**生信分析报告**

**项目标题： MYLK分子对接 ;**

**单 号： BSCL250222 ;**

**分析人员： 黄礼闯 ;**

**分析类型： 生信分析 ;**

**委 托 人： 汪光亮 ;**

**受 托 人： 杭州铂赛生物科技有限公司 .**

# 1 分析流程

# 2 材料和方法

## 2.1 数据分析平台

在 Linux pop-os x86\_64 (6.9.3-76060903-generic) 上，使用 R version 4.4.2 (2024-10-31) (<https://www.r-project.org/>) 对数据统计分析与整合分析。

## 2.2 BindingDB 药物靶点数据获取 (Dataset: MYLK)

获取 BindingDB 数据库化合物与靶点蛋白结合数据 (<https://www.bindingdb.org/rwd/bind/chemsearch/marvin/Download.jsp>) (2016, **IF:16.6**, Q1, Nucleic acids research)1。以 R 包 UniProt.ws (2.46.1) 将基因名 (symbol) 获取 UniprotKB 数据库的 Entry ID。根据 Entry ID 检索 BindingDB 数据中，与该蛋白结合的化合物信息。

## 2.3 AutoDock vina 分子对接 (Dataset: MYLK)

以 R 包 bio3d (2.4.5) 获取 PDB ID 对应的注释 (蛋白结构分辨率, resolution) 。 以 PubChem API (<https://pubchem.ncbi.nlm.nih.gov/docs/pug-rest>) 获取化合物 SDF 结构文件。使用 openbabel 的工具 (obgen) 计算 SDF 文件的 3D 构象 (转化为 3D SDF文件)。以 Python meeko 包 (mk\_prepare\_ligand.py) 转化 SDF 文件获取配体 PDBQT 用于分子对接。 以 R 包 UniProt.ws (2.46.1) 获取基因 (symbol) 的 UniProtKB-Swiss-Prot ID (Entry ID)，随后，以 Entry ID 从数据库 AlphaFold (<https://alphafold.ebi.ac.uk/>) 获取蛋白结构 (预测的结构)。以 pymol (2010, **IF:3**, Q2, Journal of Computer-Aided Molecular Design)2 删除受体 PDB 文件中的非标准残基 (例如 NAG, BMA, FUL)。以 pymol 仅保留蛋白结构 (polymer.protein) (去除了原 PDB 中的配体等其他结构)。以 ADFR (2019, **IF:3.4**, Q2, Journal of computational chemistry)3 工具组的准备受体蛋白的 PDBQT 文件 (以 prepare\_receptor 添加氢原子，并转化为 PDBQT 文件) 。请参考 <https://autodock-vina.readthedocs.io/en/latest/docking_basic.html>。 以 AutoDock-Vina 提供的工具 (prepare\_gpf.py) (<https://github.com/ccsb-scripps/AutoDock-Vina>) 创建 GPF (grid parameter file)。以 ADFR (2019, **IF:3.4**, Q2, Journal of computational chemistry)3 工具 autogrid4 计算亲和图谱 (Affinity Maps)。运行 AutoDock-Vina (2021, **IF:5.6**, Q1, Journal of Chemical Information and Modeling)4 (parameters: scoring = ad4; exhaustiveness = 32)。

## 2.4 Classyfire 获取化合物分类信息 (Dataset: MYLK)

使用 PubChem API (<https://pubchem.ncbi.nlm.nih.gov/docs/pug-rest>) 获取化合物 InChIKey。 通过 InChIKey，以 Classyfire (2016, **IF:7.1**, Q1, Journal of Cheminformatics)5 获取化合物系统分类注释。

# 3 分析结果

## 3.1 BindingDB 药物靶点数据获取 (MYLK)

在 BindingDB 数据中，检索与 MYLK 相关的化合物结合数据。

**Tab.** **1** MYLK bindingdb annotation data

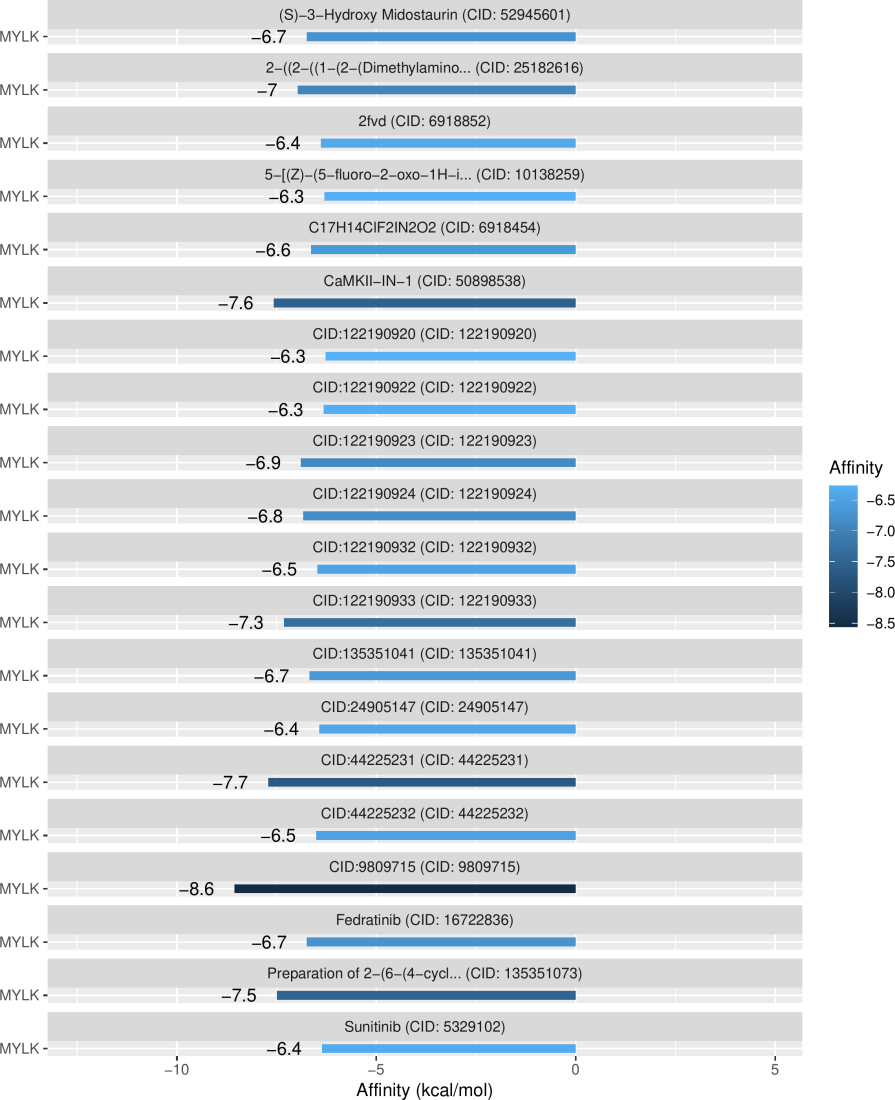
| BindingDB | PubChem | UniProt | PDB | Ligand SMILES |
| --- | --- | --- | --- | --- |
| 66810 | 9887000 | Q15746 | 2CQV,2YR3,6C6M | CN1CCN(CC1)c1ccc2... |
| 66810 | 5331000 | Q15746 | 2CQV,2YR3,6C6M | Nc1nc(Nc2ccc(cc2)... |
| 66810 | 154000 | Q15746 | 2CQV,2YR3,6C6M | CN(C)C[C@@H]1CCn2... |
| 66810 | 9885000 | Q15746 | 2CQV,2YR3,6C6M | Oc1cccc(c1)-c1nc(... |
| 66810 | 24200000 | Q15746 | 2CQV,2YR3,6C6M | CO[C@H]1[C@H](C[C... |
| ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.1\_BindingDB\_药物靶点数据获取\_(MYLK)/MYLK-bindingdb-annotation-data.xlsx)**

Tab. **[1](#MYLK-bindingdb-annotation-data)** 为 BindingDB 记录的化合物与蛋白结合信息附表。

## 3.2 AutoDock vina 分子对接 (MYLK)

将 BindingDB 预筛选的化合物与靶点蛋白，用于分子对接。从 PubChem 获取化合物 SDF 结构文件(2D)。以 openbabel 计算化合物的 3D 构象。以 meeko 从 SDF 转化得到配体的 PDBQT 文件。从数据库 AlphaFold 获取 MYLK 预测的蛋白结构 (根据 UniProtKB-Swiss-Prot ID，详见方法章节)。以 pymol 选择 Q15746 蛋白结构中的 chain A 的氨基酸残基 1-922，去除其余部分。以 pymol 去除非标准残基 (NAG, BMA, FUL)。随后，以 pymol 仅保留蛋白结构 (polymer.protein)。以 ADFR 工具给受体添加氢原子，转化为 PDBQT 文件。以 ADFR 创建 Affinity Maps (详见方法章节) 。以 Autodock-Vina 进行自动分子对接。使用 pymol 将分子对接结果可视化。



**Fig.** **1** MYLK Overall combining Affinity

**(File path: Figure+Table/3.2\_AutoDock\_vina\_分子对接\_(MYLK)/MYLK-Overall-combining-Affinity.pdf)**

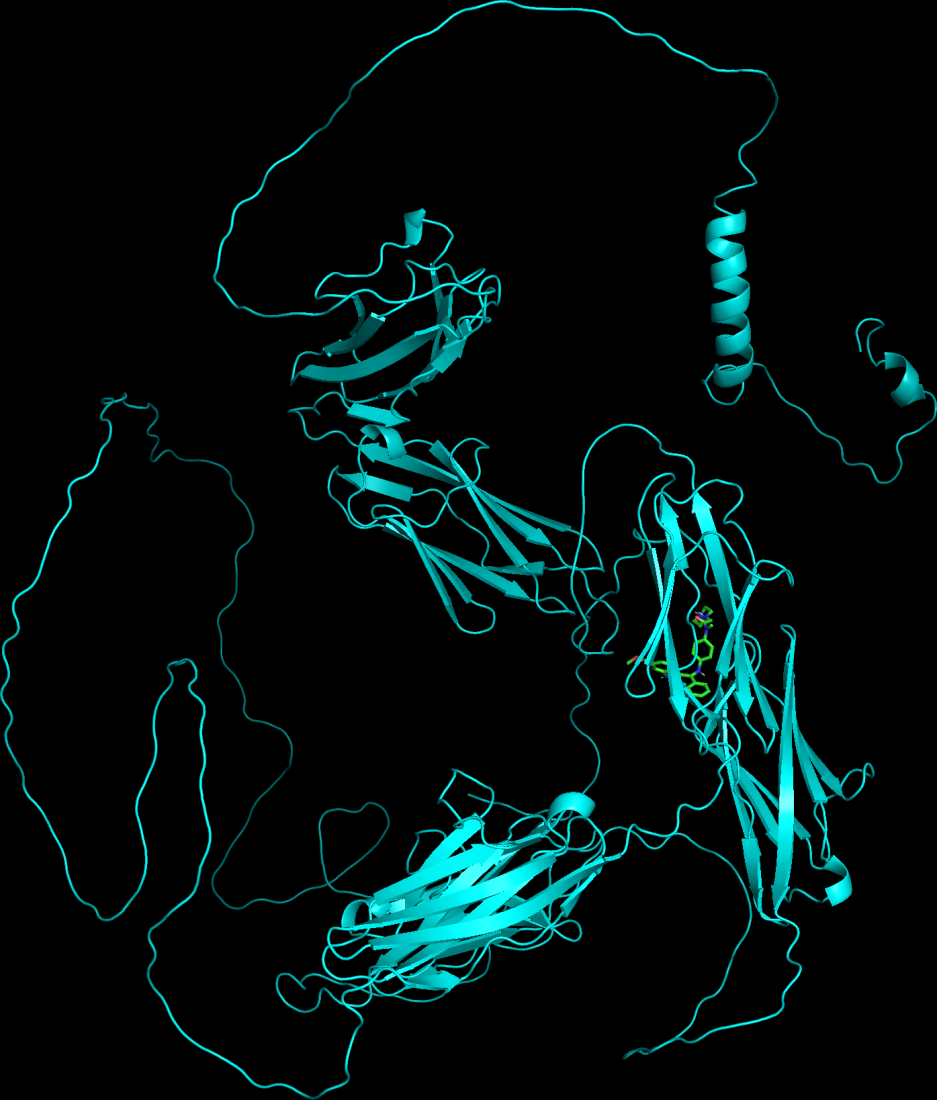
Fig. **[1](#MYLK-Overall-combining-Affinity)** 为分子对接亲和度得分可视化，能量越低，代表亲和度越高。

**Tab.** **2** MYLK All combining Affinity data

| PubChem id | PDB ID | Affinity | Dir | File |
| --- | --- | --- | --- | --- |
| 9810000 | Q15746 | -8.557 | Vina space/980971... | Vina space/980971... |
| 44230000 | Q15746 | -7.708 | Vina space/442252... | Vina space/442252... |
| 50900000 | Q15746 | -7.576 | Vina space/508985... | Vina space/508985... |
| 135400000 | Q15746 | -7.496 | Vina space/135351... | Vina space/135351... |
| 122200000 | Q15746 | -7.318 | Vina space/122190... | Vina space/122190... |
| ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.2\_AutoDock\_vina\_分子对接\_(MYLK)/MYLK-All-combining-Affinity-data.xlsx)**

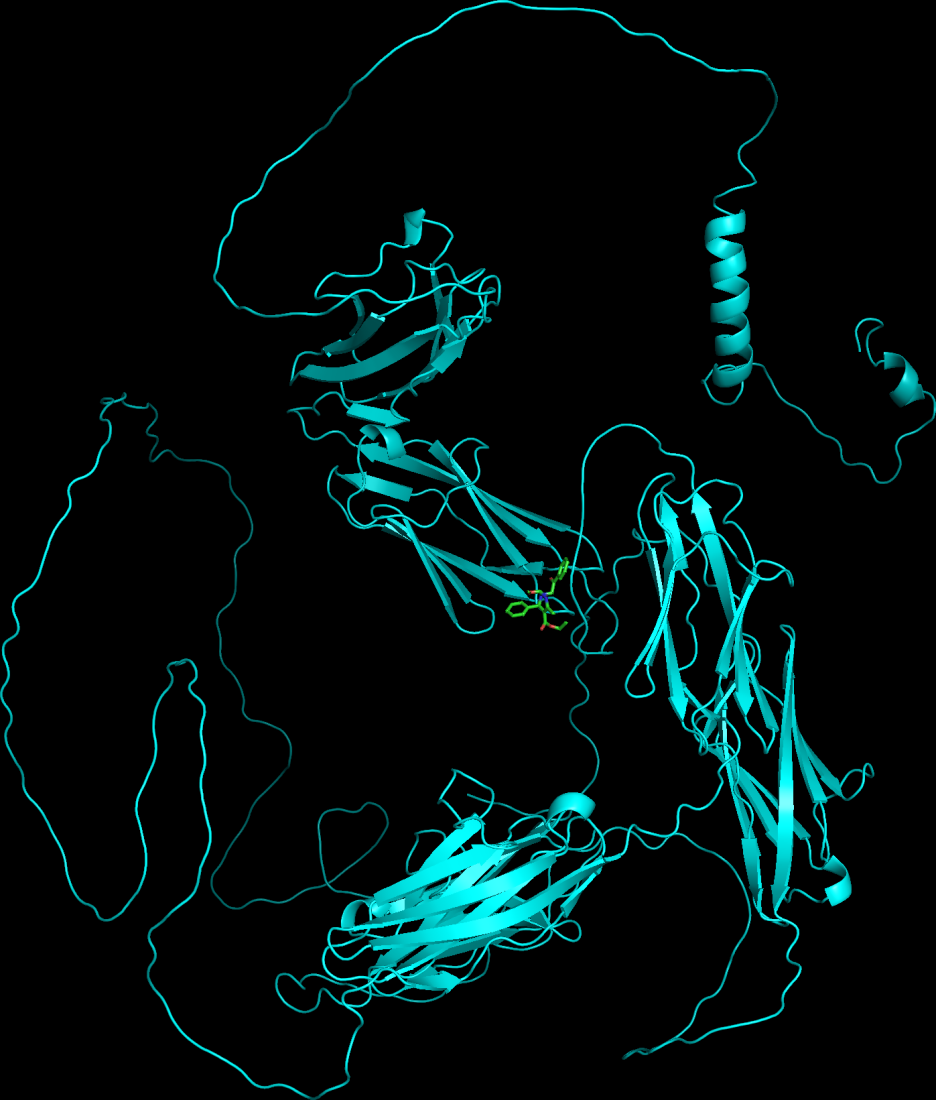
Tab. **[2](#MYLK-All-combining-Affinity-data)** 分子对接得分 (亲和度) 附表。



**Fig.** **2** Docking 9809715 into Q15746

**(File path: Figure+Table/3.2\_AutoDock\_vina\_分子对接\_(MYLK)/Docking-9809715-into-Q15746.png)**

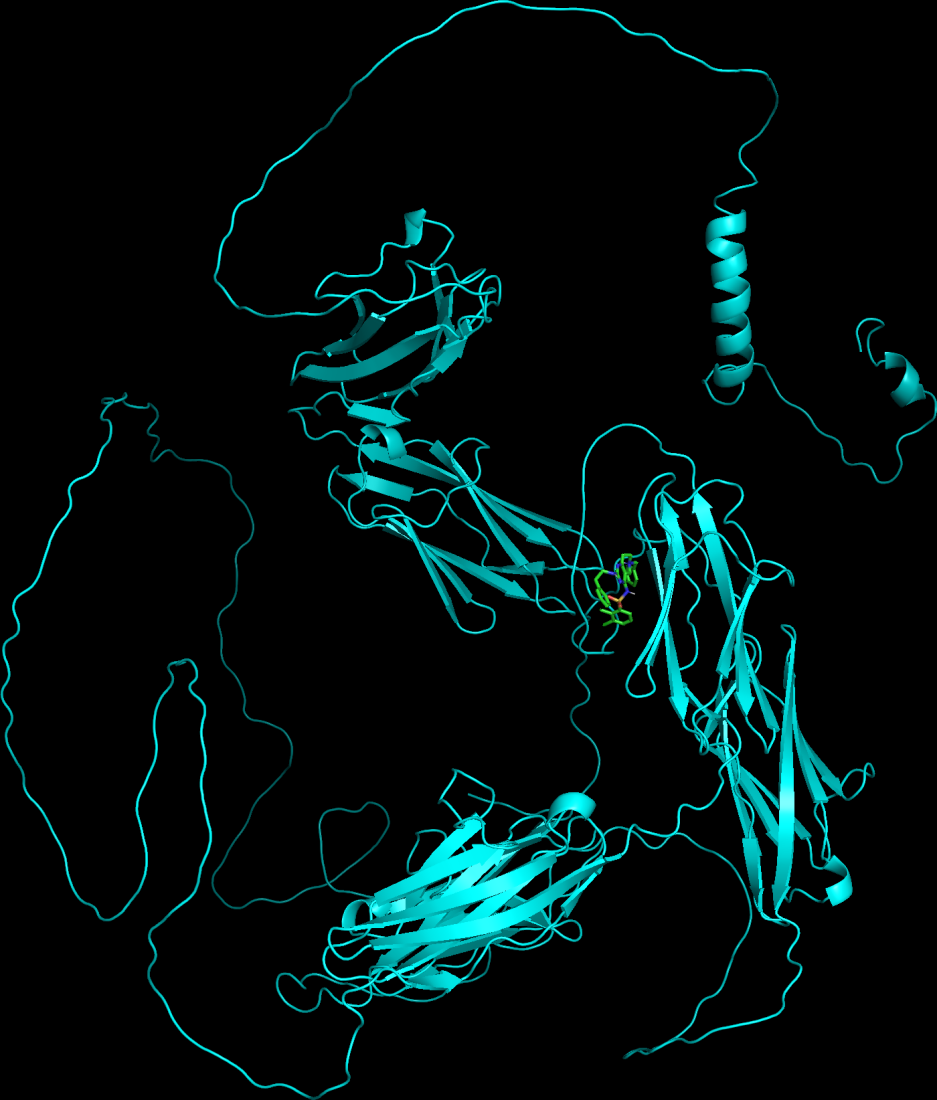
Fig. **[2](#Docking-9809715-into-Q15746)** 分子对接结果。蛋白(Symbol: MYLK) (PDB: q15746) 与化合物 (name: <CID:9809715>) (PubChem CID: 9809715)，亲和度为 -8.557。



**Fig.** **3** Docking 44225231 into Q15746

**(File path: Figure+Table/3.2\_AutoDock\_vina\_分子对接\_(MYLK)/Docking-44225231-into-Q15746.png)**

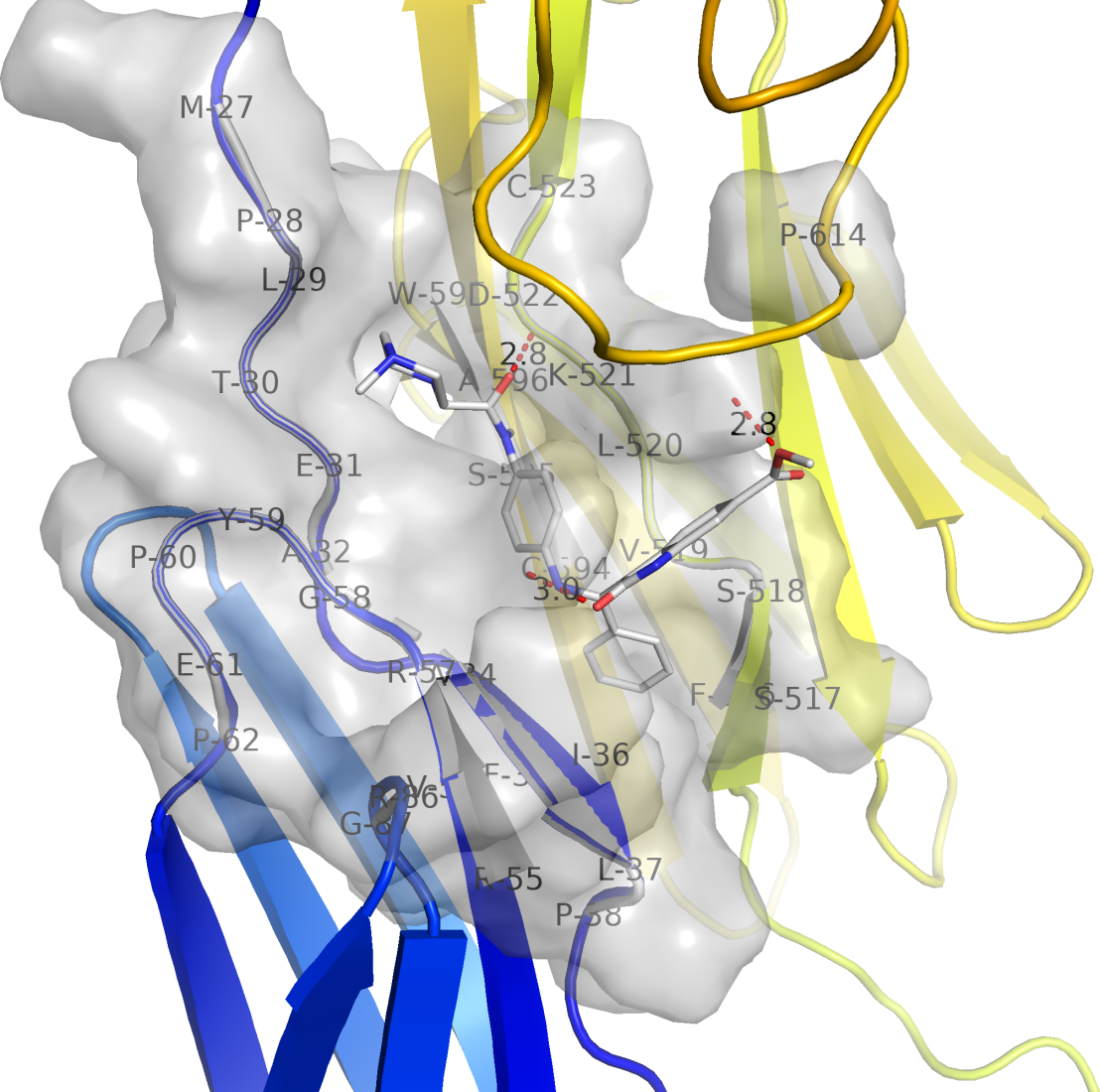
Fig. **[3](#Docking-44225231-into-Q15746)** 分子对接结果。蛋白(Symbol: MYLK) (PDB: q15746) 与化合物 (name: <CID:44225231>) (PubChem CID: 44225231)，亲和度为 -7.708。



**Fig.** **4** Docking 50898538 into Q15746

**(File path: Figure+Table/3.2\_AutoDock\_vina\_分子对接\_(MYLK)/Docking-50898538-into-Q15746.png)**

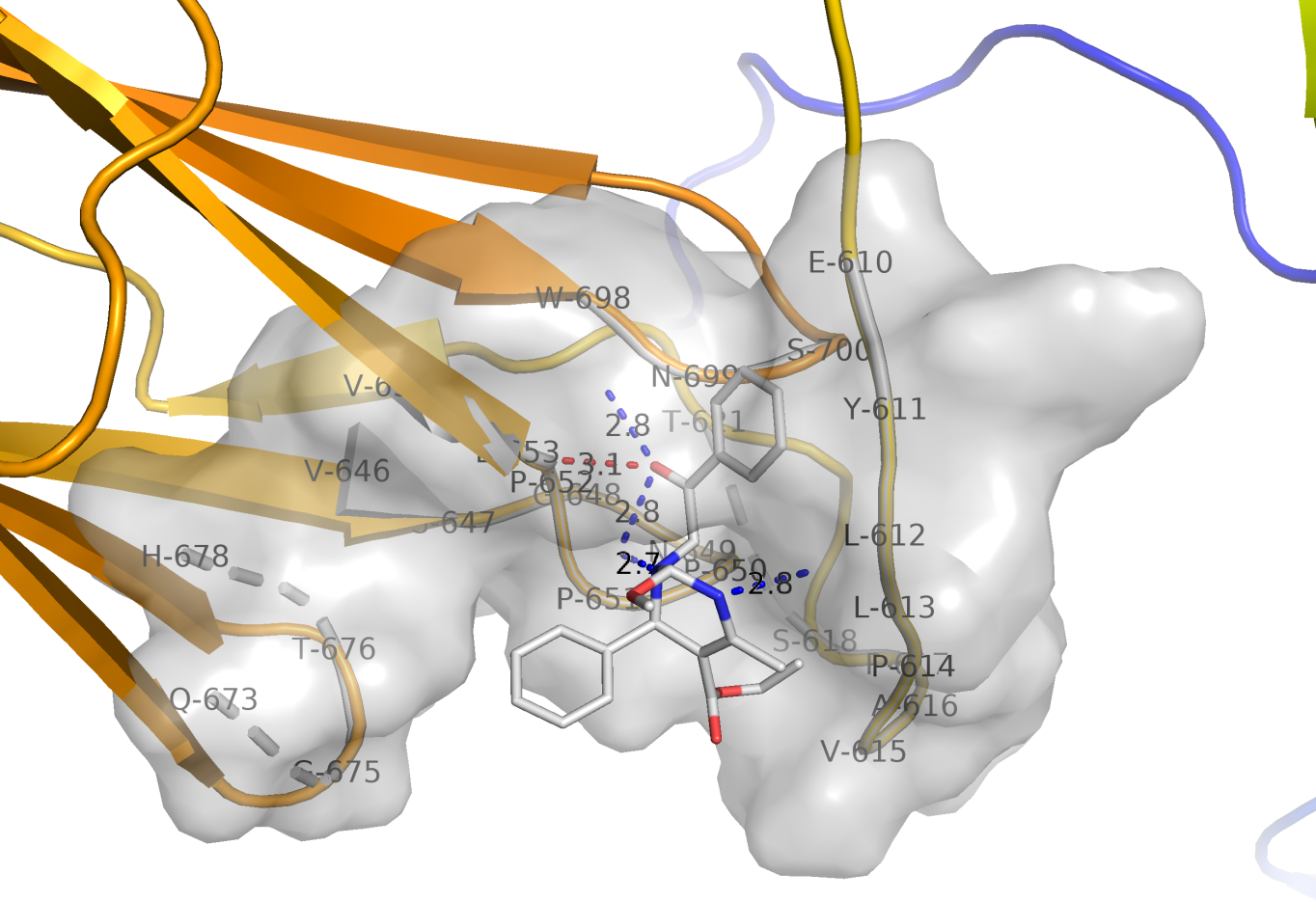
Fig. **[4](#Docking-50898538-into-Q15746)** 分子对接结果。蛋白(Symbol: MYLK) (PDB: q15746) 与化合物 (name: CaMKII-IN-1) (PubChem CID: 50898538)，亲和度为 -7.576。



**Fig.** **5** Docking 9809715 into Q15746 detail

**(File path: Figure+Table/3.2\_AutoDock\_vina\_分子对接\_(MYLK)/Docking-9809715-into-Q15746-detail.png)**

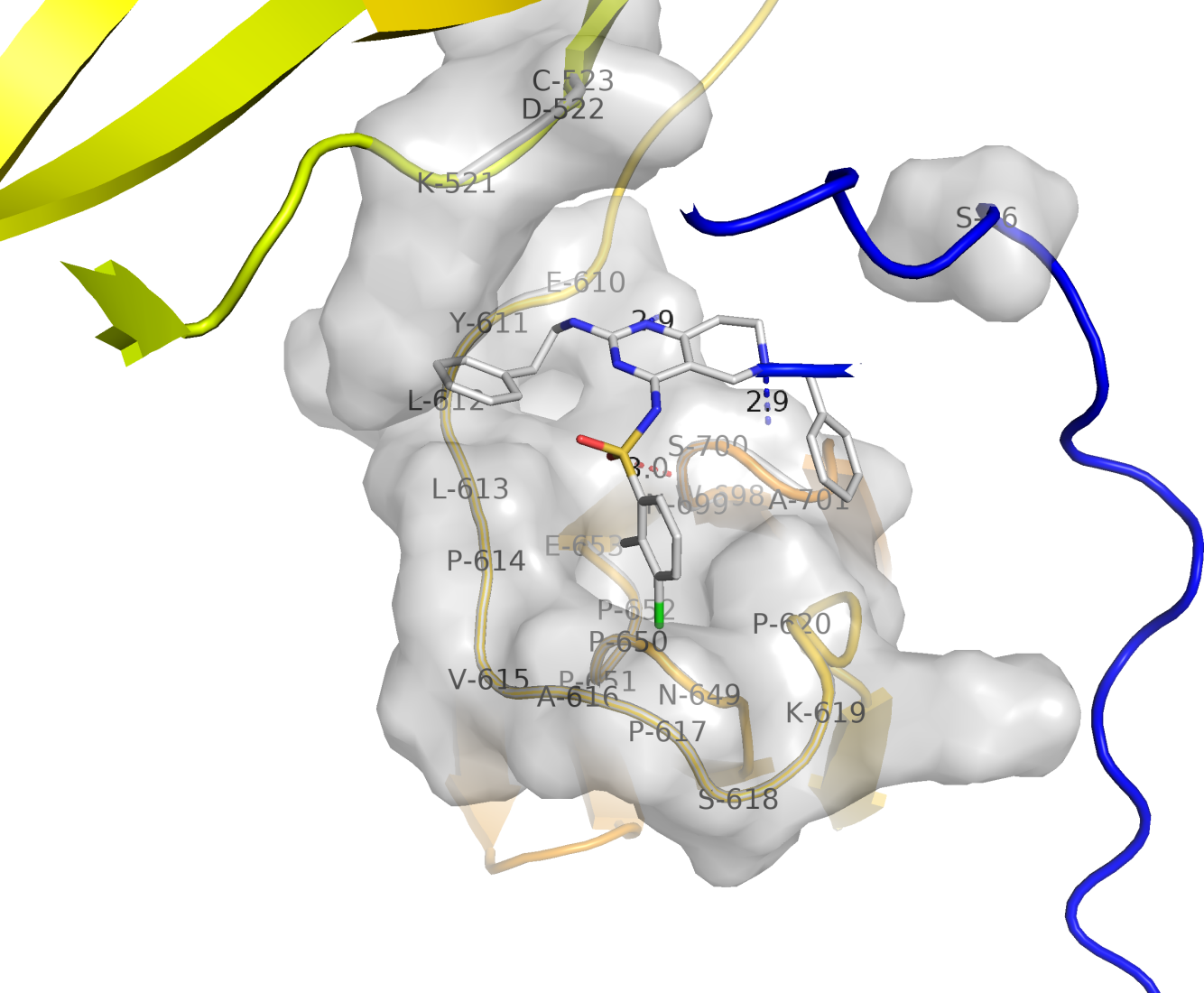
Fig. **[5](#Docking-9809715-into-Q15746-detail)** 分子对接结果局部细节。图中展示了蛋白(Symbol: MYLK) (PDB: q15746) 与化合物 (name: <CID:9809715>) (PubChem CID: 9809715) 之间的氢键结合。



**Fig.** **6** Docking 44225231 into Q15746 detail

**(File path: Figure+Table/3.2\_AutoDock\_vina\_分子对接\_(MYLK)/Docking-44225231-into-Q15746-detail.png)**

Fig. **[6](#Docking-44225231-into-Q15746-detail)** 分子对接结果局部细节。图中展示了蛋白(Symbol: MYLK) (PDB: q15746) 与化合物 (name: <CID:44225231>) (PubChem CID: 44225231) 之间的氢键结合。



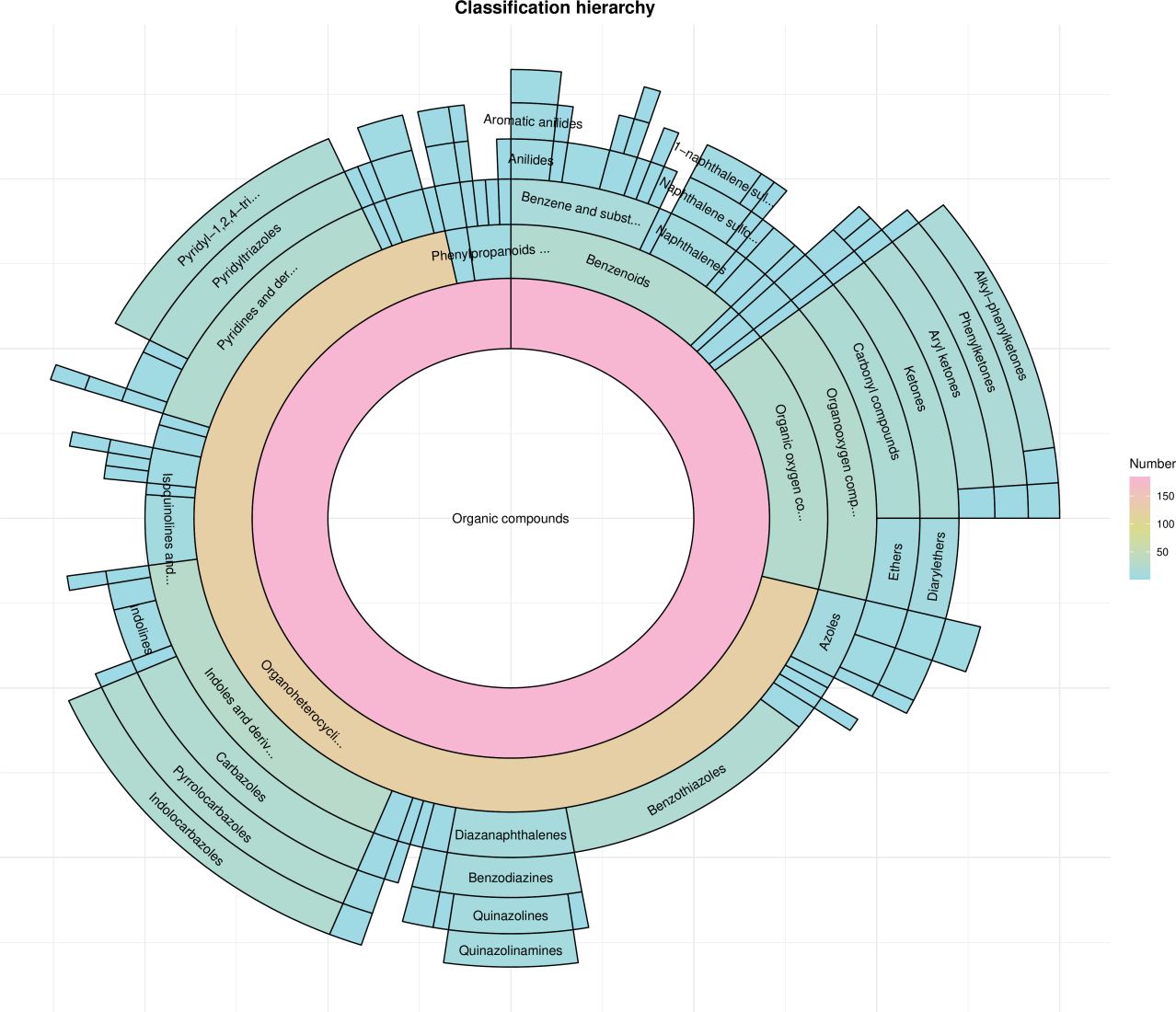
**Fig.** **7** Docking 50898538 into Q15746 detail

**(File path: Figure+Table/3.2\_AutoDock\_vina\_分子对接\_(MYLK)/Docking-50898538-into-Q15746-detail.png)**

Fig. **[7](#Docking-50898538-into-Q15746-detail)** 分子对接结果局部细节。图中展示了蛋白(Symbol: MYLK) (PDB: q15746) 与化合物 (name: CaMKII-IN-1) (PubChem CID: 50898538) 之间的氢键结合。

## 3.3 Classyfire 获取化合物分类信息 (MYLK)

对化合物 (9886808, 5330790, 153999, …[n = 194], 来自于AutoDock vina 分子对接[Section: MYLK]) 进行Classyfire 获取化合物分类信息。根据 cid 获取化合物 InChIKey。以 Classyfire 获取这些 InChIKey 的系统分类学注释。



**Fig.** **8** MYLK classification hierarchy

**(File path: Figure+Table/3.3\_Classyfire\_获取化合物分类信息\_(MYLK)/MYLK-classification-hierarchy.pdf)**

Fig. **[8](#MYLK-classification-hierarchy)** 为化合物的系统分类阶层概览。

**Tab.** **3** MYLK classification data

| Inchikey2d | Level | Classification | CHEMONT | Cid |
| --- | --- | --- | --- | --- |
| AAKJLRGGTJKAMG | Kingdom | Organic compounds | CHEMONTID:0000000 | 176900 |
| AAKJLRGGTJKAMG | Superclass | Organoheterocycli... | CHEMONTID:0000002 | 176900 |
| AAKJLRGGTJKAMG | Class | Diazanaphthalenes | CHEMONTID:0004788 | 176900 |
| AAKJLRGGTJKAMG | Subclass | Benzodiazines | CHEMONTID:0004789 | 176900 |
| AAKJLRGGTJKAMG | Level 5 | Quinazolines | CHEMONTID:0000485 | 176900 |
| ... | ... | ... | ... | ... |

**(File path: Figure+Table/3.3\_Classyfire\_获取化合物分类信息\_(MYLK)/MYLK-classification-data.xlsx)**

Tab. **[3](#MYLK-classification-data)** 为化合物系统分类附表。

# 4 总结

通过 BindingDB 数据库预筛选一批与 MYLK 结合的化合物 (Tab. **[1](#MYLK-bindingdb-annotation-data)** )， 随后，将这些化合物与 MYLK 特定的位置 (1-922) 分子对接。 结果见Fig. **[1](#MYLK-Overall-combining-Affinity)** (Top 10) 与 Tab. **[2](#MYLK-All-combining-Affinity-data)** 。

# Reference

1. Gilson, M. K. *et al.* BindingDB in 2015: A public database for medicinal chemistry, computational chemistry and systems pharmacology. *Nucleic acids research* **44**, D1045–D1053 (2016).

2. Seeliger, D. & Groot, B. L. de. Ligand docking and binding site analysis with pymol and autodock/vina. *Journal of Computer-Aided Molecular Design* **24**, (2010).

3. Zhang, Y., Forli, S., Omelchenko, A. & Sanner, M. F. AutoGridFR: Improvements on autodock affinity maps and associated software tools. *Journal of computational chemistry* **40**, 2882–2886 (2019).

4. Eberhardt, J., Santos-Martins, D., Tillack, A. F. & Forli, S. AutoDock vina 1.2.0: New docking methods, expanded force field, and python bindings. *Journal of Chemical Information and Modeling* **61**, 3891–3898 (2021).

5. Djoumbou Feunang, Y. *et al.* ClassyFire: Automated chemical classification with a comprehensive, computable taxonomy. *Journal of Cheminformatics* **8**, 61 (2016).