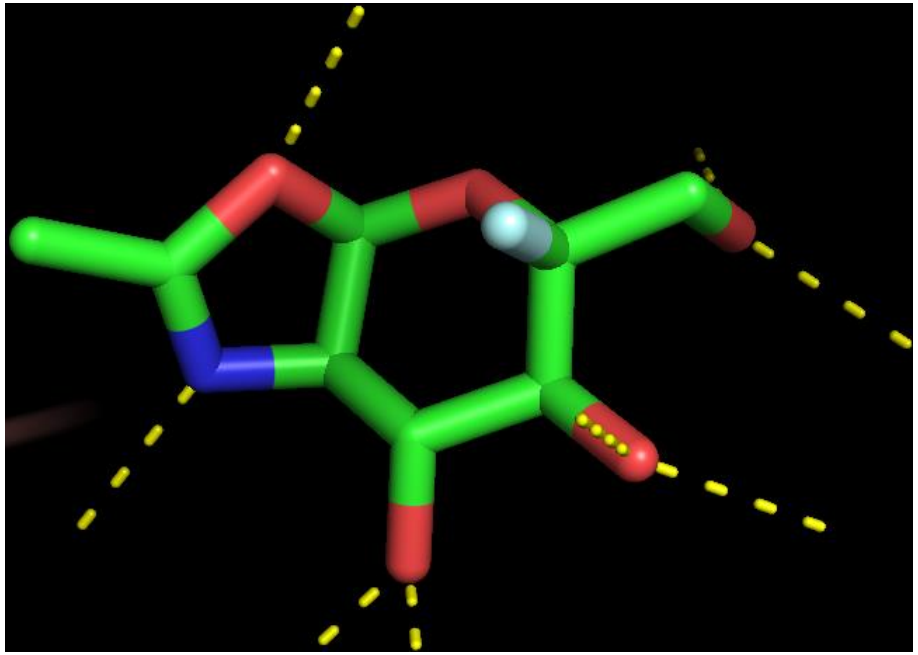
然后点击S选择sticks，点击C，选择by element，选择



但是此形式并非最佳视角范围，此时点击5FN A中的zoom或者orient

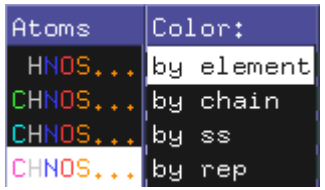


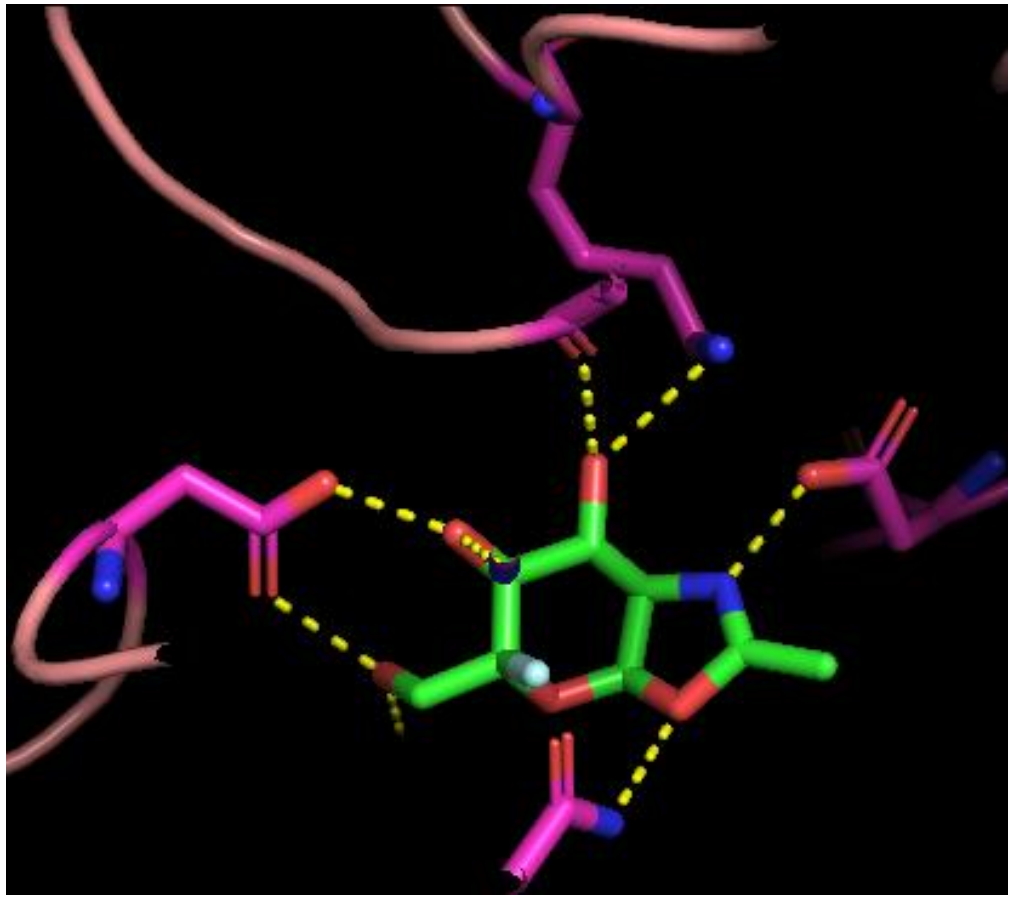
点击5FN行的A，选择find，然后选择polar contacts，再选择to other atoms in object



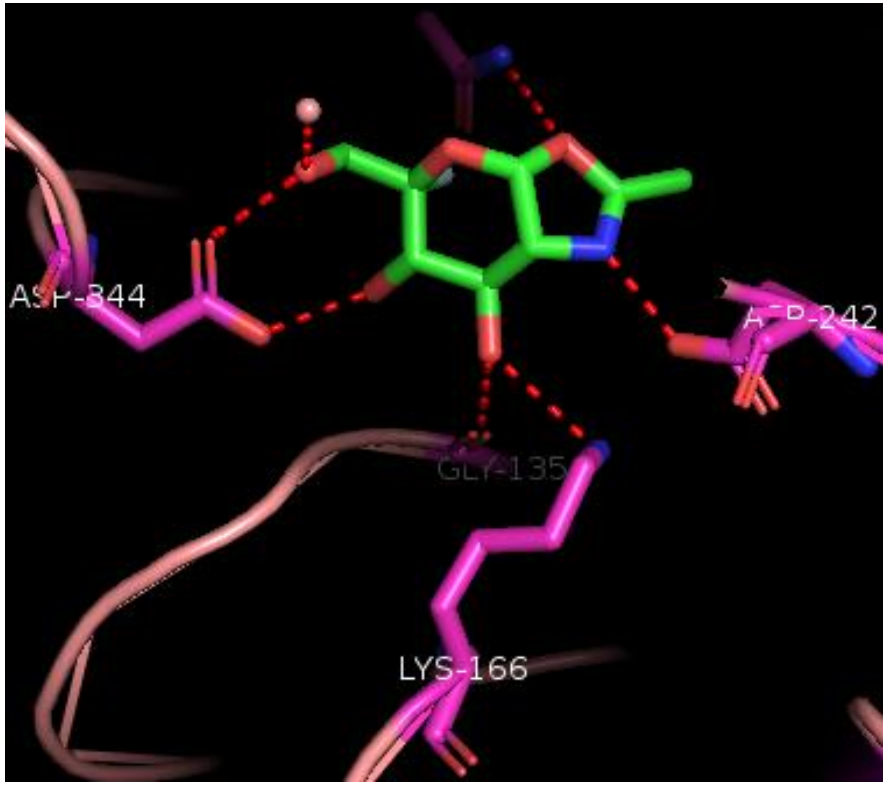
分子显示窗口中出现几个黄色的虚线，5FN行下面出现了新的一行即为5FN与蛋白质相互作用的氢键

点击2wzi行的S，选择lines，使用鼠标转动蛋白质寻找与5FN以虚线相连的残基，分别点击选择这些残基，把选择的残基（sele）改名为5FN_lig
点击2wzi行的H，选择lines
点击5FN_lig行的S，选择sticks，C选择





点击2wzi行的S，选择nb_spheres，选择相连的氧原子改名为O
点击2wzi行的H，选择nb_spheres
点击O行的S，选择nb_spheres
点击5FN_polar_conts行的C，把氢键颜色改为红色
点击5FN_lig行的L，选择residues显示残基名

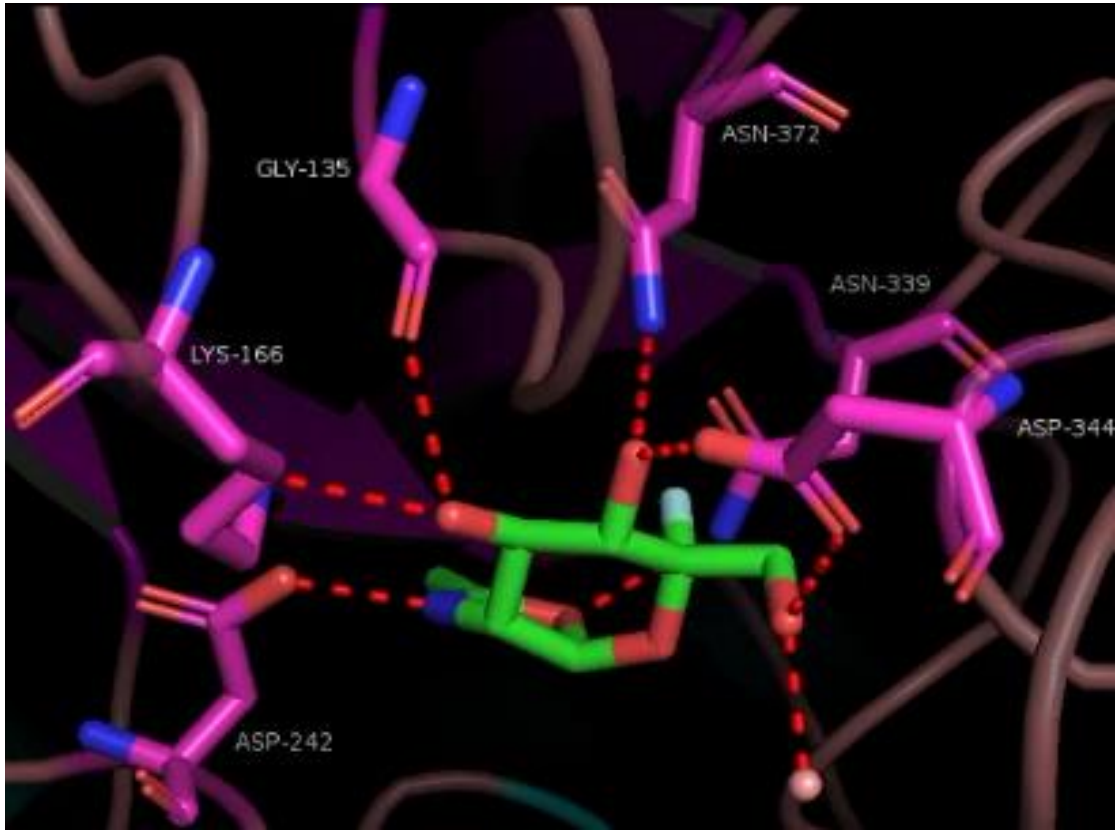


点击下方窗口，将Button Viewing改为Button Editing，摁住Ctrl键鼠标左键拖动label到合适位置，然后依次按以下步骤操作：

Setting >transparency> cartoon >调节透明度至50%

Setting > cartoon >fancy helix

Setting >cartoon>highlight color（加深对比度）



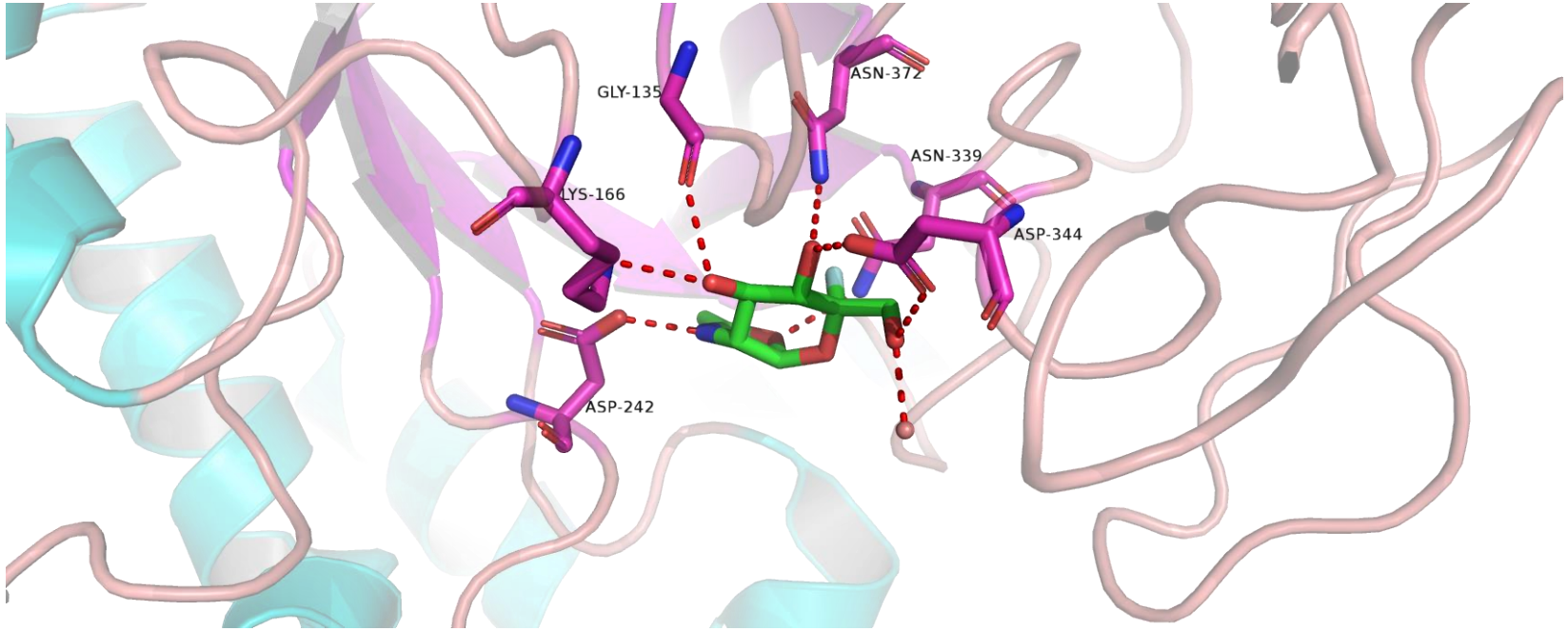
为了得到更高质量和更漂亮得图片我们可以对菜单栏按照下面的路径继续处理：

Display>background color>white

Setting >label>size>18 point

Draw/Ray>Ray (slow)

这样，我们就得到了一张展示配体与蛋白相关残基之间氢键作用的图片了



THANKS !