

Bioinformatics approaches in animal breeding

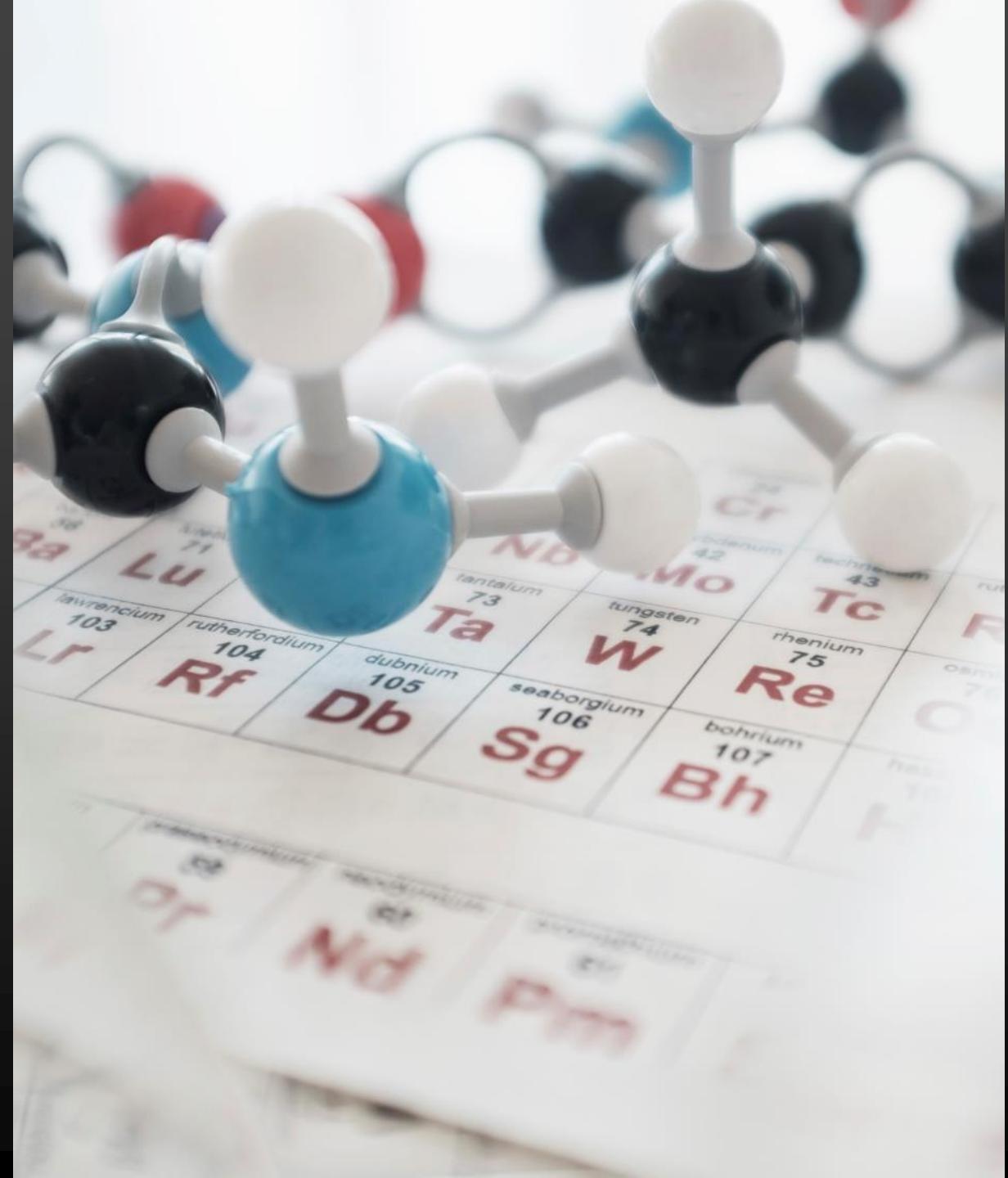
Summer School

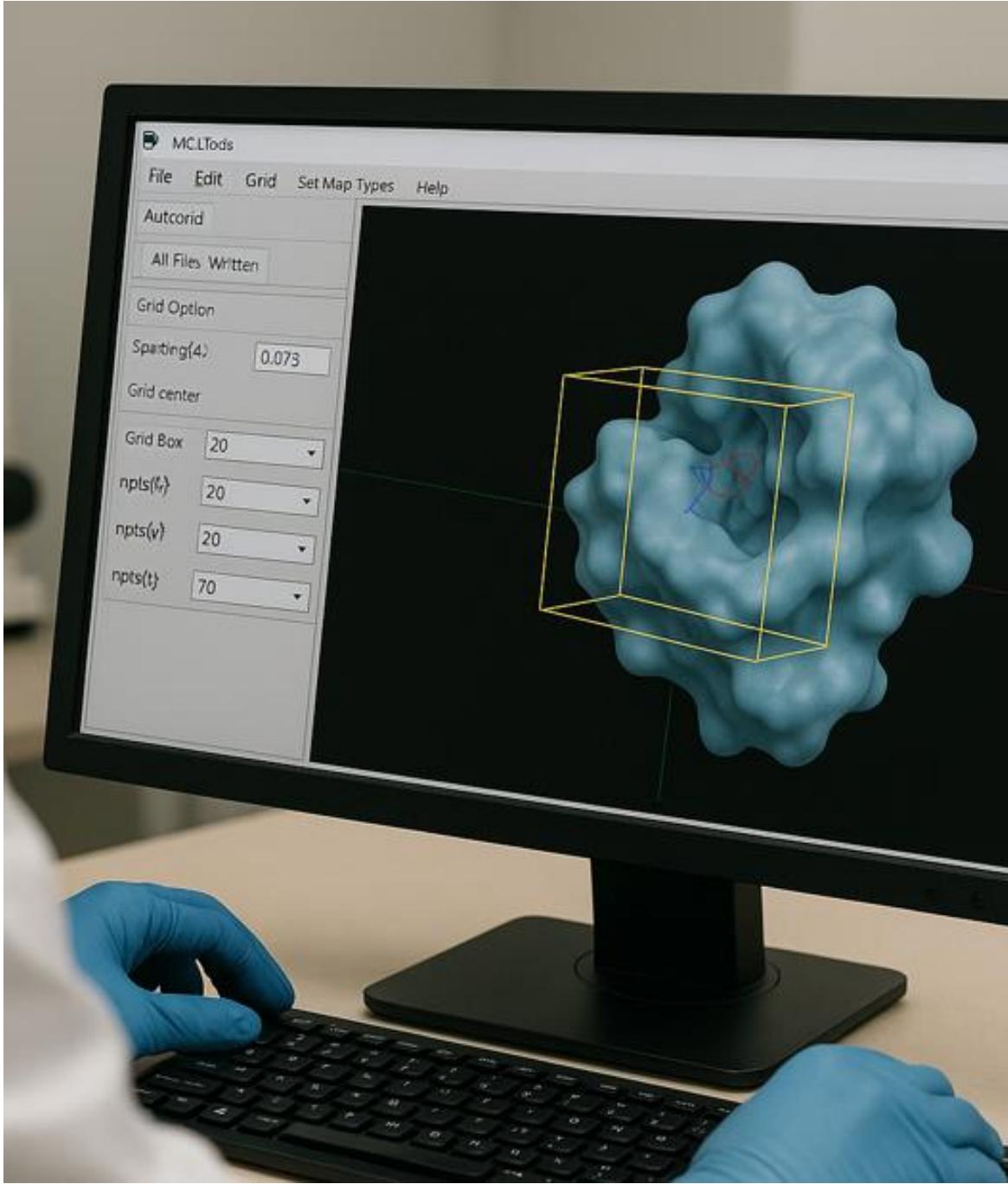
Zagreb

11.7.2025

Molecular docking

Anja Tanšek





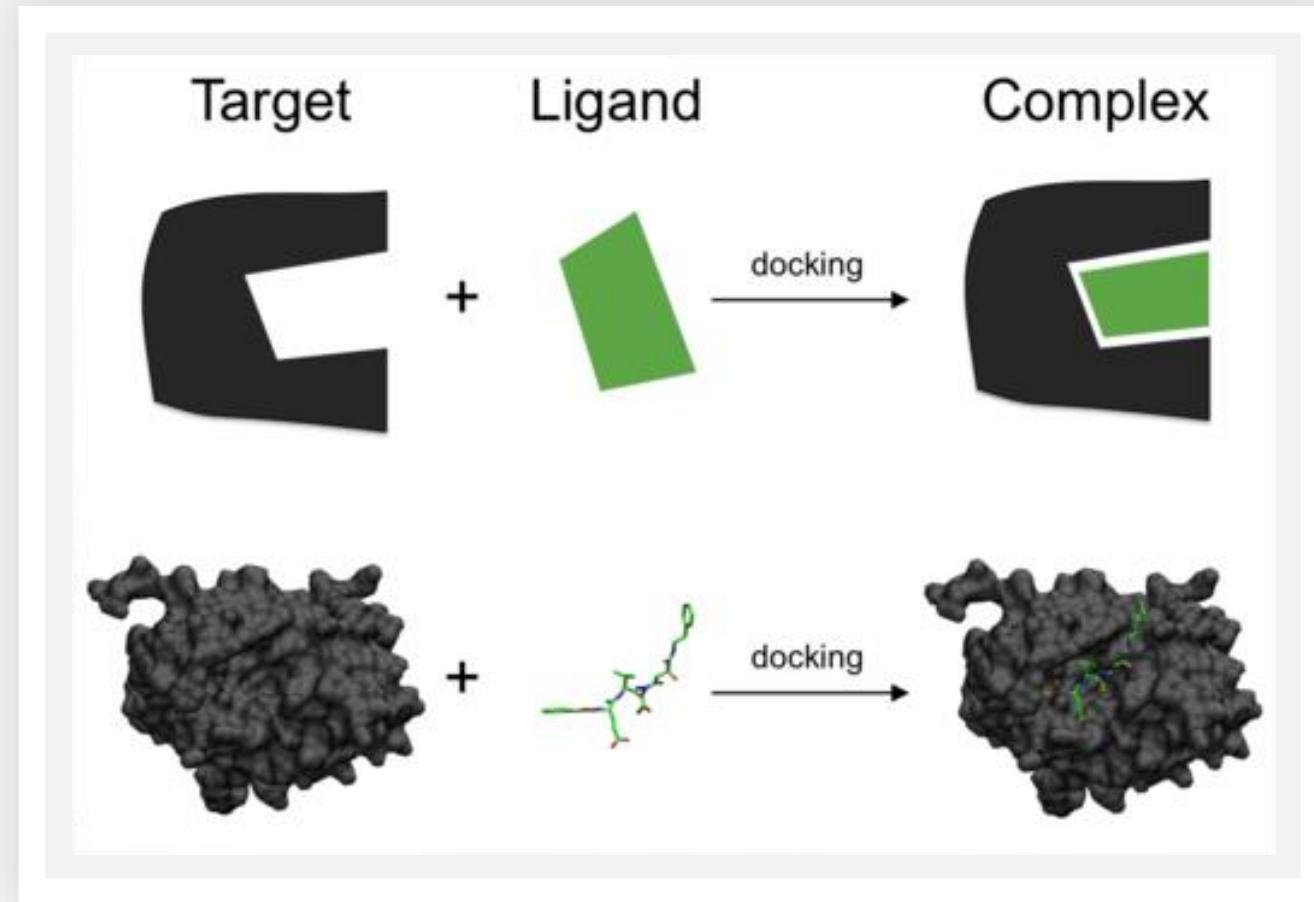
Molecular docking

What it is?



Molecular docking

- Interaction between:
 - Molecules (LIGANDS)
 - Target binding sites (RECEPTORS)
- Molecular basis of interaction:
 - Predicting binding affinity
 - Identifying binding poses
 - Screening potential ligands
- Aim:
 - Lead identification
 - Mechanism of action elucidation
 - Drug repurposing
 - Rational design
 - Cost reduction





Molecular docking

Where it can be used?



Use of molecular docking in different scientific fields

Agriculture

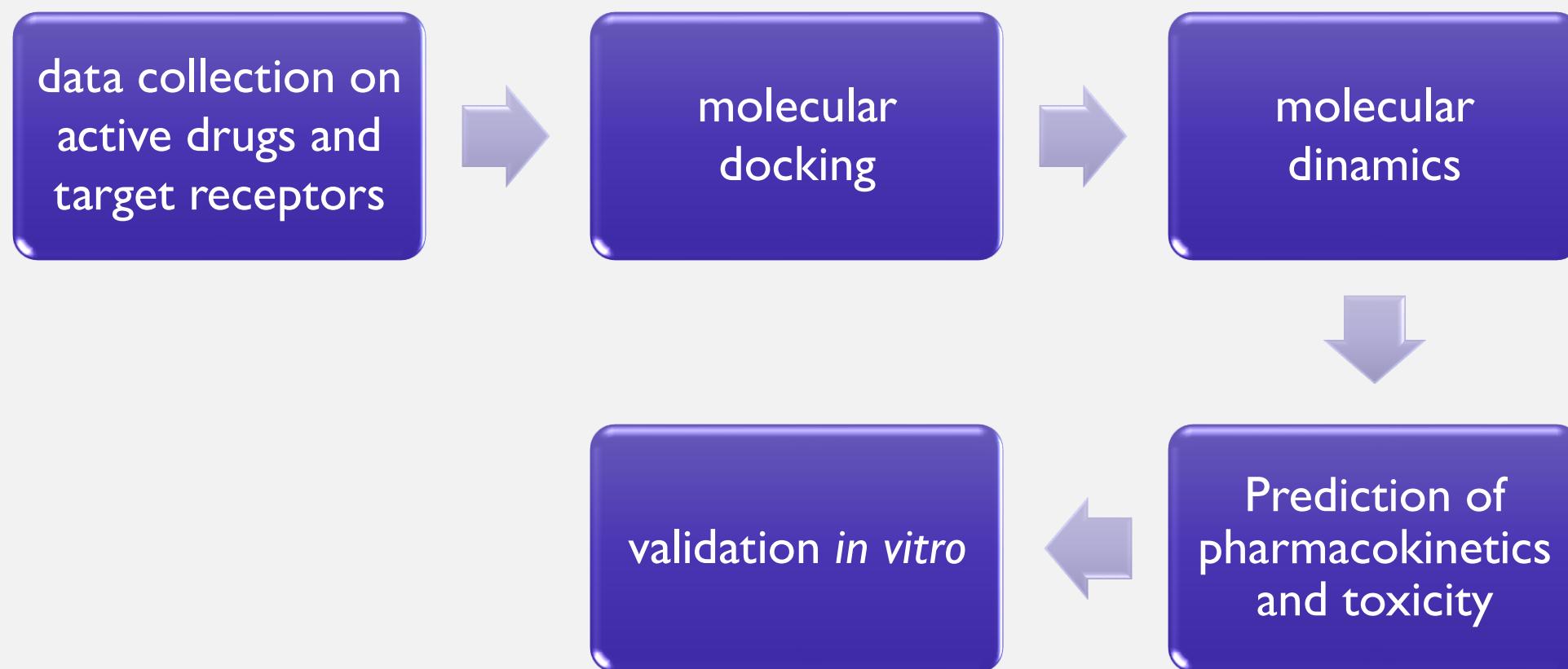
- Pesticide and herbicide development
- Crop improvement and plant biotechnology
- Soil science and microbiology

Animal Science

- Animal health and disease management
- Animal genetics and breeding
- Feed additive development
- Reducing environmental impact of animal production
- Drug discovery and development
- Diagnostics
- Toxicology
- Understanding disease mechanisms



Exploration of novel therapeutic indications for drugs through bioinformatics



Network Pharmacology and Molecular Docking: Exploring the Mechanism of Peppermint in Mastitis Prevention and Treatment in Dairy Cows

Xinyu Wang ^{1,2}, Jiaxin Lai ², Fei Xu ^{1,*} and Mingchun Liu ^{2,*}

¹ National Feed Drug Reference Laboratories, Institute of Feed Research, Chinese Academy of Agricultural Sciences, Beijing 100081, China; wangxinyu@caas.ac.cn

² College of Animal Science and Medicine, Shenyang Agricultural University, Shenyang 110866, China; laijiaxin@caas.ac.cn

* Correspondence: xufei@caas.ac.cn (F.X.); liumingchun@syau.edu.cn (M.L.)

Simple Summary: Bovine mastitis is a common condition in dairy production units that has serious productive consequences. However, due to the current limitations in the use of antibiotics, alternatives have been sought to counteract the consequences of this pathology, where mint (*Mentha* spp.) has proven to be a viable option. Therefore, this study is designed to elucidate the active ingredients, targets, and mechanisms of action of peppermint in treating bovine mastitis using network pharmacology and molecular docking methods. The findings indicate that the key targets of mint in treating bovine mastitis include Tumor Necrosis Factor (TNF), Interleukin-6 (IL-6), signal transducer and activator of transcription 3 (STAT-3), Interleukin-1 beta (IL-1 β), Recombinant Fibroblast Growth Factor 2 (PGF-2), Interferon gamma (IFNG), and Estrogen Receptor 1 (ESR-1), with the primary signaling pathways being the Advanced Glycation End-products-Receptor for Advanced Glycation End-products pathway (AGEs-RAGE), Interleukin-17 (IL-17), the Nuclear Factor kappa-light-chain-enhancer of activated B cells pathway (NF- κ B), the Toll-like receptor (TLR) pathway, the hypoxia-inducible factor-1 (HIF-1) pathway, the Transforming Growth Factor-beta pathway (TGF- β), the Phosphatidylinositol 3-kinase-Protein Kinase B pathway (PI3K-Akt), and the mitogen-activated protein kinase pathway (MAPK). Molecular docking results revealed that ursolic acid from mint has good binding activity with all core targets, and other constituents also form stable complexes with some of the core targets. These results suggest that active compounds such as ursolic acid in mint possess positive anti-inflammatory properties. Therefore, mint is considered to have a promising therapeutic effect on bovine mastitis.



Academic Editor: Keyu Zhang

Received: 18 December 2024

Revised: 3 February 2025

Accepted: 3 February 2025

Published: 5 February 2025

Citation: Wang, X.; Lai, J.; Xu, F.; Liu, M. Network Pharmacology and Molecular Docking: Exploring the Mechanism of Peppermint in Mastitis Prevention and Treatment in Dairy Cows. *Vet. Sci.* **2025**, *12*, 129. <https://doi.org/10.3390/vetsci12020129>

Copyright © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Wang et. al., (2025), DOI: 10.3390/vetsci12020129

Based on



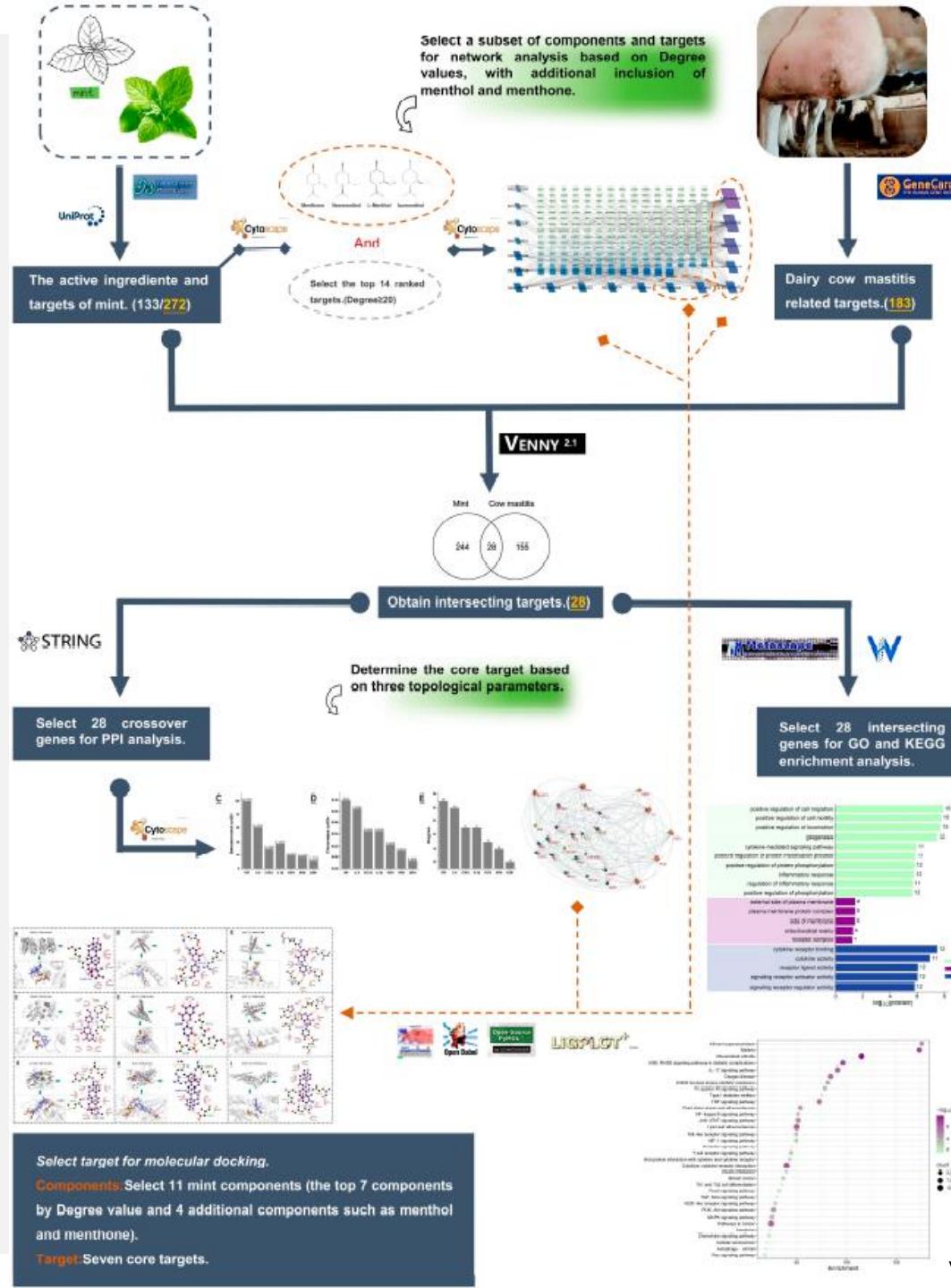
- Gather data:

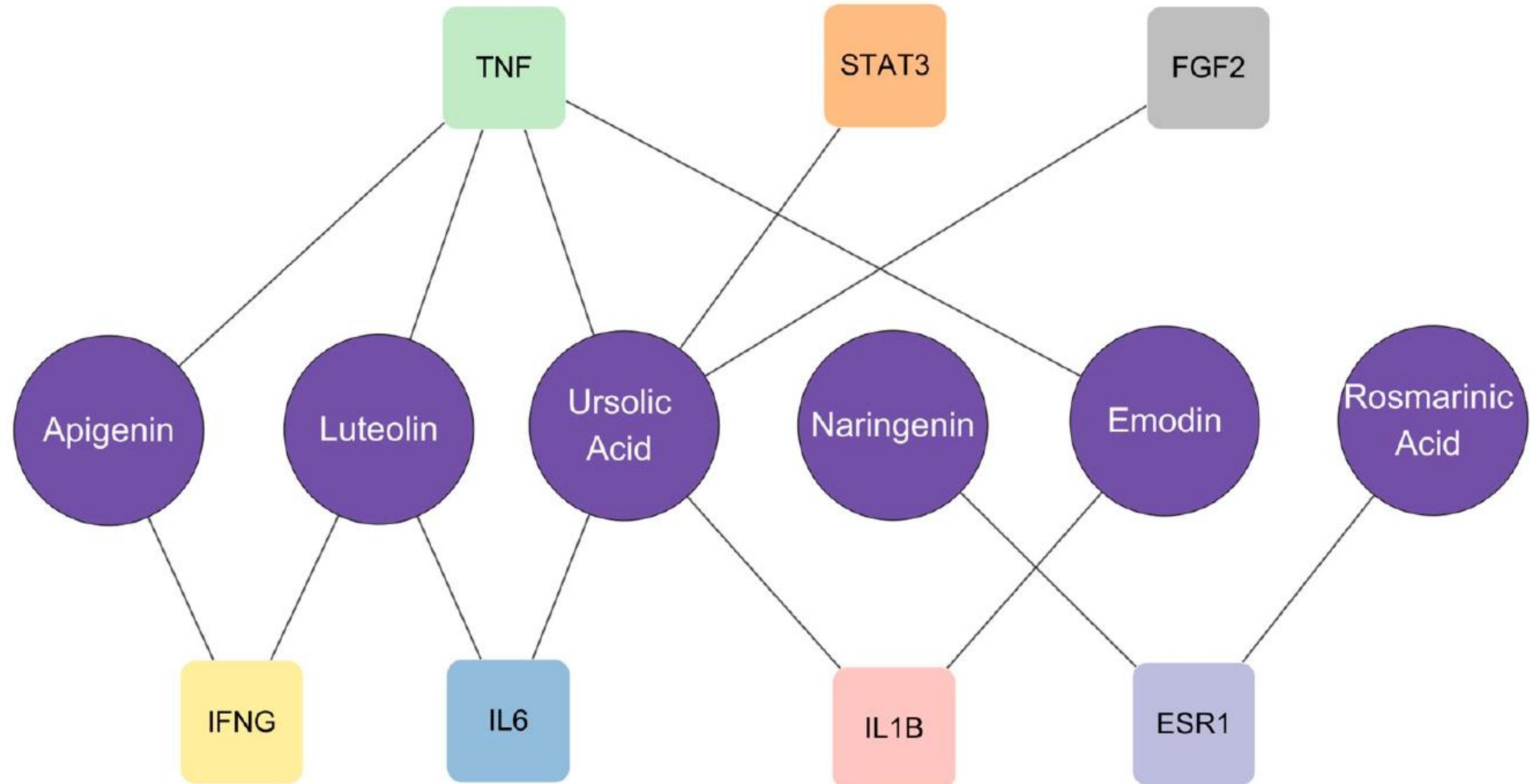
- Ligands
- Receptor

- Molecular docking

- Analyse results



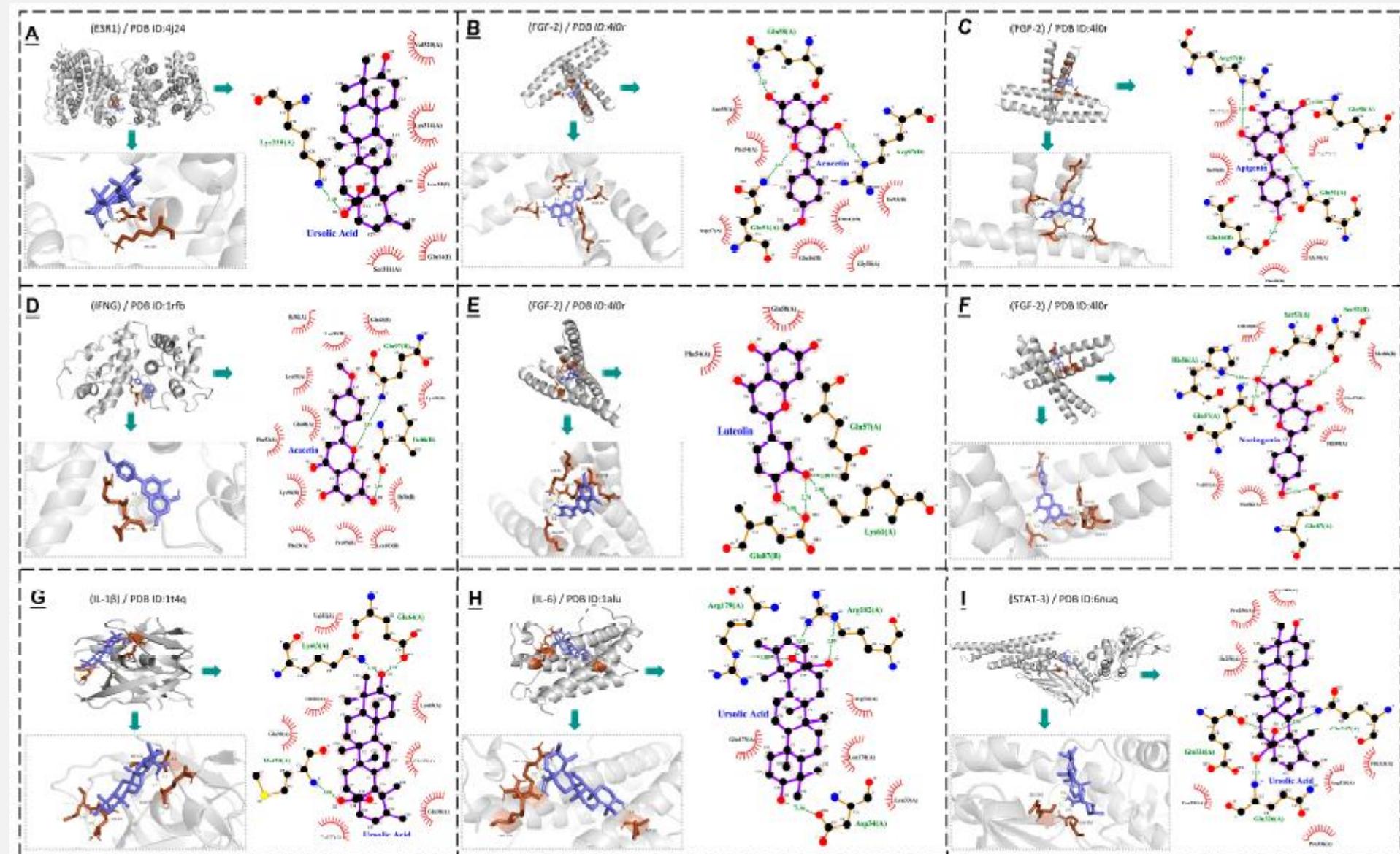




Mint Components	Core Target (PDB ID)						
	TNF (1u5y)	IL-6 (1alu)	STAT-3 (6nuq)	IL-1 β (1t4q)	FGF-2 (4l0r)	IFNG (1rfb)	ESR-1 (4j24)
Apigenin	-4.61/3	-4.3/2	-5.05/2	-5.5/2	-6.71/4	-5.63/2	-5.27/3
Luteolin	-4.27/3	-3.75/1	-4.41/3	-5.3/4	-6.32/5	-4.27/2	-4.16/3
Ursolic Acid	-5.86/2	-6.13/3	-7.45/2	-7.72/2	-5.84/1	-5.78/1	-7.16/1
Naringenin	-5.13/2	-4.09/1	-5.06/2	-5.23/2	-6.51/3	-5.38/1	-5.29/0
Emodin	-4.6/1	-4.81/2	-4.37/1	-5.35/3	-5.53/4	-4.56/1	-4.29/2
Rosmarinic Acid	-2.01/4	-3.12/3	-2.17/1	-2.76/2	-3.12/3	-4.59/2	-2.76/1
Acacetin	-4.64/2	-4.38/1	-4.58/2	-5.35/2	-6.85/3	-6.67/1	-4.38/2
Menthone	-4.72/1	-4.03/1	-4.55/1	-4.22/0	-5.18/1	-5.56/0	-4.06/1
Neomenthol	-4.43/2	-4.19/2	-4.68/2	-4.63/1	-4.97/1	-5.41/2	-5.13/2
L-Menthol	-4.58/1	-3.81/1	-4.33/2	-4.4/1	-5.54/2	-5.35/1	-4.98/1
Isomenthol	-4.39/2	-3.82/2	-4.84/2	-4.23/2	-4.96/1	-5.28/2	-4.79/2

Docking and combining energy (Kcal/mol)/number of hydrogen bonds.

Wang et. al., (2025), DOI: 10.3390/vetsci12020129



Collection of data on bioactive compounds and target receptors



Searching for compounds with similar chemical structures to those identified in the article.

- Apigenin, luteolin, **ursolic acid**, menthol, menthone, isomenthol, neomenthol



Protein Data Bank (PDB): crystallographic structure of target proteins.

- Core targets: TNF, IL-6, **STAT-3**, IL-1 β , FGF-2, IFNG, and ESR-1

Searching for suitable compounds for bioinformatic analysis:

LIGAND

- DrugBank: <https://go.drugbank.com/>
 - In the search bar, type „ursolic acid“
 - Find „Chemical Identifiers“ in the left column
 - Copy „SMILES“ chemical formula
 - Open a molecular search engine for similar structures
 - In the navigation bar: Explore → Chemical Structure
 - Click on „Marvin JS“ with the left mouse button and press Ctrl+C to paste the SMILES
 - You can see the 2D representation of the molecule
 - Set search parameters
 - Similarity
 - Similarity threshold: 0.7
 - Drug types: Approved
 - Click Search
 - We get a list of molecules that are approved for use and whose structure is 70 % similar to the structure of input molecule.
 - Repeat for all ligands.
-
- Find the desired molecules in the PubChem database: <https://pubchem.ncbi.nlm.nih.gov/>
 - Save 3D Conformer as .sdf file

RECEPTOR

- PDB: <https://www.rcsb.org/>
- In the search bar, type „STAT3“
- Select the desired/appropriate structure
 - (in)active form
 - Whole/part of the receptor
 - Monomer/multimer
 - Method and resolution
 - Missing amino acids

OR

- Model the 3D structure with AlphaFold Server: <https://alphafoldserver.com/welcome>



LIGANDA RETRIEVAL (.SDF)

Search for similar compounds in DrugBank Online and save their .sdf files from PubChem database.



AI-Powered Drug Intelligence, No Coding Required. Explore the evolution of DrugBank

REQUEST ACCESS →



Products Use Cases Our Tech Resources Company

SIGN IN

CHAT WITH AN EXPERT

Discover More

Power bigger research and better decisions
with the world's most robust drug knowledge

CHAT WITH AN EXPERT



<https://www.drugbank.com/>



Search our knowledgebase's 500,000+ drugs and drug products:

Drugs

Type your search...



The pharmaceutical data you've been looking for

DrugBank is a vital resource for your pharmaceutical research, offering comprehensive & reliable drug data, structured for immediate use or easy software integration.

SCROLL TO SEE OUR SOLUTIONS



Data management made effortless

Drug Discovery

Clinical Software

Academic Research

<https://go.drugbank.com/>

Search section headers 

Identification

Pharmacology

Interactions

Categories

Chemical Identifiers

References

Clinical Trials

Pharmacoeconomics

Properties

Spectra

Targets (7)

Ursolic acid

 Star

0

Explore a selection of our essential drug information below, or:

 SAY GOODBYE TO FRAGMENTED, OUTDATED RESEARCH TOOLS.

The Next Evolution of DrugBank is Here

DrugBank's latest evolution delivers AI-powered, scientifically validated insights to accelerate drug discovery like never before. [Request Early Access](#)

 SUBSCRIBE

Data Packages

Explore the full scope of our drug knowledge tailored for pharmaceutical research needs in our data library. [Learn more.](#)

Generic Name Ursolic acid

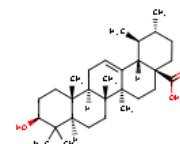
DrugBank Accession Number DB15588

Background Not Available

Modality Small Molecule

Groups Investigational

Structure



Weight
Average: 456.711
Monoisotopic: 456.360345406

Chemical Formula C₃₀H₄₈O₃

Synonyms

Malol Prunol Ursolic acid Urson

 [View As Table](#)



Search section headers 

Identification

Pharmacology

Interactions

Categories

Chemical Identifiers

UNII

CAS number

InChI Key

InChI

IUPAC Name

SMILES

References

Clinical Trials

Pharmacoconomics

Properties

Spectra

Targets (7)



CHEMICAL IDENTIFIERS

UNII P3M2575F3F

CAS number 77-52-1

InChI Key WCGUUGGRBIKTOS-GPOJBZKASA-N

InChI InChI=15/C30H48O3/c1-18-10-15-30(25(32)33)17-16-28(6)20(24(30)19(18)2)8-9-22-27(5)13-12-23(31)26(3,4)21(27)11-14-29(22,28)7/h8,18-19,21-24,31H,9-17H2,1-7H3,(H,32,33)/t18-19+,21+,22,23+,24+,27+,28-29,30+/m1/s1

IUPAC Name (15,2R,4aS,5,6aS,6bR,8aR,10S,12aR,12bR,14bS)-10-hydroxy-1,2,6a,6b,9,9,12a-heptamethyl-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-icosahydrophenene-4a-carboxylic acid

SMILES [H][C@@]12[C@@H](C)[C@H](C)CC[C@@]1(CC[C@]1(C)C2=CC[C@]2([H])[C@@]3(C)CC[C@H](O)C(C)(C)[C@]3([H])CC[C@@]12C)C(O)=O

REFERENCES

General References Not Available

External Links Human Metabolome Database HMDB0002395

CD8988

KEGG Compound

58472

ChemSpider

50148911

BindingDB

1426931

RxNav

9908

ChEBI

CHEMBL169

ChEMBL

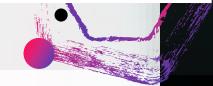
ZINC000003978827

ZINC

505

DDDB-Listed

SMILES for ursolic acid: [H][C@@]12[C@@H](C)[C@H](C)CC[C@@]1(CC[C@]1(C)C2=CC[C@]2([H])[C@@]3(C)CC[C@H](O)C(C)(C)[C@]3([H])CC[C@@]12C)C(O)=O



 Type your search...



Search section headers 

Identification

Pharmacology

Interactions

Categories

Chemical Identifiers

UNII

CAS number

InChI Key

InChI

IUPAC Name

SMILES

References

Clinical Trials

Pharmacoeconomics

Properties

Spectra

Targets (7)



 **Drug Knowledgebase**
Start here! Dig into essential drug information

 **Interaction Checker**
Check interactions in up to 5 drugs at a time

Advanced Exploration

Categories & ATC Classifications 

Pathways 

Drug Reactions 

Drug Classification 

Drug Targets 

Chemical Structure 

TOS-GPOJBZKASA-N
C13/c1-18-10-15-30(25(32)33)17-16-28(6)20(24(30)19(18)2)8-9-22-27(5)13-12-23(31)26(3,4)21(27)11-14-29(22,28)7/h8,18-19,21-24,31H,9-17H² 1-7H³ (H 32,33)/t18,19+,21+,22-,23+,24+,27+,28-,29-,30+/m1/s1
,8aR,10S,12aR,12bR,14bS)-10-hydroxy-1,2,6a,6b,9,9,12a-heptamethyl-1,2,3,4,4a,5,6,6a,6b,7,8,8a,9,10,11,12,12a,12b,13,14b-icosahydropicene-4a-carboxylic acid

Ask 

SMILES  [H] [C@H]12[C@H]([C@H]([C@H]([C@H]1(CC[C@H]1[C@C2=CC[C@H]2([H])[C@H]3[C@H]([C@H]([C@H]([C@H]3([H])CC[C@H]12C)C)C)C)C)C)C)C)C=O

REFERENCES

General References Not Available

External Links Human Metabolome Database  HMDB0002395

KEGG Compound  C08988

ChemSpider  58472

BindingDB  50148911

RxNav  1426931

ChEBI  9908

ChEMBL  CHEMBL169

ZINC  ZINCD00003978827



DRUGBANK Online

Explore ▾ For Drug Discovery ▾ For Clinical Software ▾ For Academic Research

CHAT WITH AN EXPERT

Type your search...

Chemical Structure Search

STRUCTURE SEARCH MOLECULAR WEIGHT

Search by chemical structure

Marvin JS by ChemAxon

Powered by ChemAxon

Search Options

Similarity Substructure Exact

Similarity threshold: 0.7

Minimum Weight: e.g. 100 Maximum Weight: e.g. 200

Maximum Results: 100

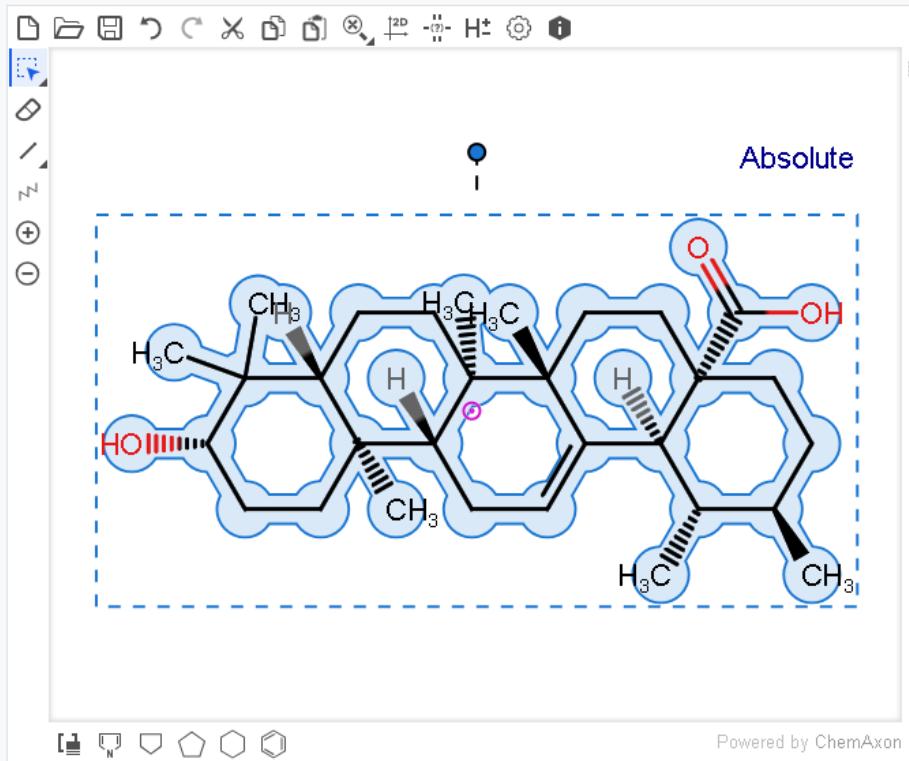
Drug Types (default all):

Approved Vet approved Nutraceutical
 Illicit Withdrawn Investigational
 Experimental

Search LOAD EXAMPLE MARVINJS TUTORIALS

STRUCTURE SEARCH MOLECULAR WEIGHT

Search by chemical structure



Search Options

Similarity Substructure Exact

Similarity threshold

0.7

Minimum Weight

e.g. 100

Maximum Weight

e.g. 200

Maximum Results

100

Drug Types (default all):

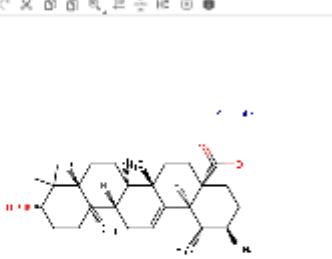
Approved Vet approved Nutraceutical
 Illicit Withdrawn Investigational
 Experimental

Search

LOAD EXAMPLE

MARVINJS TUTORIALS

Search by chemical structure



Search Options

Similarity Substructure Exact

Similarity threshold: 0.7

Molecular Weight Monoisotopic Weight

Min. MW: 100 Max. MW: 300

Molecular Ratios

100:1

Drug Type (JDK/DB/All):

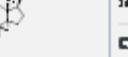
Approved Not approved Nutraceutical

Black Withdrawn Investigational

Experimental

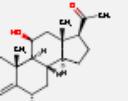
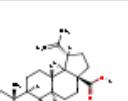
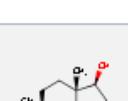
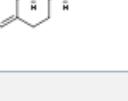
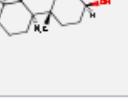
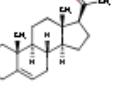
Search **LOAD EXAMPLE** **MARINE TUTORIALS**

Results 1 – 7 of approximately 7 results

DB00253 Score: 0.735		Medrysone 2668-66-8	C ₂₂ H ₃₂ O ₃ Mono mass: 344.23514489
DB12490 Score: 0.722		Betulinic Acid 472-15-1	C ₃₀ H ₄₈ O ₃ Mono mass: 456.360345406
DB00624 Score: 0.72		Testosterone 58-22-0	C ₁₉ H ₂₈ O ₂ Mono mass: 288.20893014
DB05710 Score: 0.72		Methyltestosterone 58-18-4	C ₂₀ H ₃₀ O ₂ Mono mass: 302.224580204
DB01708 Score: 0.713		Prasterone 53-43-0	C ₁₉ H ₂₈ O ₂ Mono mass: 288.20893014
DB02799 Score: 0.713		Pregnenolone 145-13-1	C ₂₁ H ₃₂ O ₂ Mono mass: 316.240230268
DB01493 Score: 0.708		Ethylestrenol 965-90-2	C ₂₀ H ₃₂ O Mono mass: 288.245315646

Results 1 – 7 of approximately 7 results

← Previous Next →

DB00253 Score: 0.735		Medrysone 2668-66-8	C ₂₂ H ₃₂ O ₃ Mono mass: 344.23514489
DB12490 Score: 0.722		Betulinic Acid 472-15-1	C ₃₀ H ₄₈ O ₃ Mono mass: 456.360345406
DB00624 Score: 0.72		Testosterone 58-22-0	C ₁₉ H ₂₈ O ₂ Mono mass: 288.20893014
DB05710 Score: 0.72		Methyltestosterone 58-18-4	C ₂₀ H ₃₀ O ₂ Mono mass: 302.224580204
DB01708 Score: 0.713		Prasterone 53-43-0	C ₁₉ H ₂₈ O ₂ Mono mass: 288.20893014
DB02799 Score: 0.713		Pregnenolone 145-13-1	C ₂₁ H ₃₂ O ₂ Mono mass: 316.240230268
DB01493 Score: 0.708		Ethylestrenol 965-90-2	C ₂₀ H ₃₂ O Mono mass: 288.245315646

Explore Chemistry

Quickly find chemical information from authoritative sources

Search

Try

aspirin

EGFR

C9H8O4

57-27-2

C1=CC=C(C=C1)C=O

InChI=1S/C3H6O/c1-3(2)4/h1-2H3

Use Entrez

Compounds

Substances

BioAssays



Draw Structure



Upload ID List



Browse Data



Periodic Table

<https://pubchem.ncbi.nlm.nih.gov/>



Ask 

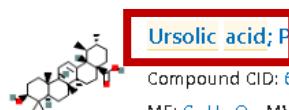
SEARCH FOR

ursolic acid



Treating this as a text search.

BEST MATCH



Ursolic acid; Prunol; Malol; Ursone; 3beta-Hydroxyurs-12-en-28-oic acid; Micromerol; (3beta)-3-Hydroxyurs-12-en-28-oic acid; CCRIS 7123; ...

Compound CID: 64945

MF: C₃₀H₄₈O₃ MW: 456.7 g/mol

IUPAC Name: (1S,2R,4aS,6aR,6aS,6bR,8aR,10S,12aR,14bS)-10-hydroxy-1,2,6a,6b,9,9,12a-heptamethyl-2,3,4,5,6,6a,7,8,8a,10,11,12,13,14b-tetradecahydro-1H-picene-4a-carboxylic acid

SMILES: C[C@H]1CC[C@H]2CC[C@H]3C(=CC[C@H]4[C@H]3CC[C@H]5[C@H]4CC[C@H](C5(C)C(=O)O)C)[C@H]2[C@H]1C)C(=O)O

InChIKey: WCGUUGGRBIKTOS-GPOJBZKASA-N

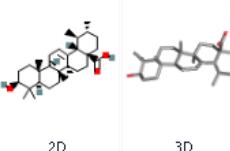
InChI: InChI=1S/C30H48O3/c1-18-10-15-30(25(32)33)17-16-28(6)20(24(30)19(18)2)8-9-22-27(5)13-12-23(31)26(3,4)21(27)11-14-29(22,28)7/h8,18-19,21-24,31H,9-17H2,1-7H3,(H,32,33)/t18-,19+,21+,22-,23+,24+,27+,28-,29-,30+/m1/s1

Create Date: 2005-06-24

[Summary](#) [Similar Structures Search](#) [Related Records](#) [PubMed \(MeSH Keyword\)](#)

COMPOUND SUMMARY

Ursolic Acid

PubChem CID	64945
Structure	
Primary Hazards	Laboratory Chemical Safety Summary (LCSS) Datasheet
Molecular Formula	$C_{30}H_{48}O_3$
Synonyms	Ursolic acid Prunol Malol Urson 3β -Hydroxyurs-12-en-28-oic acid View More...
Molecular Weight	456.7 g/mol <small>Computed by PubChem 2.2 (PubChem release 2025.04.14)</small>
Dates	Create: 2005-06-24 Modify: 2025-07-05
Description	Ursolic acid is a pentacyclic triterpenoid that is urs-12-en-28-oic acid substituted by a beta-hydroxy group at position 3. It has a role as a plant metabolite and a geroprotector. It is a pentacyclic triterpenoid and a hydroxy monocarboxylic acid. It derives from a hydride of an ursane .  ChEBI

 Cite

 Download

CONTENTS

[Title and Summary](#)

- [1 Structures](#)
- [2 Names and Identifiers](#)
- [3 Chemical and Physical Properties](#)
- [4 Spectral Information](#)
- [5 Related Records](#)
- [6 Chemical Vendors](#)
- [7 Drug and Medication Information](#)
- [8 Pharmacology and Biochemistry](#)
- [9 Use and Manufacturing](#)
- [10 Safety and Hazards](#)
- [11 Toxicity](#)
- [12 Associated Disorders and Diseases](#)
- [13 Literature](#)
- [14 Patents](#)
- [15 Interactions and Pathways](#)
- [16 Biological Test Results](#)
- [17 Taxonomy](#)
- [18 Classification](#)
- [19 Information Sources](#)

1.2 3D Conformer



Structure Search

Get Image

Download Coordinates

Interactive Chemical Structure Model

Ball and Stick

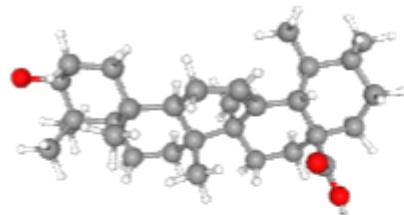
Sticks

Wire-Frame

Space-Filling

Show Hydrogens

Animate

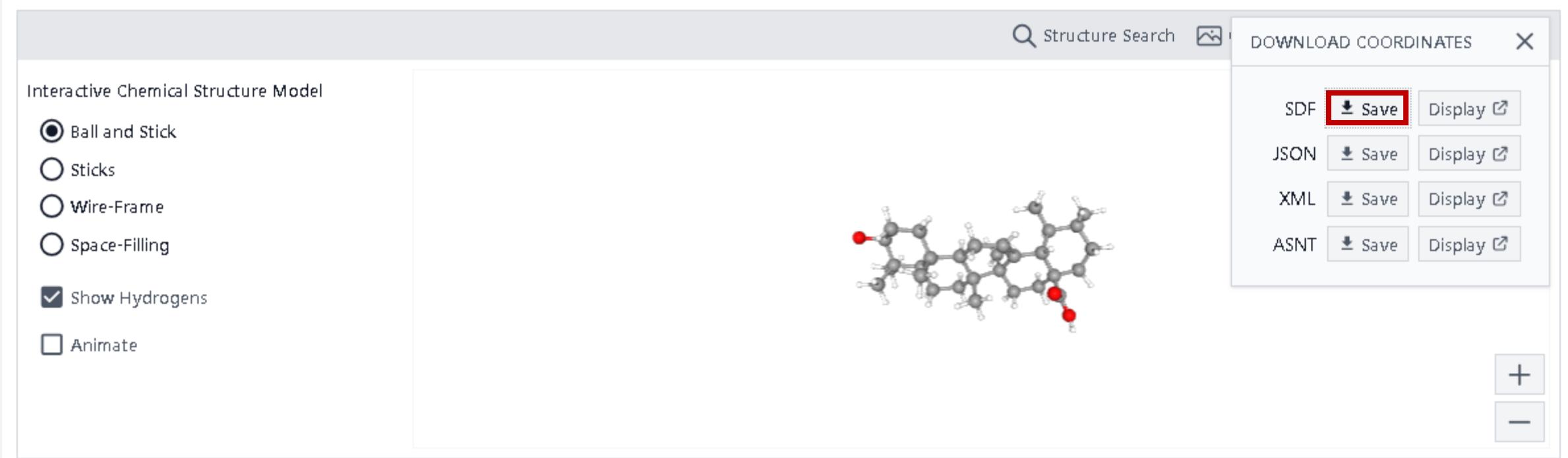


+

-



1.2 3D Conformer



RENAME FILES TO: ligand.sdf

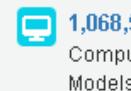
RECEPTOR RETRIEVAL (.PDB)

From PDB





238,622
Structures from the
PDB archive



1,068,577
Computed Structure
Models (CSM)

Enter search term(s), Ligand ID or sequence

Include CSM



Help



Redesigned PDB Statistics Support Enhanced Functionality

[Explore Statistics](#)

Welcome

Deposit

Search

Visualize

Analyze

Download

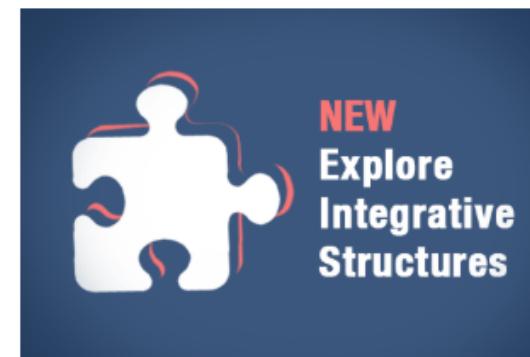
Learn

RCSB Protein Data Bank (RCSB PDB) enables breakthroughs in science and education by providing access and tools for exploration, visualization, and analysis of:

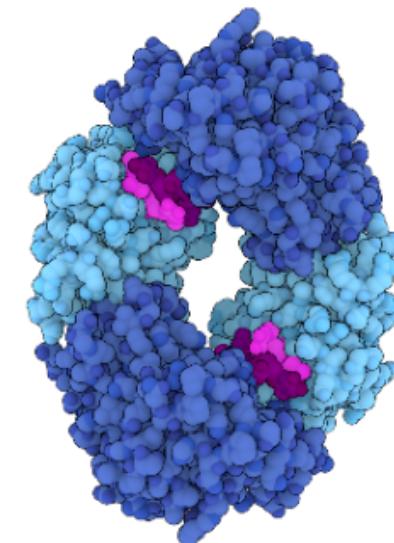
Experimentally-determined 3D structures from the **Protein Data Bank (PDB)** archive

Integrative 3D Structures from the PDB Archive

Computed Structure Models (CSM) from AlphaFold DB and ModelArchive



July Molecule of the Month



Capturing Beta-Lactamase in Action

<https://www.rcsb.org/>

Refinements ●

Structure Determination Methodology

experimental (2)

Scientific Name of Source Organism

Homosapiens (10)

Mus musculus (10)

Mus musculus 1 (0)

synthetic construct (0)

Taxonomy

Erkayata (2)

Verkandavilla (0)

other sequences (0)

Experimental Method

X-RAY DIFFRACTION (10)

ELECTRON MICROSCOPY (0)

SOLUTION NMR (0)

Polymer Entity Type

Protein (2)

DNA (0)

RNA (0)

Refinement Resolution (Å)

10-15 (0)

20-25 (0)

25-30 (0)

30-35 (0)

35-40 (0)

40-45 (0)

Release Date

1996-1999 (0)

2005-2009 (0)

2010-2014 (0)

2015-2019 (0)

2020-2024 (0)

Enzyme Classification Name

Transferases (6)

Hydrolases (0)

Symmetry Type

Asymmetric (16)

Cyclic (0)

SCOP Classification

All alpha proteins (0)

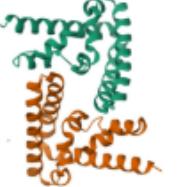
All beta proteins (0)

1 to 20 of 20 structures

Page 1 of 1 | 25

Sort by: Score

Download File | View File

 4ZIA | [pdb_00004zia](#)

Crystal Structure of STAT3 N-terminal domain

Hu, T., Chopra, R.

(2015) Mol Cell Biol 35: 3284-3300

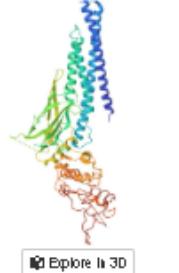
Released: 2015-07-29

Method: X-RAY DIFFRACTION 2.7 Å

Organisms: *Mus musculus*

Macromolecule: Signal transducer and activator of transcription 3 (protein)

Unique Ligands: FMT, MG, NI

 3CWG | [pdb_00003cwg](#)

Unphosphorylated mouse STAT3 core fragment

Ren, Z., Mao, X., Mertens, C., Krishnaraj, R., Qin, J., Mandal, P.K., Romanowski, M.J., McMurray, J.S.

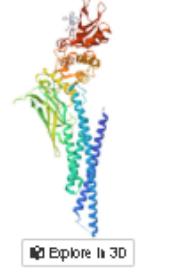
(2008) Biochem Biophys Res Commun 374: 1-5

Released: 2008-07-01

Method: X-RAY DIFFRACTION 3.05 Å

Organisms: *Mus musculus*

Macromolecule: Signal transducer and activator of transcription 3 (protein)

 6NJS | [pdb_00006njs](#)

Stat3 Core in complex with compound SD36

Meagher, J.L., Stuckey, J.A.

(2019) Cancer Cell 36: 498-511.e17

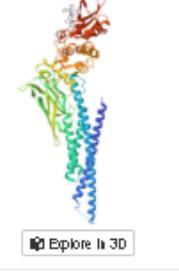
Released: 2019-12-04

Method: X-RAY DIFFRACTION 2.7 Å

Organisms: *Homo sapiens*

Macromolecule: Signal transducer and activator of transcription 3 (protein)

Unique Ligands: KQV

 6NUQ | [pdb_00006nuq](#)

Stat3 Core in complex with compound SI109

Meagher, J.L., Stuckey, J.A.

(2019) Cancer Cell 36: 498-511.e17

Released: 2019-12-04

Method: X-RAY DIFFRACTION 3.15 Å

Organisms: *Homo sapiens*

Macromolecule: Signal transducer and activator of transcription 3 (protein)

Unique Ligands: KQV

 6TLC | [pdb_00006tlc](#)

Unphosphorylated human STAT3 in complex with MS3-6 monobody

La Sala, G., Lau, K., Reynaud, A., Poier, E., Hantschel, O.







238,622
Structures from the
PDB archive

1,068,577
Computed Structure
Models (CSM)

Enter search term(s), Ligand ID or sequence

Include CSM



Help

[Advanced Search](#) | [Browse Annotations](#)

PDB-101

PDB

EMDataResource

NAKB

wwPDB
Foundation

PDB-IHM



Structure Summary

Structure

Annotations

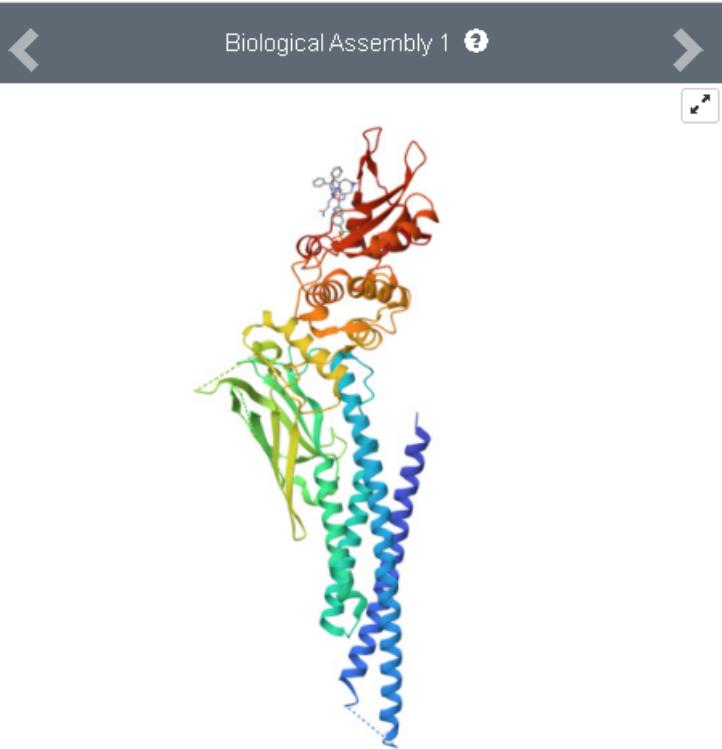
Experiment

Sequence

Genome

Ligands

Versions



Global Symmetry: Asymmetric - C1
Global Stoichiometry: Monomer - A1

[Find Similar Assemblies](#)

6NUQ | [pdb_00006nuq](#)

Stat3 Core in complex with compound SI109

PDB DOI: <https://doi.org/10.2210/pdb6NUQ/pdb>

Classification: [transcription/transcription inhibitor](#)

Organism(s): [Homo sapiens](#)

Expression System: [Escherichia coli](#)

Mutation(s): No

Deposited: 2019-02-01 Released: 2019-12-04

Deposition Author(s): [Meagher, J.L., Stuckey, J.A.](#)

Funding Organization(s): National Institutes of Health/National Human Genome Research Institute (NIH/NHGRI)

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 3.15 Å

R-Value Free:

0.260 (Depositor), 0.260 (DCC)

R-Value Work:

0.233 (Depositor), 0.240 (DCC)

R-Value Observed:

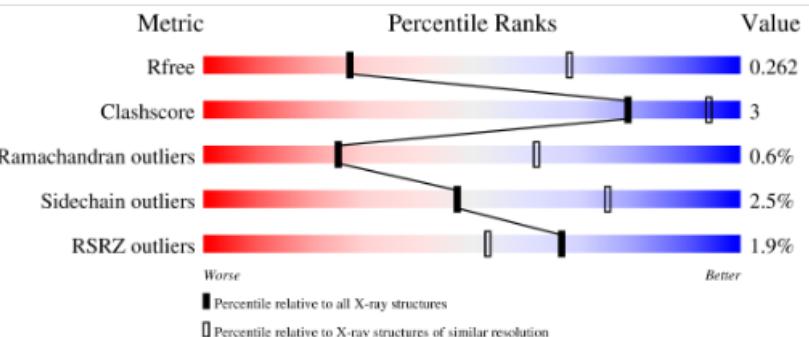
0.234 (Depositor)

Starting Model: experimental

[View more details](#)

wwPDB Validation

[3D Report](#) [Full Report](#)



Ligand Structure Quality Assessment

Molecular docking

- Prior to execution, the preparation of molecules is very important to ensure a realistic and physiologically relevant interaction between the receptor and ligand.
- We prepare the receptor by:
 - Removing water molecules
 - Removing ligands



MOLECULAR DOCKING

With AutoDock Vina



Molecular docking with AutoDock Vina

Software:

- Perl
- OpenBabel GUI
- MGL Tools
- Autodock Vina

File:

- Receptor.pdbqt
- Ligand.pdbqt
- Configuration file

Pearl, download from: <https://padre.perlide.org/>

MGL Tools, download from: <https://ccsb.scripps.edu/mgltools/downloads/>

AutoDock Vina, download from: <https://vina.scripps.edu/downloads/>

OpenBabel GUI, download from: <https://github.com/openbabel/openbabel/releases>



Open Babel GUI

Use it to convert ligand files from the .sdf to the .pdbqt format.

We already:

- Find ligands in PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>)
 - under „3D Conformer“ (Download coordinates → Save SDF)

We need to:

- Convert saved files with OpenBabel GUI (ligand.sdf → ligand.pdbqt)



1.2 3D Conformer



Structure Search

DOWNLOAD COORDINATES

Interactive Chemical Structure Model

Ball and Stick

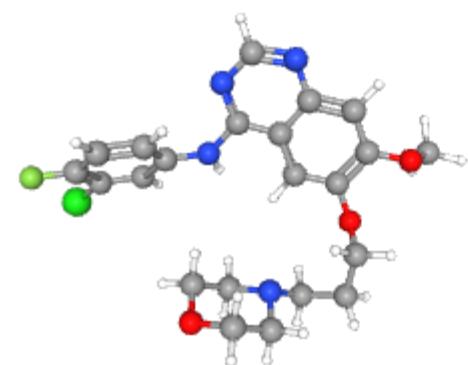
Sticks

Wire-Frame

Space-Filling

Show Hydrogens

Animate



SDF

JSON

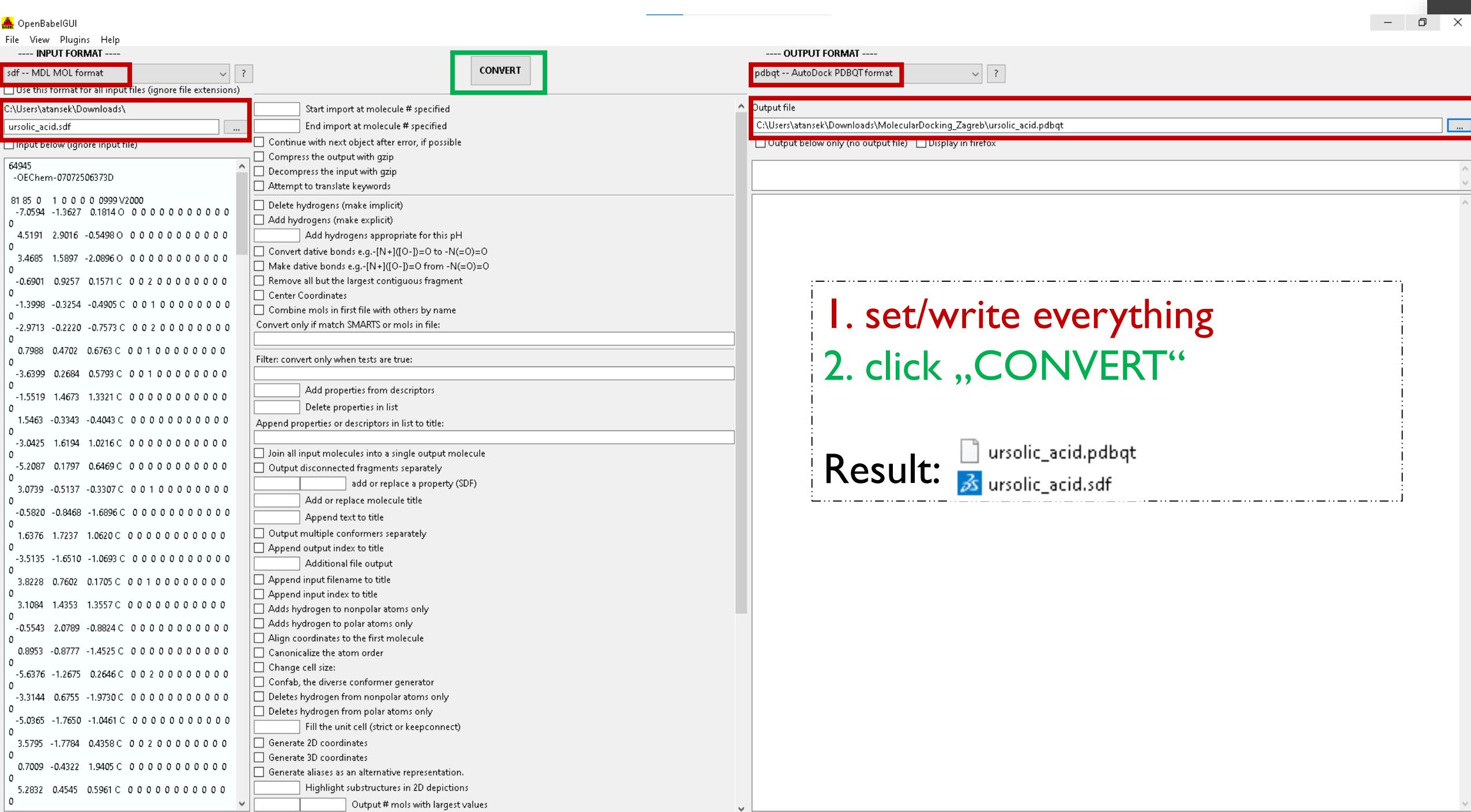
XML

ASNT



« First < Previous Conformer 1 of 10 Next > Last »





1. set/write everything
2. click „CONVERT“

Result:

MGL Tools or AutoDock Tools

With AutoDock Tools (MGL Tools), we prepare the receptor molecule for docking with AutoDock Vina.

- Input the receptor molecule: File → Read molecule (select 6nuq.pdb)
- Delete unnecessary elements from the PDB structure:
 - In case of two chains (A, B); delete all, except A (select chain B, in column „S“ – yellow circles): Edit → Delete → Delete Selected Atoms
 - Delete water molecules (HOH): Edit → Delete water
 - Delete other heteroatoms that are not a part of the receptor (select unwanted atoms in column „S“ – yellow circles): Edit → Delete → Delete Selected Atoms
- Check for missing atoms in the structure and add them:
 - Edit → Misc → Check for Missing Atoms (Select All Residues, Dismiss)
 - Edit → Misc → Repair Missing Atoms (Save as 2 sets, Dismiss)
- UNSELECT any selected elements in the structure (column „S“ – yellow circles)
- Change hydrogens on histidine to polar hydrogens:
 - Edit → Hydrogens → Edit Histidine Hydrogens (select second column for all hydrogens: »0, HE2«, Apply, Dismiss)
 - Edit → Hydrogens → Add (Polar Only, noBondOrder (for pdb files...), yes; OK)
- Add Kollman charges (only for proteins/receptors) and check if they are equally distributed:
 - Edit → Charges → Add Kollman Charges
 - Edit → Charges → Check Totals on Residues (Spread Charge Deficit over all atoms in residue, Dismiss)
 - Repeat: Edit → Charges → Check Totals on Residues (you should get the notification: no residues with non-integral charges found)
- Save receptor in file format receptor.pdbqt:
 - Grid → Macromolecule → Choose (click on 6nuq, Select Molecule, Dismiss – notification appears: initializing 6nuq.pdb:-contains no non-bonded atoms; OK) → save as »6nuq.pdbqt« file



6NUQ

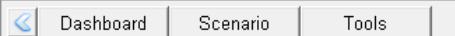
Molecular Modeling

The crystal structure of STAT3 (pY705) with DNA (PDBID: 1BG1([Becker et al., 1998](#))) was used to model the binding pose of AA-115 with STAT3. The chain A of STAT3 (pY705) from the crystal structure was extracted and the protons on the protein were added using the “protonate 3D” module in MOE ([CCG, 2018](#)) by setting the physiological pH value at 7.0. AA-115 was drawn and optimized using MOE for the docking calculation using the GOLD program (version 5.2) ([Jones et al., 1997](#)). In the docking calculation, the center of the binding site in STAT3 was set at V637 and the radius of the binding site was defined as 13 Å. A maximum number of 200,000 operations were performed on a population of 5 islands of 100 individuals in each genetic algorithm run. Operator weights for cross-over, mutation and migration were set to 95, 95 and 10 respectively. GoldScore was used as the fitness function to rank the docked poses. The top ranked pose was considered as the binding model to suggest the location of the 8-membered ring suitable for coupling to the cereblon ligand in the STAT3 degrader design.

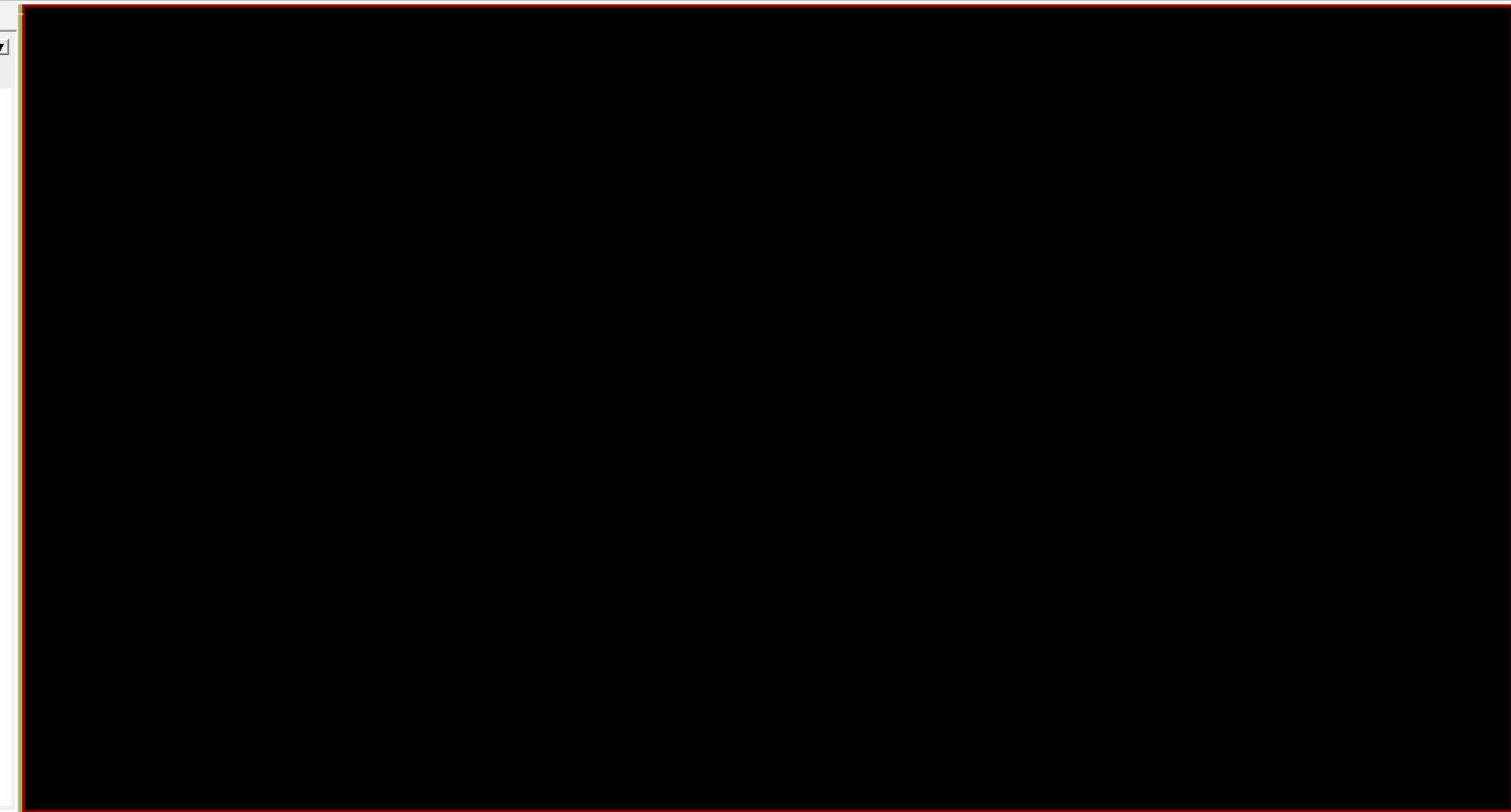




ADT4.2 Ligand Flexible Residues Grid Docking Run Analy



The image shows the VMD (Visual Molecular Dynamics) software interface. At the top, there is a toolbar with various icons for selection, transformation, and visualization. Below the toolbar, the text "Sel.: All Molecules Current Selection" is displayed. The main window shows a 3D molecular model of a protein, with atoms represented by spheres and bonds by lines. The protein structure is shown in a semi-transparent blue color.



Mod.: None Time: 1.027 Selected: 0 Atom(s)

Done 100%

Spin off FR: 333.3



- Read Molecule
- Read Session
- Import
- Recent Files
- Save
- Order Physical M...

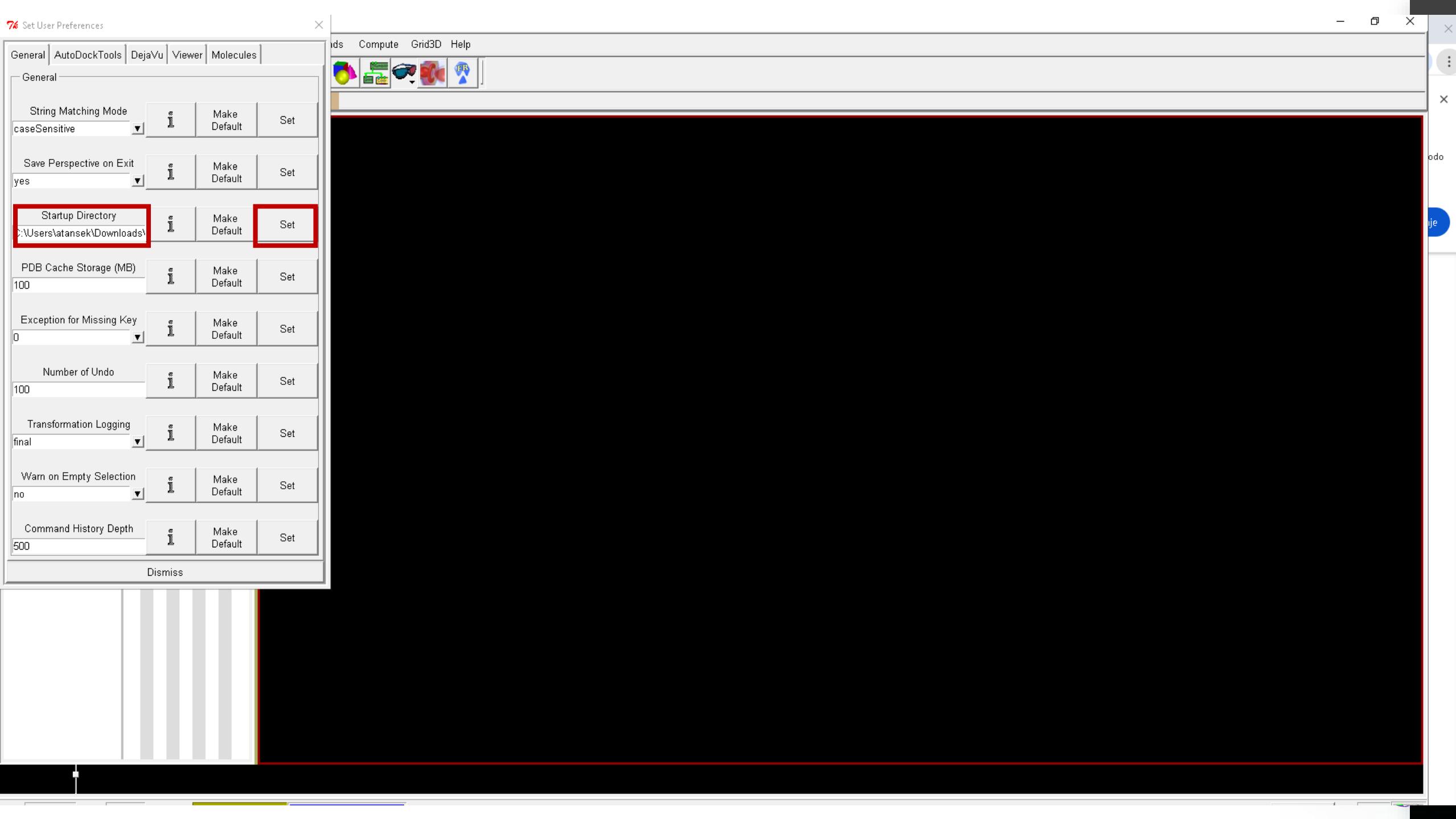
Browse Commands
Load Macros

Preferences

A screenshot of a Windows-style context menu. The menu items are: 'server', 'Exit', and a separator line followed by 'Set...', 'Set Commands to be Applied on Obj...', 'Change Font', 'Show/Hide GUI Sections', 'Bind Action to Mouse Button ...', and 'Hide VF GUI...'. The 'Set...' option is highlighted with a blue selection bar.

Set...

- Set Commands to be Applied on Object
- Change Font
- Show/Hide GUI Sections
- Bind Action to Mouse Button ...
- Hide VF GUI...



- Read Molecule
- Read Session
- Import
- Recent Files
- Save
- Order Physical Models
- Browse Commands
- Load Macros
- Preferences
- server
- Exit



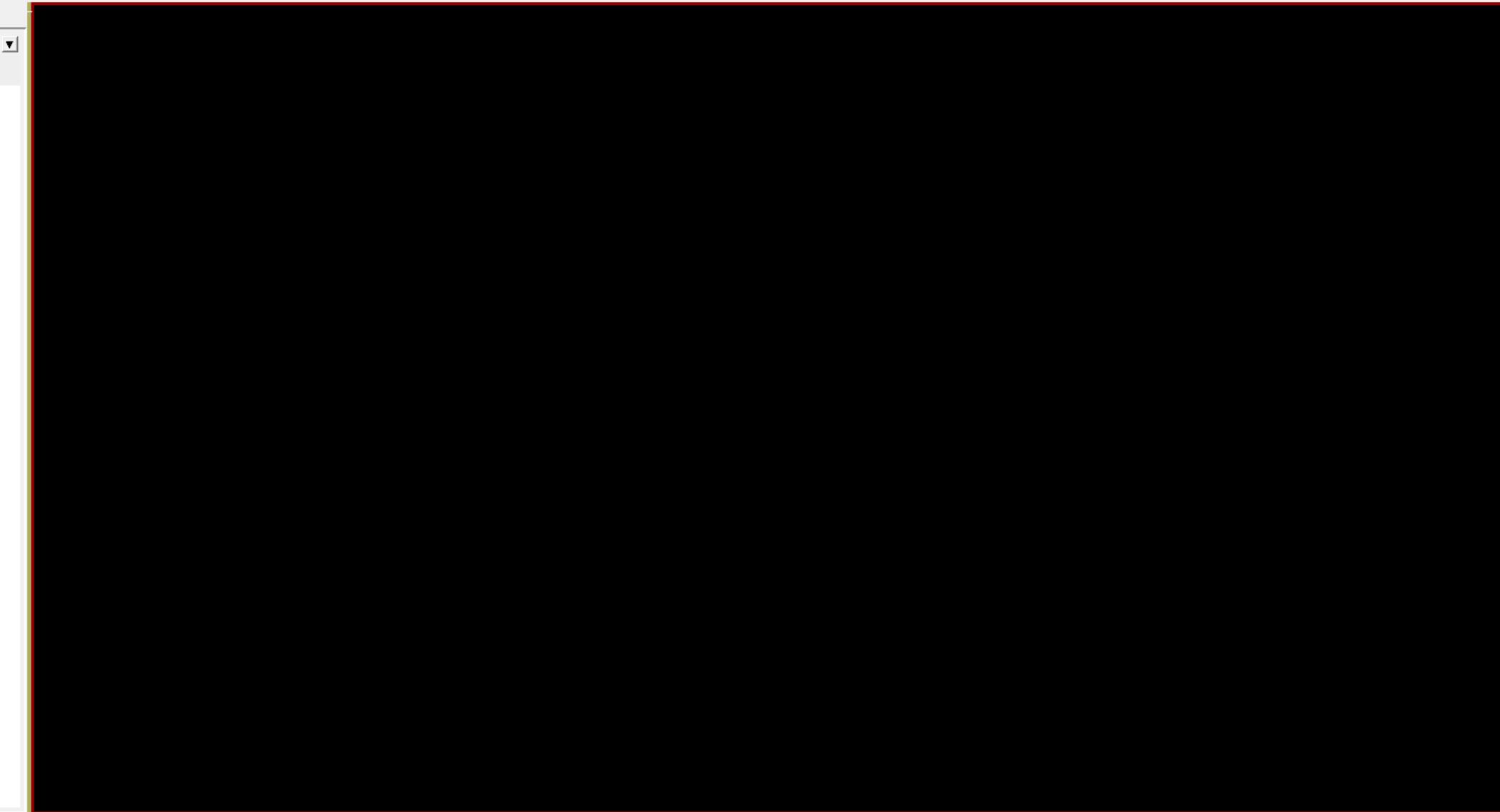
Residues Grid Docking Run Analyze

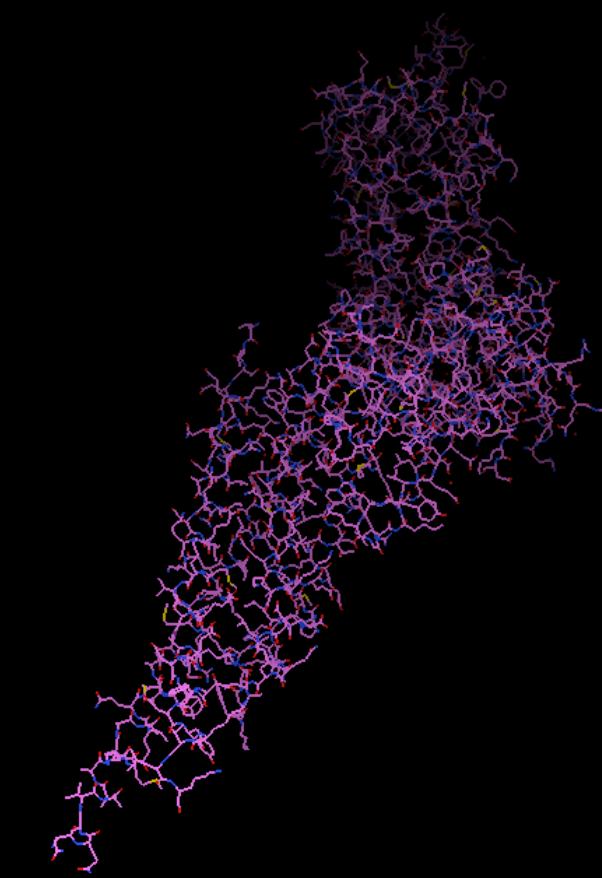
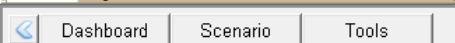
io Tools

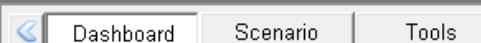
CMD

S L B C RMS L CI OH

○ ○ ○ ○ ▽ ▽ ▽ ▽

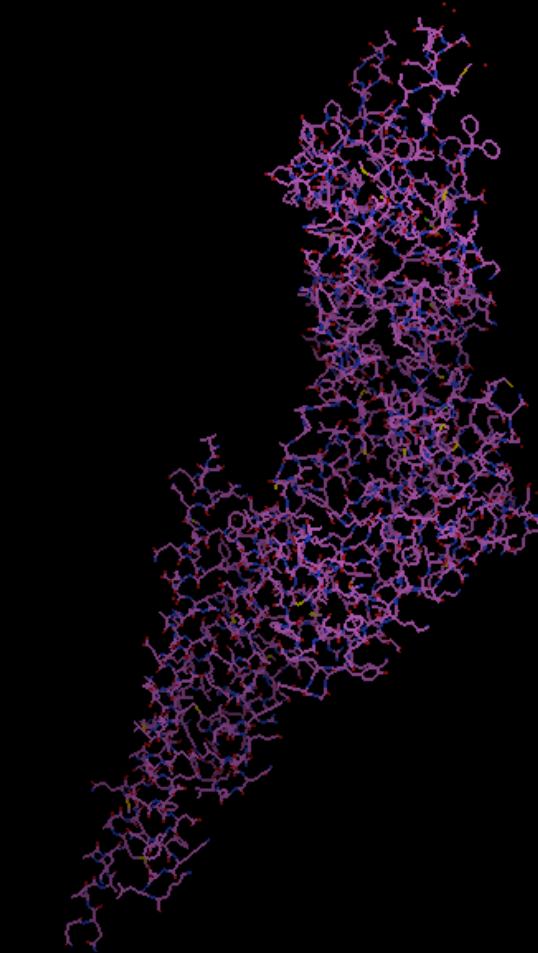


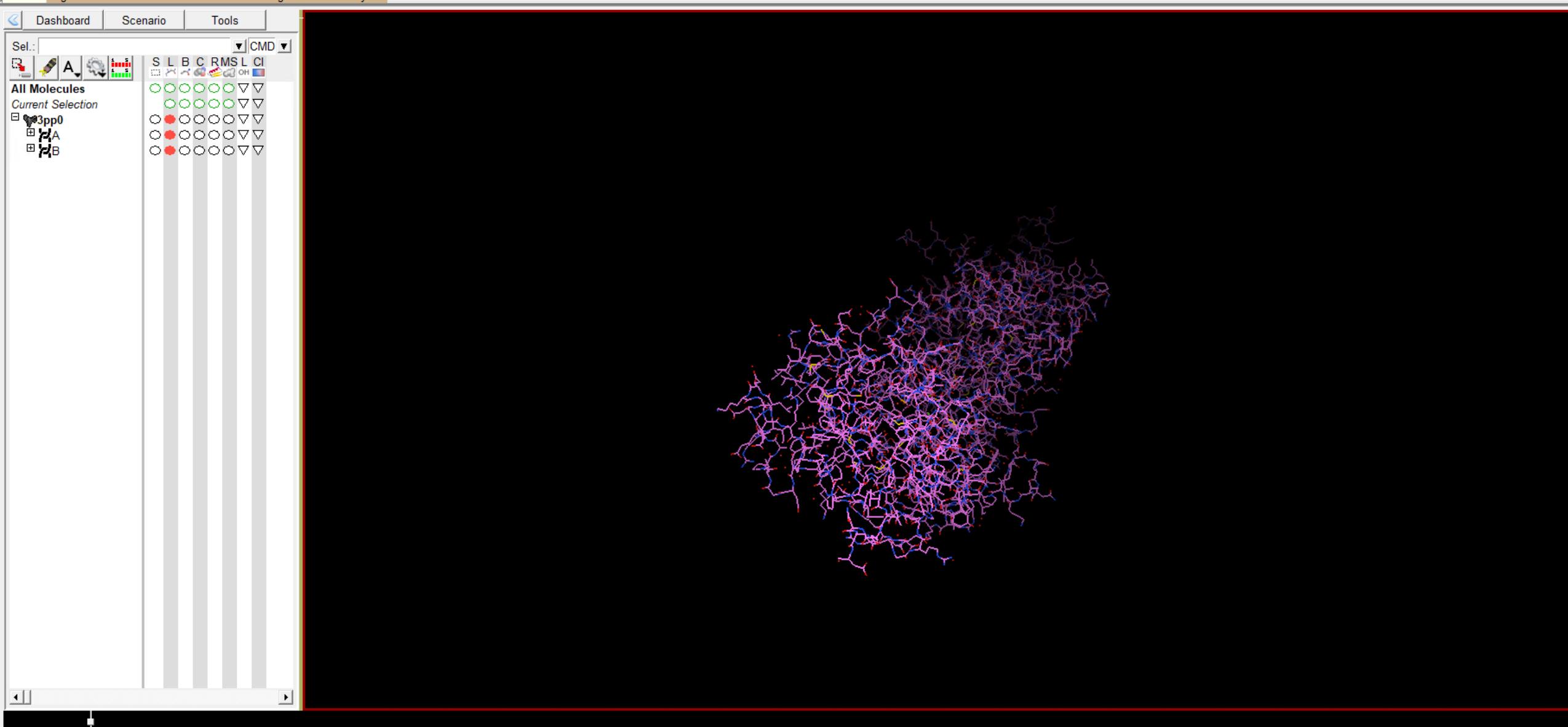




Basics of moving

- Scrolling = zoom in/out
 - Left click = rotation
 - Right click = move





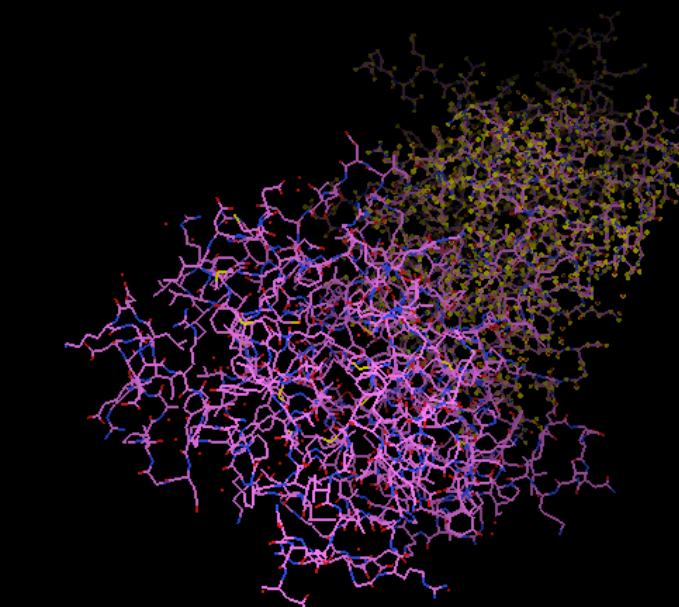


Sel.: CMD

All Molecules

Current Selection

- 3pp0
 - A
 - B



76 AutoDockTools

File 3D Graphics Edit Select Display Color Hydrogen Bonds

ADT4.2 Ligand Flexible Residues Grid Docking Run Analyze

Dashboard

Sel.: CMD

All Molecules

Current Selection

- 3pp0
 - A
 - B

Redo

Undo select

Bonds

Delete

- Atoms
- Delete Water
- Charges
- Hydrogens

Misc

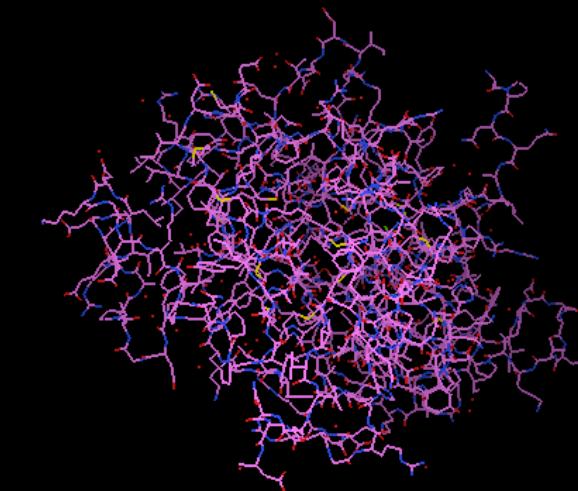
Color Palettes

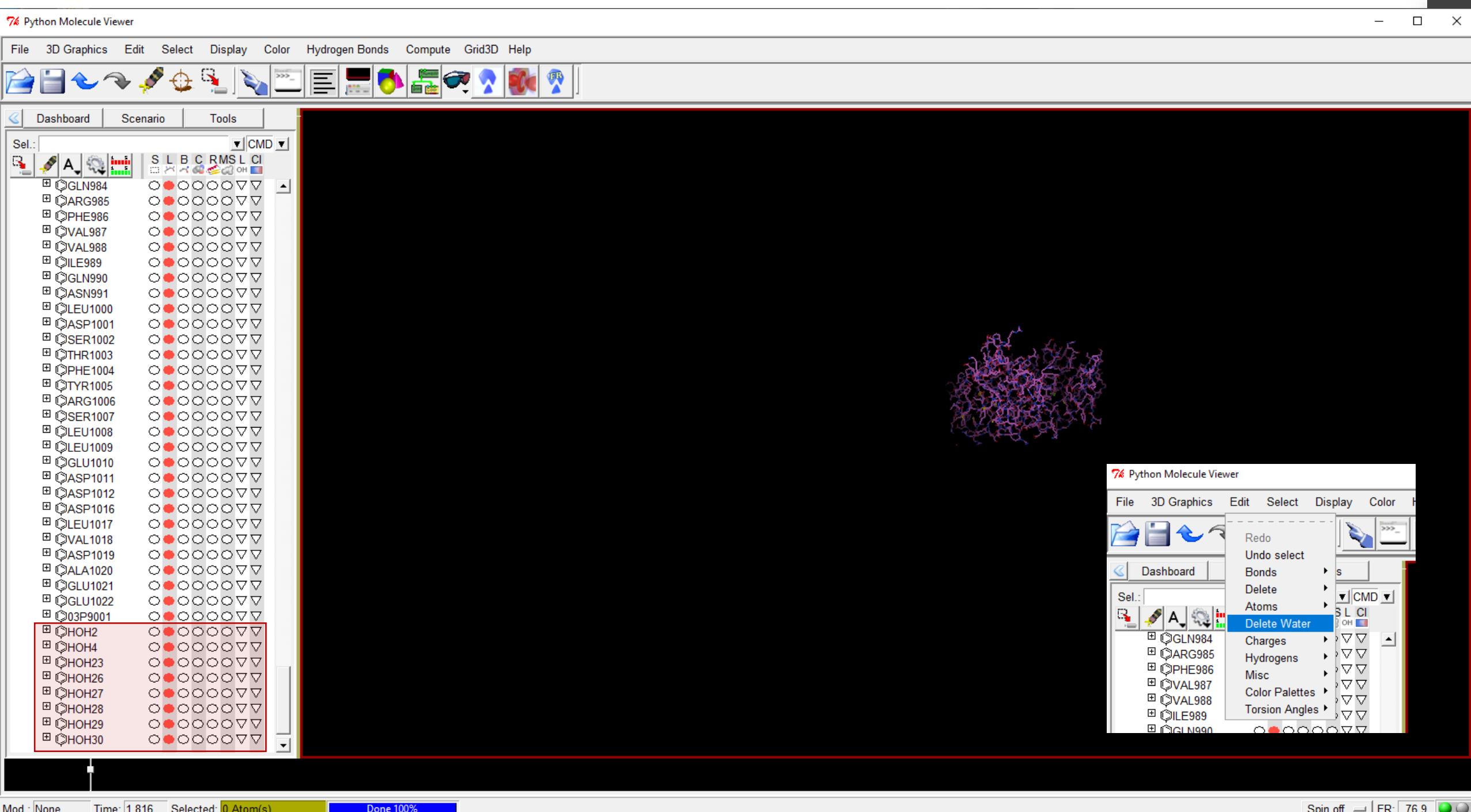
Torsion Angles

Delete Selected Atoms

Sel.: CMD

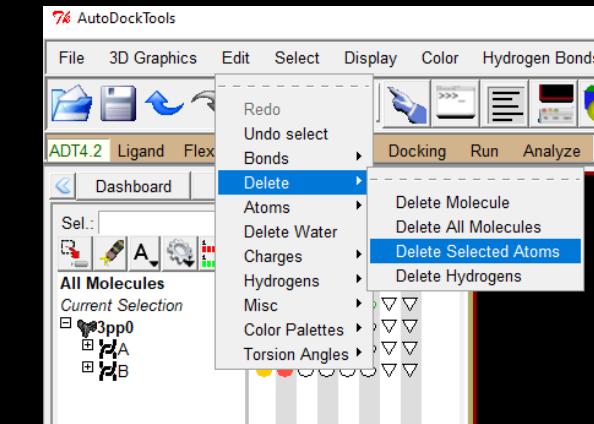
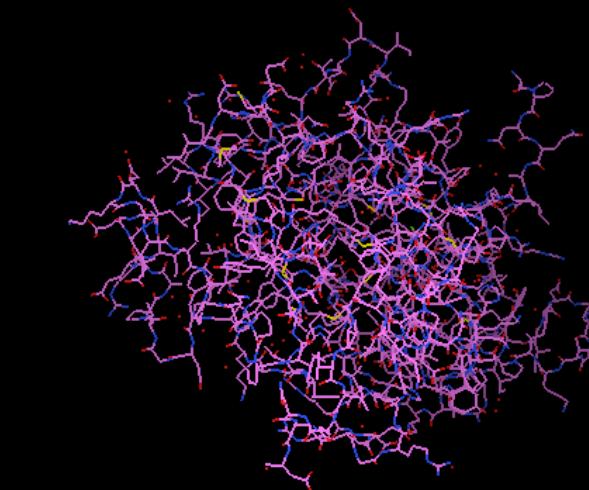
	S	L	B	C	RMS	L	Cl
All Molecules	○	○	○	○	○	○	○
Current Selection	○	○	○	○	○	○	○
012pp0	○	●	○	○	○	○	○
A	○	●	○	○	○	○	○

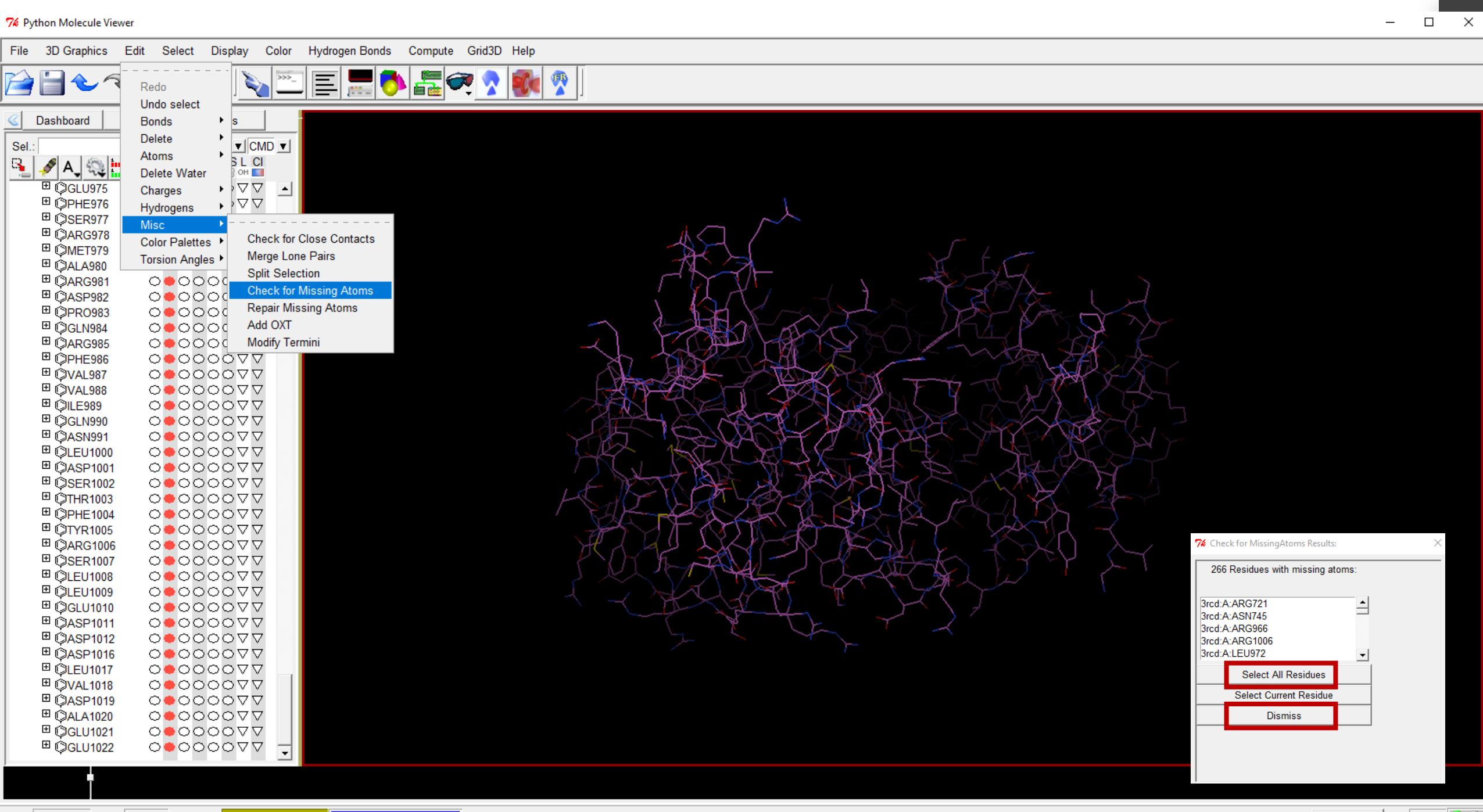


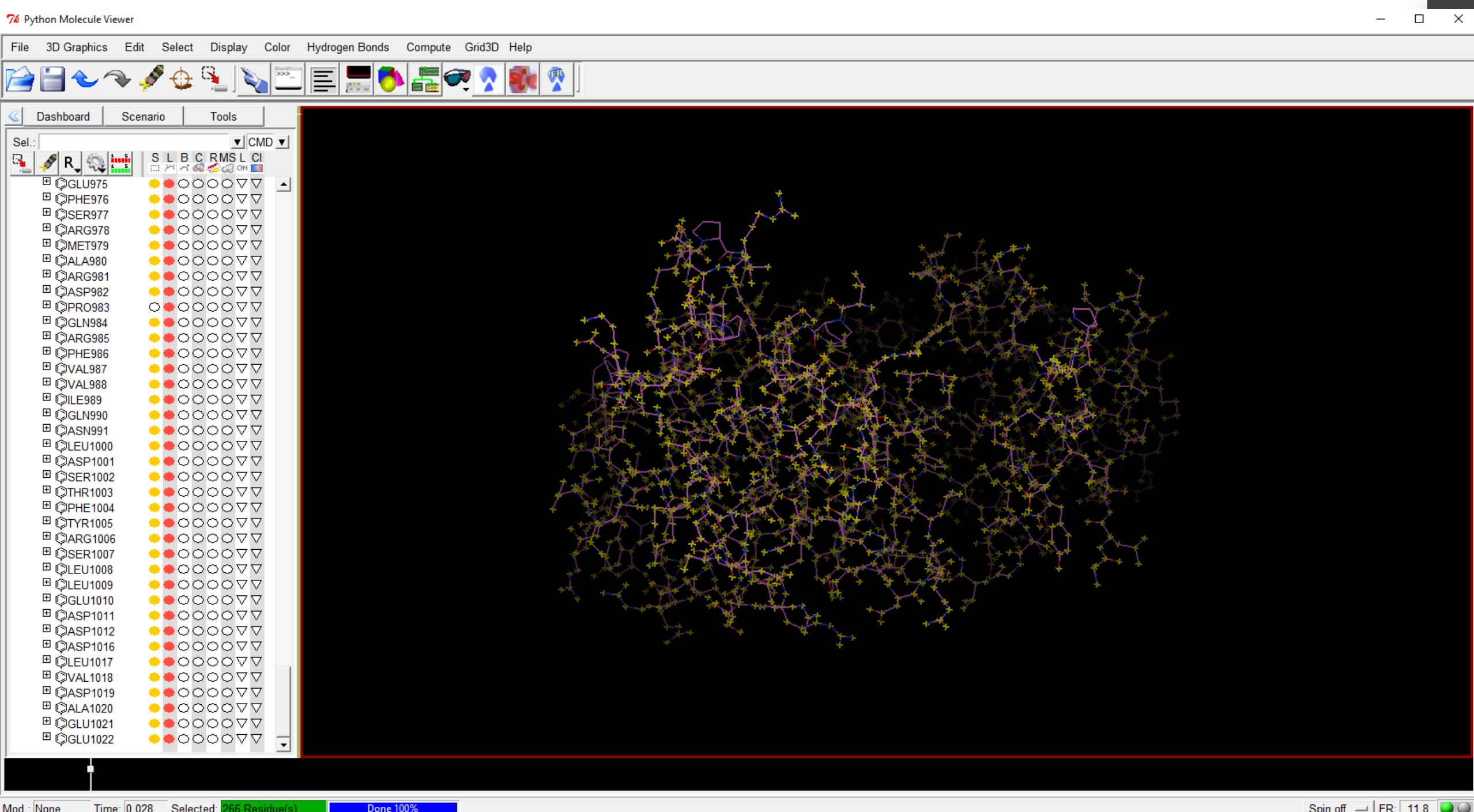


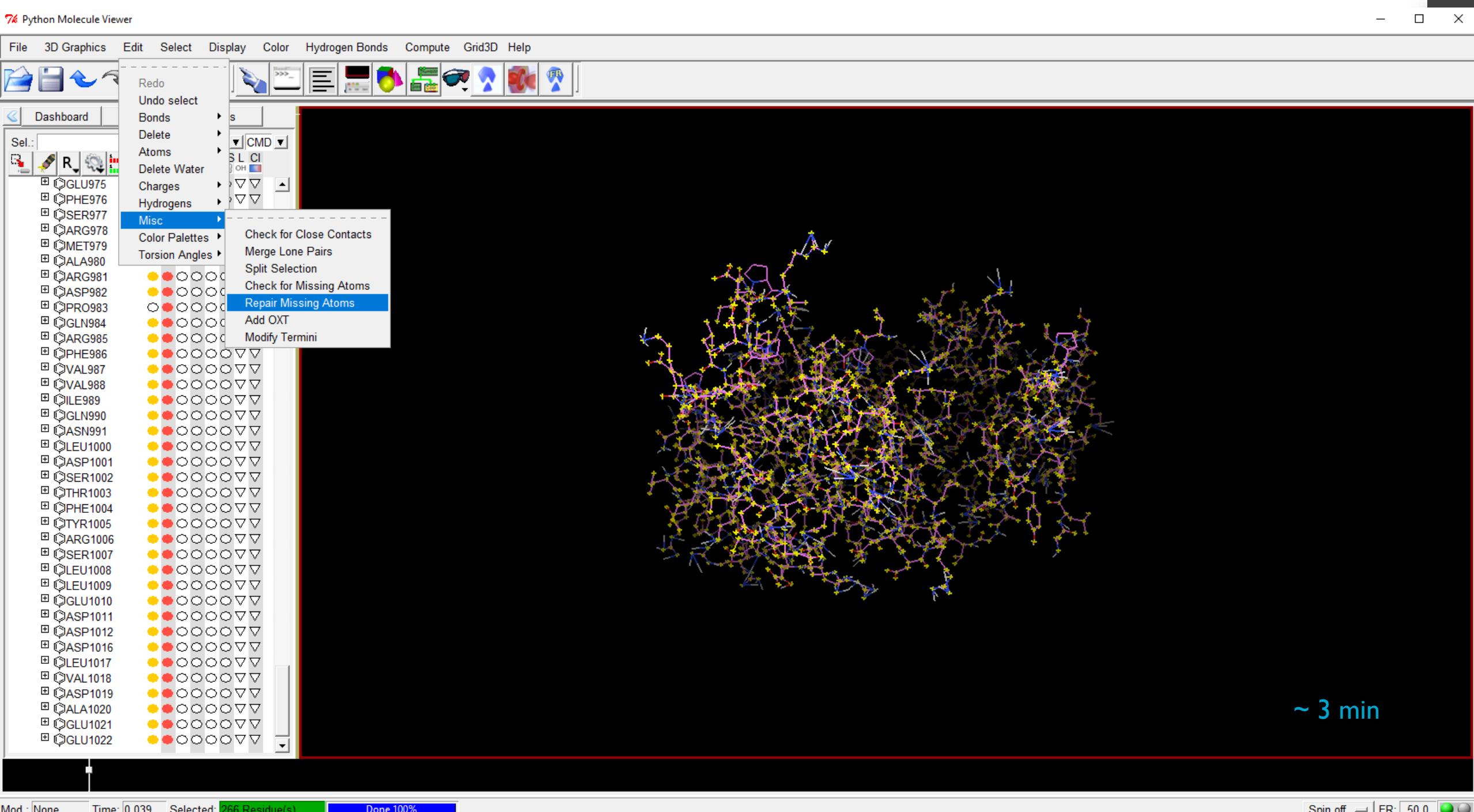
Sel.: A B C RMSL CI OH FBR

	S	L	B	C	RMSL	CI	OH	FBR
QMET955	○	●	○	○	○	○	○	○
QVAL956	○	●	○	○	○	○	○	○
QLYS957	○	●	○	○	○	○	○	○
QCYS958	○	●	○	○	○	○	○	○
QTRP959	○	●	○	○	○	○	○	○
QMET960	○	●	○	○	○	○	○	○
QILE961	○	●	○	○	○	○	○	○
QASP962	○	●	○	○	○	○	○	○
QSER963	○	●	○	○	○	○	○	○
QGLU964	○	●	○	○	○	○	○	○
QCYS965	○	●	○	○	○	○	○	○
QARG966	○	●	○	○	○	○	○	○
QPRO967	○	●	○	○	○	○	○	○
QARG968	○	●	○	○	○	○	○	○
QPHE969	○	●	○	○	○	○	○	○
QARG970	○	●	○	○	○	○	○	○
QGLU971	○	●	○	○	○	○	○	○
QLEU972	○	●	○	○	○	○	○	○
QVAL973	○	●	○	○	○	○	○	○
QSER974	○	●	○	○	○	○	○	○
QGLU975	○	●	○	○	○	○	○	○
QPHE976	○	●	○	○	○	○	○	○
QSER977	○	●	○	○	○	○	○	○
QARG978	○	●	○	○	○	○	○	○
QMET979	○	●	○	○	○	○	○	○
QALA980	○	●	○	○	○	○	○	○
QARG981	○	●	○	○	○	○	○	○
QASP982	○	●	○	○	○	○	○	○
QPRO983	○	●	○	○	○	○	○	○
QGLN984	○	●	○	○	○	○	○	○
QARG985	○	●	○	○	○	○	○	○
QPHE986	○	●	○	○	○	○	○	○
QVAL987	○	●	○	○	○	○	○	○
QVAL988	○	●	○	○	○	○	○	○
QILE989	○	●	○	○	○	○	○	○
QGLN990	○	●	○	○	○	○	○	○
QASN991	○	●	○	○	○	○	○	○
QGLU992	○	●	○	○	○	○	○	○
QASP993	○	●	○	○	○	○	○	○
Q03Q1	○	●	○	○	○	○	○	○











ADT4.2 Ligand Flex

Dashboard

Sel.:

All Molecules

Current Selection

 3pp0 3ppA

Redo

Undo select

Bonds

Delete

Atoms

Delete Water

Charges

Hydrogens

Misc

Color Palettes

Torsion Angles

Docking

Run

Analyze

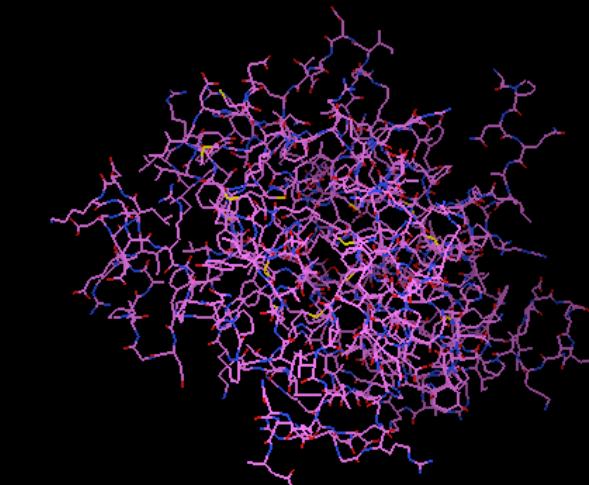
s

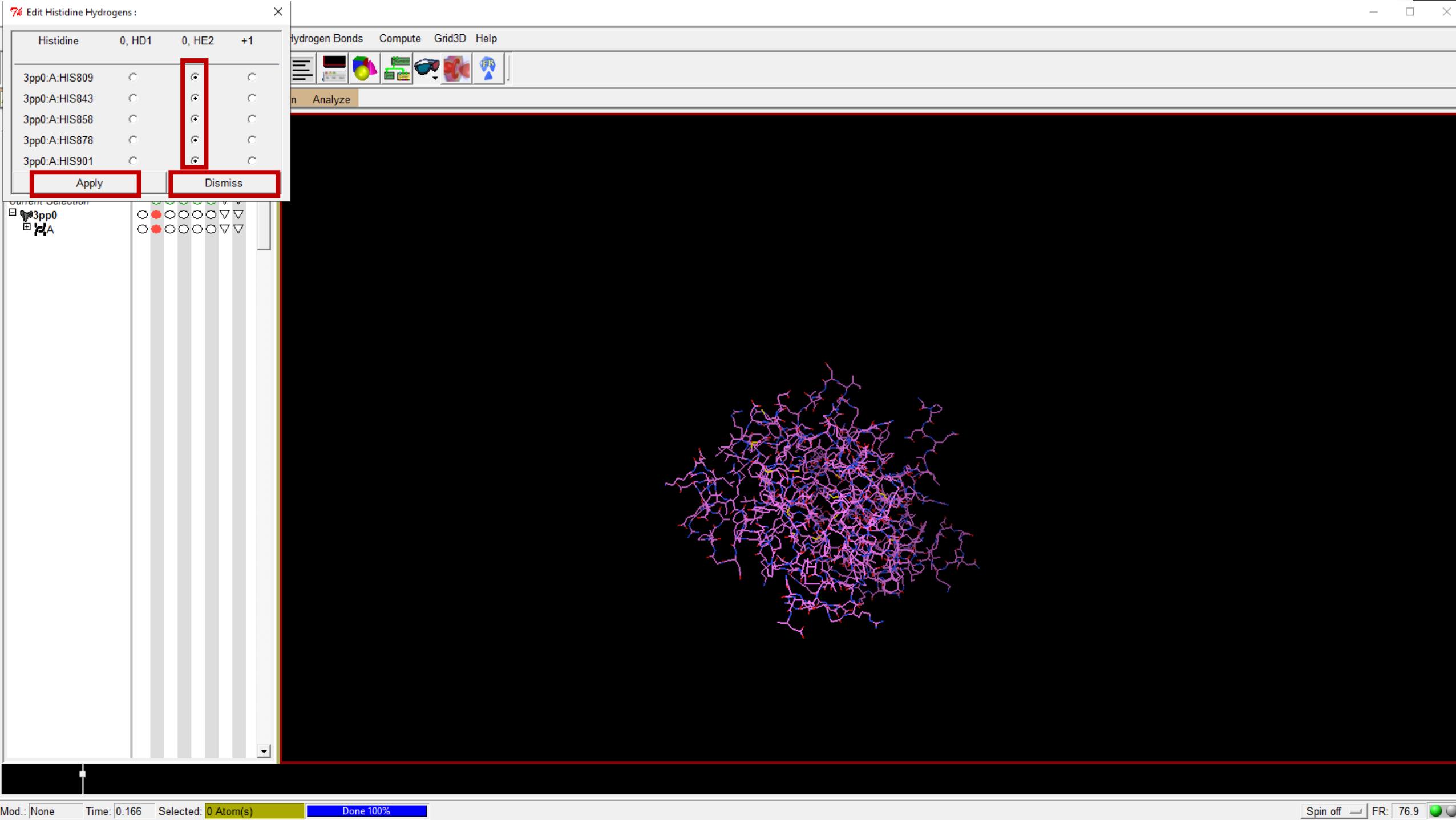
CMD

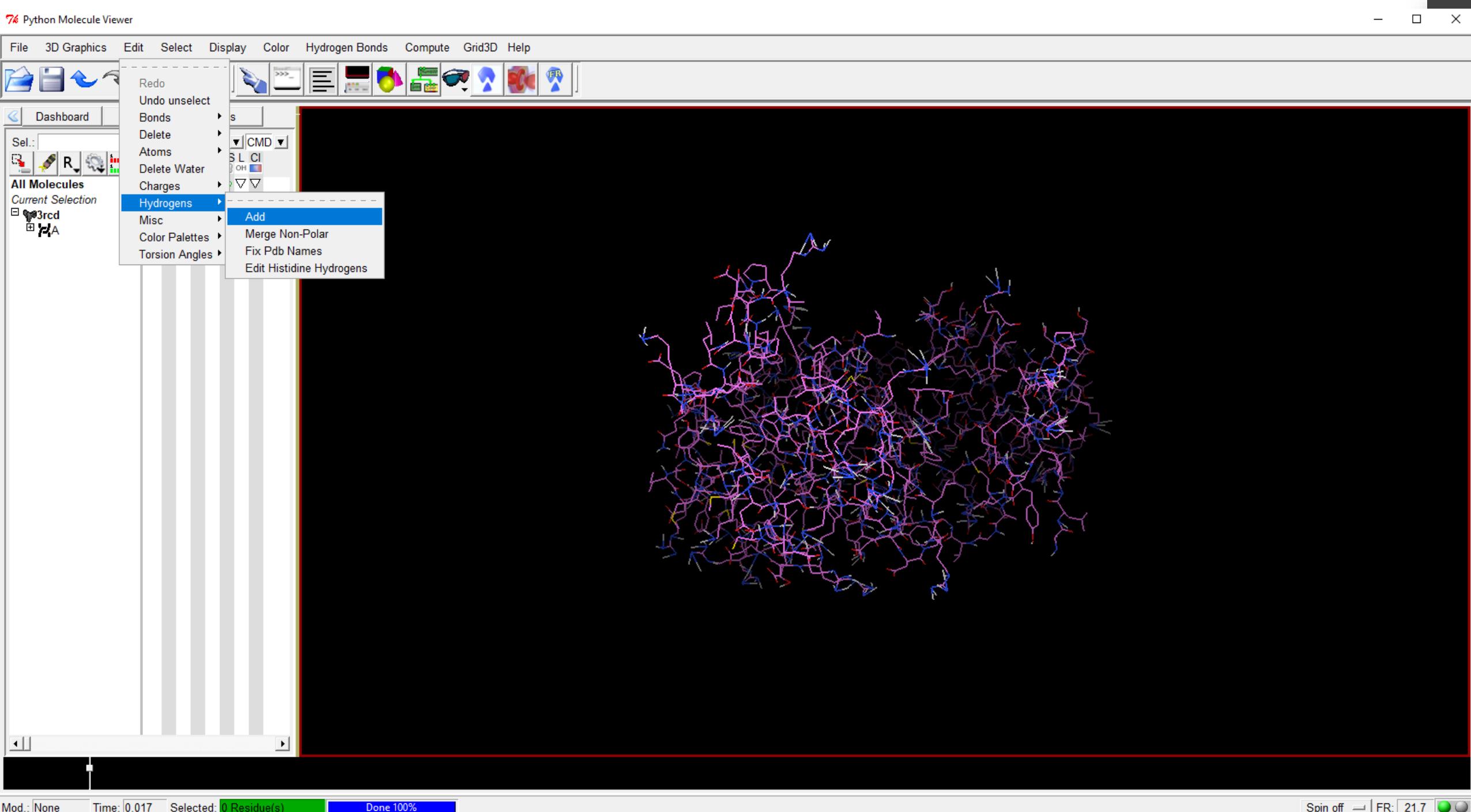
SL CI

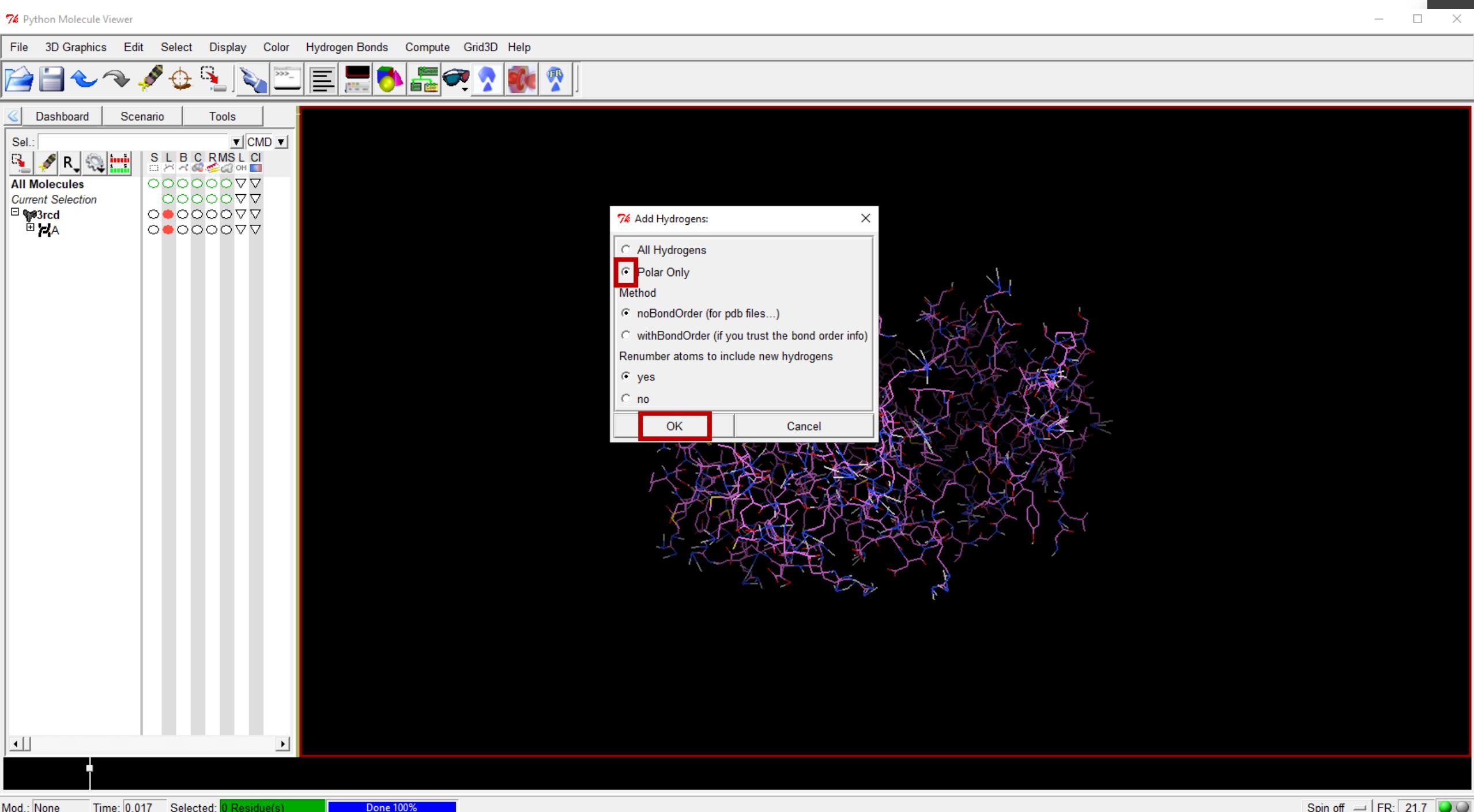
OH

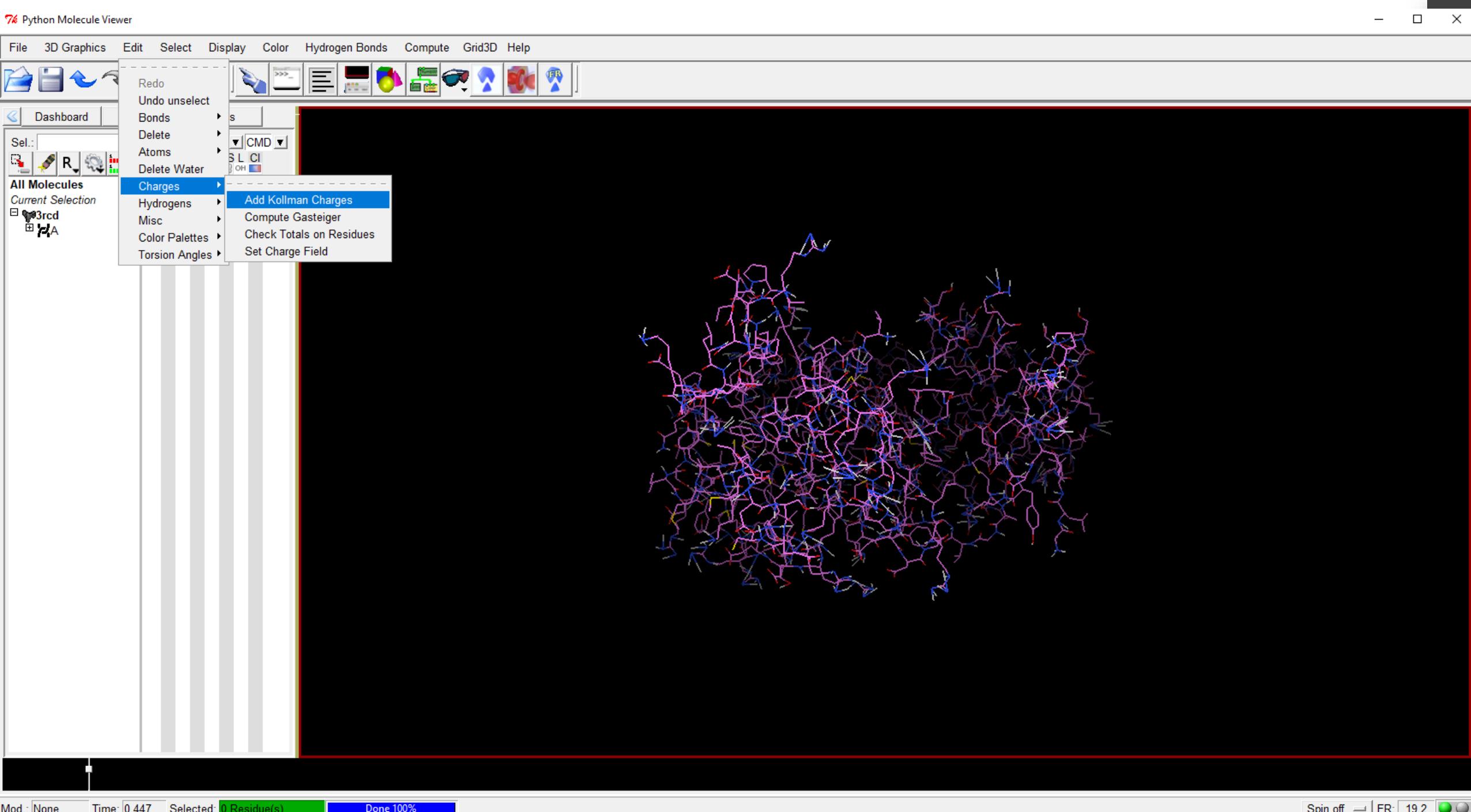
Edit Histidine Hydrogens









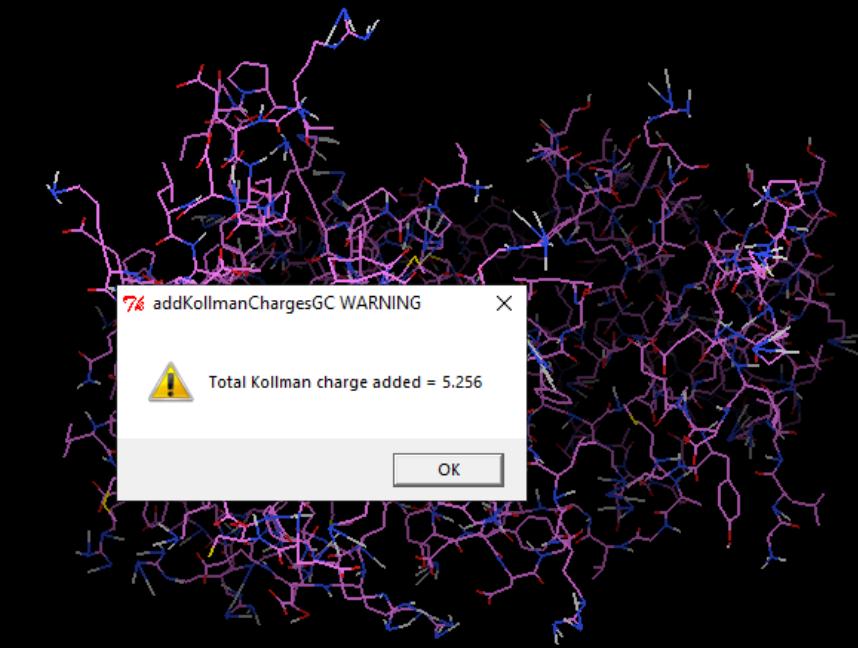




Dashboard Scenario Tools

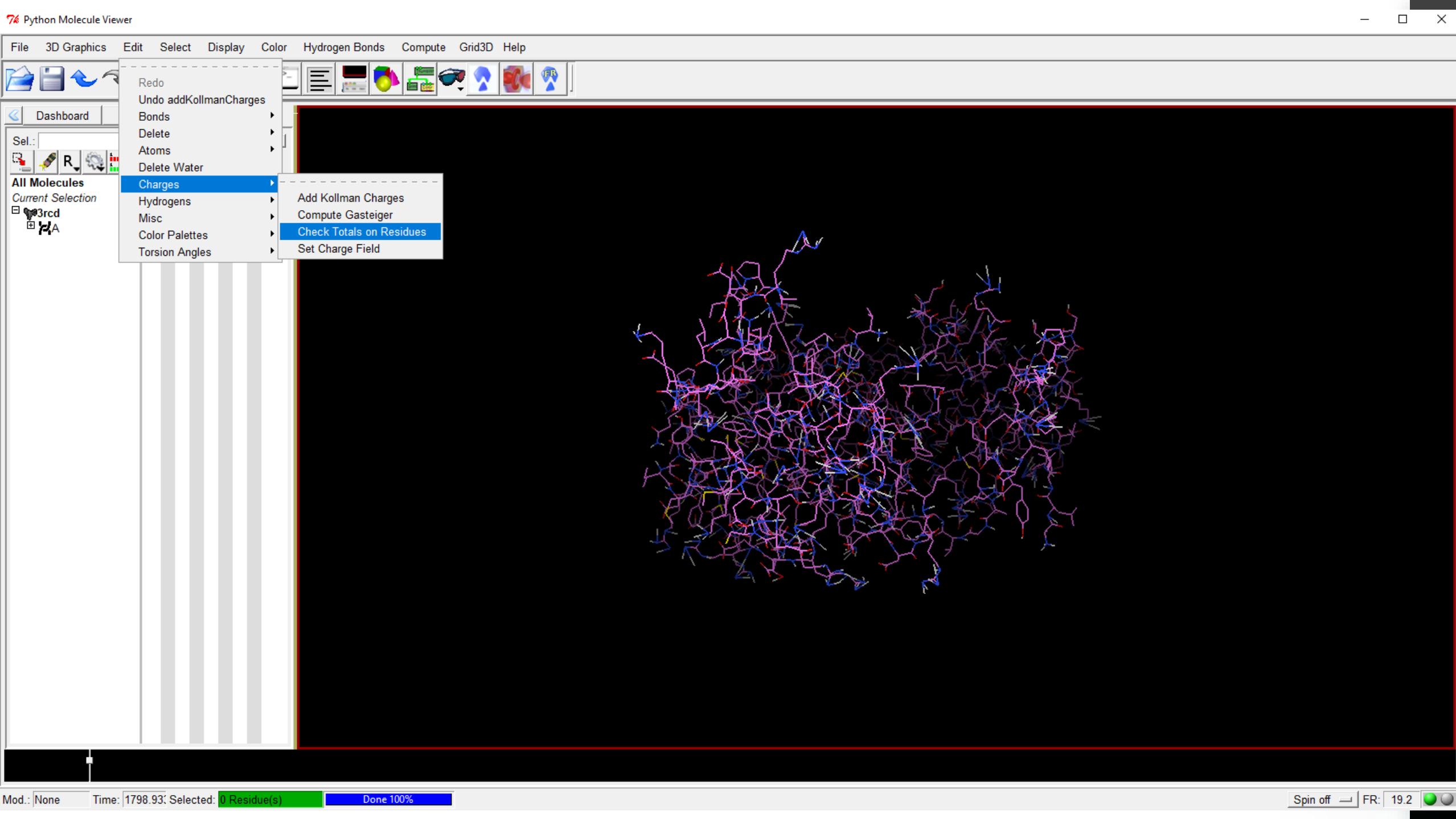
Sel.: CMD

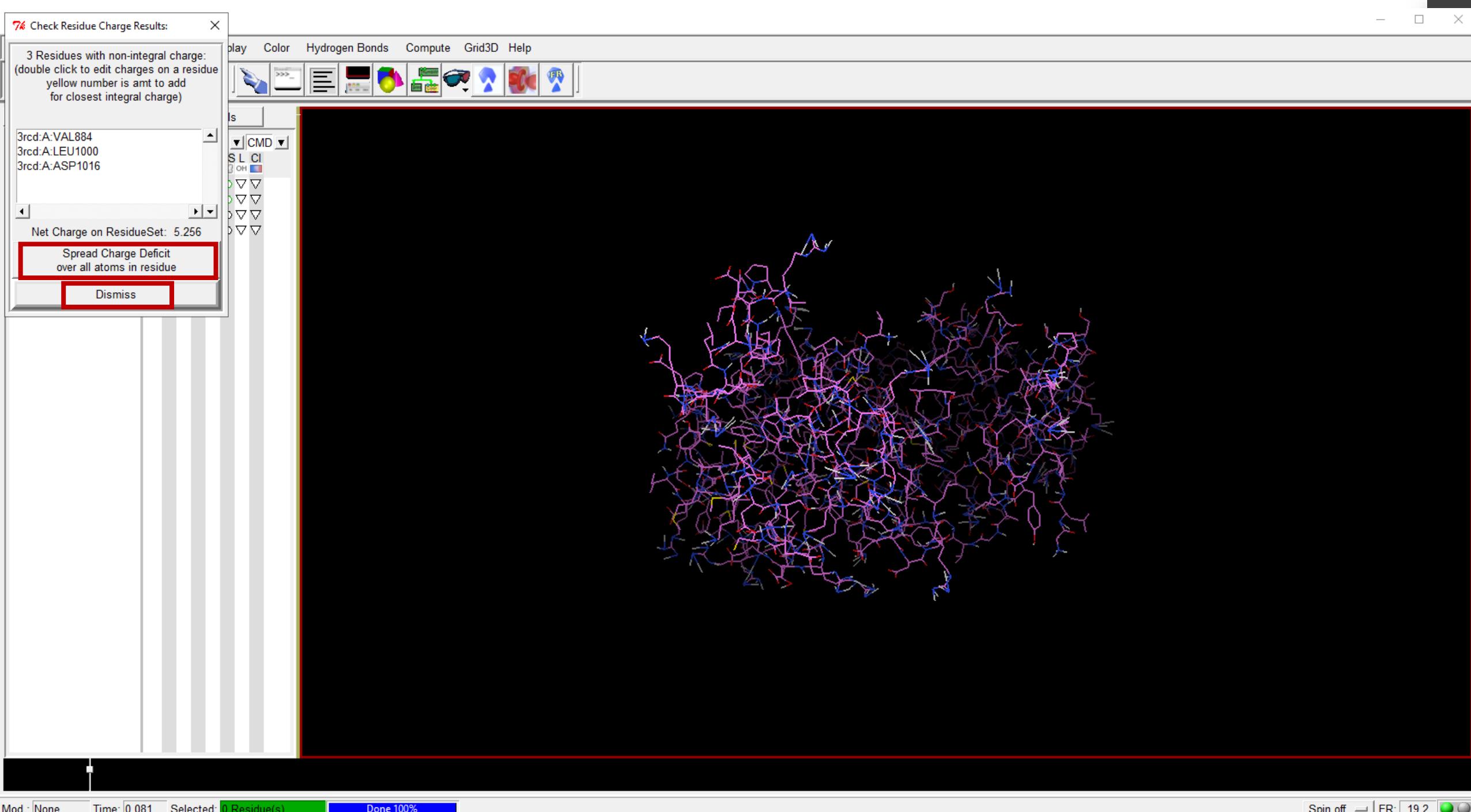
The image shows the VMD (Visual Molecular Dynamics) software interface. At the top, there is a toolbar with various icons for file operations, selection, and visualization. Below the toolbar, a list of molecules is displayed, with "All Molecules" and "Current Selection" highlighted. A selection tool panel on the right shows a 3D coordinate system with spheres and a color bar for selection ranges. The main window shows a 3D molecular structure with atoms represented by spheres and bonds by lines.

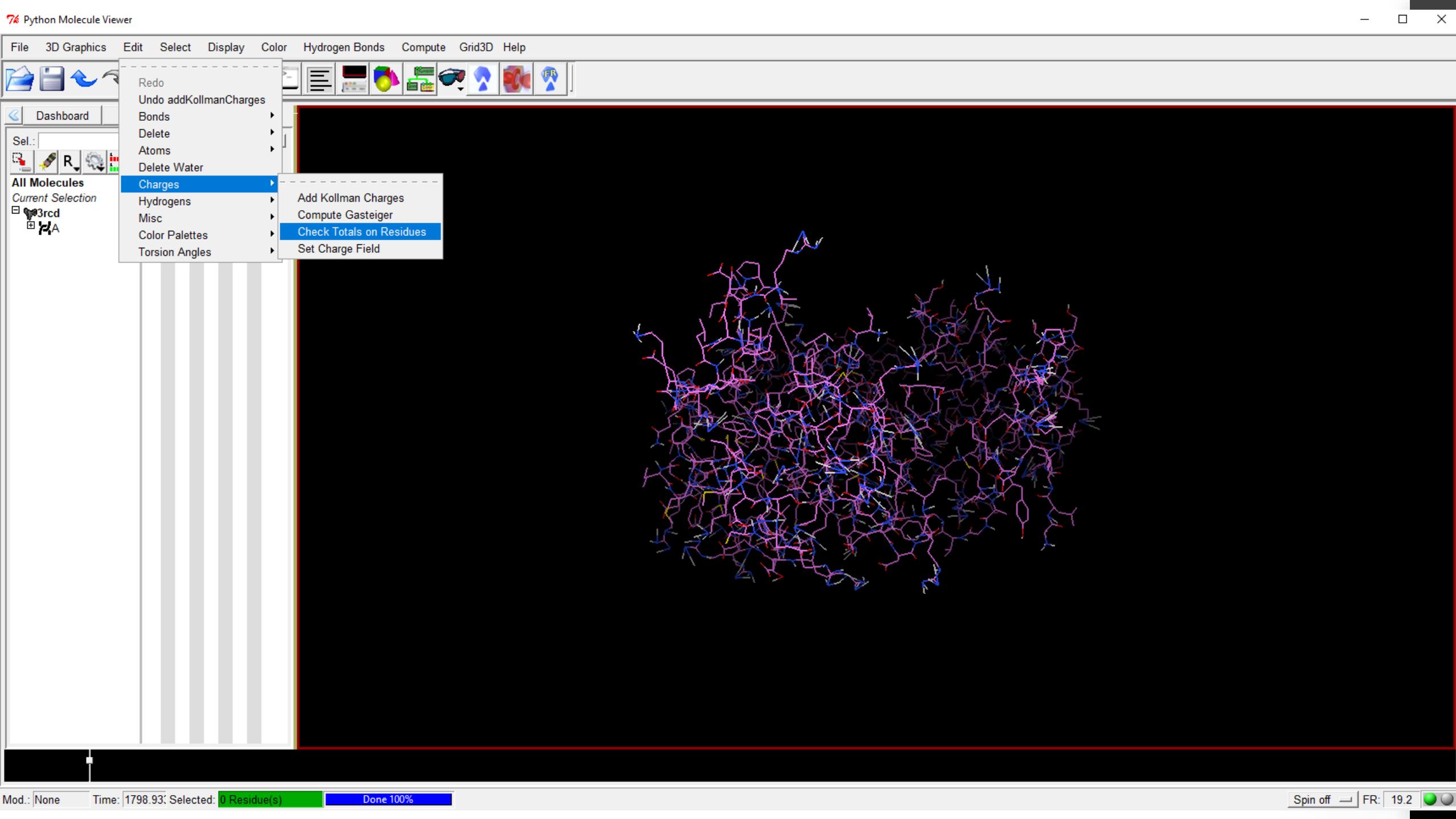


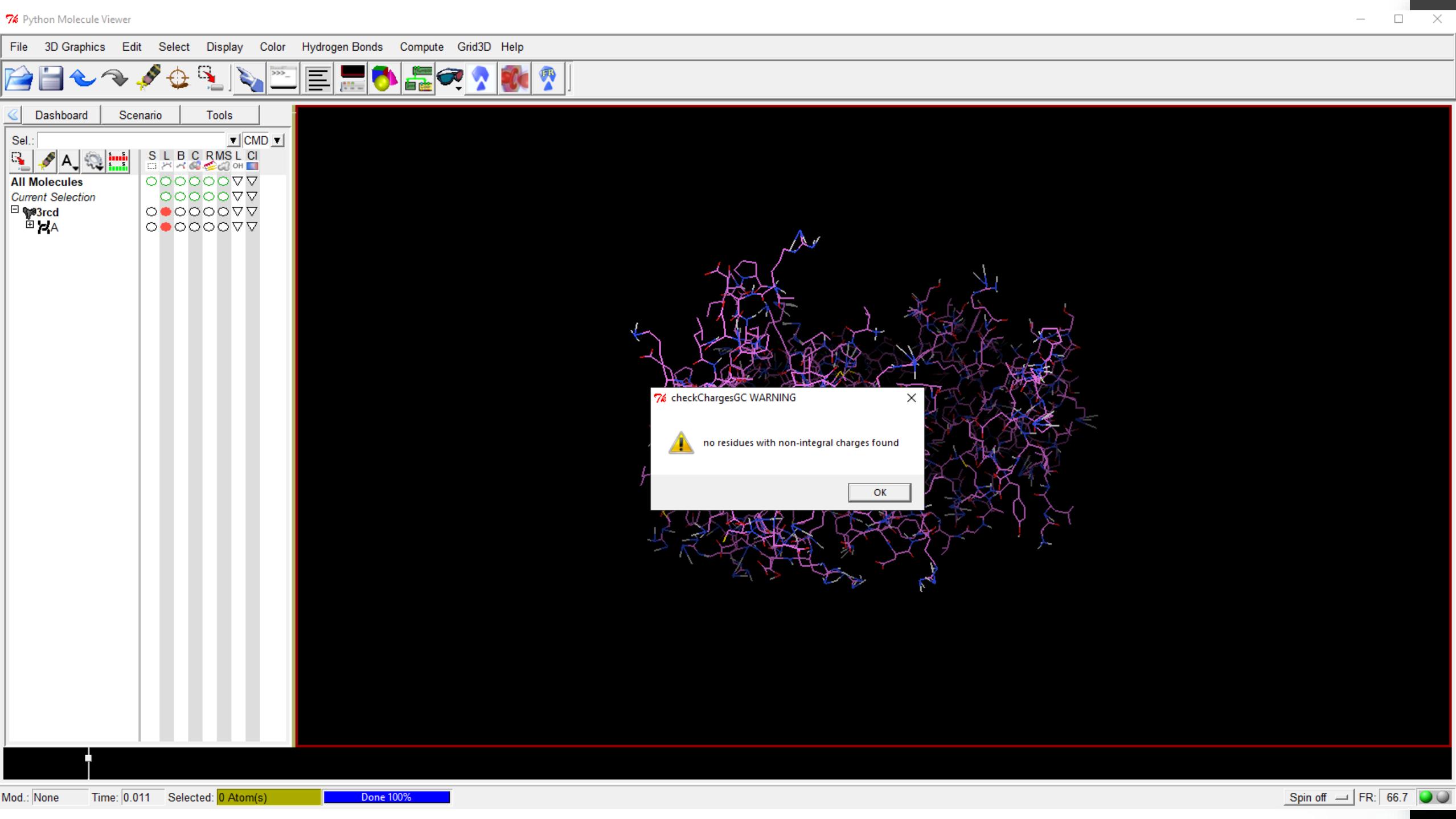
 Total Kollman charge added = 5.25

1





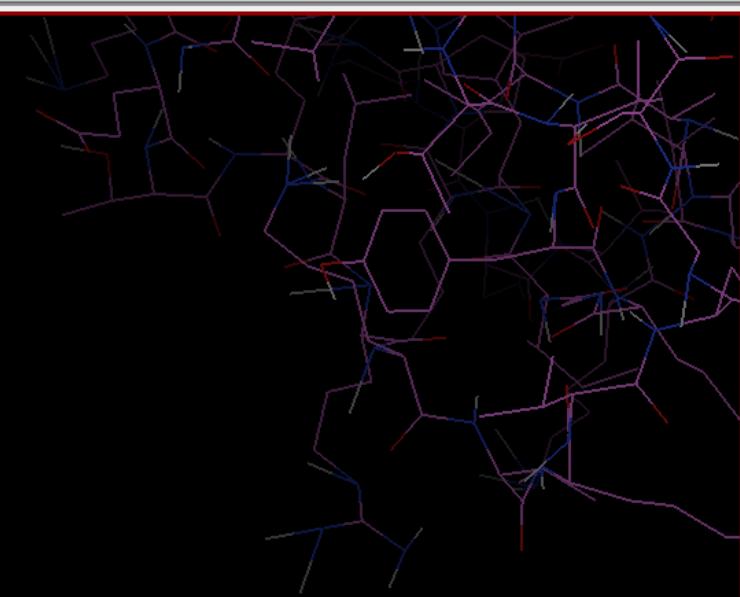
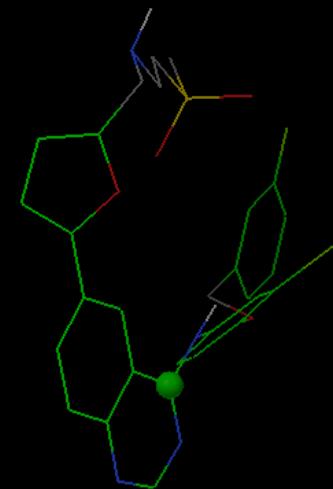


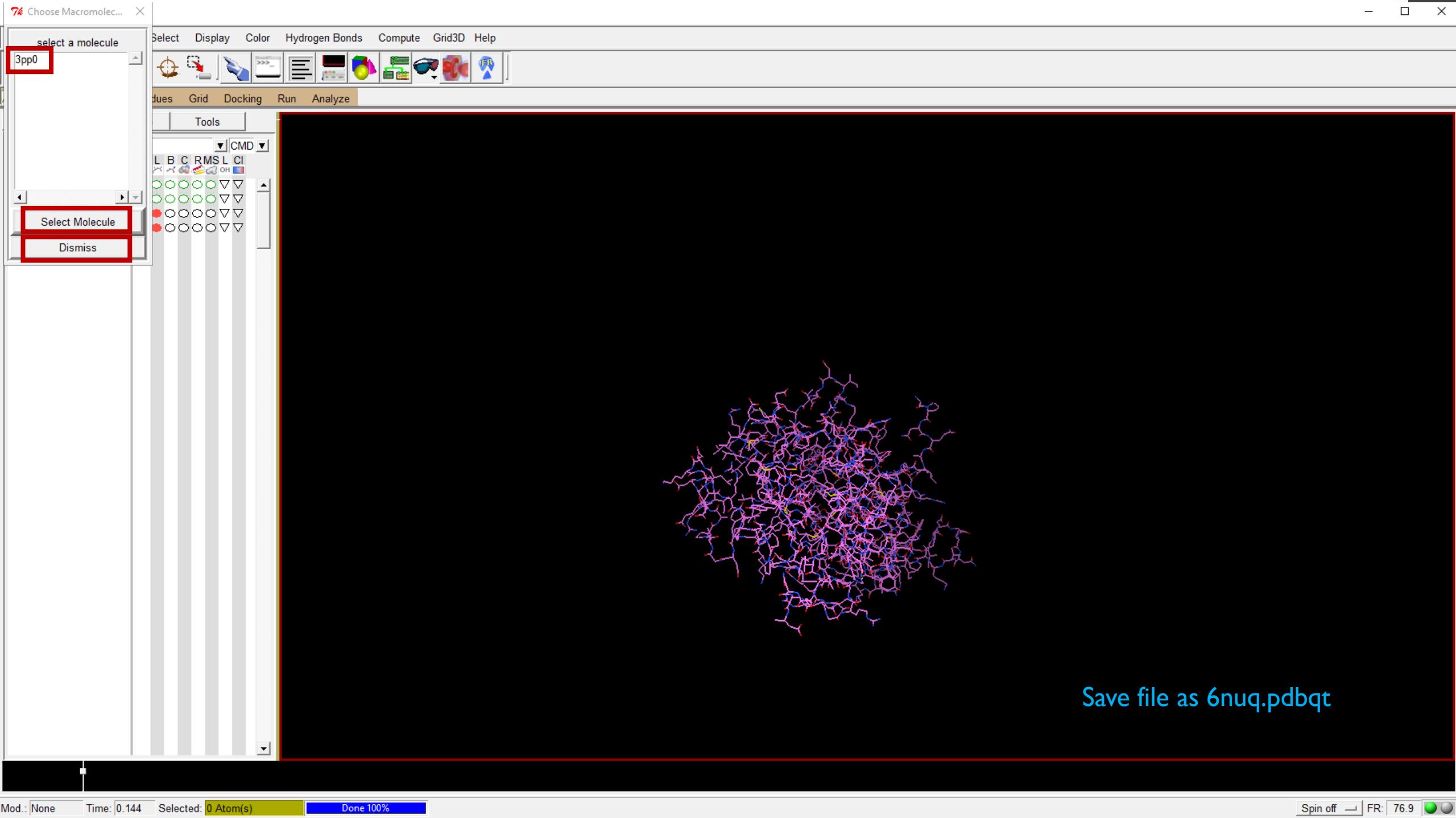




- Sel.: A B C S L B O GPF Grid Docking Run Analyze
- All Molecules
- Current Selection
- AT_3pp0
 - AT_A
 - AT_lapatinib

- Open GPF...
- Macromolecule
 - Open...
 - Choose...
- Set Map Types
 - Grid Box...
 - Other Options...
 - Output
 - Edit GPF...





MGL Tools or AutoDock Tools

With AutoDock Tools (MGL Tools), we prepare the receptor molecule for docking with AutoDock Vina.

- Select the region where ligands bind to the receptor:
 - Grid → Grid Box..(set the box in to the binding site)
- Check the amino acids that are in interaction with the ligand (article, aa&number)
- Open file 6nuq.pdb with Notepad and search for the last atom of the amino acids and write the coordinates x, y, z into the Excel file (line starts with »ATOM«)
- Calculate average values coordinates x, y, z (center of the box)
- Save values in the Configuration file (correct format!)
- Close AutoDock Tools





Open GPF...

Macromolecule >

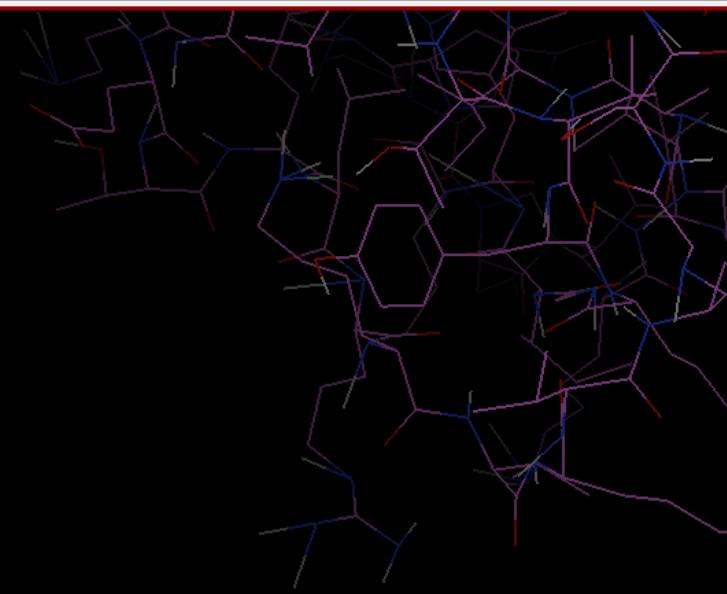
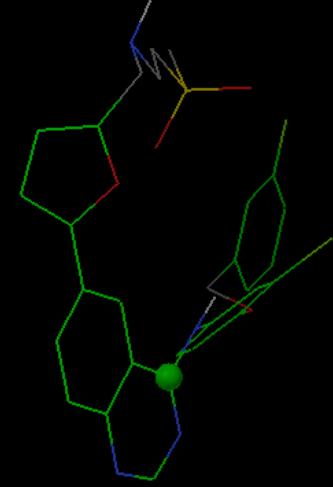
Set Map Types >

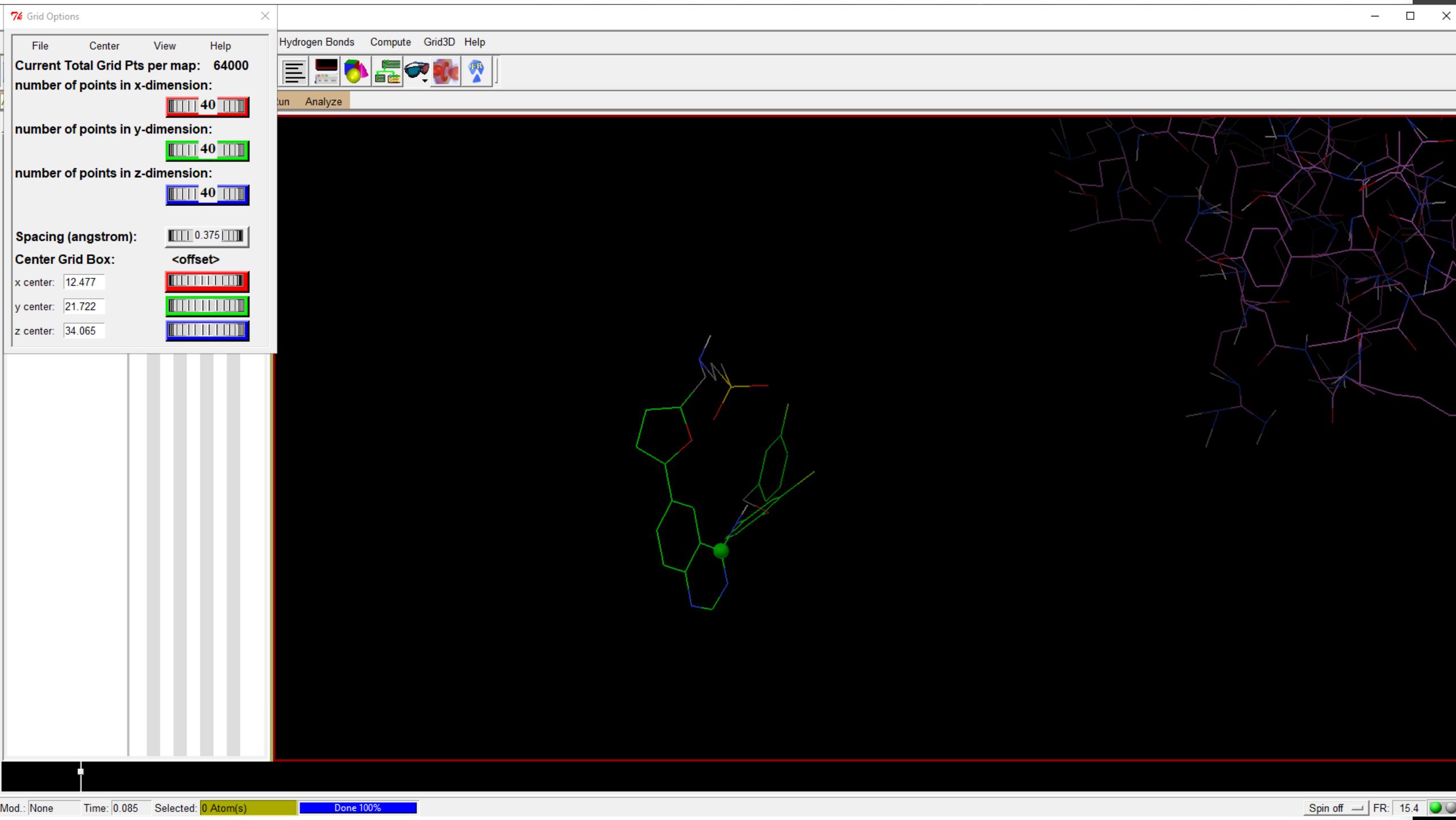
Grid Box...

Other Options... >

Output

Edit GPF...







Sel.: [A] [CMD]

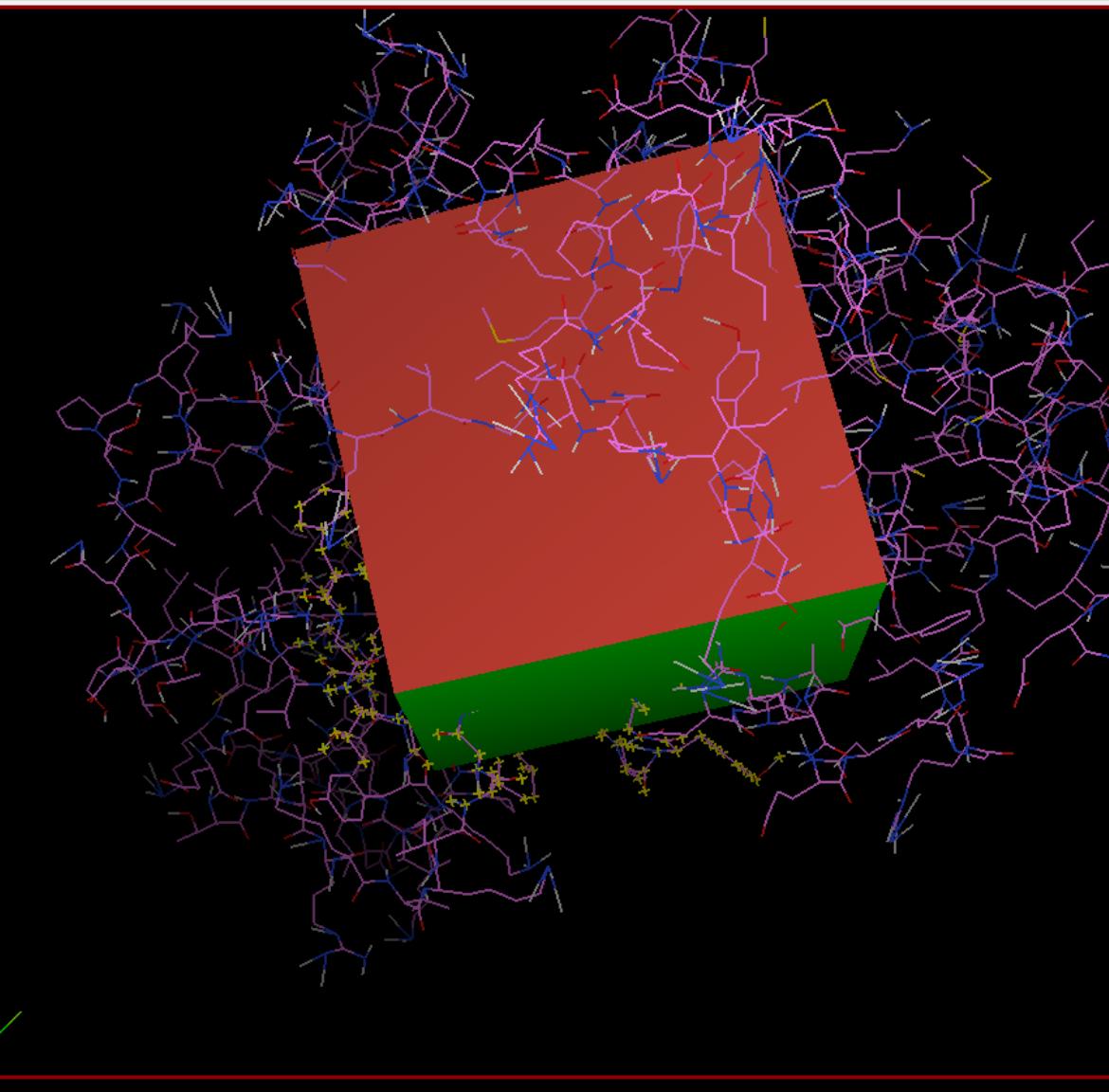
	S	L	B	C	RMSL	CI
QLYS831	○	●	○	○	○	○
QGLY832	○	●	○	○	○	○
QMET833	○	●	○	○	○	○
QSER834	○	●	○	○	○	○
QTYR835	○	●	○	○	○	○
QLEU836	○	●	○	○	○	○
QGLU837	○	●	○	○	○	○
QASP838	○	●	○	○	○	○
QVAL839	○	●	○	○	○	○
QARG840	○	●	○	○	○	○
QLEU841	○	●	○	○	○	○
QVAL842	○	●	○	○	○	○
QHIS843	○	●	○	○	○	○
QARG844	○	●	○	○	○	○
QASP845	○	●	○	○	○	○
QLEU846	○	●	○	○	○	○
QALA847	○	●	○	○	○	○
QALA848	○	●	○	○	○	○
QARG849	○	●	○	○	○	○
QASN850	○	●	○	○	○	○
QVAL851	○	●	○	○	○	○
QLEU852	●	●	○	○	○	○
QVAL853	○	●	○	○	○	○
QLYS854	○	●	○	○	○	○
QSER855	○	●	○	○	○	○
QPRO856	○	●	○	○	○	○
QASN857	○	●	○	○	○	○
QHIS858	○	●	○	○	○	○
QVAL859	○	●	○	○	○	○
QLYS860	○	●	○	○	○	○
QILE861	○	●	○	○	○	○
QTHR862	●	●	○	○	○	○
QASP863	●	●	○	○	○	○
QPHE864	●	●	○	○	○	○
QGLY865	○	●	○	○	○	○
QLEU866	○	●	○	○	○	○
QALA867	○	●	○	○	○	○
QARG868	○	●	○	○	○	○
QLEU869	○	●	○	○	○	○
QLEU870	○	●	○	○	○	○

7k Grid Options

Current Total Grid Pts per map: 226981

number of points in x-dimension: number of points in y-dimension: number of points in z-dimension: Spacing (angstrom):

Center Grid Box: <offset>

x center: y center: z center: 

Sel.: A B C RMSL CI OH CL

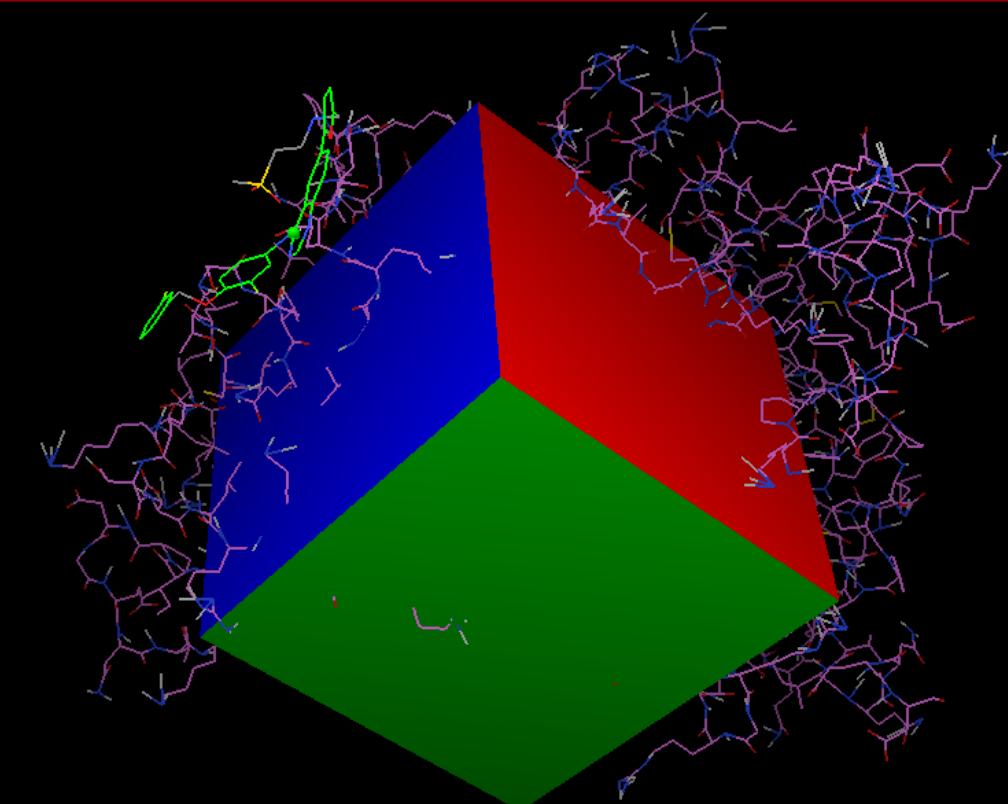
<input checked="" type="checkbox"/>	QLYS831
<input checked="" type="checkbox"/>	QGLY832
<input checked="" type="checkbox"/>	QMET833
<input checked="" type="checkbox"/>	QSER834
<input checked="" type="checkbox"/>	QTYR835
<input checked="" type="checkbox"/>	QLEU836
<input checked="" type="checkbox"/>	QGLU837
<input checked="" type="checkbox"/>	QASP838
<input checked="" type="checkbox"/>	QVAL839
<input checked="" type="checkbox"/>	QARG840
<input checked="" type="checkbox"/>	QLEU841
<input checked="" type="checkbox"/>	QVAL842
<input checked="" type="checkbox"/>	QHIS843
<input checked="" type="checkbox"/>	QARG844
<input checked="" type="checkbox"/>	QASP845
<input checked="" type="checkbox"/>	QLEU846
<input checked="" type="checkbox"/>	QALA847
<input checked="" type="checkbox"/>	QALA848
<input checked="" type="checkbox"/>	QARG849
<input checked="" type="checkbox"/>	QASN850
<input checked="" type="checkbox"/>	QVAL851
<input checked="" type="checkbox"/>	QLEU852
<input checked="" type="checkbox"/>	QVAL853
<input checked="" type="checkbox"/>	QLYS854
<input checked="" type="checkbox"/>	QSER855
<input checked="" type="checkbox"/>	QPRO856
<input checked="" type="checkbox"/>	QASN857
<input checked="" type="checkbox"/>	QHIS858
<input checked="" type="checkbox"/>	QVAL859
<input checked="" type="checkbox"/>	QLYS860
<input checked="" type="checkbox"/>	QILE861
<input checked="" type="checkbox"/>	QTHR862
<input checked="" type="checkbox"/>	QASP863
<input checked="" type="checkbox"/>	QPHE864
<input checked="" type="checkbox"/>	QGLY865
<input checked="" type="checkbox"/>	QLEU866
<input checked="" type="checkbox"/>	QALA867
<input checked="" type="checkbox"/>	QARG868
<input checked="" type="checkbox"/>	QLEU869
<input checked="" type="checkbox"/>	QLEU870

7k Grid Options

Current Total Grid Pts per map: 357911

number of points in x-dimension: number of points in y-dimension: number of points in z-dimension: Spacing (angstrom):

Center Grid Box: <offset>

x center: y center: z center: y center: z center: y center: 

76 Grid Options

X

File Center View Help

Current Total Grid Pts per map: 357911

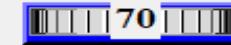
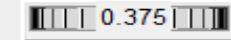
number of points in x-dimension:

 70 

number of points in y-dimension:

 70 

number of points in z-dimension:

 70 Spacing (angstrom):  0.375 

Center Grid Box:

<offset>

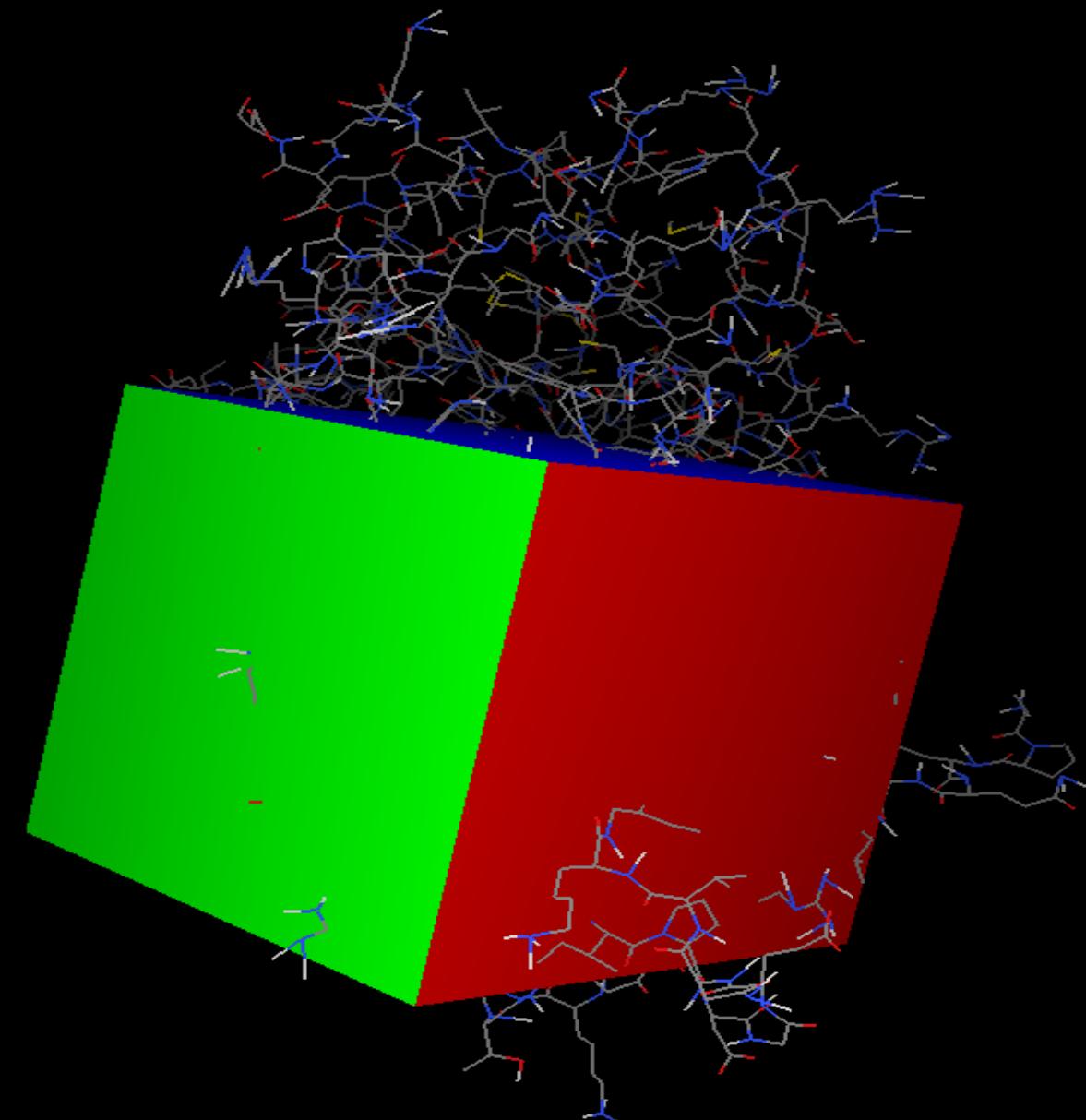
x center: 17.370

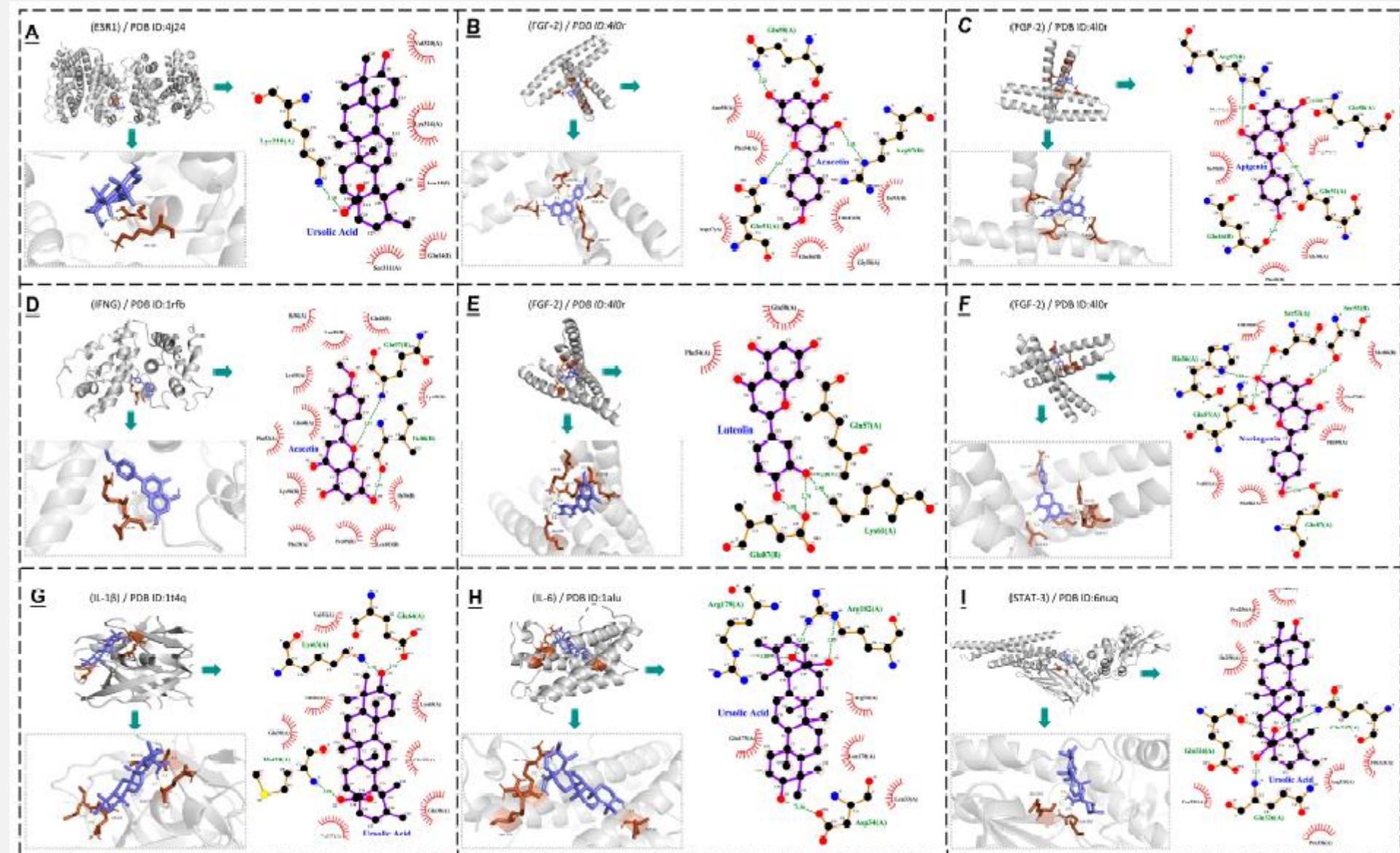


y center: 15.351

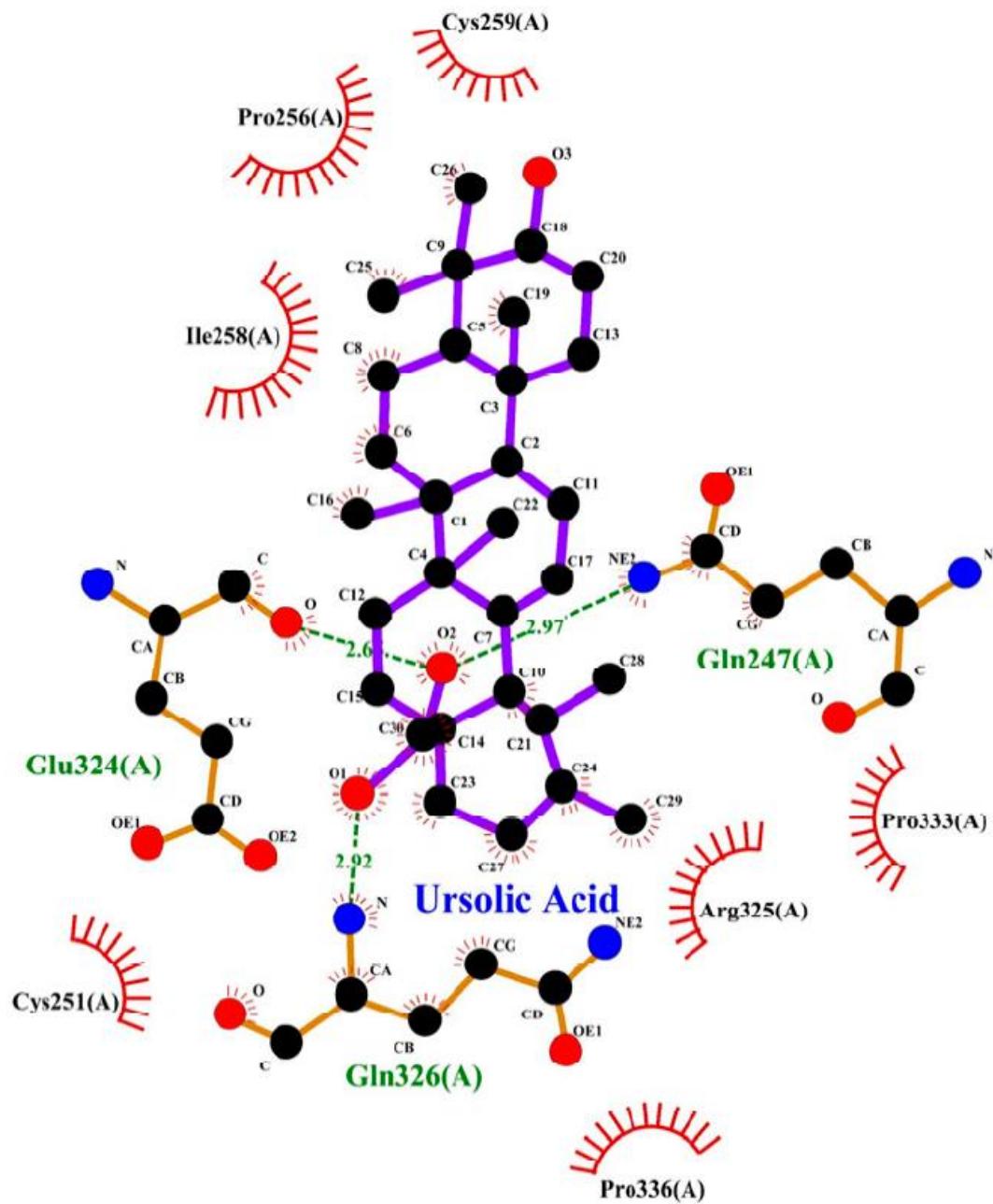
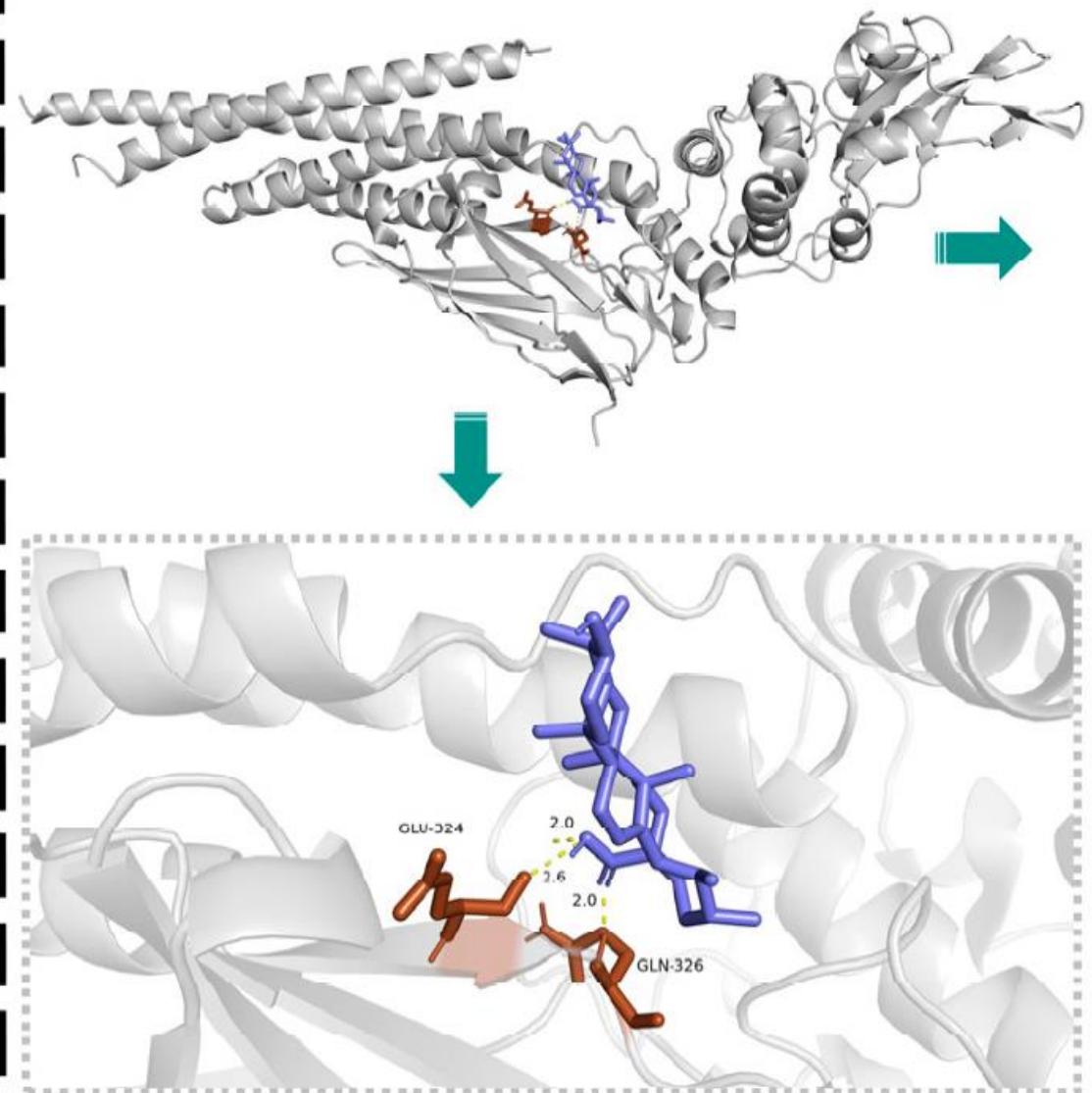


z center: 27.148





(STAT-3) / PDB ID:6nunq



ATOM	795	NE	ARG	A	245	-3.997	17.659	16.510	1.00	97.94	N
ATOM	796	CZ	ARG	A	245	-2.815	17.945	15.978	1.00	110.26	C
ATOM	797	NH1	ARG	A	245	-1.748	17.218	16.285	1.00	96.77	N
ATOM	798	NH2	ARG	A	245	-2.690	18.960	15.137	1.00	96.14	N
ATOM	799	N	ARG	A	246	-5.678	18.684	21.692	1.00	75.16	N
ATOM	800	CA	ARG	A	246	-5.987	20.079	22.019	1.00	75.50	C
ATOM	801	C	ARG	A	246	-4.775	20.763	22.673	1.00	80.72	C
ATOM	802	O	ARG	A	246	-4.593	21.965	22.486	1.00	80.79	O
ATOM	803	CB	ARG	A	246	-7.220	20.184	22.933	1.00	75.39	C
ATOM	804	CG	ARG	A	246	-8.542	20.191	22.187	1.00	84.77	C
ATOM	805	CD	ARG	A	246	-9.716	20.251	23.145	1.00	96.50	C
ATOM	806	NE	ARG	A	246	-10.156	18.914	23.548	1.00	108.60	N
ATOM	807	CZ	ARG	A	246	-10.578	18.596	24.768	1.00	123.20	C
ATOM	808	NH1	ARG	A	246	-10.619	19.515	25.725	1.00	109.28	N
ATOM	809	NH2	ARG	A	246	-10.951	17.353	25.044	1.00	110.76	N
ATOM	810	N	GLN	A	247	-3.944	19.990	23.415	1.00	77.38	N
ATOM	811	CA	GLN	A	247	-2.717	20.468	24.056	1.00	77.25	C
ATOM	812	C	GLN	A	247	-1.650	20.745	22.981	1.00	81.86	C
ATOM	813	O	GLN	A	247	-1.018	21.799	23.033	1.00	81.71	O
ATOM	814	CB	GLN	A	247	-2.213	19.459	25.110	1.00	78.45	C
ATOM	815	CG	GLN	A	247	-1.048	19.969	25.967	1.00	86.94	C
ATOM	816	CD	GLN	A	247	-0.572	18.963	26.991	1.00	96.53	C
ATOM	817	OE1	GLN	A	247	-0.220	17.819	26.677	1.00	89.95	O
ATOM	818	NE2	GLN	A	247	-0.529	19.380	28.243	1.00	86.33	N
ATOM	819	N	GLN	A	248	-1.474	19.816	22.001	1.00	78.70	N
ATOM	820	CA	GLN	A	248	-0.530	19.943	20.875	1.00	78.66	C
ATOM	821	C	GLN	A	248	-0.810	21.238	20.107	1.00	83.56	C
ATOM	822	O	GLN	A	248	0.116	21.982	19.780	1.00	83.26	O
ATOM	823	CB	GLN	A	248	-0.663	18.756	19.899	1.00	79.80	C
ATOM	824	CG	GLN	A	248	-0.094	17.437	20.380	1.00	91.02	C
ATOM	825	CD	GLN	A	248	-0.160	16.404	19.282	1.00	110.56	C
ATOM	826	OE1	GLN	A	248	-1.200	15.780	19.047	1.00	107.39	O
ATOM	827	NE2	GLN	A	248	0.948	16.203	18.578	1.00	101.20	N
ATOM	828	N	ILE	A	249	-2.103	21.498	19.840	1.00	81.04	N
ATOM	829	CA	ILE	A	249	-2.614	22.672	19.136	1.00	81.47	C
ATOM	830	C	ILE	A	249	-2.366	23.957	19.949	1.00	86.83	C
ATOM	831	O	ILE	A	249	-1.856	24.928	19.391	1.00	86.83	O
ATOM	832	CB	ILE	A	249	-4.102	22.438	18.738	1.00	84.51	C
ATOM	833	CG1	ILE	A	249	-4.187	21.519	17.502	1.00	84.92	C
ATOM	834	CG2	ILE	A	249	-4.873	23.750	18.517	1.00	85.37	C
ATOM	835	CD1	ILE	A	249	-5.450	20.687	17.406	1.00	93.44	C
ATOM	836	N	ALA	A	250	-2.689	23.943	21.264	1.00	83.75	N
ATOM	837	CA	ALA	A	250	-2.493	25.076	22.176	1.00	83.50	C
ATOM	838	C	ALA	A	250	-1.016	25.437	22.318	1.00	87.45	C
ATOM	839	O	ALA	A	250	-0.697	26.615	22.496	1.00	86.66	O
ATOM	840	CB	ALA	A	250	-3.089	24.765	23.539	1.00	84.34	C
ATOM	841	N	CYS	A	251	-0.124	24.421	22.225	1.00	84.76	N
ATOM	842	CA	CYS	A	251	1.335	24.558	22.297	1.00	84.91	C
ATOM	843	C	CYS	A	251	1.860	25.418	21.143	1.00	90.86	C
ATOM	844	O	CYS	A	251	2.736	26.254	21.368	1.00	90.32	O
ATOM	845	CB	CYS	A	251	2.016	23.191	22.329	1.00	85.26	C

6nuq.pdb opened
with Notepad

	ak iz članka (zac x	y	z
8			
9	L726	17,45	6,938
10	G727	11,614	10,162
11	F731	3,349	21,428
12	T733	11,474	9,57
13	V734	13,997	13,627
14	A751	20,078	12,681
15	K753	11,85	18,749
16	M774	17,959	27,188
17	S783	19,54	20,303
18	L785	21,622	23,62
19	L796	17,97	21,289
20	V797	22,866	15,955
21	T798	21,121	16,227
22	L800	23,676	7,893
23	M801	24,576	14,731
24	P802	24,521	10,051
25	Y803	25,121	6,459
26	G804	17,389	11,318
27	C805	12,039	11,409
28	L852	16,584	13,946
29	T862	15,773	17,663
30	D863	11,695	17,803
31	F864	17,241	24,072
32			
33			
34	VSOTA	399,505	353,082
35			
36	POVPREČJE	17,370	15,351
37			

- Type values x, y, z from 6nuq.pdb to Excel
- Calculate average values of each coordinate
- Save calculated values in to Notepad and save as „Conf_vs.txt“

conf_vs.txt - Notepad

```

File Edit Format View Help
receptor = 3pp0.pdbqt

center_x = 17.370
center_y = 15.351
center_z = 27.148

size_x = 70
size_y = 70
size_z = 70

num_modes = 10
energy_range = 4

```



conf_vs.txt

```
conf_vs.txt – Beležnica
Datoteka Uredi Oblika Pogled Pomoč
receptor = 6nuq.pdbqt

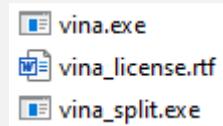
center_x = -0.15933
center_y = 19.593
center_z = 30.17333

size_x = 70
size_y = 70
size_z = 70

num_modes = 10
energy_range = 4
```

AutoDock Vina

- In the folder that we set as directory, we have to have these files:
 - All files ligand.pdbqt
 - 6nuq.pdbqt
 - Files of programme AutoDock Vina
 - Vina_windows.pl
 - conf_vs.txt
 - ligand.txt



Vina_windows.pl

```
#!/usr/bin/perl
print"Ligand_file:\t";
$ligfile=<STDIN>;
chomp $ligfile;
open (FH,$ligfile)||die "Cannot open file\n";
@arr_file=<FH>;
for($i=0;$i<@arr_file;$i++)
{
print"@arr_file[$i]\n";
@name=split(/\./,@arr_file[$i]);
}
for($i=0;$i<@arr_file;$i++)
{
chomp @arr_file[$i];
print"@arr_file[$i]\n";
system("vina.exe --config conf_vs.txt --ligand @arr_file[$i] --log @arr_file[$i]_log.log");
}
```

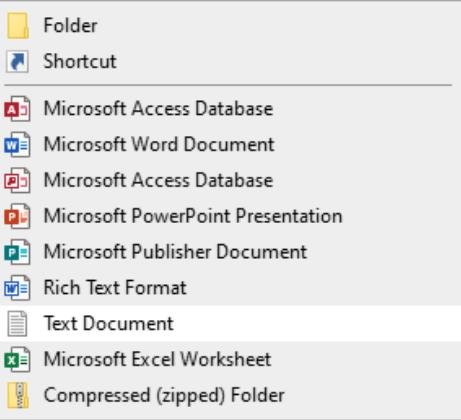
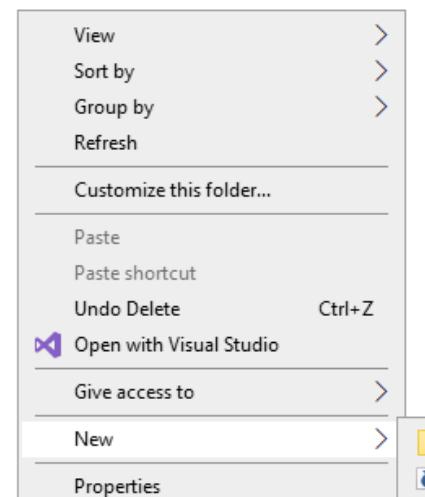


List of ligands

- Create file ligand.txt (with Notepad)
- open cmd (Command Prompt)
 - Move into your folder with all ligand files (ligand.pdbqt) with the „cd“ command
 - **dir /B > ligand.txt**
 - open ligand.txt in Notepad and delete everything that is not ligand.pdbqt



	Name	Date modified	Type	Size
	Afatinib.pdbqt	2/18/2025 4:47 AM	AutoDock Structu...	4 KB
	Alfuzosin.pdbqt	2/18/2025 4:48 AM	AutoDock Structu...	4 KB
	Belumosudil.pdbqt	2/18/2025 4:48 AM	AutoDock Structu...	4 KB
	Bosutinib.pdbqt	2/18/2025 4:49 AM	AutoDock Structu...	5 KB
	Dacomitinib.pdbqt	2/18/2025 4:49 AM	AutoDock Structu...	4 KB
	Doxazosin.pdbqt	2/18/2025 4:49 AM	AutoDock Structu...	4 KB
	Elagolix.pdbqt	2/18/2025 4:49 AM	AutoDock Structu...	6 KB
	Erlotinib.pdbqt	2/18/2025 4:50 AM	AutoDock Structu...	4 KB
	Fedratinib.pdbqt	2/18/2025 4:50 AM	AutoDock Structu...	5 KB
	Fostamatinib.pdbqt	2/18/2025 4:50 AM	AutoDock Structu...	5 KB
	Gefitinib.pdbqt	2/18/2025 4:50 AM	AutoDock Structu...	4 KB
	Inavolisib.pdbqt	2/18/2025 4:51 AM	AutoDock Structu...	4 KB
	Infigratinib.pdbqt	2/18/2025 4:51 AM	AutoDock Structu...	5 KB
	Irinotecan.pdbqt	2/18/2025 4:55 AM	AutoDock Structu...	5 KB
	Lapatinib.pdbqt	2/18/2025 4:55 AM	AutoDock Structu...	5 KB
	Linzagolix.pdbqt	2/18/2025 4:55 AM	AutoDock Structu...	4 KB
	Mobocertinib.pdbqt	2/18/2025 4:56 AM	AutoDock Structu...	6 KB
	Neratinib.pdbqt	2/18/2025 4:56 AM	AutoDock Structu...	5 KB
	Osimertinib.pdbqt	2/18/2025 4:56 AM	AutoDock Structu...	5 KB
	Prazosin.pdbqt	2/18/2025 4:57 AM	AutoDock Structu...	3 KB
	Relugolix.pdbqt	2/18/2025 4:57 AM	AutoDock Structu...	5 KB
	Sotorasib.pdbqt	2/18/2025 4:58 AM	AutoDock Structu...	5 KB
	Terazosin.pdbqt	2/18/2025 4:58 AM	AutoDock Structu...	3 KB
	Topotecan.pdbqt	2/18/2025 5:01 AM	AutoDock Structu...	4 KB
	Trametinib.pdbqt	2/18/2025 5:02 AM	AutoDock Structu...	4 KB
	Trimetrexate.pdbqt	2/18/2025 5:02 AM	AutoDock Structu...	4 KB
	Tucatinib.pdbqt	2/18/2025 5:02 AM	AutoDock Structu...	4 KB
	Vandetanib.pdbqt	2/18/2025 5:02 AM	AutoDock Structu...	4 KB
	vina.exe	5/11/2011 1:37 PM	Application	764 KB
	vina_license.rtf	5/11/2011 1:43 PM	Rich Text Format	8 KB
	vina_split.exe	5/11/2011 1:37 PM	Application	318 KB
	Vina_windows.pl	2/18/2025 1:21 PM	Perl program file	1 KB



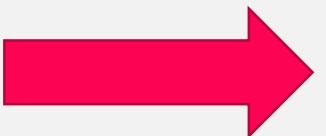
- Afatinib.pdbqt
- Alfuzosin.pdbqt
- Belumosudil.pdbqt
- Bosutinib.pdbqt
- Dacomitinib.pdbqt
- Doxazosin.pdbqt
- Elagolix.pdbqt
- Erlotinib.pdbqt
- Fedratinib.pdbqt
- Fostamatinib.pdbqt
- Gefitinib.pdbqt
- Inavolisib.pdbqt
- Infigratinib.pdbqt
- Irinotecan.pdbqt
- Lapatinib.pdbqt
- ligand.txt
- Linzagolix.pdbqt
- Mobocertinib.pdbqt
- Neratinib.pdbqt
- Osimertinib.pdbqt
- Prazosin.pdbqt
- Relugolix.pdbqt
- Sotorasib.pdbqt
- Terazosin.pdbqt
- Topotecan.pdbqt
- Trametinib.pdbqt
- Trimetrexate.pdbqt
- Tucatinib.pdbqt
- Vandetanib.pdbqt
- vina.exe
- vina_license.rtf
- vina_split.exe
- Vina_windows.pl

```
cmd Command Prompt
Microsoft Windows [Version 10.0.20348.3207]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Anja>cd C:\Users\Anja\Desktop\BIC_1

C:\Users\Anja\Desktop\BIC_1>dir /B > ligand.txt

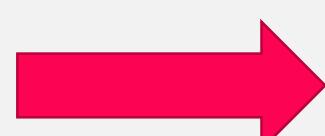
C:\Users\Anja\Desktop\BIC_1>_
```



ligand.txt - Notepad

File Edit Format View Help

Afatinib.pdbqt
Alfuzosin.pdbqt
Belumosudil.pdbqt
Bosutinib.pdbqt
Dacomitinib.pdbqt
Doxazosin.pdbqt
Elagolix.pdbqt
Erlotinib.pdbqt
Fedratinib.pdbqt
Fostamatinib.pdbqt
Gefitinib.pdbqt
Inavolisib.pdbqt
Infigratinib.pdbqt
Irinotecan.pdbqt
Lapatinib.pdbqt
ligand.txt
Linzagolix.pdbqt
Mobocertinib.pdbqt
Neratinib.pdbqt
Osimertinib.pdbqt
Prazosin.pdbqt
Relugolix.pdbqt
Sotorasib.pdbqt
Terazosin.pdbqt
Topotecan.pdbqt
Trametinib.pdbqt
Trimetrexate.pdbqt
Tucatinib.pdbqt
Vandetanib.pdbqt
vina.exe
vina_license.rtf
vina_split.exe
Vina_windows.pl



*ligand.txt - Notepad

File Edit Format View Help

Afatinib.pdbqt
Alfuzosin.pdbqt
Belumosudil.pdbqt
Bosutinib.pdbqt
Dacomitinib.pdbqt
Doxazosin.pdbqt
Elagolix.pdbqt
Erlotinib.pdbqt
Fedratinib.pdbqt
Fostamatinib.pdbqt
Gefitinib.pdbqt
Inavolisib.pdbqt
Infigratinib.pdbqt
Irinotecan.pdbqt
Lapatinib.pdbqt
Linzagolix.pdbqt
Mobocertinib.pdbqt
Neratinib.pdbqt
Osimertinib.pdbqt
Prazosin.pdbqt
Relugolix.pdbqt
Sotorasib.pdbqt
Terazosin.pdbqt
Topotecan.pdbqt
Trametinib.pdbqt
Trimetrexate.pdbqt
Tucatinib.pdbqt
Vandetanib.pdbqt

```
Ukazni poziv - perl Vina_windows.pl
```

```
Microsoft Windows [Version 10.0.19045.5965]
(c) Microsoft Corporation. Vše pravice pridržane.

C:\Users\atansek>cd C:\Users\atansek\Downloads\Molecular docking_files_test2
```

```
C:\Users\atansek\Downloads\Molecular docking_files_test2>dir /B > ligand.txt
```

```
C:\Users\atansek\Downloads\Molecular docking_files_test2>perl Vina_windows.pl
```

```
Ligand_file:
```

```
Ukazni poziv - perl Vina_windows.pl
```

```
Microsoft Windows [Version 10.0.19045.5965]
(c) Microsoft Corporation. Vše pravice pridržane.
```

```
C:\Users\atansek>cd C:\Users\atansek\Downloads\Molecular docking_files_test2
```

```
C:\Users\atansek\Downloads\Molecular docking_files_test2>dir /B > ligand.txt
```

```
C:\Users\atansek\Downloads\Molecular docking_files_test2>perl Vina_windows.pl
```

```
Ligand_file: ligand.txt
```



```
(base) C:\Users\atansek\Downloads\MolecularDocking_Zagreb>perl Vina_windows.pl
Ligand_file:    ligand.txt
ursolic_acid.pdbqt
ursolic_acid.pdbqt
#####
# If you used AutoDock Vina in your work, please cite:      #
#                                                       #
# O. Trott, A. J. Olson,                                     #
# AutoDock Vina: improving the speed and accuracy of docking  #
# with a new scoring function, efficient optimization and    #
# multithreading, Journal of Computational Chemistry 31 (2010) #
# 455-461                                              #
#                                                       #
# DOI 10.1002/jcc.21334                                     #
#                                                       #
# Please see http://vina.scripps.edu for more information.  #
#####
```

WARNING: The search space volume > 27000 Angstrom^3 (See FAQ)

Output will be ursolic_acid_out.pdbqt

Detected 8 CPUs

Reading input ... done.

Setting up the scoring function ... done.

Analyzing the binding site ... done.

Using random seed: 569215864

Performing search ...

0% 10 20 30 40 50 60 70 80 90 100%
|---|---|---|---|---|---|---|---|---|

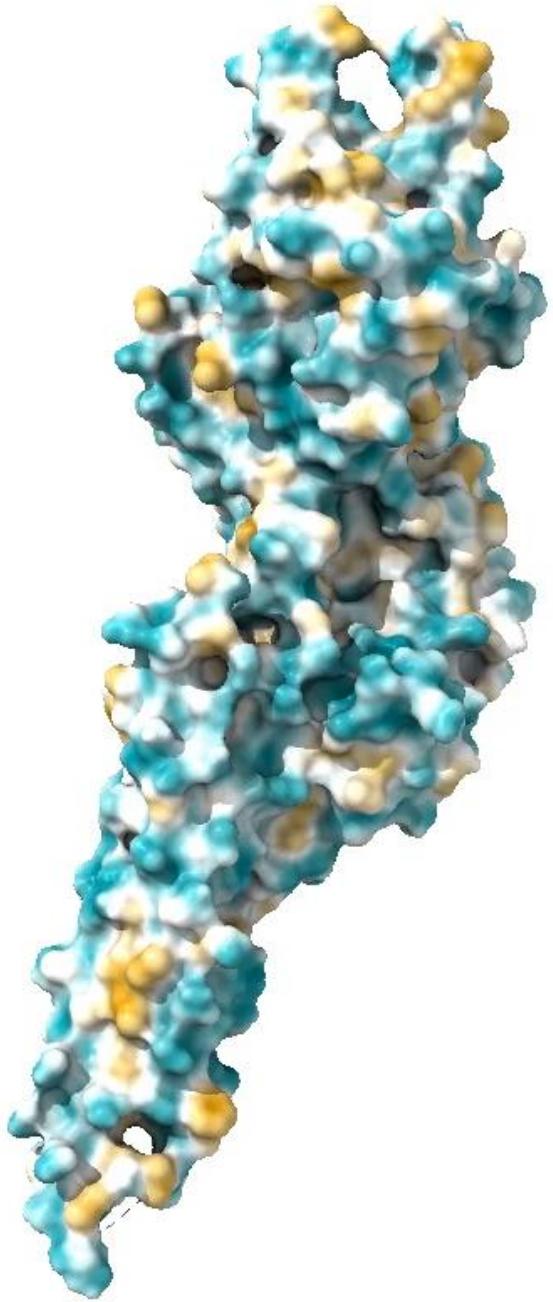
done.

Refining results ... done.

mode	affinity	dist from best mode
	(kcal/mol)	rmsd l.b. rmsd u.b.
1	-8.7	0.000 0.000
2	-8.6	1.935 8.177
3	-8.4	24.323 27.541
4	-8.2	1.709 8.078
5	-8.0	36.359 39.621
6	-7.9	24.787 28.048
7	-7.7	22.425 26.813
8	-7.7	37.301 40.483
9	-7.6	25.865 28.783
10	-7.5	34.732 37.758

Writing output ... done.

```
(base) C:\Users\atansek\Downloads\MolecularDocking_Zagreb>
```

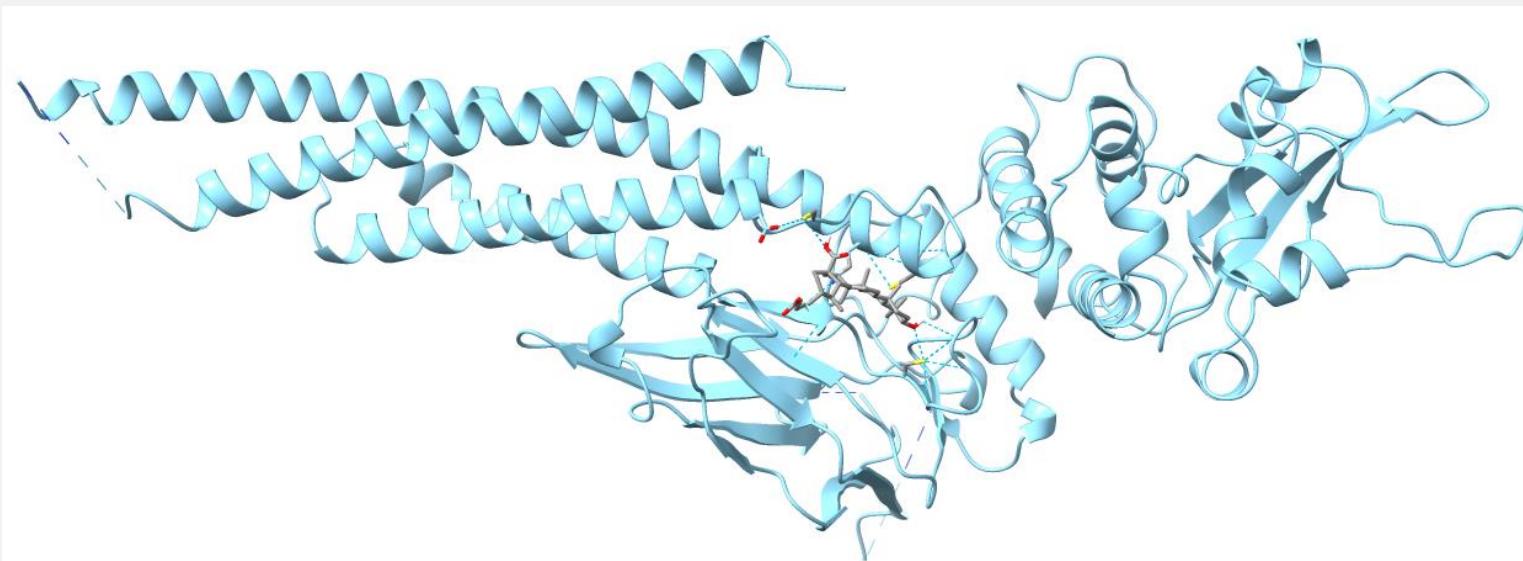
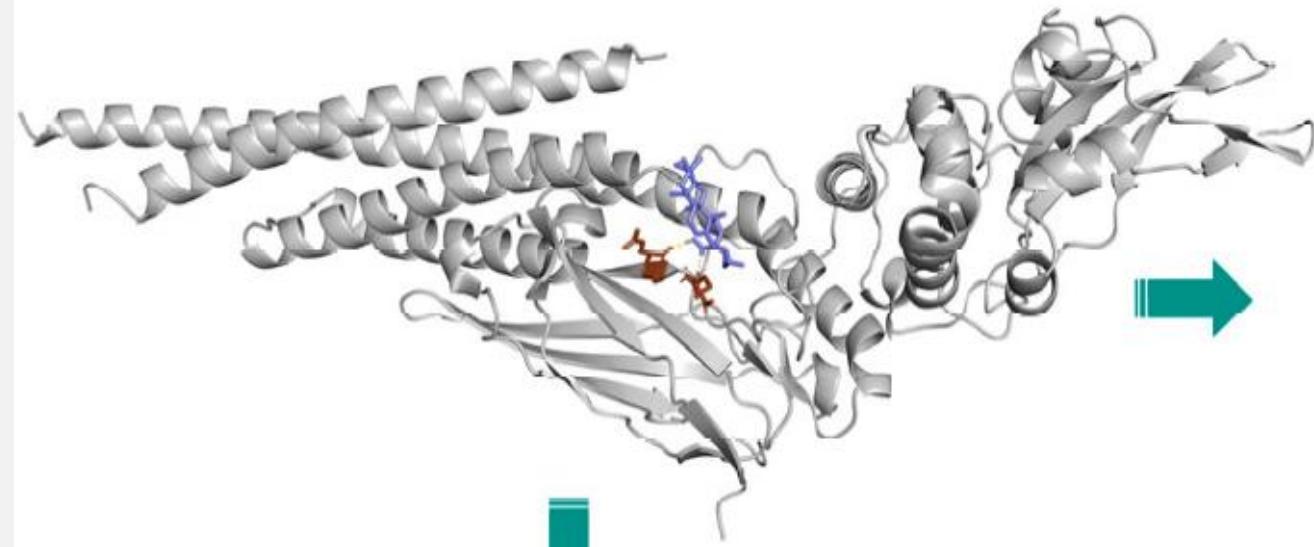


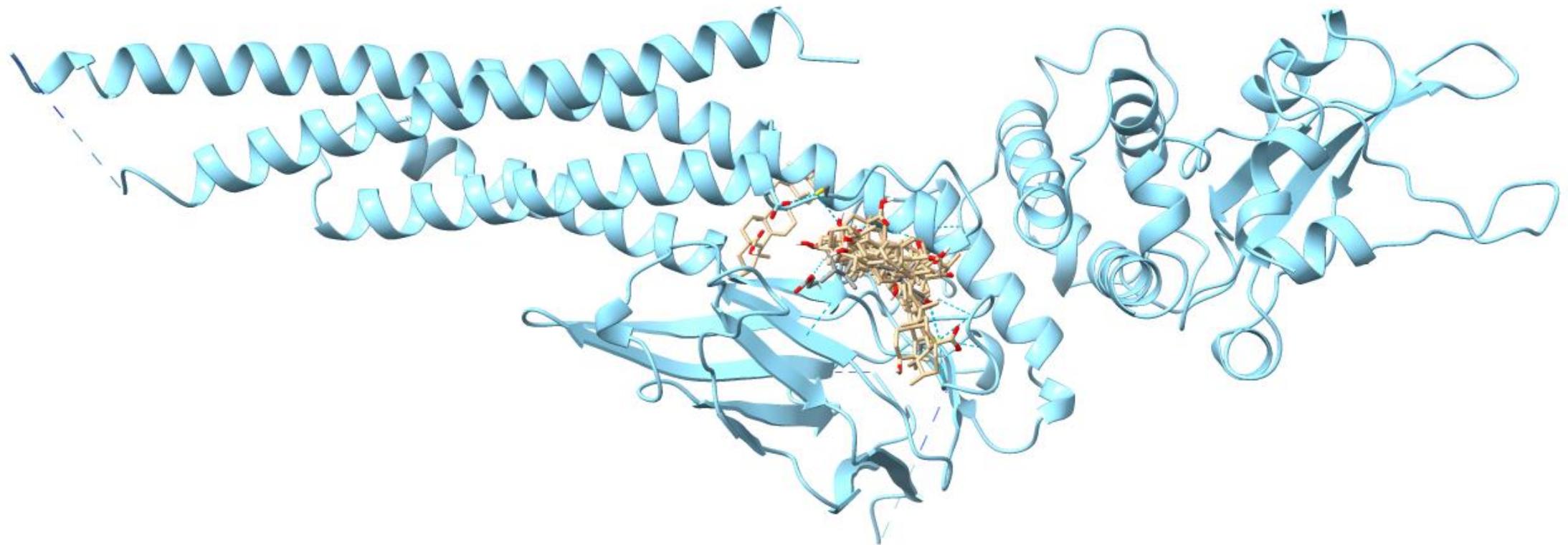
Analysis

ChimeraX



I (STAT-3) / PDB ID:6nuaq

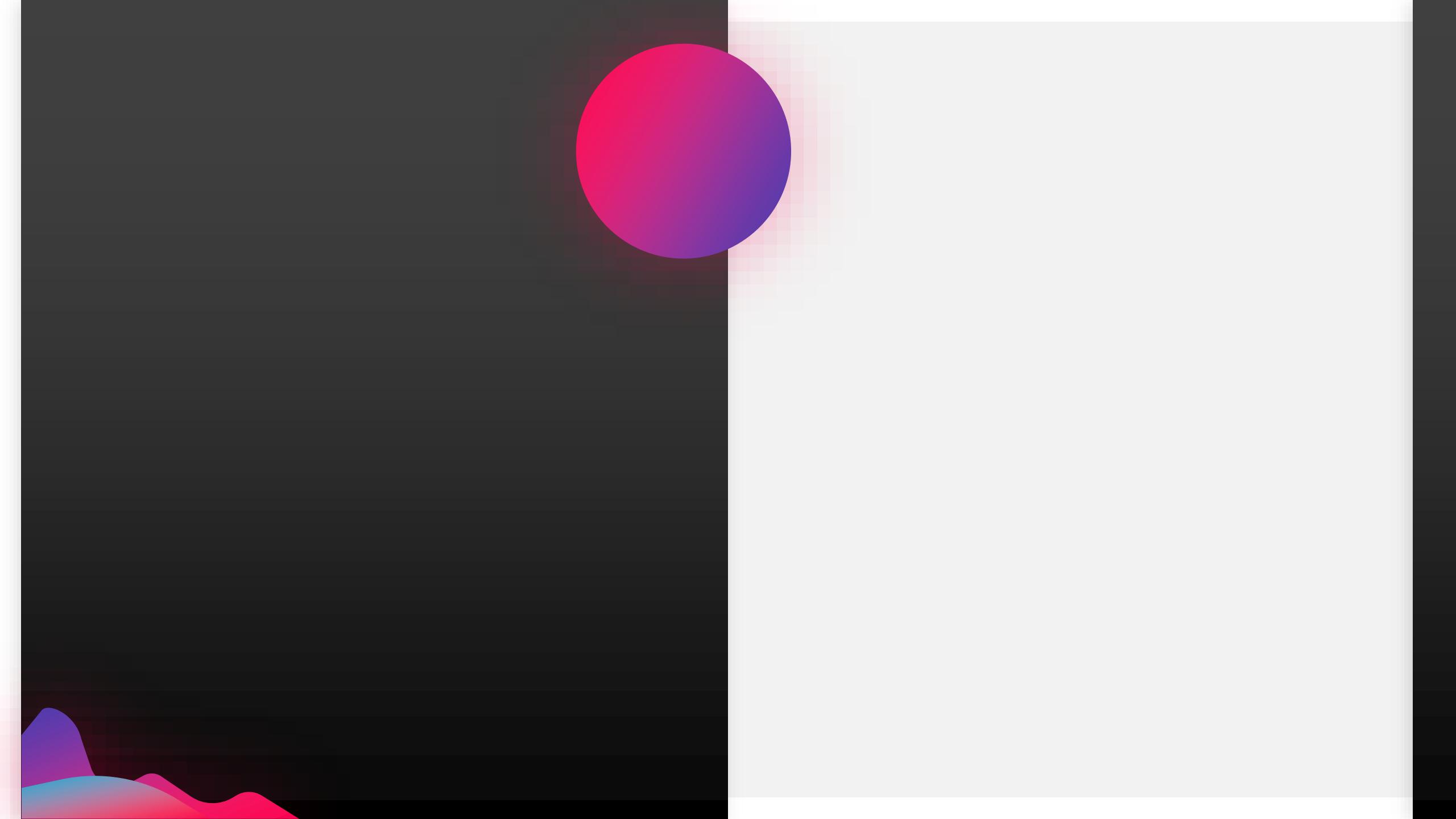




Thank you

Anja Tanšek

✉ anja.tansek@bf.uni-lj.si



RECEPTOR RETRIEVAL (.PDB)

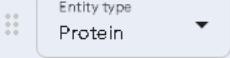
With AlphaFold Server



Remaining jobs: 20

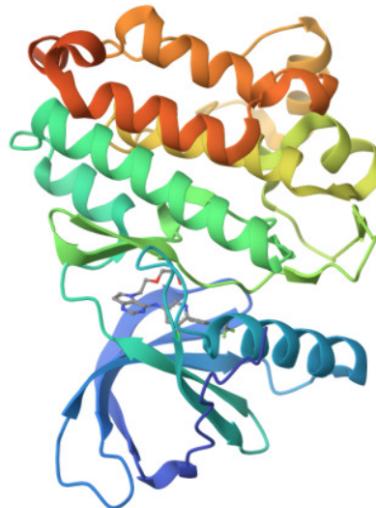
AlphaFold Server allows you to model a structure consisting of many biological molecules

[Learn more](#)  [Upload JSON](#)  [Clear](#)

 Entity type	Protein	 Copies	1	 
	10	20	30	40
PSGAMPNQAQ	MRILKETELR	KVKVLGSQAF	GTVYKGIVIP	DGENVKIPVA
70	80	90	100	110
KANKEILDEA	YVMAGVGSPY	VSRLLGICLT	STVQLVTQLM	PYGCLLDHVR
130	140	150	160	170
LLNWCMQIAK	GMSYLEDVRL	VHRDLAARNV	LVKSPNHVKI	TDFGLARLLD
190	200	210	220	230
GKVPPIKWMAL	ESILRRRFTH	QSDVWSYGV	VWELMTFGAK	PYDGIPAREI
250	260	270	280	290
PQPPPICTIDV	YMI MVKWCWMI	DSECRPRFRE	LVSEFSRMR	PDLLEKGERL
310	320	330	332	300
STFYRSLLED	DDMGDLVDAE	EYLVPQQGFF	CP	EDLGPA

 [Add entity](#) [Save job](#)[Continue and preview job](#)

<input checked="" type="checkbox"/>	Name	Modified	
<input checked="" type="checkbox"/>	2025-02-12_13:52	2025-02-12 13:54	

Biological Assembly 1  Explore in 3D: [Structure](#) | [Sequence Annotations](#) | [Electron Density](#) | [Validation Report](#) | [Ligand Interaction \(03Q\)](#)Global Symmetry: Asymmetric - C1 
Global Stoichiometry: Monomer - A1 [Find Similar Assemblies](#)

Biological assembly 1 assigned by authors and generated by PISA (software)

 3PPO

Crystal Structure of the Kinase domain of Human HER2 (erbB2).

PDB DOI: <https://doi.org/10.2210/pdb3PPO/pdb>

Classification: TRANSFERASE/TRANSFERASE INHIBITOR

Organism(s): Homo sapiens

Expression System: *Antherea*Mutation(s): Yes 

Deposited: 2010-11-23 Released: 2011-03-30

Deposition Author(s): [Skene, R.J.](#), [Aertgeerts, K.](#), [Sogabe, S.](#)

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 2.25 Å

R-Value Free:

0.260 (Depositor), 0.260 (DCC) 

R-Value Work:

0.185 (Depositor), 0.190 (DCC) 

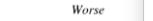
R-Value Observed:

0.189 (Depositor) 

Starting Model: experimental

[View more details](#)wwPDB Validation 

Metric

Rfree Clashscore Ramachandran outliers Sidechain outliers RSRZ outliers Worse Percentile red Percentile red Better 

Ligand Structure Quality As

Worse 0

 1 Better

Ligand structure goodness of fit to experimental data

Datoteka

Osnovno

Pogled



Prilepi

Odložišče

Izreži



Kopiraj



Prilepi



B



I



U



abe



x

x²

A



-



11

A⁺A⁻

-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



-



Slika

Risba v

Datum

Vstavi

predmet

Najdi

ab

Zamenjaj

Izberi vse



3

2

1

-

1

-

2

-

3

-

4

-

5

-

6

-

7

-

8

-

9

-

10

-

11

-

12

-

13

-

14

-

15

-

16

-

17

-

18

-

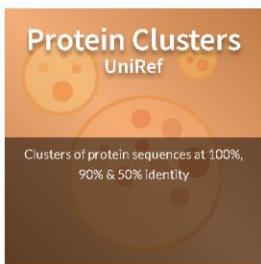
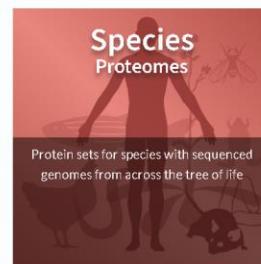
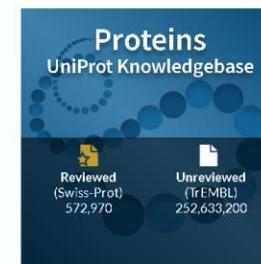
>3PP0_1|Chains A, B|Receptor tyrosine-protein kinase erbB-2|Homo sapiens (9606)
MSGAAPNQALLRILKETELRKVVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANK
EILDEAYVMAGVGSPYVSRLLIGICLTSTVQLVTQLMPYGCLLDHVRENRRGLGSQDLLNWCMQI
AKGMSYLEDVRLVHRDLAARNVLVKSPNHWKITDFGLARLLIDIDETEYHADGGKVPIKWMAL
SILRRRFTHQSDVWWSYGVTWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPPICTIDVYMIMVK
CWMDIDSECPRPRFELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLLEDMMGDLVDAE
EYLVPQQGAAASHHHHHH

Find your protein

UniProtKB Advanced | List

Examples: Insulin, APP, Human, P05067, organism_id:9606

UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information. [Cite UniProt](#) •



Status

Reviewed (Swiss-Prot) (40)

Unreviewed (TrEMBL) (484)

Popular organisms

Human (19)

Mouse (9)

Rat (4)

A. thaliana (3)

Zebrafish (3)

Taxonomy

Filter by taxonomy

Group by

Taxonomy

Keywords

Gene Ontology

Enzyme Class

Proteins with

3D structure (19)

Active site (259)

Activity regulation (3)

Alternative products (isoforms) (12)

Alternative splicing (12)

More items

Protein existence

Homology (283)

Predicted (183)

Protein level (35)

Transcript level (23)

UniProtKB 524 results or search "her2" as a Gene Name, Protein Name, Strain, Organism, or Taxonomy

Tools Download (524) Add View: Cards Table Customize columns Share

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
<input type="checkbox"/> P04626	ERBB2_HUMAN	Receptor tyrosine-protein kinase erbB-2[...]	ERBB2, HER2, MLN19, NEU, NGL	Homo sapiens (Human)	1,255 AA
<input type="checkbox"/> Q03557	GATA_YEAST	Glutamyl-tRNA(Gln) amidotransferase subunit A, mitochondrial[...]	HER2, GEP6, LRC6, YMR293C	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast)	464 AA
<input type="checkbox"/> O75368	SH3L1_HUMAN	Adapter SH3BGR[...]	SH3BGR	Homo sapiens (Human)	114 AA
<input type="checkbox"/> P33893	GATB_YEAST	Glutamyl-tRNA(Gln) amidotransferase subunit B, mitochondrial[...]	PET112, YBL080C, YBL0724	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast)	541 AA
<input type="checkbox"/> P53260	GATF_YEAST	Glutamyl-tRNA(Gln) amidotransferase subunit F, mitochondrial[...]	GTF1, YGR102C	Saccharomyces cerevisiae (strain ATCC 204508 / S288c) (Baker's yeast)	183 AA
<input type="checkbox"/> Q90464	Q90464_DANRE	HER-2 protein[...]	her2, HER-2	Danio rerio (Zebrafish) (Brachydanio rerio)	108 AA
<input type="checkbox"/> Q6ZN17	LN28B_HUMAN	Protein lin-28 homolog B[...]	LIN28B, CSDD2	Homo sapiens (Human)	250 AA
<input type="checkbox"/> Q6C0M4	GATA_YARLI	Glutamyl-tRNA(Gln) amidotransferase subunit A, mitochondrial[...]	HER2, YAL10F23441g	Yarrowia lipolytica (strain CLIB 122 / E 150) (Yeast) (Candida lipolytica)	459 AA
<input type="checkbox"/> Q75D84	GATA_EREGS	Glutamyl-tRNA(Gln) amidotransferase subunit A, mitochondrial[...]	HER2, ABR140C	Eremothecium gossypii (strain ATCC 10895 / CBS 109.51 / FGSC 9923 / NRRL Y-1056) (Yeast) (Ashbya gossypii)	463 AA
<input type="checkbox"/> P98177	FOXO4_HUMAN	Forkhead box protein O4[...]	FOXO4, AFX, AFX1, MLLT7	Homo sapiens (Human)	505 AA
<input type="checkbox"/> Q01973	ROR1_HUMAN	Inactive tyrosine-protein kinase transmembrane receptor ROR1[...]	ROR1, NTRKR1	Homo sapiens (Human)	937 AA
<input type="checkbox"/> C4YRY0	GATA_CANAW	Glutamyl-tRNA(Gln) amidotransferase subunit A, mitochondrial[...]	HER2, CAWG_04838	Candida albicans (strain WO-1) (Yeast)	450 AA
<input type="checkbox"/> Q9NUP9	LIN7C_HUMAN	Protein lin-7 homolog C[...]	LIN7C, MALS3, VELI3	Homo sapiens (Human)	197 AA
<input type="checkbox"/> Q9H9Z2	LN28A_HUMAN	Protein lin-28 homolog A[...]	LIN28A, CSDD1, LIN28, ZCCHC1	Homo sapiens (Human)	209 AA
<input type="checkbox"/> Q9HAP6	LIN7B_HUMAN	Protein lin-7 homolog B[...]	LIN7B, MALS2, VELI2, UNQ3116/PRO10200	Homo sapiens (Human)	207 AA
<input type="checkbox"/> O14910	LIN7A_HUMAN	Protein lin-7 homolog A[...]	LIN7A, MALS1, VELI1	Homo sapiens (Human)	233 AA
<input type="checkbox"/> P24593	IBP5_HUMAN	Insulin-like growth factor-binding protein 5[...]	IGFBP5, IBP5	Homo sapiens (Human)	272 AA
<input type="checkbox"/> Q92673	SORL_HUMAN	Sortilin-related receptor[...]	SORL1, C11orf32	Homo sapiens (Human)	2,214 AA
<input type="checkbox"/> O95905	ECD_HUMAN	Protein ecdysoneless homolog[...]	ECD	Homo sapiens (Human)	644 AA

Function

★ P04626 · ERBB2_HUMAN

Names & Taxonomy

Proteinⁱ Receptor tyrosine-protein kinase erbB-2

Amino acids 1255 (go to sequence)

Subcellular Location

Geneⁱ ERBB2Protein existenceⁱ Evidence at protein level

Disease & Variants

Statusⁱ UniProtKB reviewed (Swiss-Prot)Annotation scoreⁱ 5/5

PTM/Processing

Organismⁱ Homo sapiens (Human)

Expression

Entry Variant viewer 5,670 Feature viewer Genomic coordinates Publications External links History

Interaction

Tools Download Add Community curation (2) Add a publication Entry feedback

Structure

Family & Domains

Sequence & Isoforms

Similar Proteins

Functionⁱ

Protein tyrosine kinase that is part of several cell surface receptor complexes, but that apparently needs a coreceptor for ligand binding. Essential component of a neuregulin-receptor complex, although neuregulins do not interact with it alone. GP30 is a potential ligand for this receptor. Regulates outgrowth and stabilization of peripheral microtubules (MTs). Upon ERBB2 activation, the MEMO1-RHOA-DIAPH1 signaling pathway elicits the phosphorylation and thus the inhibition of GSK3B at cell membrane. This prevents the phosphorylation of APC and CLASP2, allowing its association with the cell membrane. In turn, membrane-bound APC allows the localization of MACF1 to the cell membrane, which is required for microtubule capture and stabilization. Curated

In the nucleus is involved in transcriptional regulation. Associates with the 5'-TCAAATTC-3' sequence in the PTGS2/COX-2 promoter and activates its transcription. Implicated in transcriptional activation of CDKN1A; the function involves STAT3 and SRC. Involved in the transcription of rRNA genes by RNA Pol I and enhances protein synthesis and cell growth. 3 Publications

Catalytic activityⁱ

Rhea 10596

L-tyrosyl-[protein] + ATP = O-phospho-L-tyrosyl-[protein] + ADP + H⁺ PROSITE-ProRule Annotation 1 Publication
EC:2.7.10.1 (UniProtKB | ENZYME | Rhea)

Hide Rhea reaction ^

L-tyrosyl-[protein]

RHEA-COMP:10136

L-tyrosine residue

CHEBI:44458

ATP

CHEBI:30616

O-phospho-L-tyrosyl-[protein]

RHEA-COMP:10137

ADP

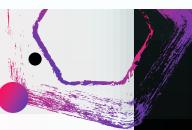
CHEBI:456216

L-tyrosine-phosphate residue

CHEBI:456216

H⁺

CHEBI:15378

H⁺

Function

Entry Variant viewer 5,670 Feature viewer Genomic coordinates Publications External links History

Names & Taxonomy

Subcellular Location

Disease & Variants

PTM/Processing

Expression

Interaction

Structure

Family & Domains

Sequence & Isoforms

Similar Proteins

Sequence & Isoformsⁱ

Align Isoforms (6) Add Isoforms

Sequence statusⁱ CompleteSequence processingⁱ The displayed sequence is further processed into a mature form.This entry describes 6 isoformsⁱ produced by Alternative splicing & Alternative initiation.

P04626-1

This isoform has been chosen as the canonical sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.

Name 1

Synonyms ERBB2, HER2

See also sequence in UniParc or sequence clusters in UniRef

Tools Download Add Copy sequence

Length 1,255

Last updated 1987-08-13 v1

Mass (Da) 137,910

MD5 Checksumⁱ D89A90920775435179A0D5281A800F78

10	20	30	40	50	60	70	80	90	100	110	120	130	140
MELAALCRWG	LLLALLPPGA	ASTQVCTGTD	MKLRLPASPE	THLDMLRHLY	QGCQVQQGNL	ELTYLPTNAS	LSFLQDIDQEV	QGYVLIAHNQ	VRQVPLQRLR	IVRGTQLFED	NYALAVLDNG	DPLNNTPVT	GASPGLREL
150	160	170	180	190	200	210	220	230	240	250	260	270	280
QLRSLTEILK	GGVLIQRNQ	LCYQDTILWK	DIFHKNNQLA	LTLIDTNRSR	ACHPCSPMCK	GSRCWGESSE	DCQSLTRTVC	AGGCARCKGP	LPTDCCHEQC	AAGCTGPKHS	DCLACLHFNH	SGICELHCPA	LVTYNTDTFE
290	300	310	320	330	340	350	360	370	380	390	400	410	420
SMPNPEGRT	FGASCVTAC	YNYLSTDVGS	CTLVCPHLNQ	EVTAEDGTQR	CEKCSKPCAR	VCYGLGMHEH	REVRAVTSAN	IQEFAGCKKI	FGSLAFLPES	FDGDPASNTA	PLQPEQLQVF	ETLEEITGYL	YISAWPDSL
430	440	450	460	470	480	490	500	510	520	530	540	550	560
DLSVFQNLQV	IRGRILHNGA	YSLTLQGLGI	SWLGLRSLRE	LGSGLALIHH	NTHLCFVHTV	PWDQLFRNPH	QALLHTANRP	EDECVGEGLA	CHQLCARGHC	WGPGBTQCVN	CSQFLRGQEC	VEECRVLQGL	PREYVNARHC
570	580	590	600	610	620	630	640	650	660	670	680	690	700
LPCHPECQPO	NGSVCFCGP	ADQCVACAHY	KDPPFCVARC	PSGVKPDLSY	MPIWKFPDDE	GACQPCPCIN	THSCVLDLDDK	GCPAEQRASP	LTSIIISAVVG	ILLVVVVLGVV	FGILIKRQQ	KIRKYTMRRRL	LQETELVEPL
710	720	730	740	750	760	770	780	790	800	810	820	830	840
TPSGAMPNQA	QMRILKETEL	RKVVLGSGA	FGTVYKGWI	PDGENVKIPV	AIKVLRENTS	PKANEKILDE	AYVMAGVGSP	YVSRLLLGICL	TSTVQLVTQL	MPYGCLLDHV	RENGRGLGSQ	DLLNWCMQIA	KGMSYLEDV
850	860	870	880	890	900	910	920	930	940	950	960	970	980
LVHRDLAARN	VLVKSPNIVK	ITDFGLARLL	DIDETEYHAD	GGKVPPIKWM	LESLIIRRFT	HQSDVWSYGV	TVWELMTFGA	KPYDGIPARE	IPDLLEKGER	LPQPPICITD	VYMIMVKCWM	IDSECPRFR	ELVSEFSRMA
990	1000	1010	1020	1030	1040	1050	1060	1070	1080	1090	1100	1110	1120
RDPQRFFVVIQ	NEDLGPASPL	DSTFYRSLL	DDDMGDLVDA	EEYLVPQQGF	FCDPDPAPGAG	GMVHHRRSS	STRSGGGDLT	LGLPSEEEA	PRSLPLASEG	AGSDVFDGDL	GMGAAGLQS	LPTHDPSPLO	RYSEDPTVPL
1130	1140	1150	1160	1170	1180	1190	1200	1210	1220	1230	1240	1250	1260
PSETDGYVAP	LTCSPQPEYV	NQPDVPRQPP	SPREGPLPAA	RPAGATLERP	KTLSPGKNGV	VKDVFAGGA	VENPEYLTPQ	GGAAPQPHPP	PAFSPAFDNL	YYWDQDPPER	GAPPSTFKGT	PTAENPEYLG	LDVPL

P04626-2

>sp|P04626|ERBB2_HUMAN Receptor tyrosine-protein kinase erbB-2 OS=Homo sapiens OX=9606 GN=ERBB2 PE=1 SV=1
MELAALCRWGLLALLPPGAASTQVCTGTDMLRLPASPETHLDMLRHLYQGCQVVQGNL
ELTYLPTNASLSFLQDIQEYQGYVLIAHNQRQVPLQRLRIVRGTQLFEDNYALAVLDNG
DPLNNNTTPVTGASPGLRELQLRSLTEILKGGLIQRNPQLCYQDTILWKDIFHKNNQLA
LTLIDTNRSRACHPCSPMCKGSRWGESSEDCQLTRTCAGGCARCKGPLPTDCCHEQC
AAGCTGPKHSDCLACLHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP
YNYLSTDVGSCTLVCPLHNQEVTAEDEGTQRCEKCSKPCARVCYGLGMEHLREVRAVTSAN
IQEFAGCKKIFGSLAFLPESFDGDPASNTAPLQPEQLQVFETLEEITGYLYISAWPDSLP
DLSVFQNLQVIRGRILHNGAYSLTLQGLGISWLGLRSLRELGSGLAIHHNTHLCFVHTV
PWDQLFRNPHQALLHTANRPEDECVGEGLACHQLCARGHCWPGPTQCVNCSQFLRGQEC
VEECRVLQGLPREYVNARHCLPCHPECQPQNGSVTCFGPEADQCVACAHYKDPPFCVARC
PSGVKPDLSYMPIWKFDEEGACQPCPINCTHSCVDLDDKGCPAEQRASPLTSIISAVVG
ILLVVVLGVVFGILIKRRQQKIRKYTMRRLLQETELVEPLTPSGAMPNQAQMRLIKETEL
RKVKVLGSGAFGTVYKGWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP
YVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRGRLGSQDLLNWCMQIAKGMSYLEDVR
LVHRDLAARNVLVKSPNHVKITDFGLARLLDIDETEHADGGKVPIKWMALESILRRRFT
HQSDVWSYGVTVWELMTFGAKPYDGPAREIPDLLEKGERLPQPPPICTIDVYMIMVKCWM
IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDDDMGDLVDA
EEYLVPQQGFVCPDPAPGAGGMVHHRHRSSTRSGGGDLTLGLEPSEEAPRSPLAPSEG
AGSDVFDGDLGMGAAKGLQSLPTHDPPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYV
NQPDVVRPQPPSPREGPLPAARPAGATLERPKTLSPGKNGVVKDVFAFGGAVENPEYLTPQ
GGAAPQPHPPPAFSPAfdnLYYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVPV

Explore Sequence Analysis Tools with

Job Dispatcher

EMBL's European Bioinformatics Institute

Job Dispatcher Help & Privacy Your Jobs

Feedback

Welcome to the Job Dispatcher website! If you need assistance or have feedback, please [contact us](#).



The Job Dispatcher at EMBL-EBI offers free access to a range of bioinformatics tools and biological datasets through its web and programmatic interfaces. It also powers various popular sequence analysis services hosted at the EMBL-EBI, including InterProScan, UniProt, and Ensembl Genomes.

Retrieve Your Job

Retrieve the job results

Example searches: [fasta-R20230421-121005-0548-43453433-p1m](#)

Search

Your Recent Jobs

Job ID	Title	Status	Last update
clustalo-I20250216-182615-0457-56121608-p1m	Clustal Omega's job		12 hours ago
clustalo-I20250212-124733-0062-24432412-p1m	Clustal Omega's job		4 days ago
clustalo-I20250212-124647-0344-23821832-p1m	Clustal Omega's job		4 days ago
clustalo-I20250212-124437-0000-74130528-p1m	Clustal Omega's job		4 days ago
clustalo-I20250212-124107-0392-75054802-p1m	Clustal Omega's job		4 days ago

More

Tool Categories



Pairwise Sequence Alignment

Identify regions of similarity between two biological sequences.

[Needle](#) | [Stretcher](#) | [GGSEARCH2SEQ](#) | [Water](#) | [Matcher](#) | More...



Sequence Similarity Search

Find sequences in databases based on similarity.

[NCBI BLAST](#) | [PSI-BLAST](#) | [FASTA](#) | [SSSEARCH](#) | [PSI-Search](#) | More...



Multiple Sequence Alignment

Identify conserved sequence patterns from multiple related sequences.

[Clustal Omega](#) | [MUSCLE](#) | [Kalign](#) | [MAFFT](#) | [T-Coffee](#) | More...



Sequence Translation

Emboss sequence translation and back translation tools.

[Transdecoder](#) | [Sixpack](#) | [Backtransdecoder](#) | [Backtranambig](#)

<https://www.ebi.ac.uk/jdispatcher/>

Clustal Omega

Multiple Sequence Alignment (MSA)

Job Dispatcher Help & Privacy Your Jobs Input form

Feedback

Welcome to the Job Dispatcher website! If you need assistance or have feedback, please [contact us](#).

X

Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more sequences. For the alignment of two sequences please instead use our pairwise sequence alignment tools. This tool can align up to 4000 sequences or a maximum file size of 4 MB.

Input sequence

Sequence Type

Protein DNA RNA

Paste your sequence here - or use the example sequence

Izberite datoteko Nobena datoteka ni izbrana

[Use the example](#) [Clear sequence](#) [More example inputs](#)

Parameters

OUTPUT FORMAT

ClustalW with character counts

[More options](#)

Submit

Title

Clustal Omega's job

Submit

If you use this service, please consider citing the following publication: [The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024](#). More information about this bioinformatics application can be found in its [bio.tools](#) record.

Please read the provided [Help & Privacy](#) before seeking help from our support staff. If you have any feedback or experienced any issues please let us know via [EMBL-EBI Support](#). Read our [Privacy Notice](#) if you are concerned with your privacy and how we handle personal information.

FASTA format

>name_of_protein

RKVVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP

>name_of_DNA

ACTGGTCTTAAAGCGATTACTGGTCTTAAAGCGATTACTGGTCTTAAAGC

sp P04626 ERBB2_HUMAN 3PP0_1 Chains	MELAALCRWGLLLALLPPGAESTQVCTGTDMLRLPASPETHLDMLRHLYQGCQVVQNL 60	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	ELTYLPTNASLFLQDIEVQGYVILAHNQVRQVPLQLRLRIVRGTLQFLEDNYALAVLDNG 120	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	DPLNNTTPVTGASPGLRELQLRLSLEILKGGVLIQRNQPLCYQDTILWKDIFHKNQLA 180	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	LTLIOTNRSRACHPCSPMCKGSRCWGESEDQCSLRTVCAGGCCARCKGPLPTDCHEQC 240	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	AAGCTGPKHSOCLALHFNHSGTCELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACP 300	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	YNYLSTDVGSTLVCPHLNQEVTAEDGTQRCEKCSKPCARVCYGLGMEHLREVRAVTSAN 360	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	IQEFAKGCKKIFGSLAFLPKESFDGDPASNTAPLQPEQLQFETLEEITGYLYISANPDSP 420	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	DLSVFQNLQVIRGRILHNGAYSLTLQGLGISMGLRSLRELGSGLAIIHHNTHLCFVHTV 480	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	PWDQLFRNPHQALLHTANRPEDECVGEGLACHQLCARGHCWGPPTQCVNCQFRLGQEC 540	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	VEECRVLQGLPREYYNARHCLPCHECPQPNQSVTCFGEADQCVAACAHYKDPFCVAC 600	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	PSGVKPDLSYMPWKFDDEEGACQPCPINCTHSCVLDOKGCPAEQRASPLTSIISAVG 660	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	ILLVVVLGVVFGILIKRRQQKIRKYTMRRLQETELVEPLTPSGAMPNQAQMRLKETEL 720	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	MSGAAAPNQALLRLKETEL *** *** :***** 19	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	RKVVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 780	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	RKVVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 79	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	YVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 840	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	YVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 139	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	LVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFT 900	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	LVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFT 199	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPPICTIDVYIMVKCWM 960	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPPICTIDVYIMVKCWM 259	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLLEDDDMGDLVDA 1020	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLLEDDDMGDLVDA 319	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	EEYLVPQQGFFCPDPAPGAGGMVHHRHRSSTSRSGGDLTLGLEPSEEAPRSPLAPSEG 1080	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	EEYLVPQQGAAASHHHHHH----- 338	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	***** .	0
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	AGSDVFDGLGMGAAGKGLQSLPTHQSPQLQRYSEOPTV/PLPSETDGYVAPLTCSPQPEYV 1140	338
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	NQPDVRPQPPSPREGLPAAAGTALERPKTLSPGKNGVVKDVFAGGAPEVPEYLTQ 1200	338
sp P04626 ERBB2_HUMAN 3PP0_1 Chains	GGAAPQPHPPPAFSPAFDNLYWQDQPPERGAPPSTFKGPTAENPEYGLDVPV 1265	338

sp|P04626|ERBB2_HUMAN
3PP0_1|ChainsILLVVVLGVVFGILIKRRQQKIRKYTMRRLQETELVEPLTPSGAMPNQAQMRLKETEL 720
-----MSGAAAPNQALLRLKETEL 19
*** *** :*****sp|P04626|ERBB2_HUMAN
3PP0_1|ChainsRKVVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 780
RKVVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 79
*****sp|P04626|ERBB2_HUMAN
3PP0_1|ChainsYVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 840
YVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 139
*****sp|P04626|ERBB2_HUMAN
3PP0_1|ChainsLVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFT 900
LVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVPIKWMALESILRRRFT 199
*****sp|P04626|ERBB2_HUMAN
3PP0_1|ChainsHQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPPICTIDVYIMVKCWM 960
HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPPICTIDVYIMVKCWM 259
*****sp|P04626|ERBB2_HUMAN
3PP0_1|ChainsIDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLLEDDDMGDLVDA 1020
IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLLEDDDMGDLVDA 319
*****sp|P04626|ERBB2_HUMAN
3PP0_1|ChainsEEYLVPQQGFFCPDPAPGAGGMVHHRHRSSTSRSGGDLTLGLEPSEEAPRSPLAPSEG 1080
EEYLVPQQGAAASHHHHH----- 338

>HER2_alphaFold

PSGAMPNQAQMRLKETELRKVVLGSGAFGTVYKGIWIPDGENVKIPVAIKVLRENTSPK
ANKEILDEAYVMAGVGSPYVSRLLGICLTSTVQLVTQLMPYGCLLDHVRENRRGLGSQDLL
NWCMQIAKGMSYLEDVRLVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGG
KVPIKWMALESILRRRFTHQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQ
PICTIDVYIMVKCWMIDSECRPRFRELVSEFSRMARDPQRFVVIQNEDLGPASPLDSTFYR
SLLEDDDMGDLVDAEYLYVPQQGFFCP

We use this sequence for modeling with AlphaFold Server.



3PP0_1|Chains
sp|P04626|ERBB2_HUMAN
chain

-----MSGAAPNQALLRILKETEL 19
ILVVVLGVVFGILIKRRQQKIRKYTMRRLLQETELVEPLTPSGAMPNQAMRILKETEL 720
-----PSGAMPNQAMRILKETEL 19
*** **** :*****

3PP0_1|Chains
sp|P04626|ERBB2_HUMAN
chain

RKVVLGSGAFTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 79
RKVVLGSGAFTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 780
RKVVLGSGAFTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 79

3PP0_1|Chains
sp|P04626|ERBB2_HUMAN
chain

YVSRLLGICLTSTVQLTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 139
YVSRLLGICLTSTVQLTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 840
YVSRLLGICLTSTVQLTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 139

3PP0_1|Chains
sp|P04626|ERBB2_HUMAN
chain

LVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVKPIKWMALESILRRRFT 199
LVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVKPIKWMALESILRRRFT 900
LVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVKPIKWMALESILRRRFT 199

3PP0_1|Chains
sp|P04626|ERBB2_HUMAN
chain

HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICHTDVMIMVKCWM 259
HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICHTDVMIMVKCWM 960
HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICHTDVMIMVKCWM 259

3PP0_1|Chains
sp|P04626|ERBB2_HUMAN
chain

IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLEDDDMGDLVDA 319
IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLEDDDMGDLVDA 1020
IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLEDDDMGDLVDA 319

3PP0_1|Chains
sp|P04626|ERBB2_HUMAN
chain

EEYLVPQQGAAASHHHHH----- 338
EEYLVPQQGFFCPDPAPGAGGMVHHRHSSSTRSGGDLTGLEPSEEAPRSPLAPSEG 1080
EEYLVPQQGFFCP----- 332
***** .

sp|P04626|ERBB2_HUMAN
HER2_alphafold

ILLVVVLGVVFGILIKRRQQKIRKYTMRRLLQETELVEPLTPSGAMPNQAMRILKETEL 720
-----PSGAMPNQAMRILKETEL 19

sp|P04626|ERBB2_HUMAN
HER2_alphafold

RKVVLGSGAFTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 780
RKVVLGSGAFTVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVMAGVGSP 79

sp|P04626|ERBB2_HUMAN
HER2_alphafold

YVSRLLGICLTSTVQLTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 840
YVSRLLGICLTSTVQLTQLMPYGCLLDHVRENRRGLGSQDLLNWCQIAKGMSYLEDVR 139

sp|P04626|ERBB2_HUMAN
HER2_alphafold

LVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVKPIKWMALESILRRRFT 900
LVHRDLAARNVLVKSPNHHVKITDFGLARLLDIDETEYHADGGKVKPIKWMALESILRRRFT 199

sp|P04626|ERBB2_HUMAN
HER2_alphafold

HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICHTDVMIMVKCWM 960
HQSDVWSYGVTVWELMTFGAKPYDGIPAREIPDLLEKGERLPQPPICHTDVMIMVKCWM 259

sp|P04626|ERBB2_HUMAN
HER2_alphafold

IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLEDDDMGDLVDA 1020
IDSECRPRFRELVSEFSRMARDPQRFVVIQNEDELGPASPLDSTFYRSLEDDDMGDLVDA 319

sp|P04626|ERBB2_HUMAN
HER2_alphafold

EEYLVPQQGFFCPDPAPGAGGMVHHRHSSSTRSGGDLTGLEPSEEAPRSPLAPSEG 1080
EEYLVPQQGFFCP----- 332

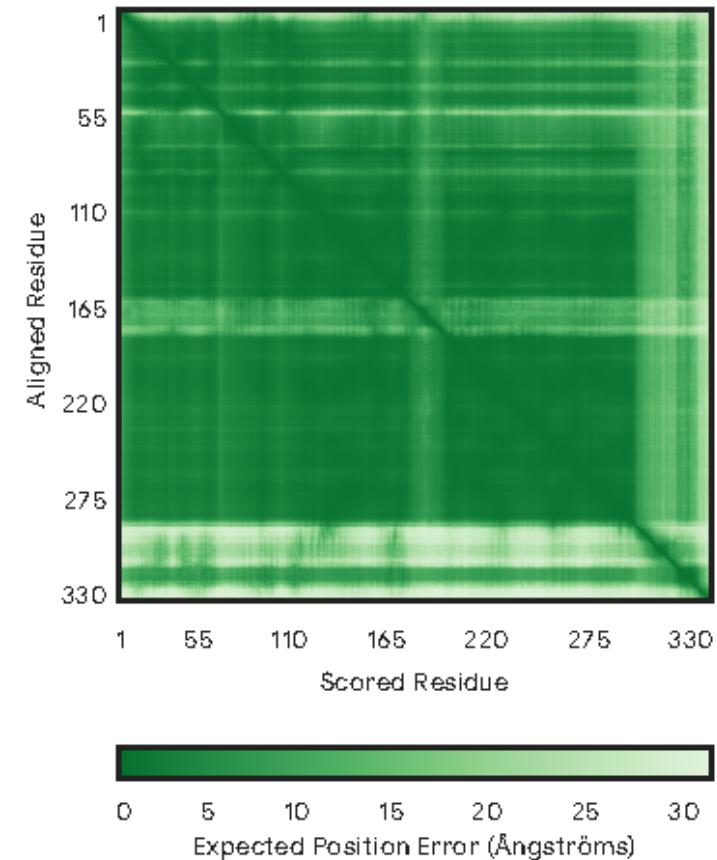
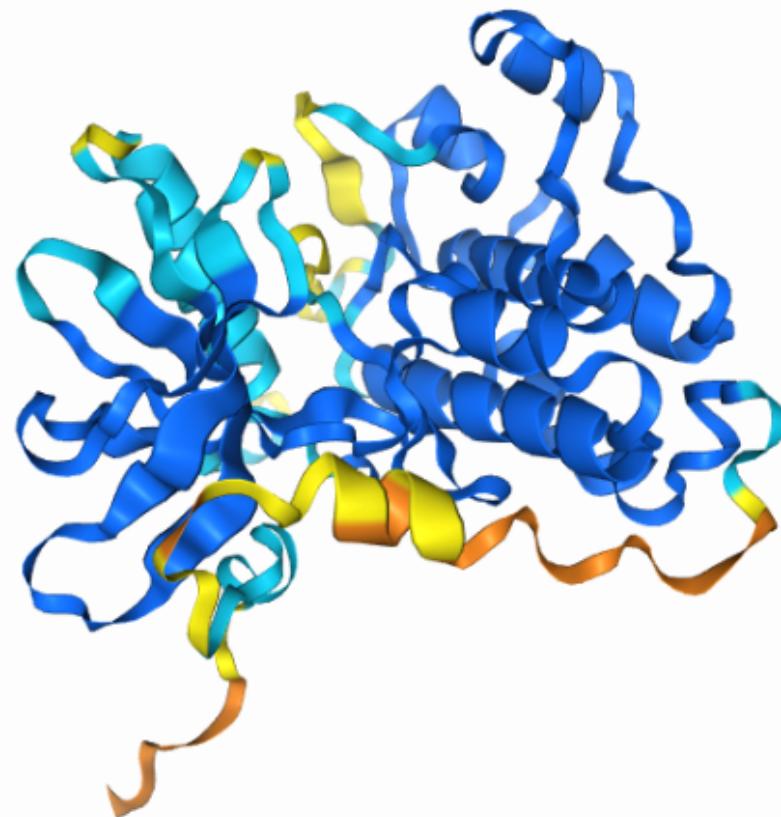
[← Back](#)[Download](#)[Clone and reuse](#)[Feedback on structure](#)

Very high (pIDDT > 90)

Confident (90 > pIDDT > 70)

Low (70 > pIDDT > 50)

Very low (pIDDT < 50)

ipTM = - pTM = 0.87 [learn more](#)

0 5 10 15 20 25 30
Expected Position Error (Ångströms)

Save .zip file → unzip in to the folder with this files:

-  fold_2025_02_12_13_52_full_data_0.json
-  fold_2025_02_12_13_52_full_data_1.json
-  fold_2025_02_12_13_52_full_data_2.json
-  fold_2025_02_12_13_52_full_data_3.json
-  fold_2025_02_12_13_52_full_data_4.json
-  fold_2025_02_12_13_52_job_request.json
-  fold_2025_02_12_13_52_model_0.cif
-  fold_2025_02_12_13_52_model_1.cif
-  fold_2025_02_12_13_52_model_2.cif
-  fold_2025_02_12_13_52_model_3.cif
-  fold_2025_02_12_13_52_model_4.cif
-  fold_2025_02_12_13_52_summary_confidences_0.json
-  fold_2025_02_12_13_52_summary_confidences_1.json
-  fold_2025_02_12_13_52_summary_confidences_2.json
-  fold_2025_02_12_13_52_summary_confidences_3.json
-  fold_2025_02_12_13_52_summary_confidences_4.json
-  terms_of_use.md

MOLECULAR DOCKING

Online with CB-Dock2





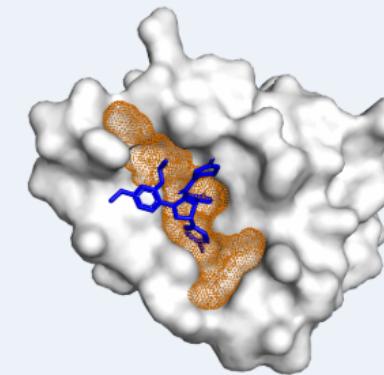
CB-Dock2

Cavity-detection guided Blind Docking

[Home](#)[Dock](#)[Example](#)[Manual](#)[Contact](#)

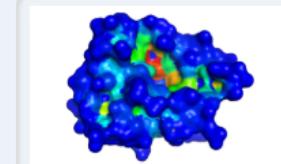
CB-Dock2 is an improved version of the [CB-Dock](#) server for protein-ligand blind docking, integrating cavity detection, docking and homologous template fitting.

Given the three-dimensional (3D) structure of a protein and a ligand, we can predict their binding sites and affinity for computer-aided drug discovery.

[Get Started >>](#)

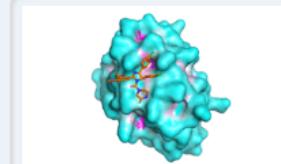
Features

CB-Dock2 combines multiple features to facilitate molecular docking.



Search Cavity

Detecting cavities on proteins based on clustering of solvent-accessible surface.



Structure-based Docking

Molecular docking at the detected candidate pockets with AutoDock



Template-based Docking

Molecular docking based on the pattern of homologous templates



Upload Protein

Select a protein file (pdb format) [Example: MDM2.pdb \(PDB: 4HG7\)](#) [Submitted Protein: HER2.pdb](#)

SEQRES	1	A	332	PRO	SER	GLY	ALA	NET	PRO	ASN	GLN	ALA	GLN	MET	ARG	ILE
SEQRES	2	A	332	LEU	LYS	GLY	THR	GLU	LEU	ARG	LYS	VAL	LYS	VAL	LEU	GLY
SEQRES	3	A	332	SER	GLY	ALA	PHE	GLY	THR	VAL	TYR	LYS	GLY	ILE	TRP	ILE
SEQRES	4	A	332	PRO	ASP	GLY	GLU	ASN	VAL	LYS	ILE	PRO	VAL	ALA	ILE	LYS
SEQRES	5	A	332	VAL	LEU	ARG	GLU	ASN	THR	SER	PRO	LYS	ALA	ASN	LYS	GLU
SEQRES	6	A	332	ILE	LEU	ASP	GLU	ALA	TYR	VAL	MET	ALA	GLY	VAL	GLY	SER
SEQRES	7	A	332	PRO	TYR	VAL	SER	ARG	LEU	LEU	GLY	ILE	CYS	LEU	THR	SER
SEQRES	8	A	332	THR	VAL	GLN	LEU	VAL	THR	GLN	LEU	MET	PRO	TYR	GLY	CYS
SEQRES	9	A	332	LEU	LEU	ASP	HIS	VAL	ARG	GLU	ASN	ARG	GLY	ARG	LEU	GLY
SEQRES	10	A	332	SER	GLN	ASP	LEU	LEU	ASN	TRP	CYS	MET	GLN	ILE	ALA	LYS

Protein uploaded!

Remove file

Upload Ligand

Select a ligand file (mol2, mol, sdf, pdb format) [Example: Nutlin-3a.sdf] [Submitted Ligand: Gefitinib.sdf]

[Upload Local](#) [Draw Ligand](#)

123631
-OEChem-02162513423D

Draw Ligand

-OChem-02162513423D

```

55 58 0 0 0 0 0 0 0999 V2000
-5.7668 0.3767 2.1969 C1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-6.5134 -0.4026 -0.5639 F 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.5666 3.9138 -1.2756 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3.4187 0.9841 0.7617 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5.1811 -0.9472 -0.1795 O 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.8065 3.1706 0.1093 N 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```

Ligand uploaded!

Remove file

 More parameters 

Other customizable parameters

Enter your email address to receive the results (optional)

Search Cavities

Auto Blind Docking



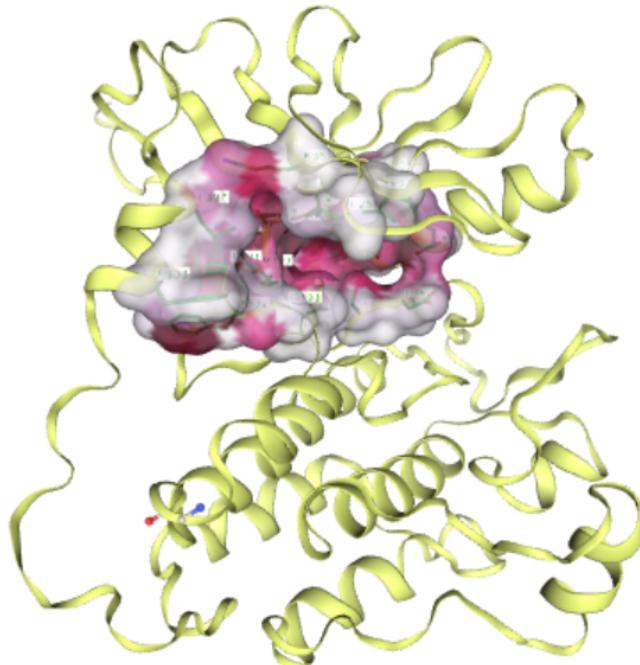
Structure-based cavity detection

Detected pockets from the input structure (Submitted Protein: [HER2.pdb](#))

Tips: Click the table below to view the pocket information. Select the checkbox in the table below to mark the pockets you are interested in for blind docking. Click the residue in the sequence list to view the residue you are interested in.

Query

A 10 20 30 40 50 60 70 80 90 100 110 120
PSGAMPNQAQ MRILKETELR KVKV**L**GSGAF GT**V**KGIWIP DGENVKIP**V**A IK**V**LRENTSP KANKEIL**E**A YV**M**AGVGSPY VS**R**LLGICLT ST**V**QL**T**Q**L**M PYG**C**LLDHVR ENR**G**R**L**GSQ**D** LLNWC**Q**I
.130 140 150 160 170 180 190 200 210 220 230 240
AK GMSYLEDV**R**L VRHD**L**AARN**V** LV**K**SPN**H**V**K**I TDFGLARLLD IDETEYHADG GKVPIKWMAL ESILRRRFTH QSDVWSYG**V**T V**W**ELMTFGAK PYDGIPAREI PDLLEKGERL PQPICTID
250 260 270 280 290 300 310 320 330
V YMIMVKC**W**MI DSEC**R**PRF**R**E LV**S**FSRM**A**R DPQRFV**V**IQ**N** EDLG**P**ASPLD ST**F**YRS**L**ED DDMGDLVDAE EYLVPQQGFF CP



Center Fullscreen Show Surface Hide Pocket Hide Label

Docking at the selected pockets

Tips: Select the checkbox in the table above to mark the pockets you are interested in for blind docking.

Submitted Protein
[HER2.pdb](#)

Submitted Ligand
[Gefitinib.sdf](#)

Selected CurPockets
 C1 C2 C3 C4 C5

BlindDock

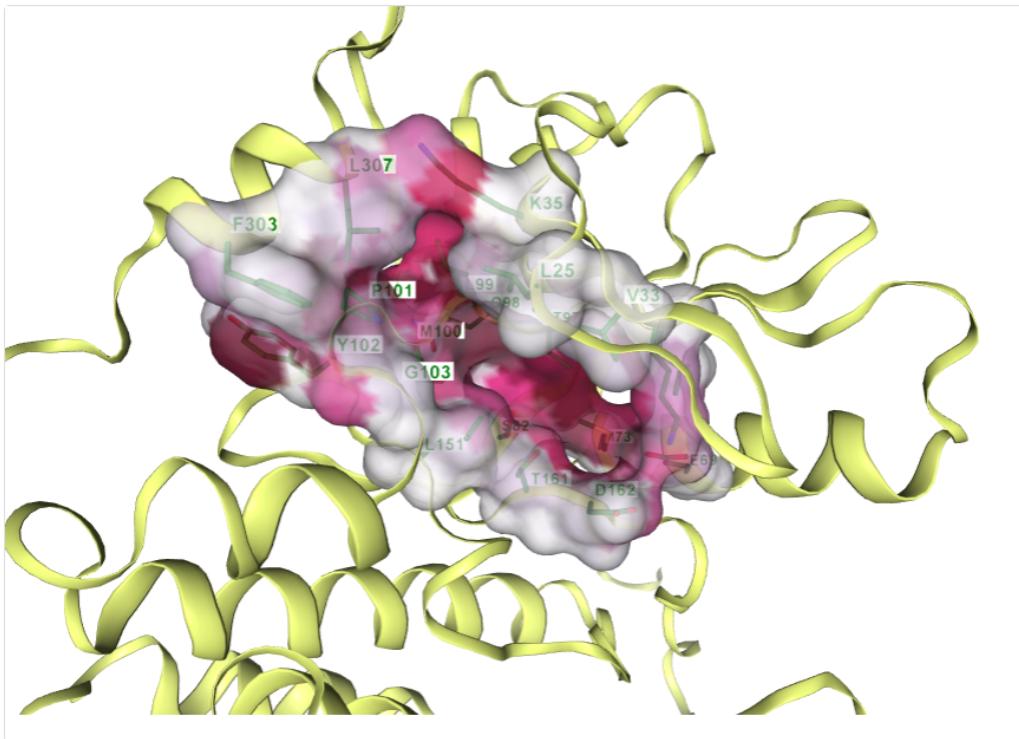


Structure-based cavity detection

Detected pockets from the input structure (Submitted Protein: [HER2.pdb](#))

Tips: Click the table below to view the pocket information. Select the checkbox in the table below to mark the pockets you are interested in for blind docking. Click the residue in the sequence list to view the residue you are interested in.

Query10.....20.....30.....40.....50.....60.....70.....80.....90.....100.....110.....120.....
A PSGAMPNQAQ MRILKETELR KVVKLIGSGAF GTVKGIWIP DGENVKIPVA IKVLRENTSP KANKEILDEA YV¹³⁰MAGVGSPY VSRLLGICLT STVQ¹⁴⁰VTQLM PYGC¹⁵⁰CLLDHVR ENRGRLGSQD LLNWCMQIA
130.....140.....150.....160.....170.....180.....190.....200.....210.....220.....230.....240.....
K GMSYLEDVRL VHRDLAARNV LV²⁵⁰KSPNKHVKI TD²⁶⁰GLARLLD IDETEYHADG GK²⁷⁰PIKWMAL ESILRRRFTH QSDVWSYGV²⁸⁰T V²⁹⁰WELMTFGAK PYDGIPAREI PDLLEKGERL PQPICTID
250.....260.....270.....280.....290.....300.....310.....320.....330..
V YMIMVKCWM³⁴⁰I DSEC³⁵⁰PRPFRE LVSEFSRM³⁶⁰AR DPQRFVVIQN EDLG³⁷⁰PASPLD ST³⁸⁰YRS³⁹⁰LLED DDMGDLVDAE EYLVPQQGFF CP



CurPocket ID	Cavity volume (Å ³)	Center (x, y, z)	Cavity size (x, y, z)
C1	2106	-15, -3, -7	22, 18, 21
C2	934	10, 9, 5	14, 14, 20
C3	616	5, -7, 18	13, 17, 16
C4	445	3, 3, -8	16, 8, 13
C5	328	15, -6, 1	11, 13, 14

[Download CurPockets](#)

Docking at the selected pockets

Tips: Select the checkbox in the table above to mark the pockets you are interested in for blind docking.

Submitted Protein
[HER2.pdb](#)

Submitted Ligand
[Gefitinib.sdf](#)

Selected CurPockets

C1 C2 C3 C4 C5

BlindDock





Upload Protein

Select a protein file (pdb format) [Example: MDM2.pdb (PDB: 4HG7)] [Submitted Protein: 3PP0_he2_ok.pdb]

```

REMARK BIOVIA PDB file
REMARK Created: 2025-02-16T21:12:53Z
CRYST1 48.705 78.951 152.675 90.00 90.00 90.00 P212121
ATOM   1 N  ALA A 706      28.242 41.311 27.511 1.00 63.83  N1+
ATOM   2 CA  ALA A 706     29.299 40.259 27.561 1.00 63.84  C
ATOM   3 C   ALA A 706     29.455 39.562 26.209 1.00 62.05  C
ATOM   4 O   ALA A 706     28.452 39.154 25.611 1.00 60.43  O
ATOM   5 CB  ALA A 706     28.973 39.234 28.639 1.00 63.85  C
ATOM   6 N  PRO A 707     30.707 39.435 25.715 1.00 60.78  N
ATOM   7 CA  PRO A 707     30.951 38.595 24.535 1.00 59.11  C

```

Protein uploaded!

Attention: The chains of the uploaded protein are broken at position A879(ALA), A882(GLY)!

Remove file

Upload Ligand

Select a ligand file (mol2, mol, sdf, pdb format) [Example: Nutlin-3a.sdf] [Submitted Ligand: LAPATINIB.sdf]

Upload Local Draw Ligand

208908
-OChem-02162513423D

66 70 0 0 0 0

Ligand uploaded!

Remove file

 More parameters 

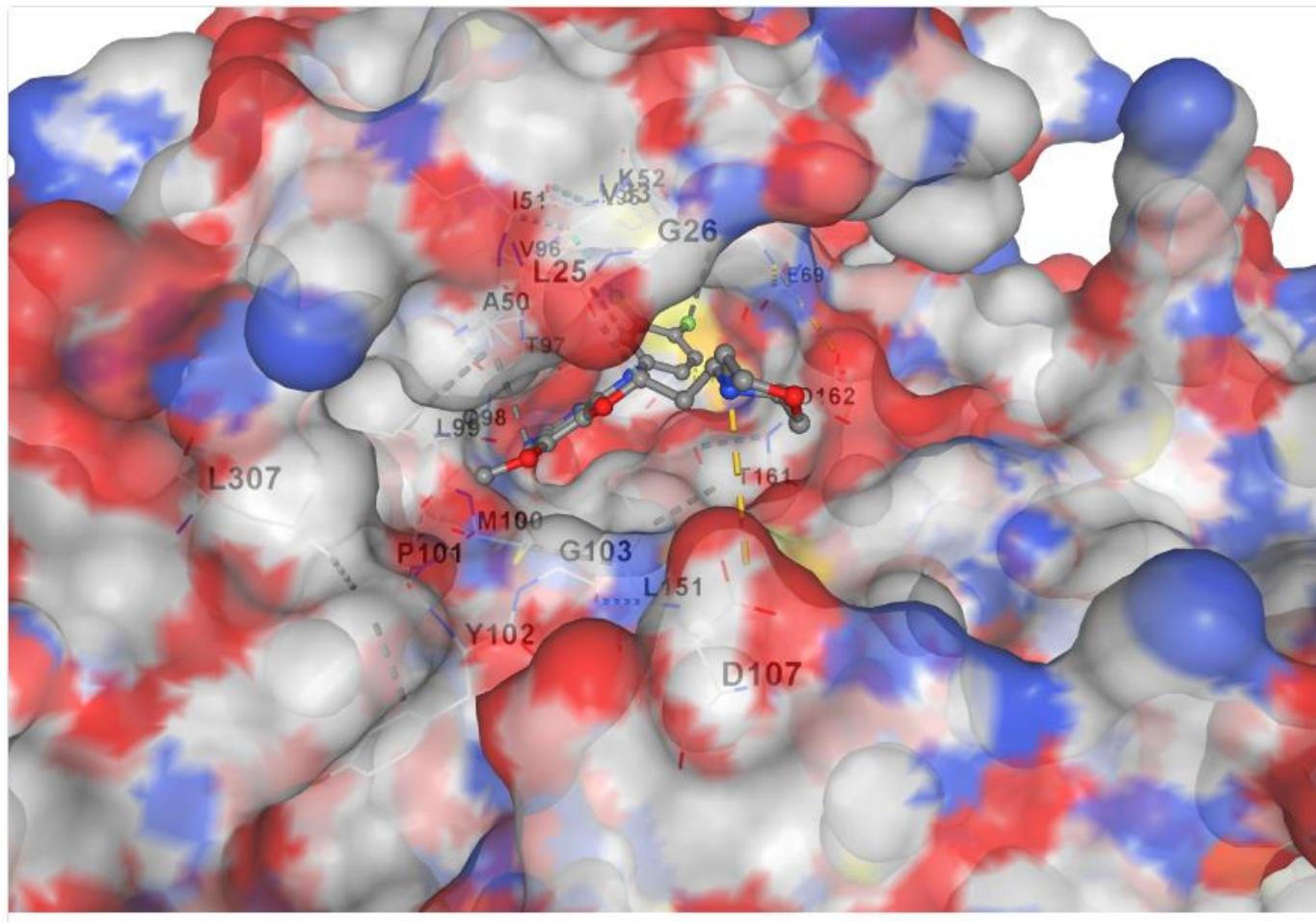
Other customizable parameters

Enter your email address to receive the results (optional)

Search Cavities

Auto Blind Docking



Docking with homologous protein complex guidanceSubmitted Protein
[HER2.pdb](#)Submitted Ligand
[Gefitinib.sdf](#)Detected FitPockets
1[Hide](#)

Center [Fullscreen](#) [Show Template\(Lig\)](#) [Ligand Style](#) [Color Ligand](#) [Receptor Style](#)
[Color Receptor](#)

FitPocket ID	FitDock score	Template ID	Contact residues	Download
○ F1	-7.3	t5	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
○ F1	-7.2	t1	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
○ F1	-7.1	t2	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
○ F1	-7.0	t3	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
○ F1	-6.8	t4	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]

Pocket: F1 & Template: t5 & Score: -7.3 x

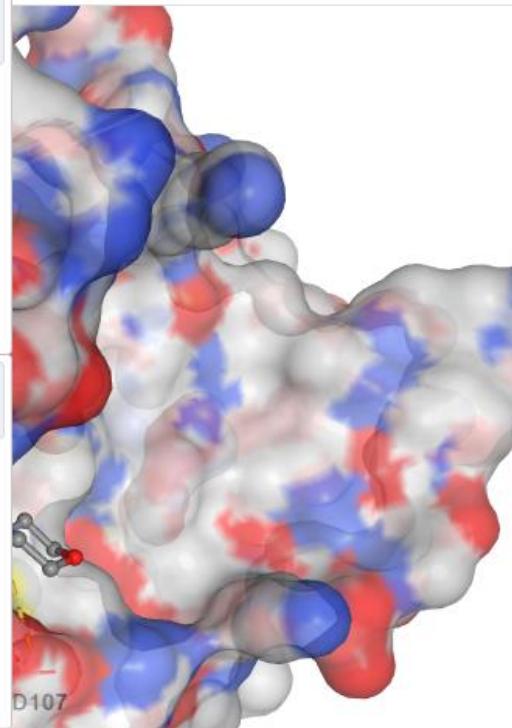
Chain A: LEU25 GLY26 VAL33 ALA50 ILE51 LYS52
GLU69 MET73 LEU95 VAL96
THR97 GLN98 LEU99 MET100 PRO101 TYR102
GLY103 ASP107 LEU151 THR161
ASP162 LEU307

Submitted Protein
[HER2.pdb](#)Submitted Ligand
[Gefitinib.sdf](#)Detected FitPockets
1

Hide

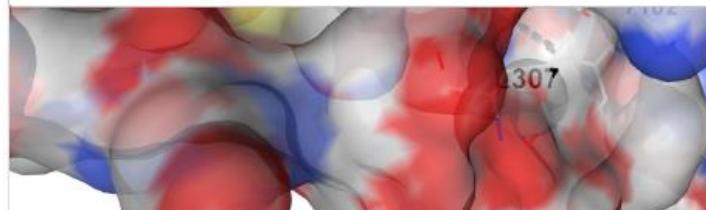
Pocket: F1 & Template: t1 & Score: -7.2

Chain A: LEU25 GLY26 VAL33 ALA50 ILE51 LYS52
 GLU69 MET73 LEU95 VAL96
 THR97 GLN98 LEU99 MET100 PRO101 GLY103
 CYS104 ASP107 LEU151 THR161
 ASP162



Pocket: F1 & Template: t2 & Score: -7.1

Chain A: LEU25 GLY26 VAL33 ALA50 ILE51 LYS52
 GLU69 MET73 LEU95 VAL96
 THR97 GLN98 LEU99 MET100 PRO101 GLY103
 CYS104 ASP107 LEU151 THR161
 ASP162



Pocket: F1 & Template: t3 & Score: -7.0

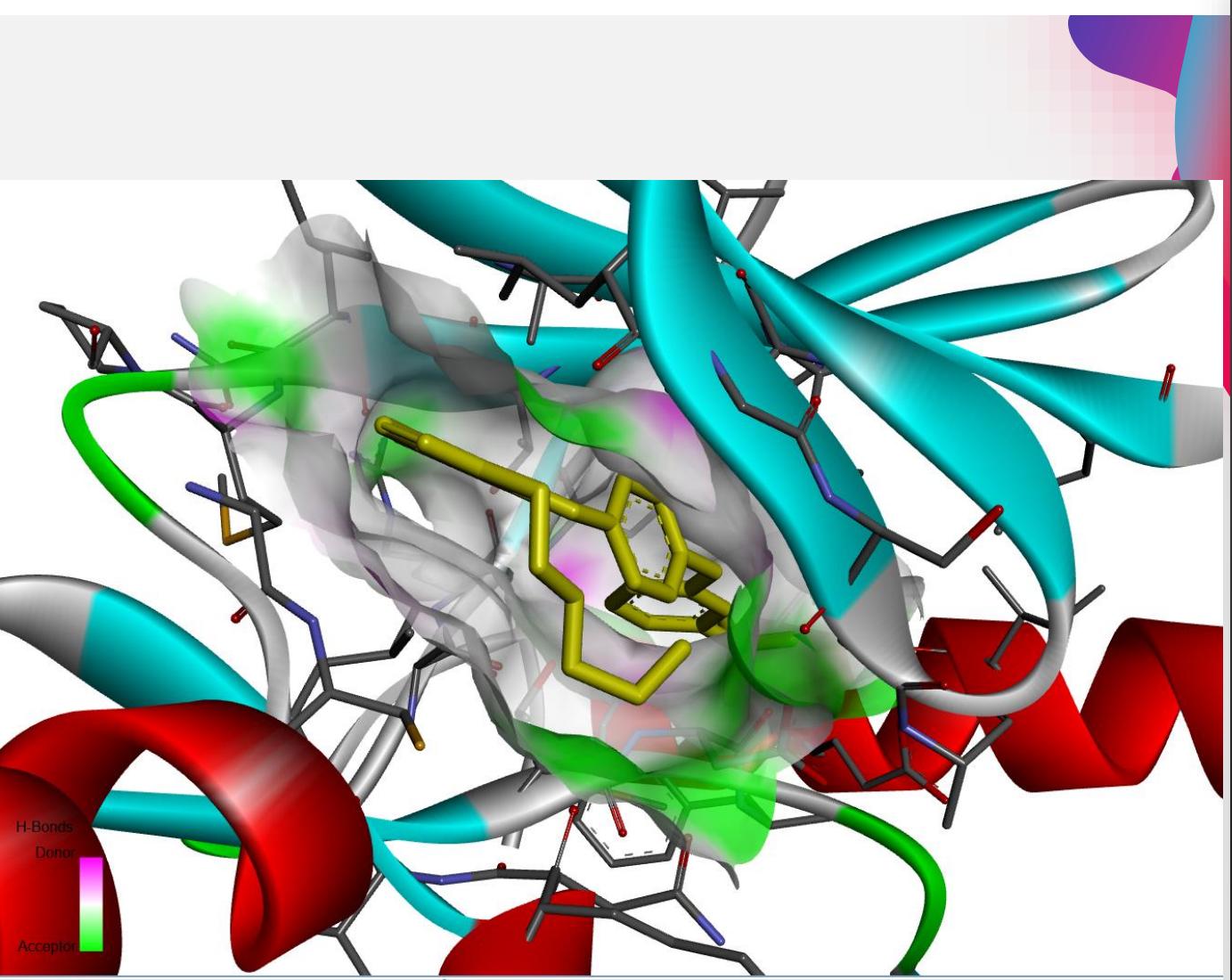
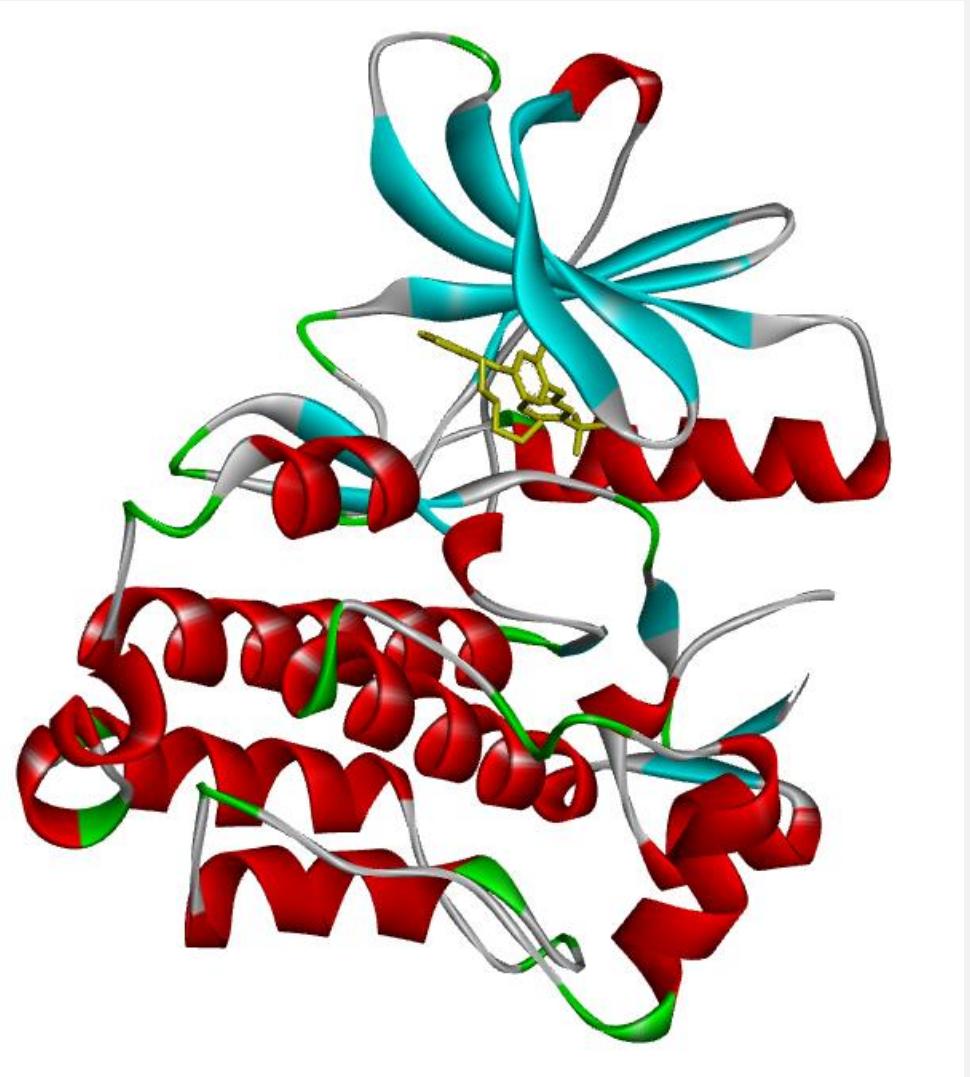
Chain A: LEU25 GLY26 VAL33 ALA50 ILE51 LYS52
 GLU69 MET73 LEU95 VAL96
 THR97 GLN98 LEU99 MET100 PRO101 TYR102
 GLY103 CYS104 ASP107 LEU151
 THR161 ASP162

Center Fullscreen Show Template(Lig) Ligand Style
 Color Receptor

FitPocket ID	FitDock score	Template ID	Contact residues	Download
F1	-7.3	t5	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F1	-7.2	t1	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F1	-7.1	t2	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F1	-7.0	t3	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]
F1	-6.8	t4	View	Ligand [MOL2] , [PDB] Protein-Ligand [PDB]

Pocket: F1 & Template: t4 & Score: -6.8

Chain A: LEU25 GLY26 VAL33 ALA50 LYS52
 GLU69 MET73 LEU95 THR97 GLN98
 LEU99 MET100 PRO101 TYR102 GLY103 CYS104
 ASP107 LEU151 THR161 ASP162
 LEU307



3PP00

