Protein Data Bank and Structure Display with PyMOL

Ching-Shu Suen (孫慶姝)
Supervisor: Dr. Ming-Jing Hwang (黃明經)

Bioinformatics Service Support
BioIT, IBMS (生醫所)
Academia Sinica

Outline

- About the RCSB PDB and the PDB Archive
- Searching for your structures
- Looking at structures
- Advanced applications in PDB
  - Structure visualization – NGL & JSmol viewers
  - Structure comparison -- jFATCAT-rigid
  - Drug target mapping --DrugBank
  - Protein-ligand interaction -- Ligand Explorer
- PyMOL Basic Applications
- PyMOL Advanced Applications

Worldwide Protein Data Bank (wwPDB)

- RCSB PDB (USA) http://www.rcsb.org/pdb/

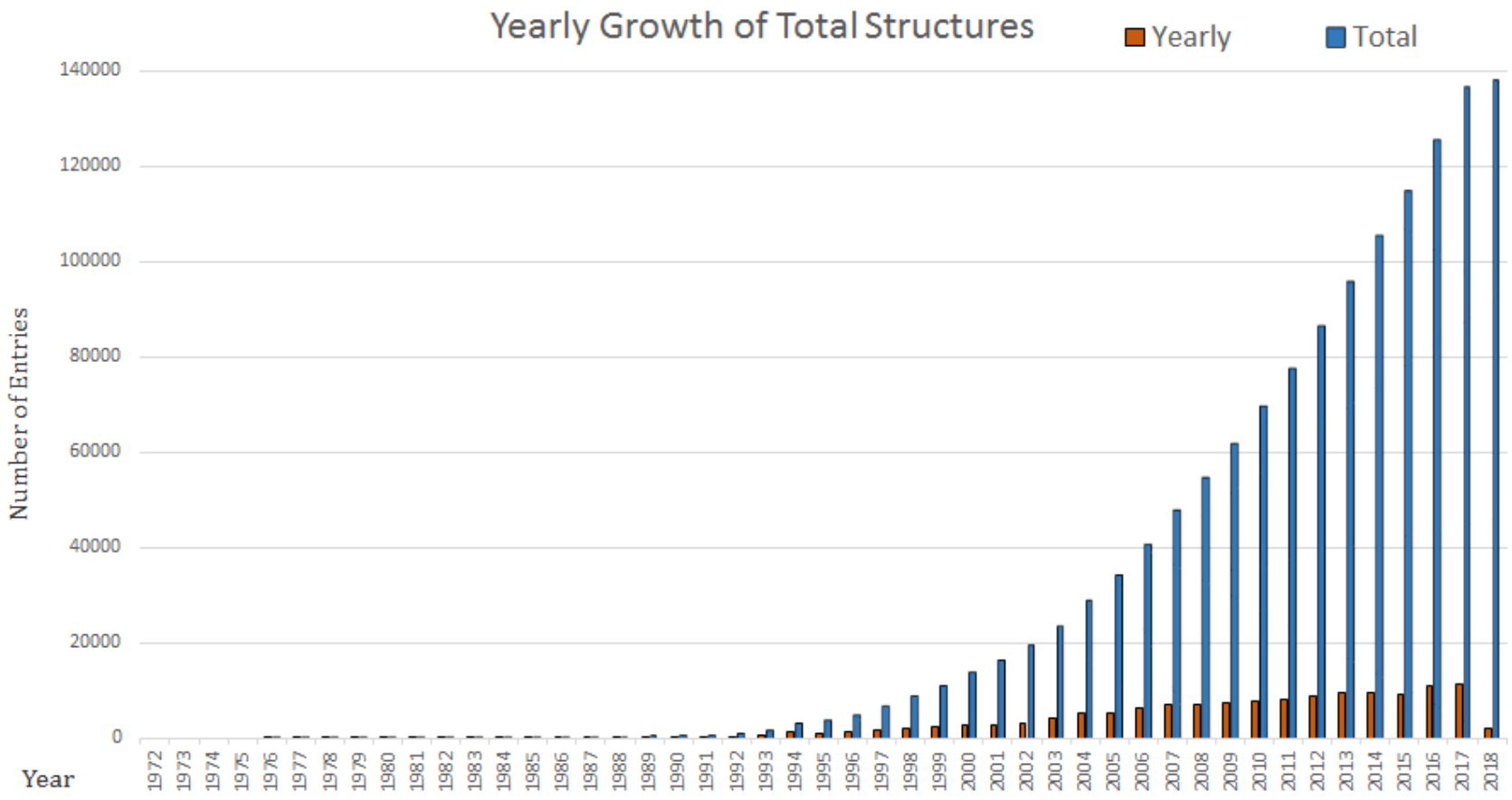
  (*Nucleic Acids Res.* 2017 **45**:D271-D281)

  - Established in 1971 at Brookhaven National Laboratory (BNL)
  - Taken over in 1998 by Research Collaboratory for Structural Bioinformatics
- PDBe (Europe) http://www.ebi.ac.uk/pdbe/

  (*Nucleic Acids Res.* 2018 **46**:D486-D492)
- PDBj (Japan) http://pdbj.org/

  (*Nucleic Acids Res.* 2017 **45**:D282-D288)
- BMRB (USA) http://www.bmrb.wisc.edu/
  - Biological Magnetic Resonance Bank (BioMagResBank)

RCSB PDB Statistics (1)



RCSB PDB Statistics (2)

- Statistics

Exp. Method	Proteins	Nucleic Acids	Protein / NA complexes	Other	Total
X-ray	116,115	1,916	5,922	10	123,963
NMR	10,660	1,236	249	8	12,153
Electron Microscope	1,459	31	506	0	1,996
Other	210	4	6	13	233
Multi Method	112	4	2	1	119
Total	128,556	3,191	6,685	32	138,464

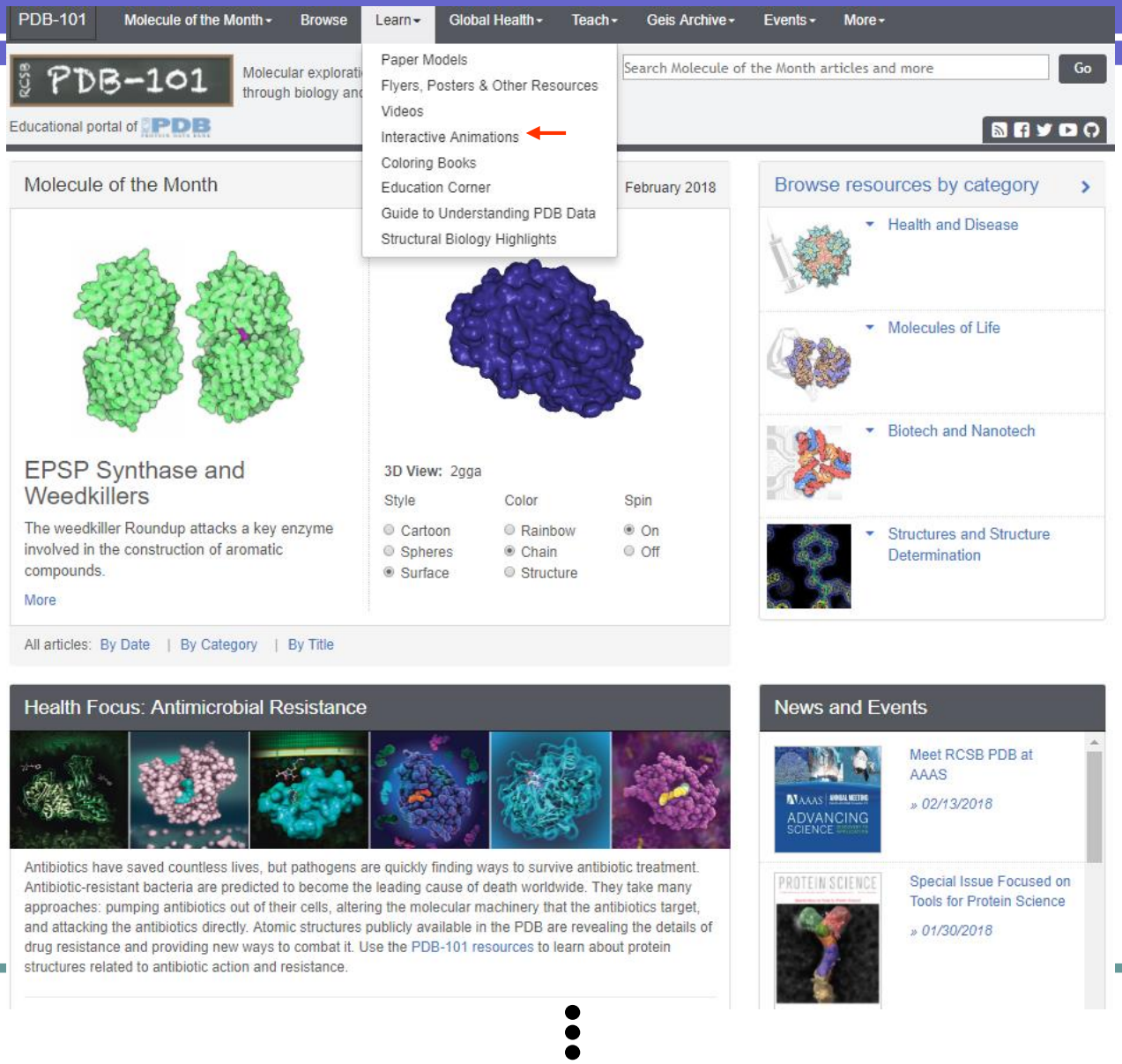
RCSB PDB (http://www.rcsb.org/)

Top bar menu

Education
PDB-101

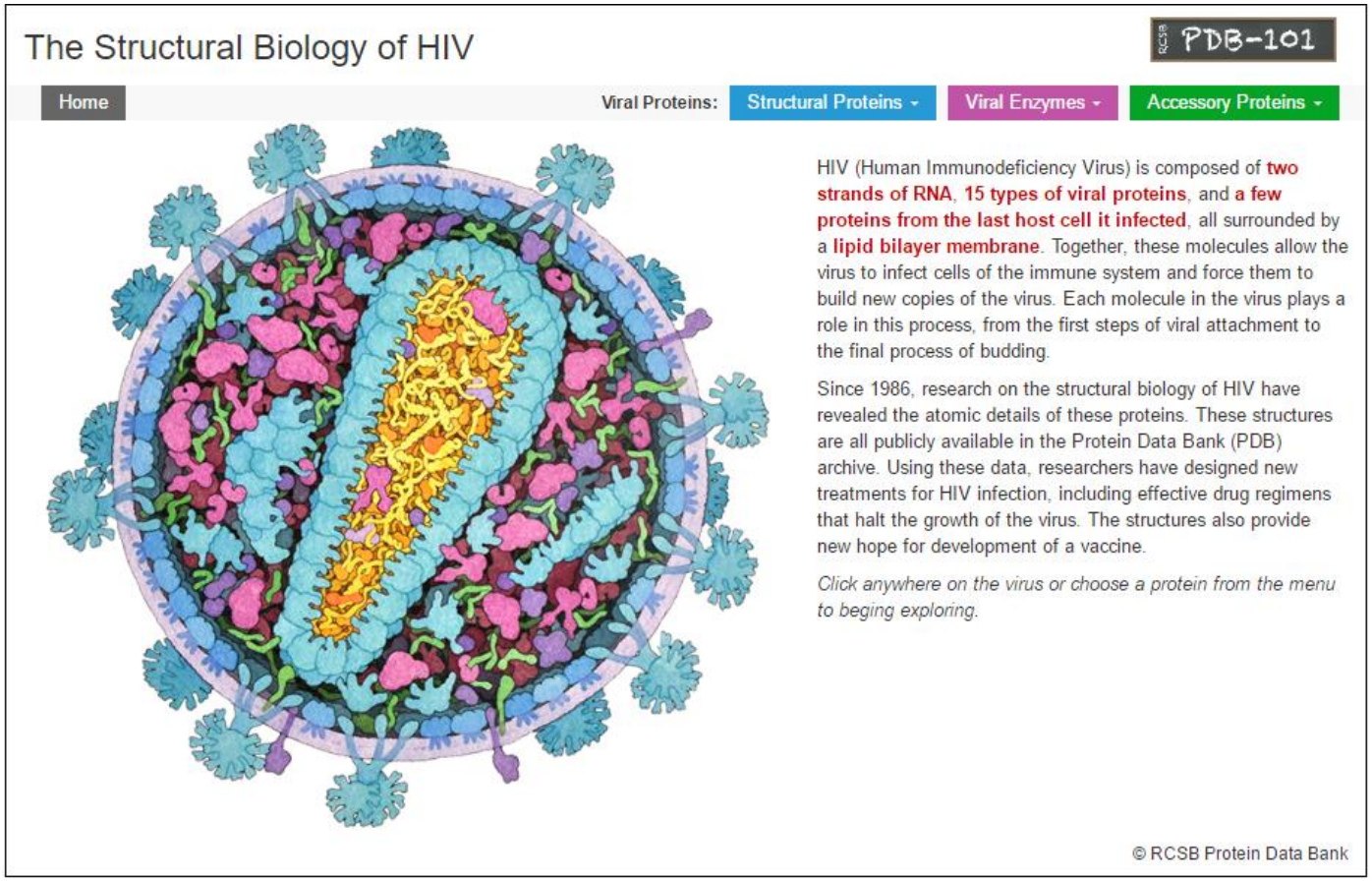
Query area
Enter what you want to know …

Education portal - PDB-101

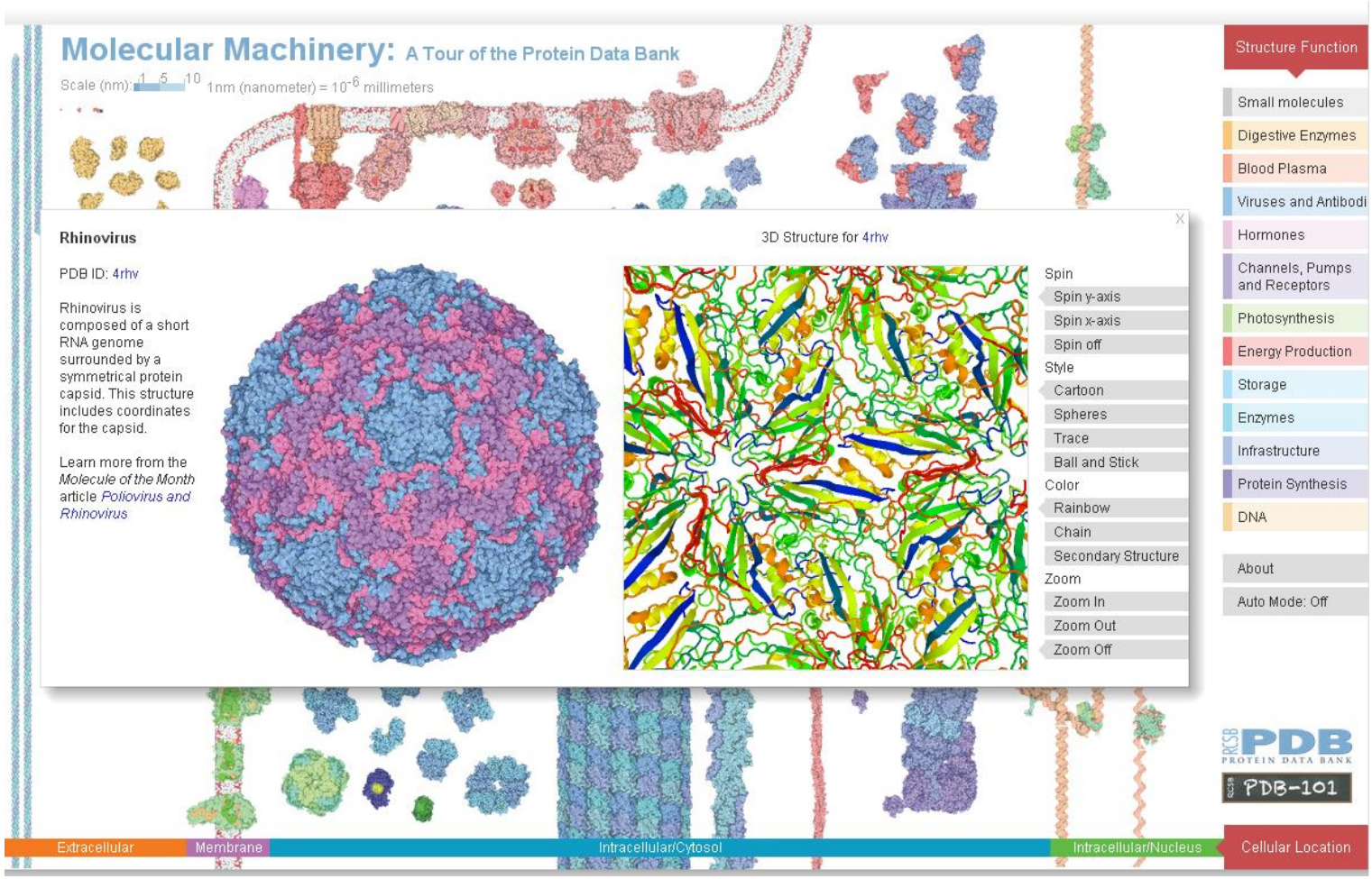


Take an Interactive Tour of the PDB (1)

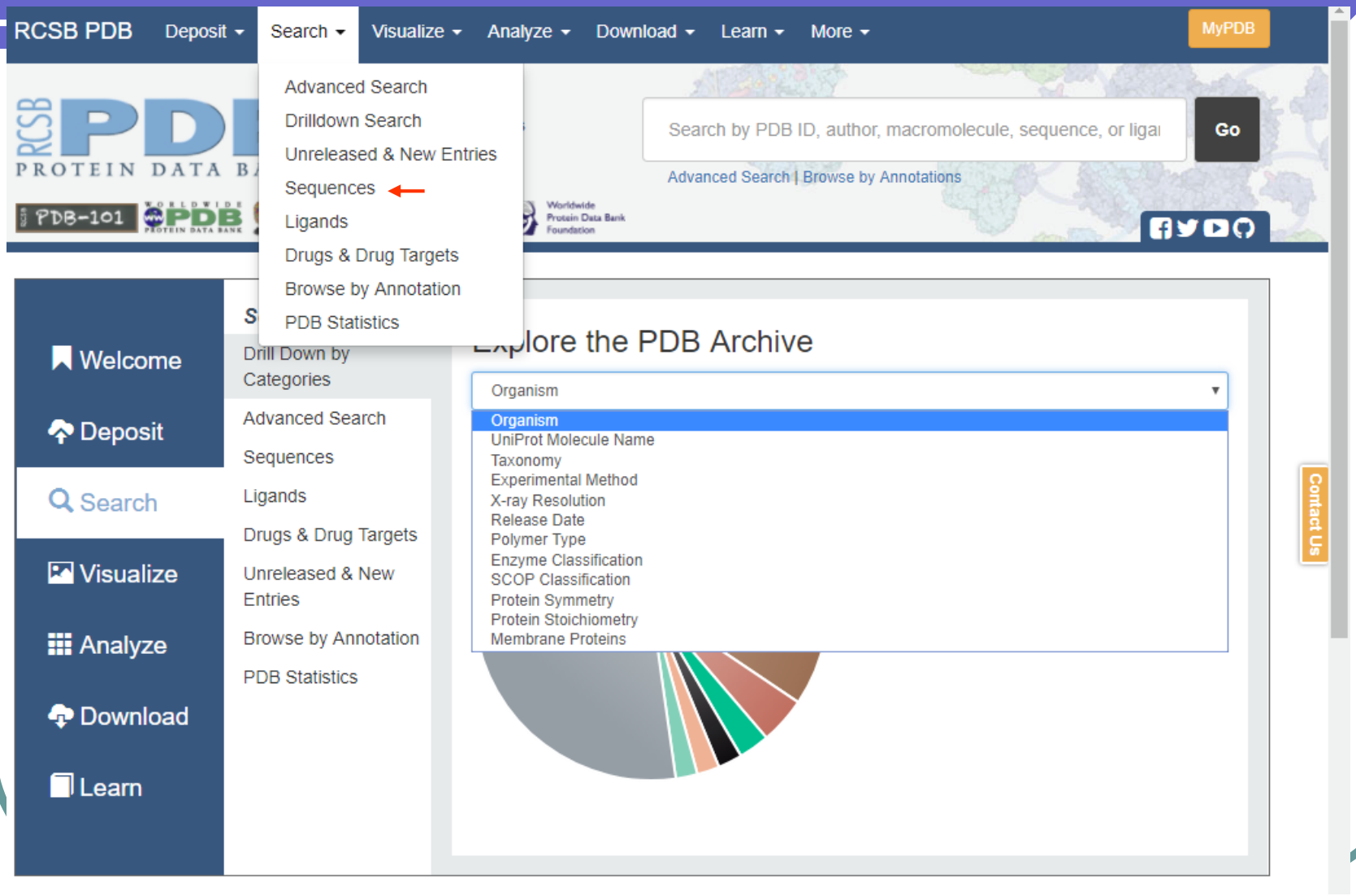
- Interactive Animations



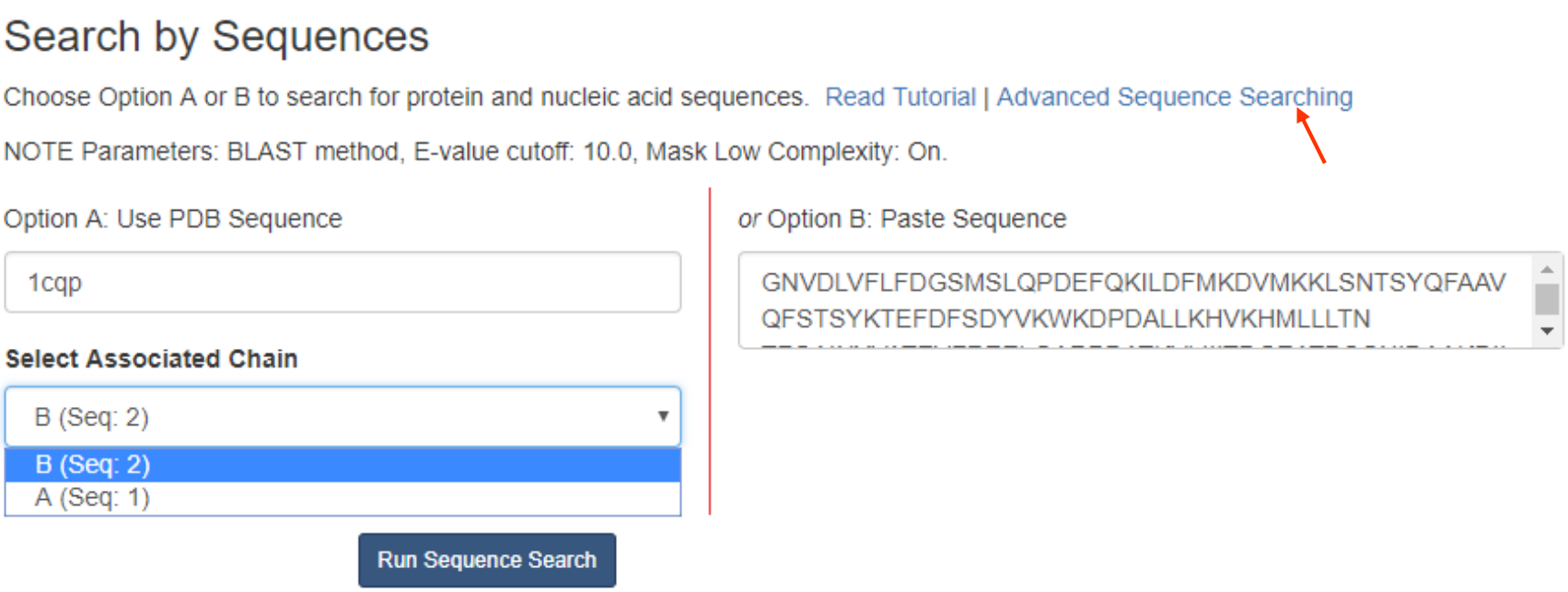
Take an Interactive Tour of the PDB (2)



Searching for your structure (1)

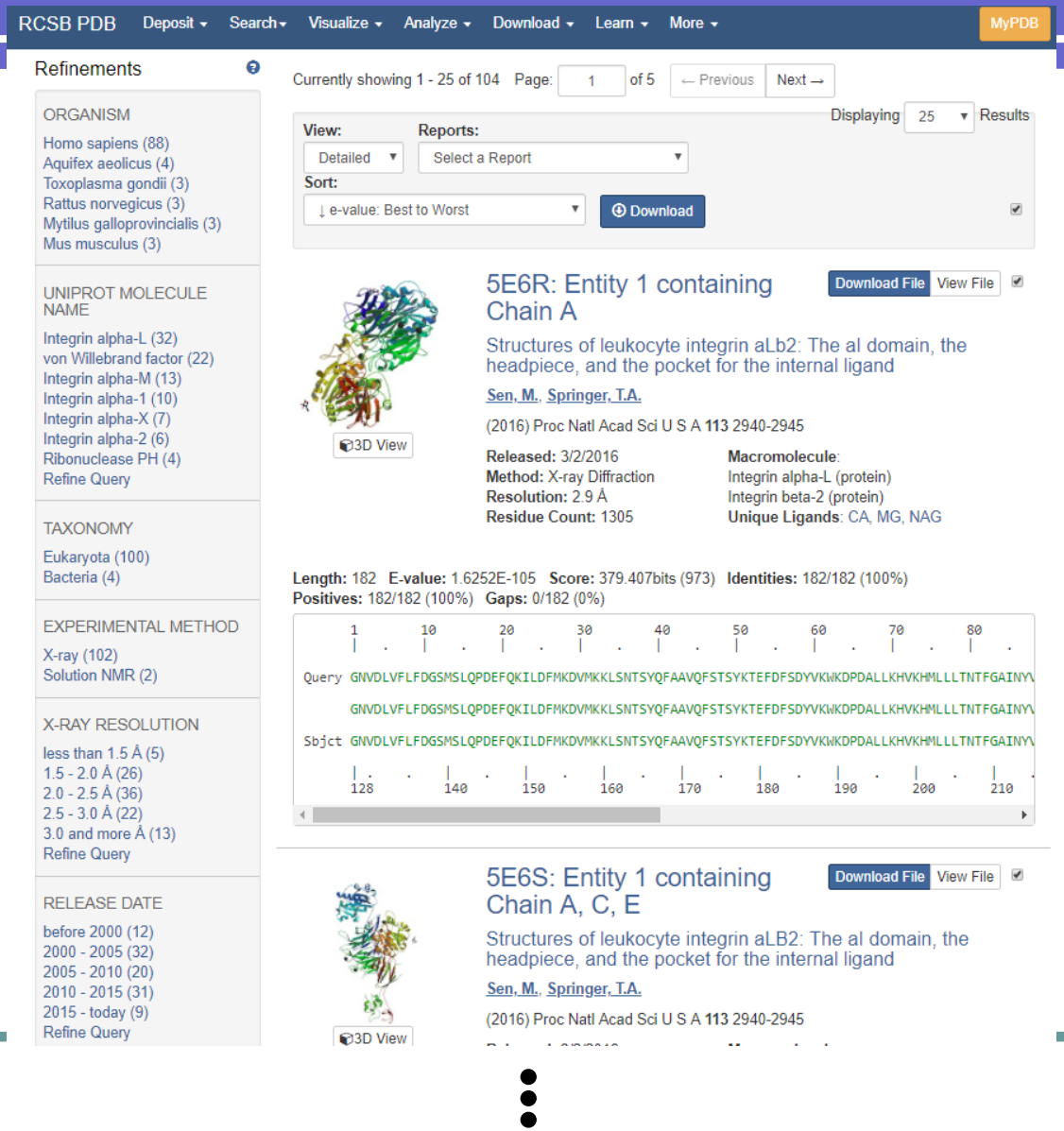


Searching for your structure (2)



Searching for your structure (3)

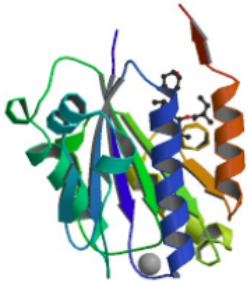
Query result browser



Retrieving your structure – 1CQP

Structure Summary | 3D View | Annotations | Sequence | Sequence Similarity | Structure Similarity | Experiment

Biological Assembly 1



3D View: Structure | Electron Density | Ligand Interaction

Standalone Viewers
Protein Workshop | Ligand Explorer

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

Biological assembly 1 assigned by authors.

Macromolecule Content

- Total Structure Weight: 42499.41
- Atom Count: 2998
- Residue Count: 364
- Unique protein chains: 1

Display Files | Download Files

- FASTA Sequence
- PDB Format
- PDB Format (gz)
- PDBx/mmCIF Format
- PDBx/mmCIF Format (gz)
- PDBML/XML Format (gz)
- Biological Assembly 1
- Biological Assembly 2
- Structure Factors (CIF)
- Structure Factors (CIF - gz)

1CQP

CRYSTAL STRUCTURE ANALYSIS OF THE COMPLEX
LOVASTATIN AT 2.6 A RESOLUTION

DOI: 10.2210/pdb1CQP/pdb

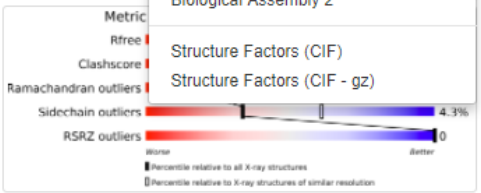
Classification: IMMUNE SYSTEM
Organism(s): Homo sapiens

Deposited: 1999-08-10 **Released:** 2000-08-07
Deposition Author(s): Kallen, J., Welzenbach, K., Ramage, P., Ge
S., Weitz-Schmidt, G., Hommel, U.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 2.6 Å
R-Value Free: 0.257
R-Value Work: 0.190

wwPDB Validation



This is version 1.3 of the entry. See complete history.

Literature — Download Primary Citation

Structural basis for LFA-1 inhibition upon lovastatin binding to the CD11a I-domain.

Kallen, J., Welzenbach, K., Ramage, P., Geyl, D., Kriwacki, R., Legge, G., Cottens, S., Weitz-Schmidt, G., Hommel, U.
(1999) J.Mol.Biol. **292**: 1-9

PubMed: 10493852 Search on PubMed
DOI: 10.1006/jmbi.1999.3047

PubMed Abstract:
The lymphocyte function-associated antigen (LFA-1) belongs to the family of beta2-integrins and plays an important role in T-cell activation and leukocyte migration to sites of inflammation. We report here that lovastatin, a drug clinically used for ...

PDB format (1) – 1CQP

Header, title and author records

```
HEADER    IMMUNE SYSTEM                                10-AUG-99   1CQP
TITLE     CRYSTAL STRUCTURE ANALYSIS OF THE COMPLEX LFA-1 (CD11A) I-DOMAIN /
TITLE    2 LOVASTATIN AT 2.6 A RESOLUTION
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: ANTIGEN CD11A (P180);
COMPND   3 CHAIN: A, B;
COMPND   4 FRAGMENT: I-DOMAIN, RESIDUES 153-334;
COMPND   5 SYNONYM: INTEGRIN ALPHA L, LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 1;
COMPND   6 ALPHA POLYPEPTIDE;
COMPND   7 ENGINEERED: YES;
COMPND   8 OTHER_DETAILS: COMPLEXED WITH LOVASTATIN WHICH OCCURS NATURALLY IN
COMPND   9 FUNGI
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: HOMO SAPIENS;
SOURCE   3 ORGANISM_COMMON: HUMAN;
SOURCE   4 ORGANISM_TAXID: 9606
KEYWDS    ROSSMANN FOLD, STRUCTURAL BASIS FOR LFA-1 INHIBITION, IMMUNE SYSTEM
EXPDTA    X-RAY DIFFRACTION
AUTHOR    J.KALLEN,K.WELZENBACH,P.RAMAGE,D.GEYL,R.KRIWACKI,G.LEGGE,S.COTTENS,
AUTHOR   2 G.WEITZ-SCHMIDT,U.HOMMEL
REVDAT   3   12-NOV-14 1CQP    1       KEYWDS
REVDAT   2   24-FEB-09 1CQP    1       VERSN
REVDAT   1   07-AUG-00 1CQP    0
JRNL        AUTH   J.KALLEN,K.WELZENBACH,P.RAMAGE,D.GEYL,R.KRIWACKI,G.LEGGE,
JRNL        AUTH 2 S.COTTENS,G.WEITZ-SCHMIDT,U.HOMMEL
JRNL        TITL   STRUCTURAL BASIS FOR LFA-1 INHIBITION UPON LOVASTATIN
JRNL        TITL 2 BINDING TO THE CD11A I-DOMAIN.
JRNL        REF    J.MOL.BIOL.                   V. 292     1 1999
```

PDB format (2)– 1CQP

Remark records

```
…
REMARK   2 RESOLUTION.    2.60 ANGSTROMS.
…
REMARK 290 CRYSTALLOGRAPHIC SYMMETRY
REMARK 290 SYMMETRY OPERATORS FOR SPACE GROUP: P 21 21 21
…
REMARK 290 CRYSTALLOGRAPHIC SYMMETRY TRANSFORMATIONS
REMARK 290 THE FOLLOWING TRANSFORMATIONS OPERATE ON THE ATOM/HETATM
REMARK 290 RECORDS IN THIS ENTRY TO PRODUCE CRYSTALLOGRAPHICALLY
REMARK 290 RELATED MOLECULES.
REMARK 290   SMTRY1   1  1.000000  0.000000  0.000000        0.00000
REMARK 290   SMTRY2   1  0.000000  1.000000  0.000000        0.00000
REMARK 290   SMTRY3   1  0.000000  0.000000  1.000000        0.00000
REMARK 290   SMTRY1   2 -1.000000  0.000000  0.000000       36.35000
REMARK 290   SMTRY2   2  0.000000 -1.000000  0.000000        0.00000
REMARK 290   SMTRY3   2  0.000000  0.000000  1.000000       45.90000
…
```

Heteroatom records

(non-standard residues e.g. ligands, ions and water)

```
HET     MG  A 310       1
HET     MG  B 310       1
HET    803  A 311      29
HET    803  B 311      29
HETNAM      MG MAGNESIUM ION
HETNAM     803 LOVASTATIN
HETSYN     803 MK-803; LOVALIP; MEVACOR
FORMUL   3   MG    2(MG 2+)
FORMUL   5  803    2(C24 H36 O5)
FORMUL   7  HOH   *86(H2 O)
```

PDB format (3)– 1CQP

Atom records

```
...
ATOM      1  N   GLY A 128      44.810  32.209  24.312  1.00 60.46           N
ATOM      2  CA  GLY A 128      43.318  32.144  24.502  1.00 59.56           C
ATOM      3  C   GLY A 128      42.879  31.077  25.499  1.00 56.37           C
ATOM      4  O   GLY A 128      43.308  31.073  26.654  1.00 57.03           O
...
ATOM   1462  CA  ILE A 309      52.536  35.917  28.130  1.00 58.16           C
ATOM   1463  C   ILE A 309      51.906  37.303  28.426  1.00 65.50           C
ATOM   1464  O   ILE A 309      52.269  38.298  27.740  1.00 70.09           O
ATOM   1465  CB  ILE A 309      51.655  35.083  27.164  1.00 55.20           C
ATOM   1466  CG1 ILE A 309      52.420  33.837  26.734  1.00 50.60           C
ATOM   1467  CG2 ILE A 309      51.255  35.876  25.948  1.00 67.96           C
ATOM   1468  CD1 ILE A 309      53.803  34.136  26.244  1.00 39.29           C
ATOM   1469  OXT ILE A 309      51.061  37.396  29.345  1.00 70.38           O
TER    1470      ILE A 309
ATOM   1471  N   GLY B 128      62.823  33.550  24.044  1.00 67.05           N
ATOM   1472  CA  GLY B 128      64.227  33.694  24.572  1.00 66.00           C
ATOM   1473  C   GLY B 128      64.358  33.241  26.017  1.00 61.57           C
ATOM   1474  O   GLY B 128      63.468  32.572  26.556  1.00 67.10           O
...
ATOM   2931  N   ILE B 309      54.545  27.250  24.474  1.00 43.38           N
ATOM   2932  CA  ILE B 309      54.748  27.915  23.202  1.00 50.44           C
ATOM   2933  C   ILE B 309      55.615  27.111  22.257  1.00 56.16           C
ATOM   2934  O   ILE B 309      56.646  26.621  22.747  1.00 63.24           O
ATOM   2935  CB  ILE B 309      55.413  29.255  23.479  1.00 53.14           C
ATOM   2936  CG1 ILE B 309      54.447  30.132  24.261  1.00 58.81           C
ATOM   2937  CG2 ILE B 309      55.846  29.919  22.204  1.00 63.82           C
ATOM   2938  CD1 ILE B 309      55.101  31.303  24.905  1.00 69.98           C
ATOM   2939  OXT ILE B 309      55.286  27.012  21.049  1.00 62.52           O
TER    2940      ILE B 309
```

PDB format (4)– 1CQP

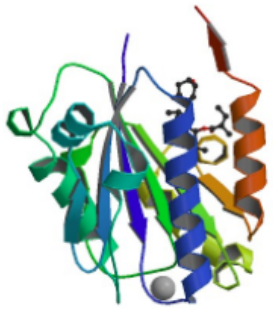
HETAM records

```
...
HETATM 2941 MG    MG A 310      48.079  -2.336  22.333  1.00 23.16          MG
HETATM 2942  C1   803 A 311      52.497  19.025  23.052  1.00 43.13           C
HETATM 2943  C2   803 A 311      52.393  17.482  22.971  1.00 36.38           C
HETATM 2944  C3   803 A 311      52.359  16.834  24.377  1.00 33.15           C
HETATM 2945  C21  803 A 311      53.777  16.458  24.857  1.00 27.12           C
...
HETATM 2971 MG    MG B 310      59.074  50.051  55.046  1.00 42.60          MG
HETATM 2972  C1   803 B 311      55.082  39.941  35.863  1.00 39.50           C
HETATM 2973  C2   803 B 311      55.211  40.642  37.236  1.00 34.66           C
HETATM 2974  C3   803 B 311      55.219  39.639  38.406  1.00 26.04           C
HETATM 2975  C21  803 B 311      53.761  39.334  38.807  1.00 28.07           C
...
HETATM 3001  O    HOH A   1      46.704   3.370  17.939  1.00 21.72           O
HETATM 3002  O    HOH A   2      37.676  -1.658  26.021  1.00 24.22           O
HETATM 3003  O    HOH A   3      47.452  -0.486  21.226  1.00 26.23           O
HETATM 3004  O    HOH A   4      46.992  -3.309  17.690  1.00 26.75           O
...
CONECT   90 2941
CONECT  104 2941
CONECT  905 2941
CONECT 1560 2971
CONECT 1574 2971
CONECT 2375 2971
...
MASTER      309    0    4   20   14    0    6    6 3084    2   70   28
END
```

Looking at your structure (1) – 1CQP

Structure Summary | 3D View | Annotations | Sequence | Sequence Similarity | Structure Similarity | Experiment

Biological Assembly 1



3D View: Structure | Electron Density | Ligand Interaction

Standalone Viewers
Protein Workshop | Ligand Explorer

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

Biological assembly 1 assigned by authors.

Macromolecule Content

- Total Structure Weight: 42499.41
- Atom Count: 2998
- Residue Count: 364
- Unique protein chains: 1

Display Files | Download Files

1CQP

CRYSTAL STRUCTURE ANALYSIS OF THE COMPLEX LFA-1 (CD11A) I-DOMAIN / LOVASTATIN AT 2.6 A RESOLUTION

DOI: 10.2210/pdb1CQP/pdb

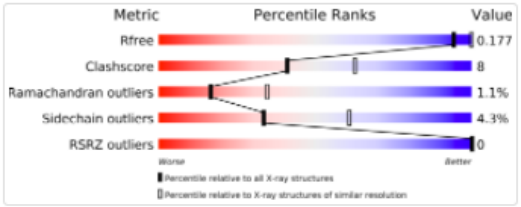
Classification: IMMUNE SYSTEM
Organism(s): Homo sapiens

Deposited: 1999-08-10 **Released:** 2000-08-07
Deposition Author(s): Kallen, J., Welzenbach, K., Ramage, P., Geyl, D., Kriwacki, R., Legge, G., Cottens, S., Weitz-Schmidt, G., Hommel, U.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 2.6 Å
R-Value Free: 0.257
R-Value Work: 0.190

wwPDB Validation 3D Report | Full Report



This is version 1.3 of the entry. See complete history.

Literature Download Primary Citation

Structural basis for LFA-1 inhibition upon lovastatin binding to the CD11a I-domain.

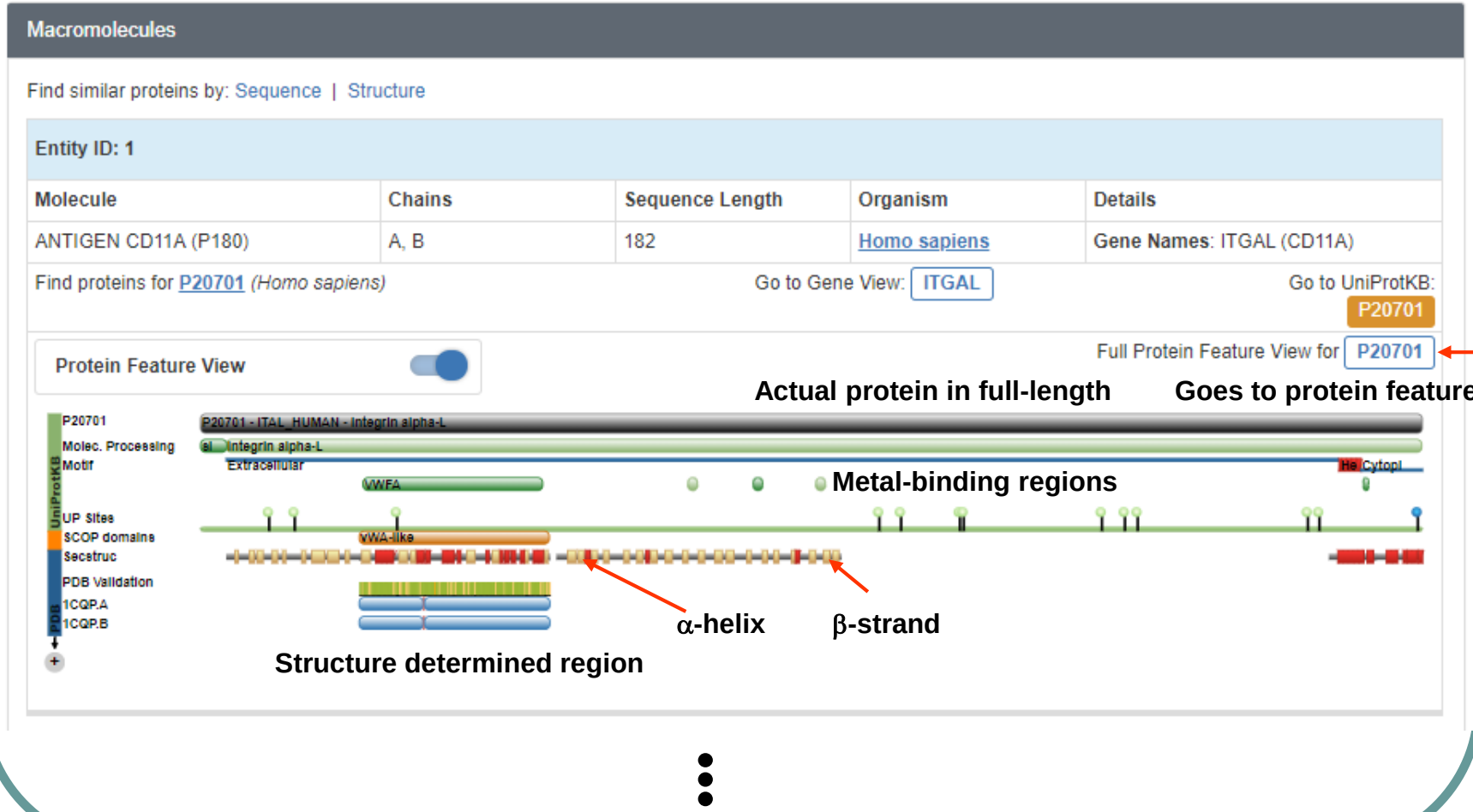
Kallen, J., Welzenbach, K., Ramage, P., Geyl, D., Kriwacki, R., Legge, G., Cottens, S., Weitz-Schmidt, G., Hommel, U.
(1999) J.Mol.Biol. **292**: 1-9

PubMed: 10493852 Search on PubMed
DOI: 10.1006/jmbi.1999.3047

PubMed Abstract:
The lymphocyte function-associated antigen (LFA-1) belongs to the family of beta2-integrins and plays an important role in T-cell activation and leukocyte migration to sites of inflammation. We report here that lovastatin, a drug clinically used for ...

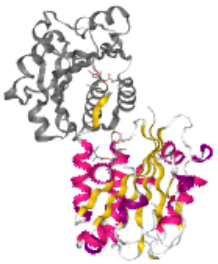
Click it to launch Ligand Explorer

Looking at your structure (2) – 1CQP



Protein feature view – 1CQP

Viewer: NGL



Organism: **Human** Length:1170 UniProt

Isoforms: 3 , currently showing only the 'canonical' sequence.

Gene View for ITGAL

Chromosome Location

Other Gene names: CD11A

This protein in other organisms (by gene name):

P20701 - Homo sapiens 34

- P24063 - Mus musculus no matching PDB entries
- O43746 - Homo sapiens no matching PDB entries
- Q6TYB8 - Bos taurus no matching PDB entries
- Q9UBC8 - Homo sapiens no matching PDB entries
- Q45H73 - Homo sapiens no matching PDB entries
- Q96HB1 - Homo sapiens no matching PDB entries
- P61625 - Bos taurus no matching PDB entries

PDB ID 1CQP.A

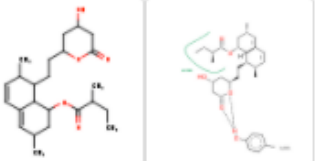
Predicted possible disorder region

hydrophobic

Looking at your structure (3) – 1CQP

Small Molecules

Ligands 2 Unique

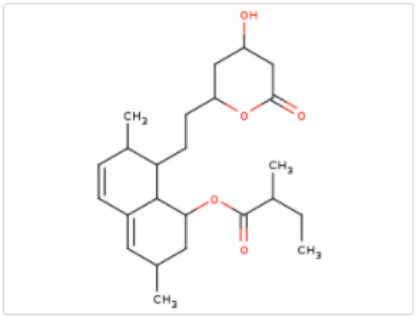
ID	Chains	Name / Formula / InChI Key	2D Diagram & Interactions	3D Interactions
MG Query on MG Download SDF File Download CCD File	A, B	**MAGNESIUM ION** Mg JLVVSXFLKOJNIY-UHFFFAOYSA-N	Mg^{2+}	Ligand Interaction
803 Query on 803 Download SDF File Download CCD File	A, B	**LOVASTATIN** *MK-803; LOVALIP; MEVACOR* $C_{24}H_{36}O_5$ PCZOHLXUXFIOCF-BXMDZJJMSA-N		Ligand Interaction

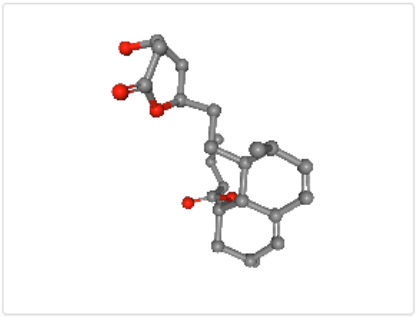
External Ligand Annotations

ID	Binding Affinity (Sequence Identity %)
803	Kd: 12900 nM (99) BINDINGDB
803	IC50: 2400 nM PDBBIND

⋮

Browsing your ligand





Rotate | Hydrogens | Labels

View / Download Files

803

LOVASTATIN

803 as a free ligand exists in 1 entry. Examples include: 1CQP

Find related ligands: Stereoisomers Similar ligands Chemical Structure Search

View summary at Ligand Expo

Related Data Resources for 803 eg. PubChem, DrugBank,…

Chemical information

Chemical Component Summary	
Name	LOVASTATIN
Identifiers	[(1S,3R,7S,8S,8aR)-8-[2-[(2R,4R)-4-hydroxy-6-oxo-oxan-2-yl]ethyl]-3,7-dimethyl-1,2,3,7,8,8a-hexahydronaphthalen-1-yl] (2S)-2-methylbutanoate
Formula	$C_{24}H_{36}O_5$
Molecular Weight	404.54
Type	NON-POLYMER
Isomeric SMILES	CC[C@H](C)C(=O)O[C@H]1C[C@@H](C)C=C2C=C[C@H](C)[C@H](CC[C@@H]3C[C@@H](O)CC(=O)O3)[C@@H]12
InChI	InChI=1S/C24H36O5/c1-5-15(3)24(27)29-21-11-14(2)10-17-7-6-16(4)20(23(17)21)9-8-19-12-18(25)13-22(26)28-19/h6-7,10,14-16,18-21,23,25H,5,8-9,11-13H2,1-4H3/t14-,15-,16-,18+,19+,20-,21-,23-/m0/s1
InChIKey	PCZOHLXUXFIOCF-BXMDZJJMSA-N

Chemical Details	
Formal Charge	0
Atom Count	65
Chiral Atom Count	8
Chiral Atoms	C1, C12, C14, C16, C3, C7, C8, C9
Bond Count	67
Aromatic Bond Count	0
Leaving Atoms	n/a

Drug Info: DrugBank

DrugBank ID	DB00227 **Different stereochemistry**
Name	Lovastatin
Groups	• approved • investigational

Looking at your structure (4) – 1CQP

Experimental Data & Validation

Experimental Data

Method: X-RAY DIFFRACTION
Resolution: 2.6 Å
R-Value Free: 0.257
R-Value Work: 0.190
Space Group: P 2_1 2_1 2_1

Unit Cell:

Length (Å)	Angle (°)
a = 72.700	α = 90.00
b = 77.700	β = 90.00
c = 91.800	γ = 90.00

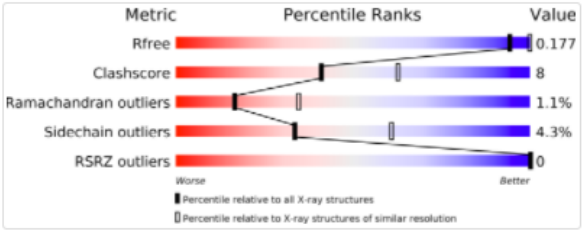
Software Package:

Software Name	Purpose
X-PLOR	phasing
DENZO	data reduction
X-PLOR	refinement
X-PLOR	model building
SCALEPACK	data scaling

View more in-depth experimental data

Structure Validation

View Full Validation Report or Ramachandran Plots



Entry History

Deposition Data

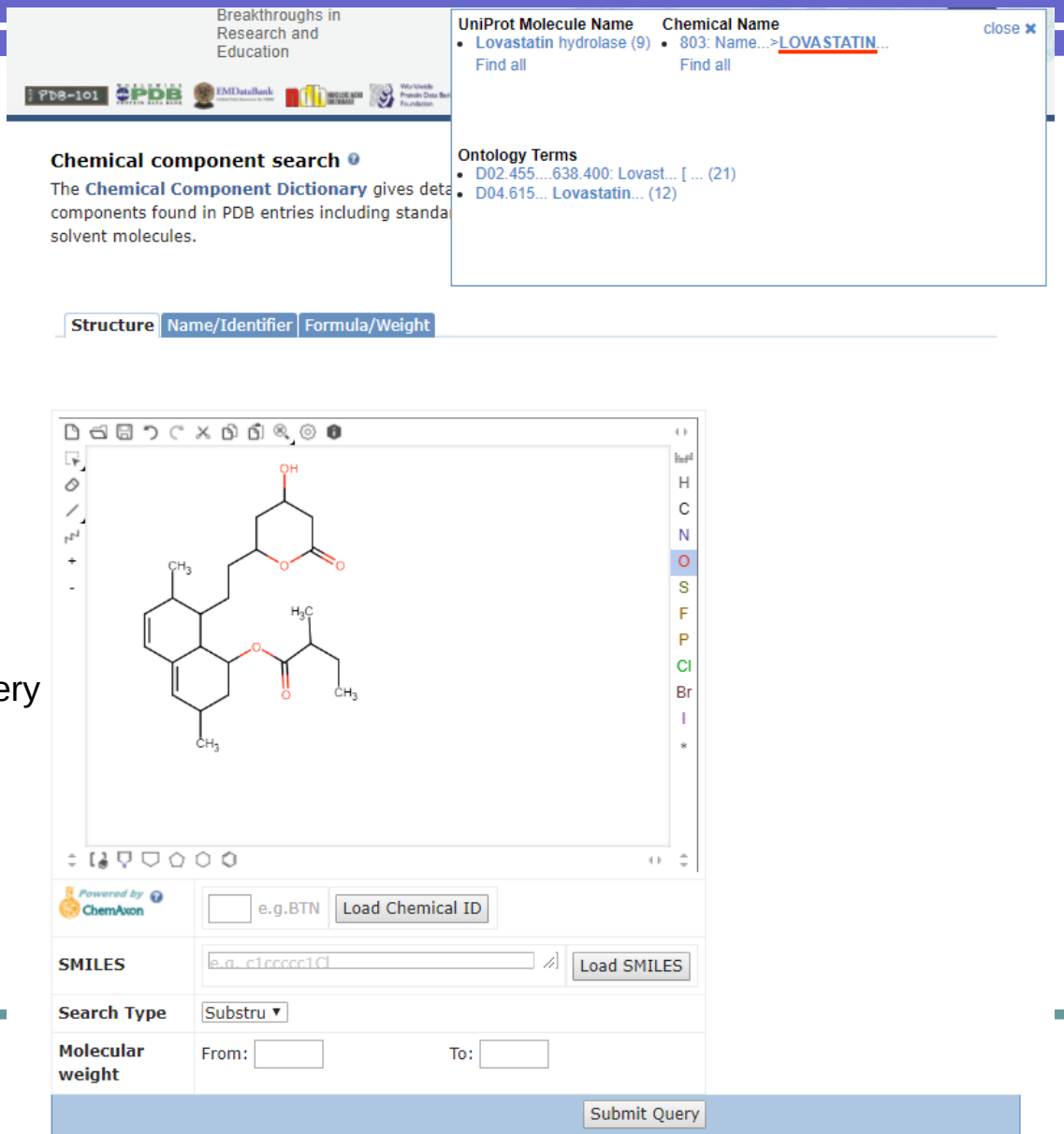
Deposited Date: 1999-08-10
Released Date: 2000-08-07
Deposition Author(s): Kallen, J., Welzenbach, K., Ramage, P., Geyl, D., Kriwacki, R., Legge, G., Cottens, S., Weitz-Schmidt, G., Hommel, U.

Revision History

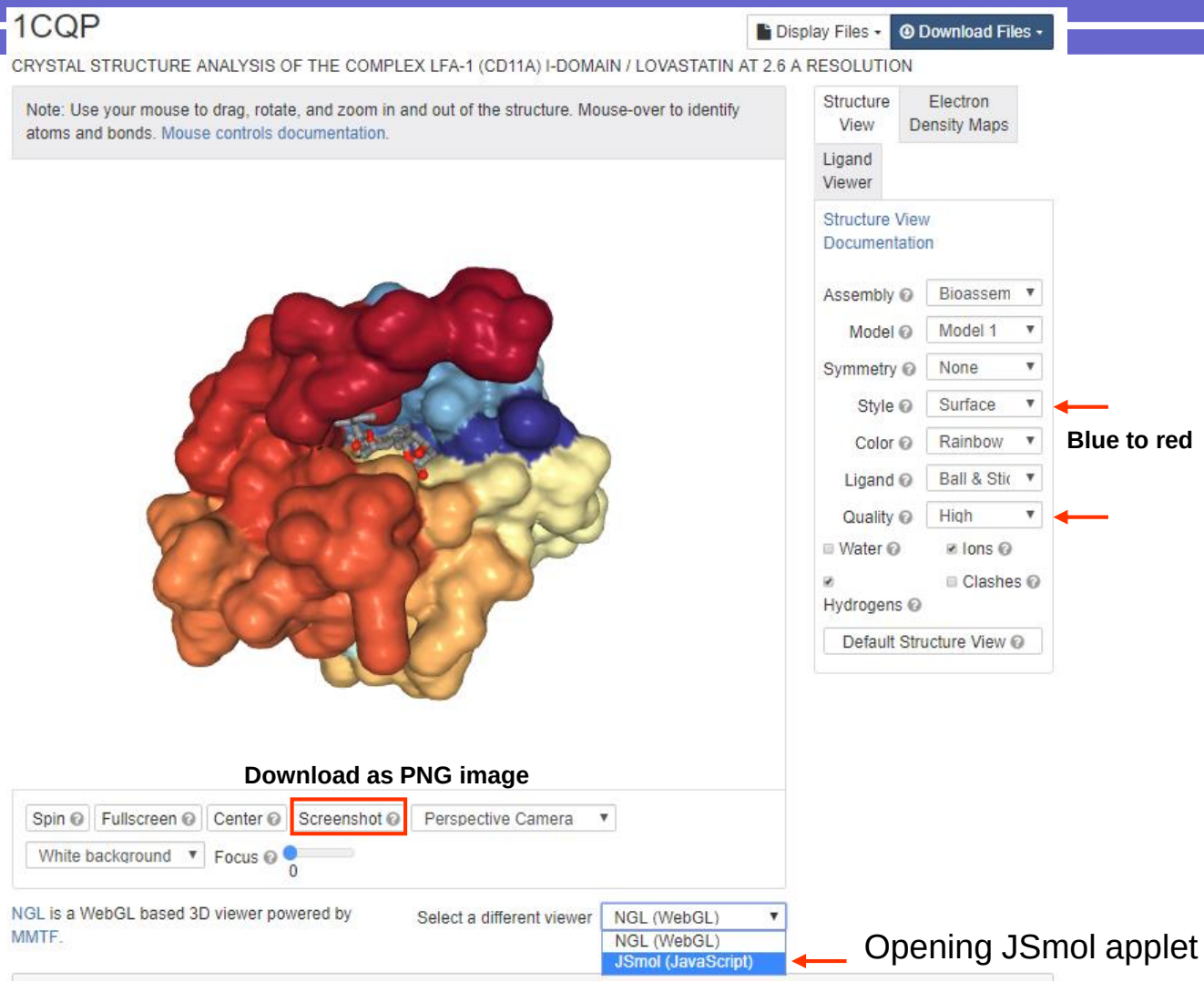
- **Version 1.0: 2000-08-07**
  Type: Initial release
- **Version 1.1: 2008-04-27**
  Type: Version format compliance
- **Version 1.2: 2011-07-13**
  Type: Version format compliance
- **Version 1.3: 2014-11-12**
  Type: Structure summary

Searching for ligand

Chemical sketch →
click on the Submit Query

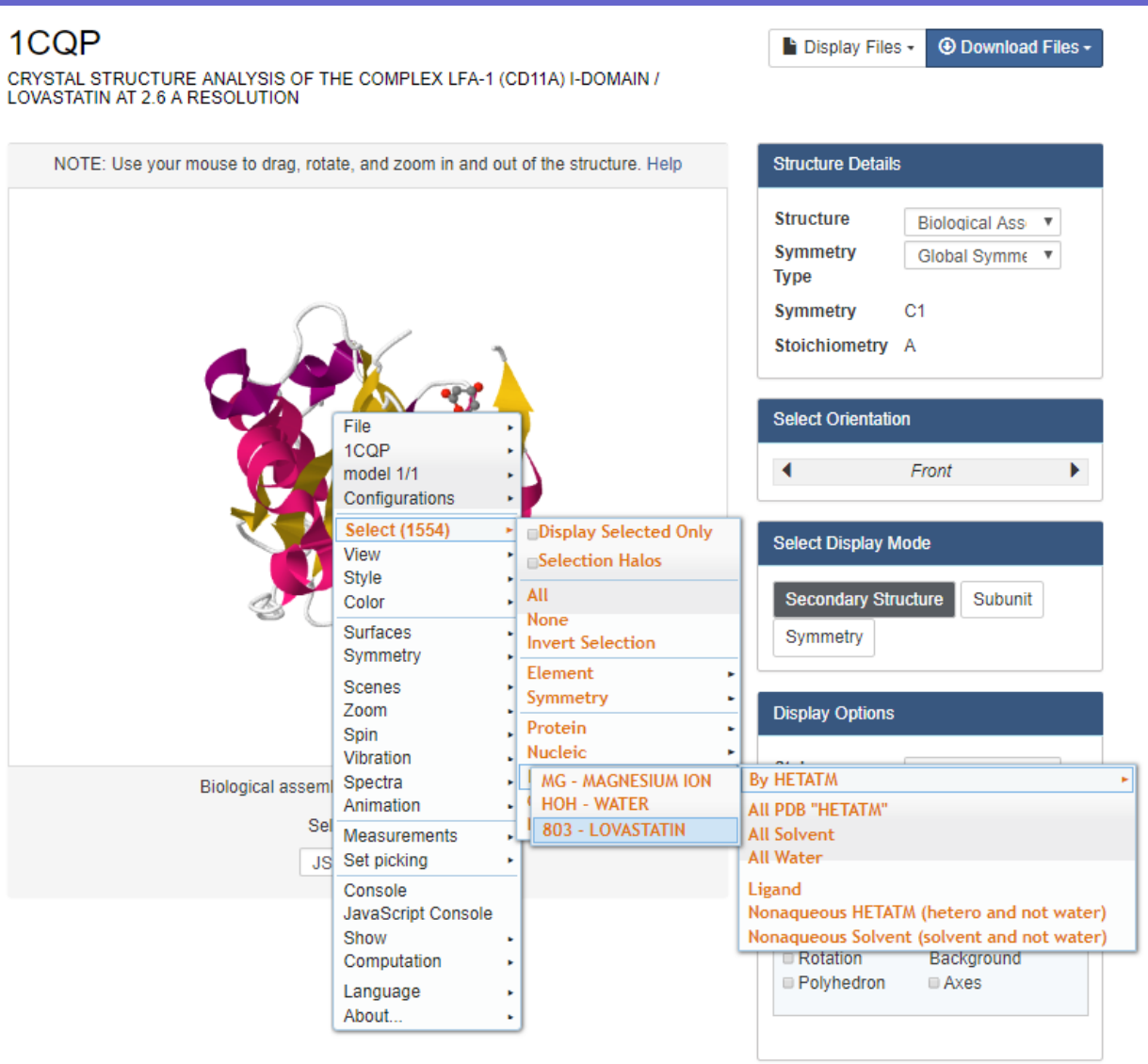


Structure visualization – NGL



Structure visualization – JSmol (1)

Press on the right bottom of mouse to show the drop-down menu



Structure visualization – JSmol (2)

Select → Hetero → By HETATM→ HEM

Style → Scheme → Ball and Stick

Surface → Dot surface

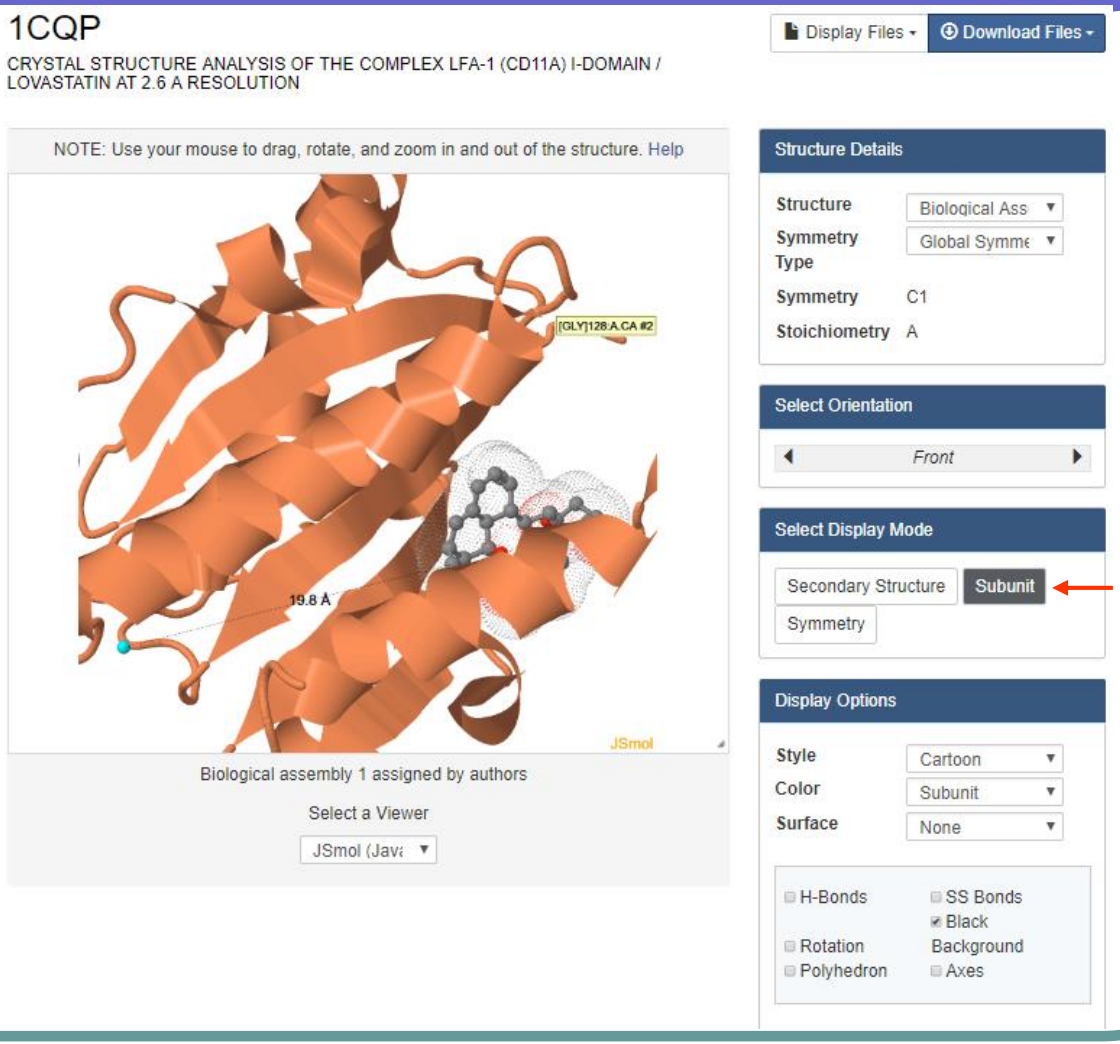
Select → Hetero → By HETATM→ MG

Color → Atoms → Cyan

Measurements → Click for distance measurement

on the screen: Click on MG and [803]311:A C21 #1474 Distance: 19.8Å

File → Export → Export PNG image



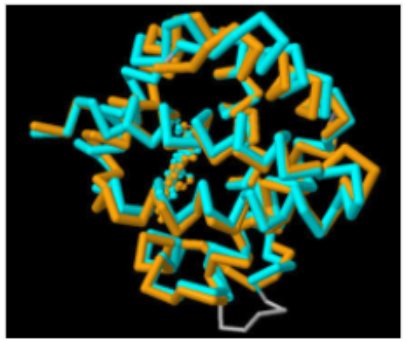
Structure comparison -- jFATCAT-rigid (1)

Analyze Related Features

Sequence & Structure Alignment

RCSB PDB's Comparison Tool calculates pairwise sequence (blast2seq, Needleman-Wunsch, and Smith-Waterman) and structure alignments (FATCAT, CE, Mammoth, TM-Align, TopMatch).

Comparisons can be made for any protein in the PDB archive and for customized or local files not in the PDB. Special features include support for both rigid-body and flexible alignments and detection of circular permutations.



1CQP ⟷ 5E6S

Select Associated Chain ID

A (Seq: 1)

Select Associated Chain ID

A (Seq: YNLDVRGARSFSPPRAGRHFGYRVLQV(

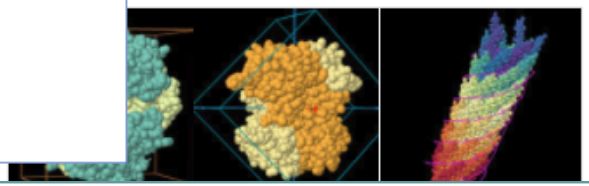
- Select Comparison Method -

Align | More Options

- Select Comparison Method -
Sequence Alignment
blast2seq
Needleman-Wunsch
Smith-Waterman
Structure Alignment
jFATCAT - rigid
jFATCAT - flexible
jCE algorithm
jCE Circular Permutation
External: FatCat
External: Mammoth
External: TM-Align
External: TopMatch

al symmetry among subunits. The view displays the symmetry axes, a polyhedron asizes the symmetry.

bin Streptavidin Inovirus



Structure comparison -- jFATCAT-rigid (2)

Structure Alignment View

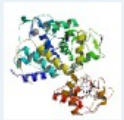
Pre-calculated jFATCAT_rigid results for 4GED.A vs. 2WD4.A .

This page provides a summary view of the protein structure alignment.

Structure Alignment Results

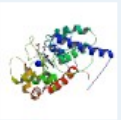
Alignment Details: **Query:** (orange/dark grey) ***Ascorbate peroxidase***

P-value: 0.00e+00
Score: 639.52
RMSD: 1.51
%Id: 38.7%



PDB ID: 4GED
Chain ID: A
Length: 268
Similarity: 89%
EC number: 1.11.1.11

Subject: (cyan/light grey) ***ASCORBATE PEROXIDASE***



PDB ID: 2WD4
Chain ID: A
Length: 248
Similarity: 96%
EC number: 1.11.1.11

Comparison Method

Select these two chains for other comparison: --- Select Comparison Method ---

Click **here** to align other protein chains. Back to the all vs. all search results for **4GED.A** or **2WD4.A**

Jmol

Structure comparison -- jFATCAT-rigid (3)

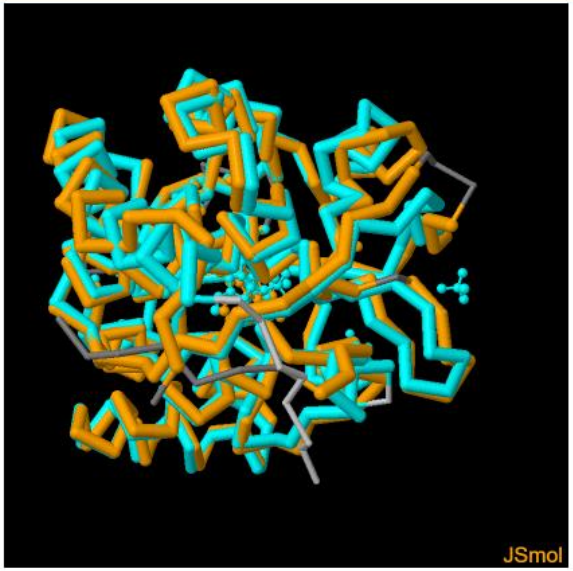
Comparison Method

Select these two chains for other comparison:

--- Select Comparison Method ---

Click here to align other protein chains. Back to the all vs. all search results for 4GED.A or 2WD4.A

Jmol



Tip: right-mouse click on Jmol to get access to additional Jmol functionality.

Reset Display

Jmol Script

Display Query & Target Show Both Style Backbone Color Secondary Structure

Toggle Selection H-Bonds SS Bonds Rotation Antialias Display (nicer) Black Background

Color Legend 4GED.A 2WD4.A

It is also possible to view this alignment using the stand-alone Java Web Start application.

Launch Web Start

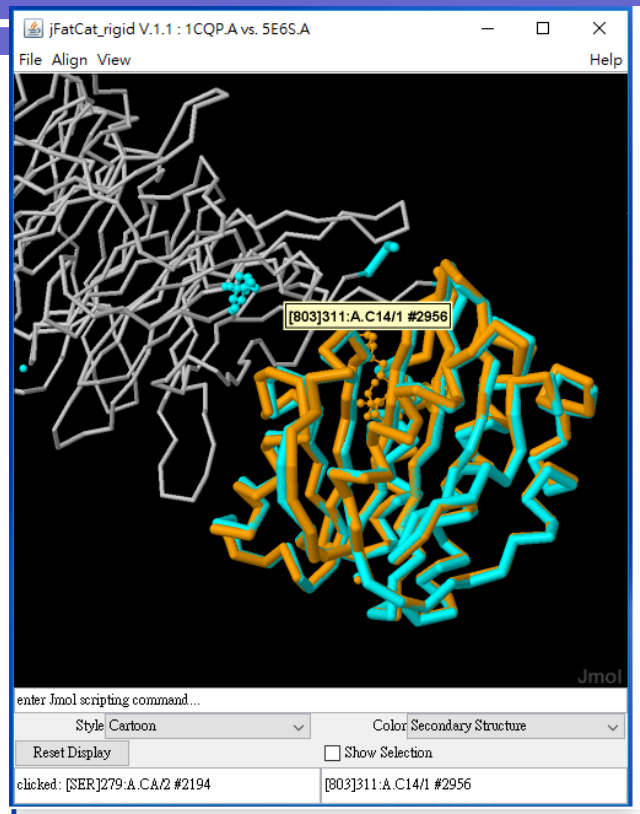
Structure comparison -- jFATCAT-rigid (4)

Alignment Block(s) | Alignment with Sequence Conservation

```
Align 4GED.A.pdb Length1: 268 with 2WD4.A.pdb Length2: 248
P-value: 0.00e+00
Equ: 239
RMSD: 1.51
Score: 639.52
Align-len: 261
Gaps: 22 (8.43%)

Identity: 38.70%
Similarity: 53.64%

35:A          50:A              70:A              90:A
|    |    .    |    .    |    .    |    .    |    .    |    .    |
EEPPFDIRALRADIEDMISEKLELGPSLIRLAWHEAASYDCFKKDGSPNSASMRFKPECLYAGNKGLDIP
.:....:...:..:...|.||. ..|.::|||.|.|..:|...|.| ......:...|..:..|.||||.
ADYQKAVEKAKKKLRGFIAEKR-CAPLMLRLAAHSAGTFDKGTKTG-GPFGTIKHPAELAHSANNGLDIA
||        |    .    |    .    |    .    |    .    |    .    |    .
10:A              30:A              50:A              70:A


     110:A             130:A             150:A             170:A
.    |    .    |    .    |    .    |    .    |    .    |    .    |
RKALETLKKKYPQISYADLWVLAAYVAIEYMGGPTIPFCWGRVDAKDGSVCGPDGRLPDGSKTQSHVREV
.:.||.||...:|.:||||.:.||..||:|..|||.:|||..||.|..    ...|:|||||..|...|.|:|
VRLLEPLKAEFPILSYADFYQLAGVVAVEVTGGPEVPFHPGREDKP---EPPPEGRLPDATKGSDHLRDV
   |    .    |    .    |    .    |    .    |      .    |    .    |
           90:A              110:A             130:A


               190:A             210:A             230:A
.    |    .    |    .    |    .    |    .    |    .    |    .    |
FR-RLGFNDQETVALIGAHTCGECHIEFSGYHGPWTHDKNGFDNSFFTQLLDEDWVLNPKVEQMQLMDRA
|. .:|..||:.|||.|.||.|..|.|.||:.||||.....||||:||:||...:
FGKAMGLTDQDIVALSGGHTIGAAHKERSGFEGPWTSNPLIFDNSYFTELLSGE----------------
 .   |    .    |    .    |    .    |    .    |    .
     150:A             170:A             190:A


      250:A             270:A             290:A
.    |    .    |    .    |    .    |    .    |
TTKLMMLPSDVCLLLDPSYRKYVVELYAKDNDRFNKDFANAFKKLTELGTRN
...|:.||||..||.||.:|..|:.||.|.|.|..|:|.|..||.|||...
KEGLLQLPSDKALLSDPVFRPLVDKYAADEDAFFADYAEAHQKLSELGFAD
  |    .    |    .    |    .    |    .    |
```

Sequence-Structure alignment

Structure comparison -- jFATCAT-rigid (5)

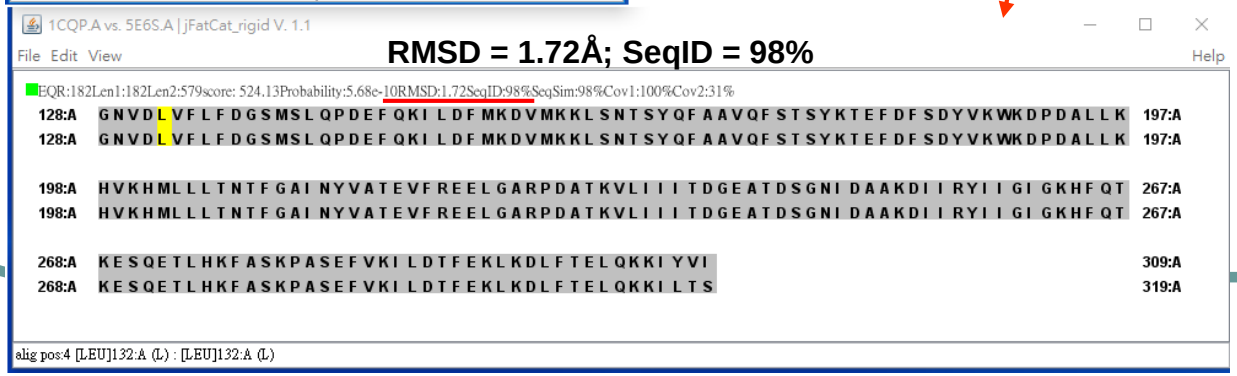


This structure alignment also can be viewed using the stand-alone Java Web Start application.

Click on Launch Web Start

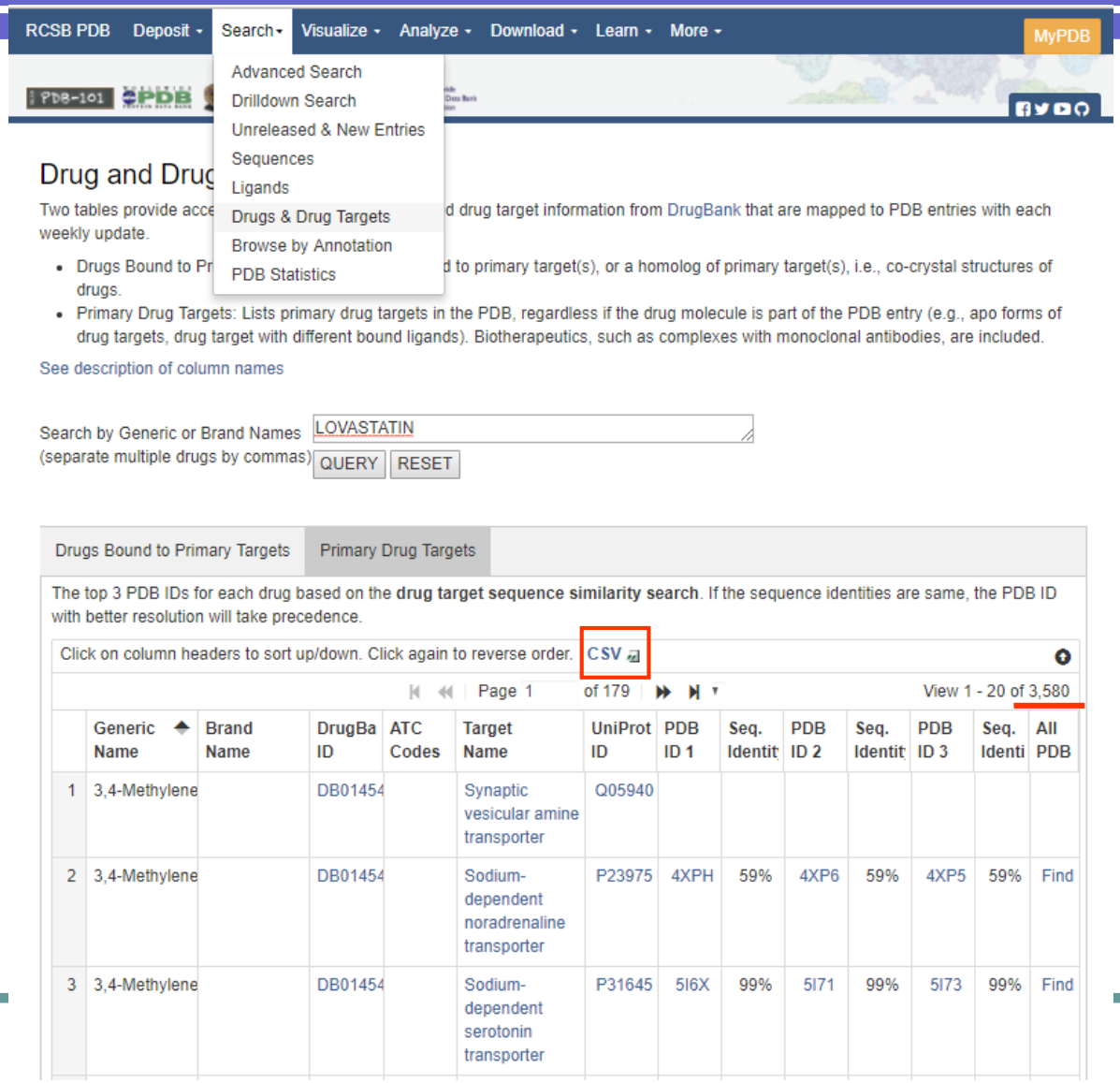
Structure alignment

Sequence alignment



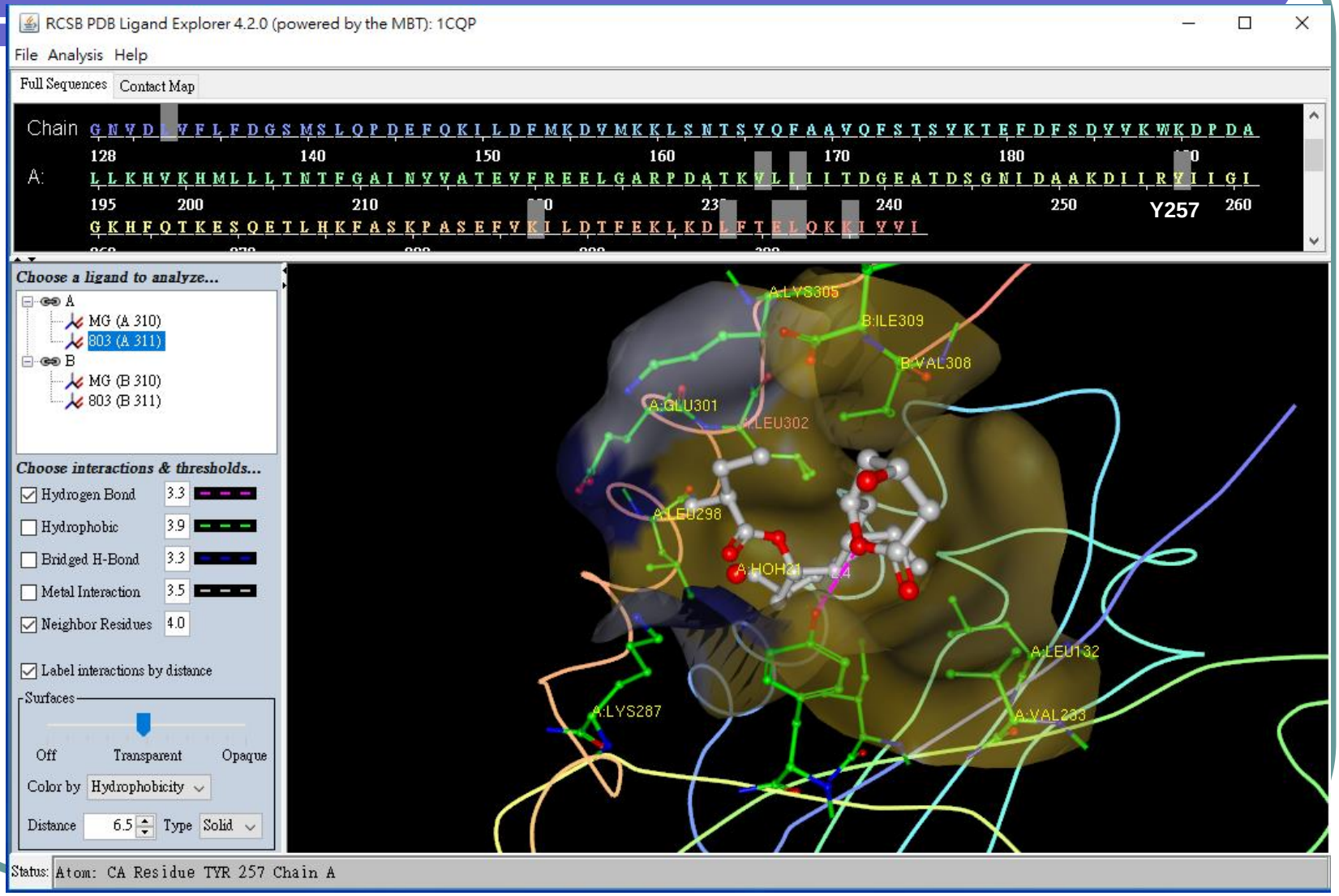
RMSD = 1.72Å; SeqID = 98%

Drug and Drug Target Mapping

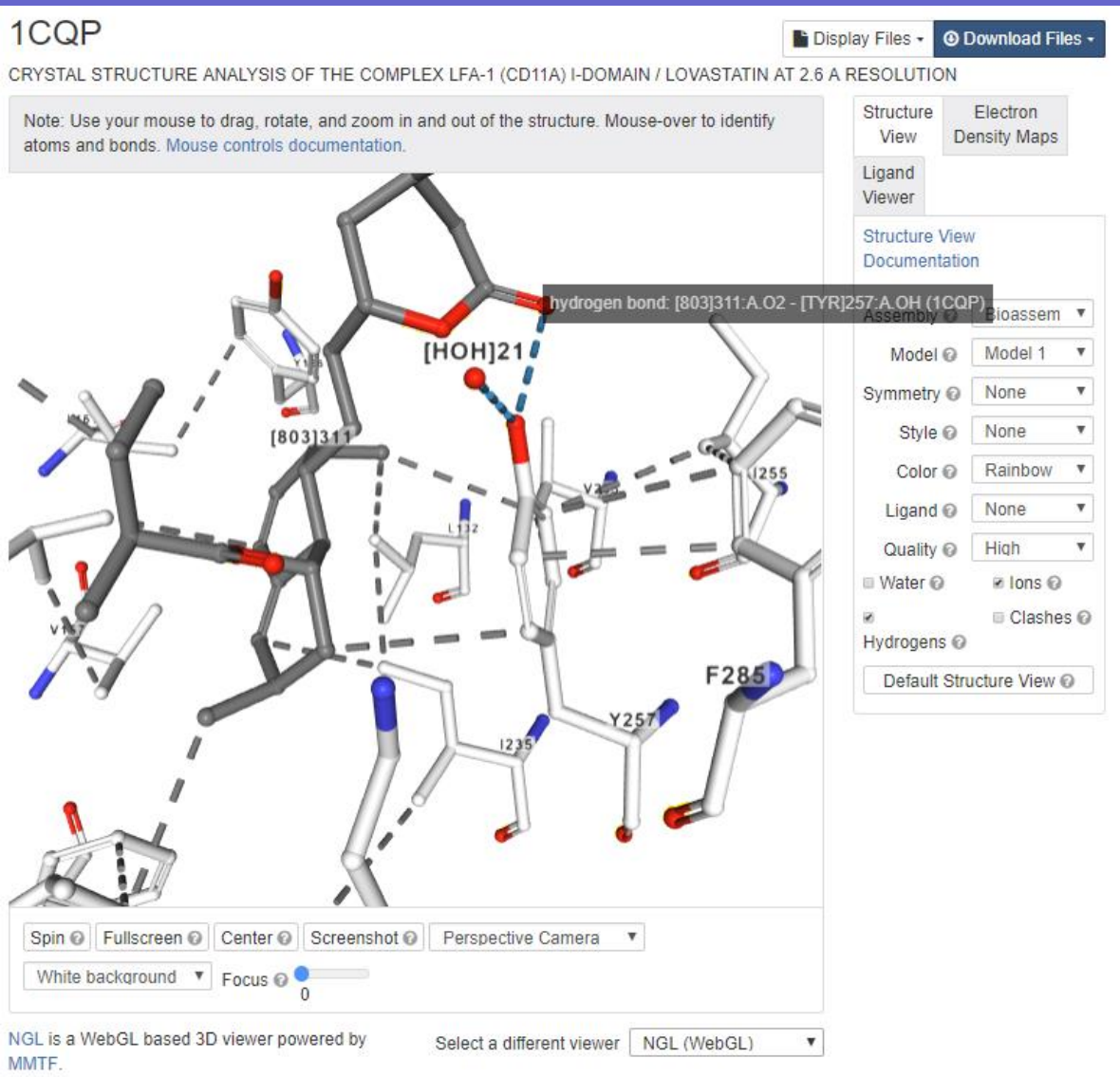


Protein-ligand interaction (1)

Save the image

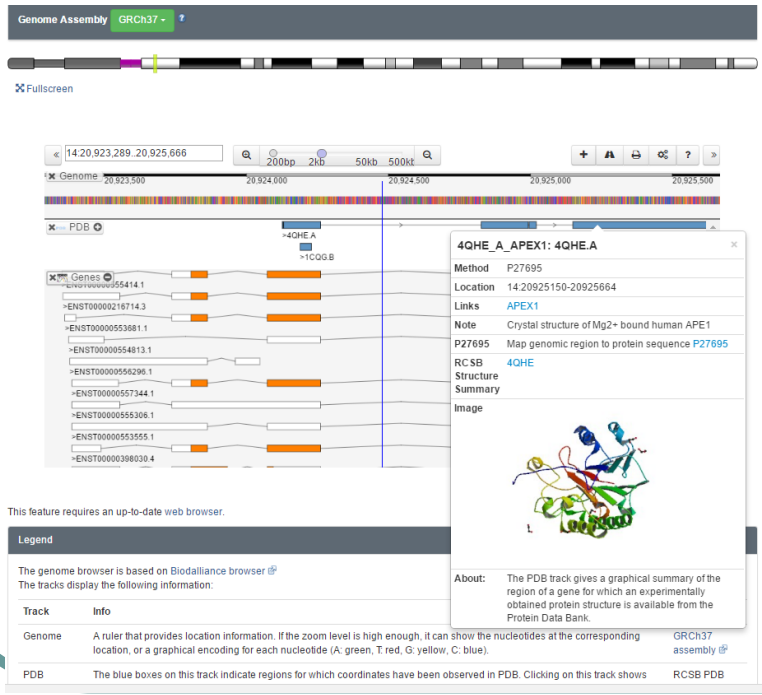


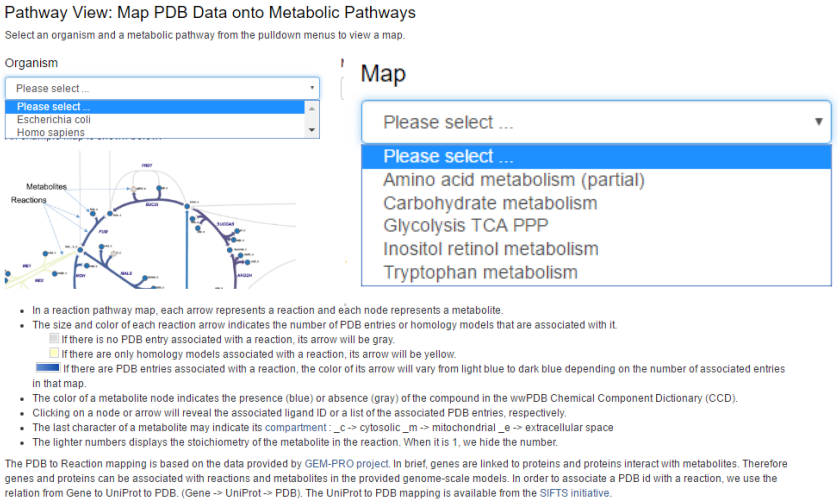
Protein-ligand interaction (2)

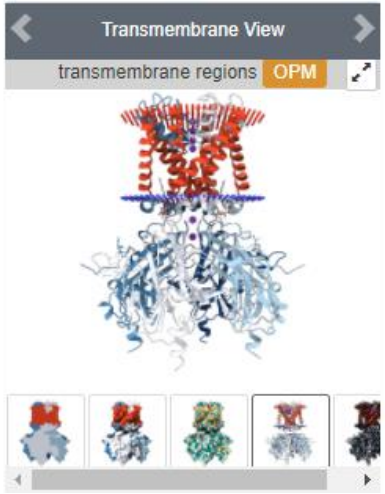


Other features

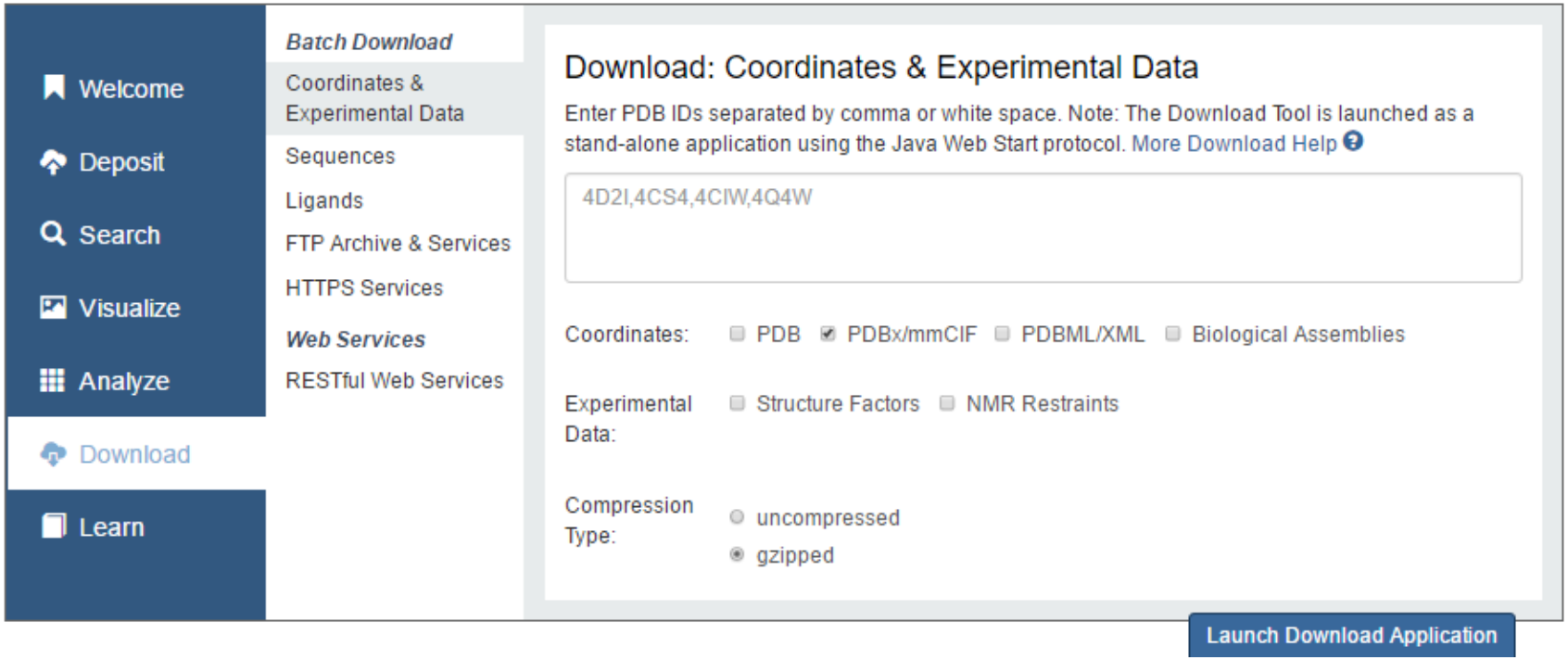
- Human Gene View
- Pathway View
- Transmembrane Proteins







Multiple file downloading



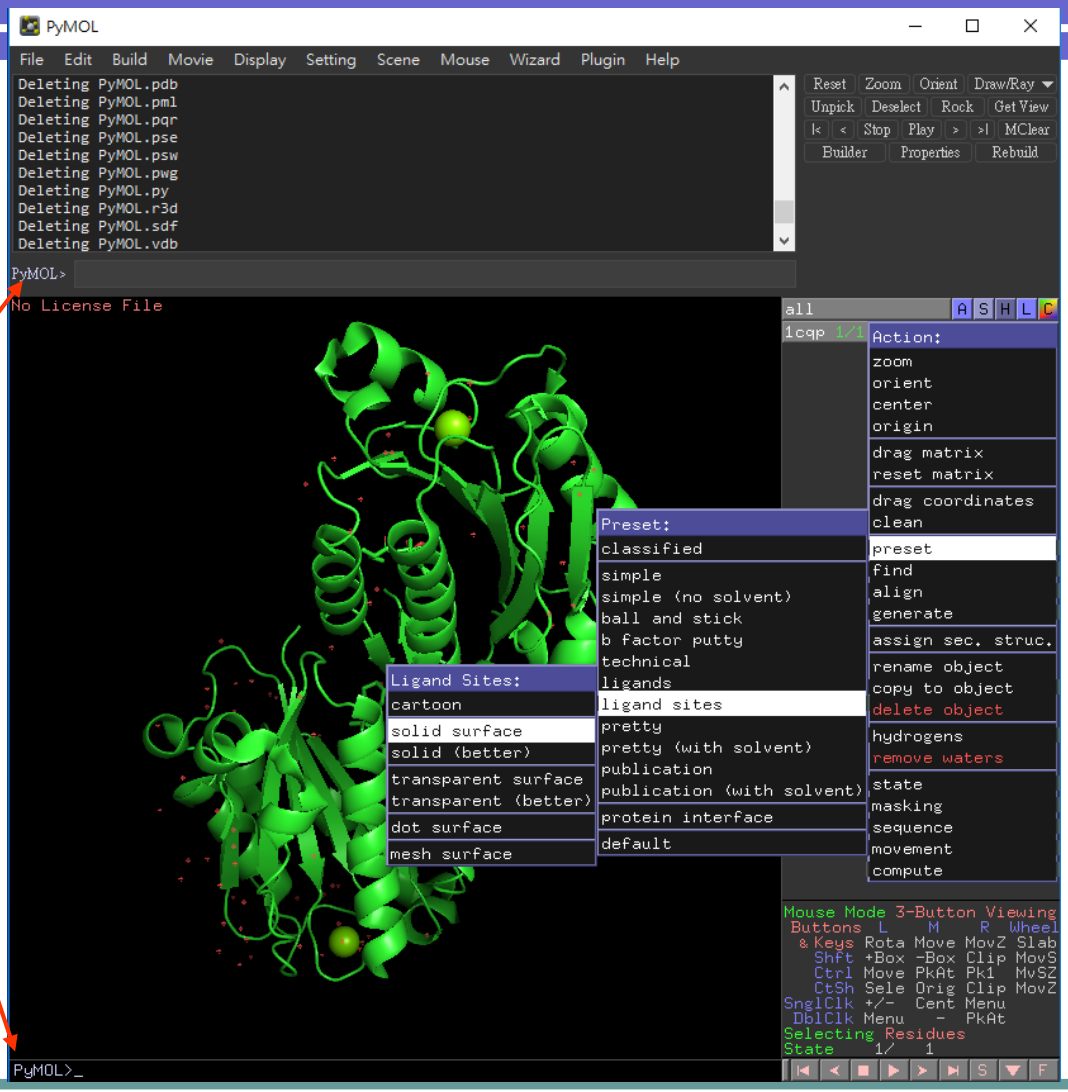

PyMOL

What PyMOL Can Do …

- Open source
- Support multiple file formats including pdb, mol2, sdf,...
- Manipulate multiple molecules
- High quality rendering
- Read in density maps in CCP4 or X-PLOR format
- Van der Waals surface rendering
- Extensive animation generation
- Written in C and Python languages

- Get the free source code or purchase licenses
  (https://pymol.org/ and https://sourceforge.net/projects/pymol/)

Visualize 1CQP using PyMOL



External GUI window

Command lines

Name panel
A: Active S: Show
H: Hide L: Label
C: Colour

Viewer window

Mouse Mode
3-Button Viewing
3-Button Editing

Frame indicator

The command language

- The majority of simple functions, such as open, save…, are available via the external GUI menu. Commands can also be used to interact with PyMOL.

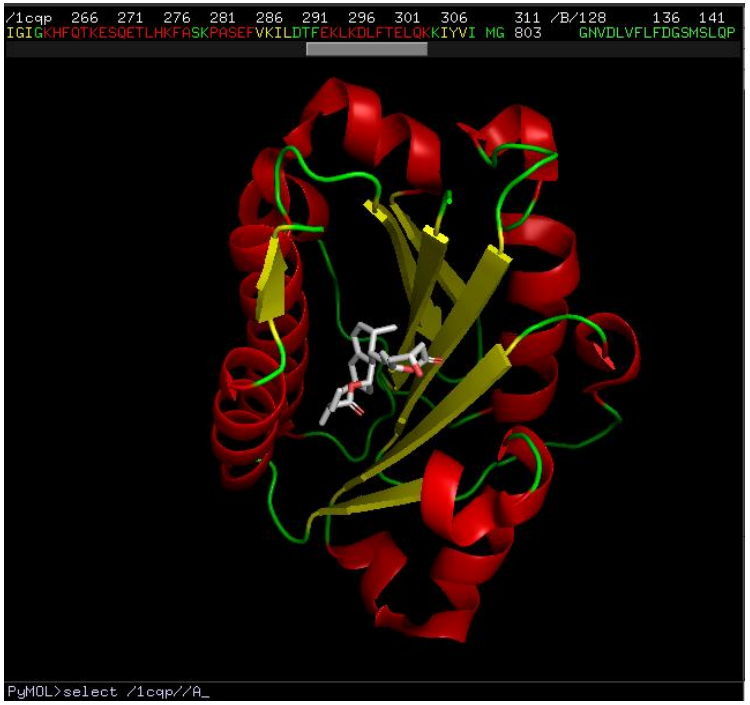
- Commonly-used commands:
  - load <$PYMOL_PATH/filename>
  - select </object_name/segiment_identifier/chain_identifier/residue_identifier/name_identifier>
  - color <color>
  - show <representation type>
  - hide < representation type >
  - set <parameter>
  - zoom <select or object>
  - distance <two atoms>
  - ⋮

- Arguments are separated by one or more commas.
  - show cartoon, chain A

PyMOL Basic Application (1)

- File → Open 1cqp.pdb
- Display → Sequence
- PyMOL>select /1cqp//A
- (sele) A → rename selection → chain_a
- 1cqp H → everything
- (chain_a) S → cartoon
- (chain_a) C → by ss
- click on 803
- (sele) S → sticks
- (sele) C → by element
- File → Export Image As → PNG

[or]Draw/Ray → Save Image to File

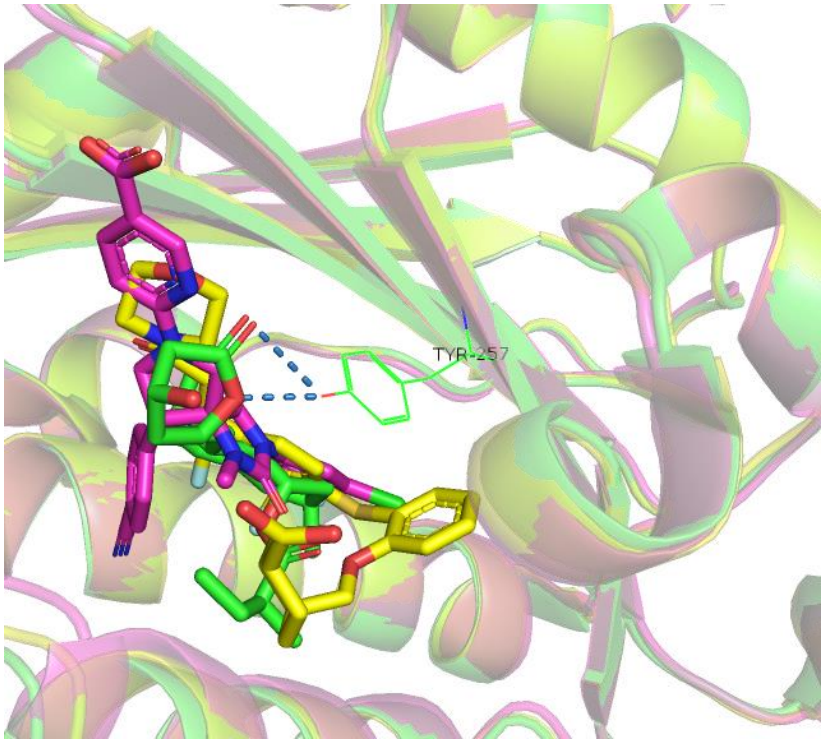


Exercise:

PyMOL>select /1cqp//B
(sele) A → remove atoms
Display → Sequence
click on 803 and Y257
(sele) S → sticks
PyMOL>distance 257/OH, 311/O1

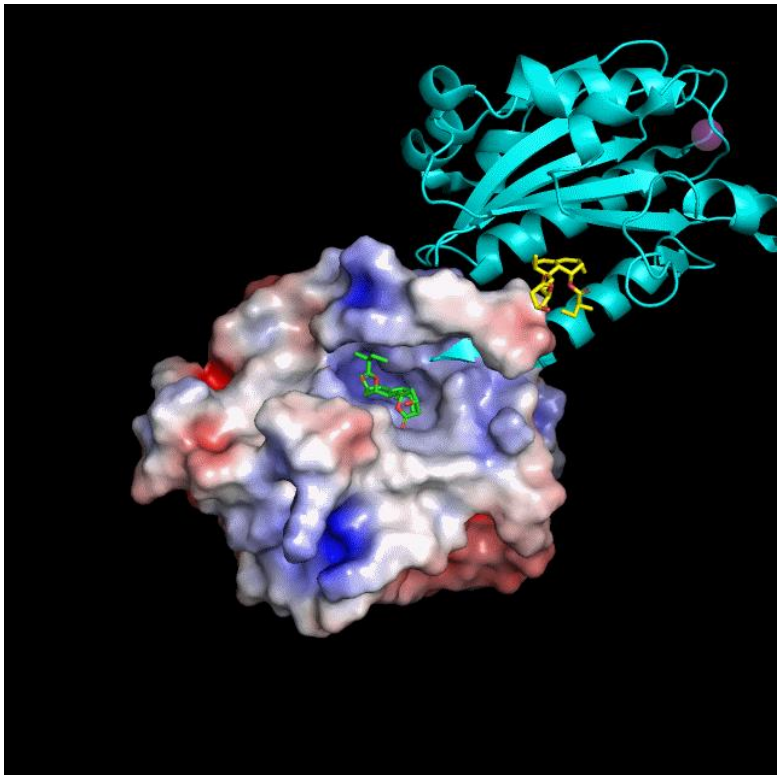
PyMOL Basic Application (2)

- Display → Background → white
- File → Open 1cqp.pdb (3e2m.pdb & 3m6f.pdb)
- PyMOL>select /1cqp//A (select /3e2m//A)
- (sele) A → rename selection → 1cqp_a (3e2m_a)
- (3e2m_a) C → yellows; A → align → to selection → 1cqp_a (RMS = 0.317Å)
- 3m6f A → align → to selection → 1cqp_a (RMS = 0.37Å)
- all H → everything
- 3m6f (1cqp_a & 3e2m_a) S → cartoon
- Click on 803 (E2M & BJZ) S → sticks; C → by element
- Click on 1cqp_a Y257 (sele) S → lines; C → by element; S → label
- Setting → Transparency → Cartoon → 50%
- Wizard → Measurement Distances; click on oxygen atoms of Y257 & 803
- measure01 C → skyblue; H → labels
- measure02 C → skyblue; H → labels
- File → Export Image As → PNG
- File → Save Session As → *.pse



PyMOL Advanced Applications

- Electrostatic surface
- Plugin
- Molecular movement
- Animation
- Script



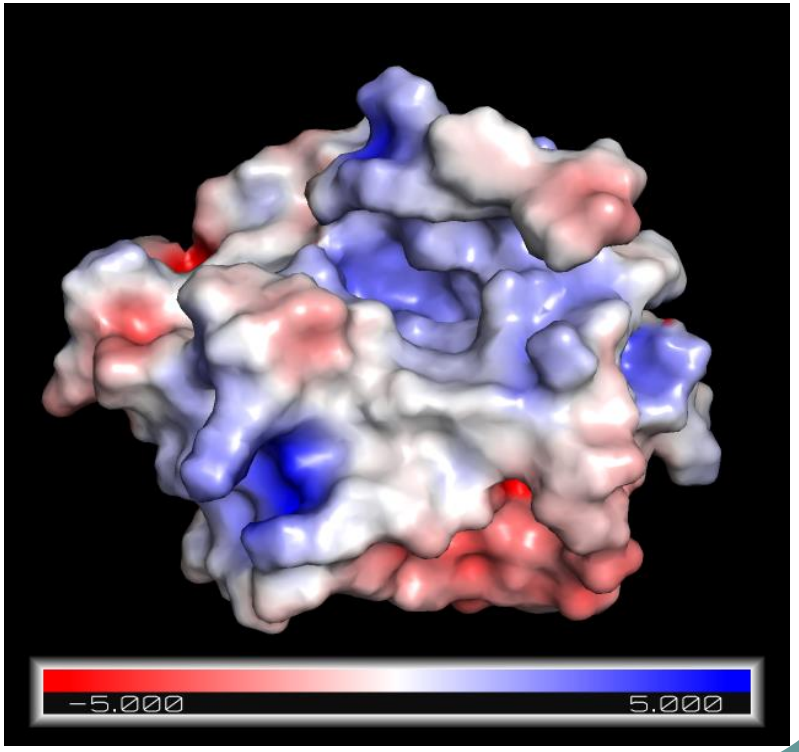
Electrostatic surface and plugin

- An input file of APBS (Adaptive Poisson-Boltzmann Solver) is required to generate the electrostatic surface for the structure. This file can be obtained from PDB2PQR web server. (http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/)

- File → Open 1cqpa.pdb
- File → Open 15209457729-pot-PE0.dx
- Plugin → Initialize Plugin System
- Plugin → APBS Electrostatics → run

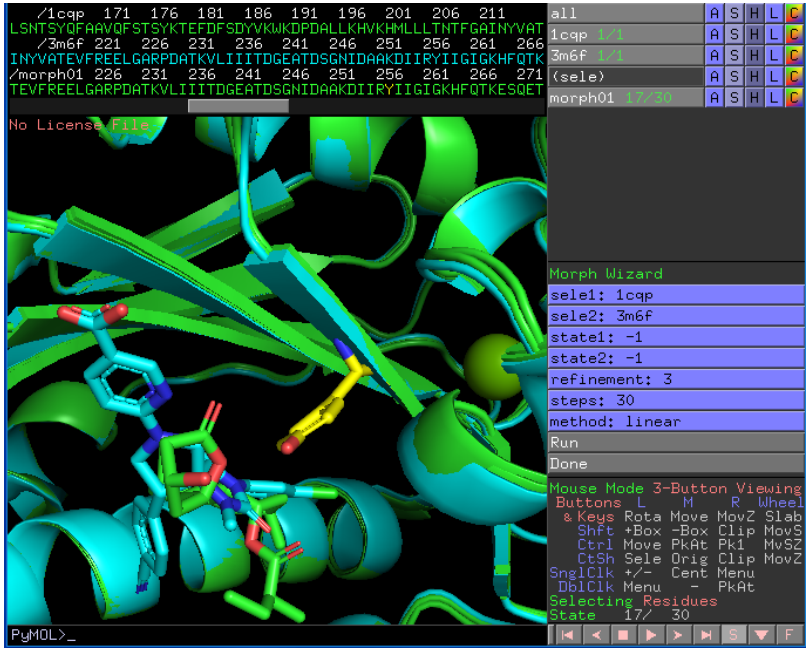
Exercise:

File → Open 1cqp.pdb
1cqp C → by chain; H → waters
Mouse Mode → Residues
click on 803 of chain B
(sele) C → by element → carbon atoms in yellow
click on MG C → magentas → purple
Setting → Transparency → Sphere → 50%



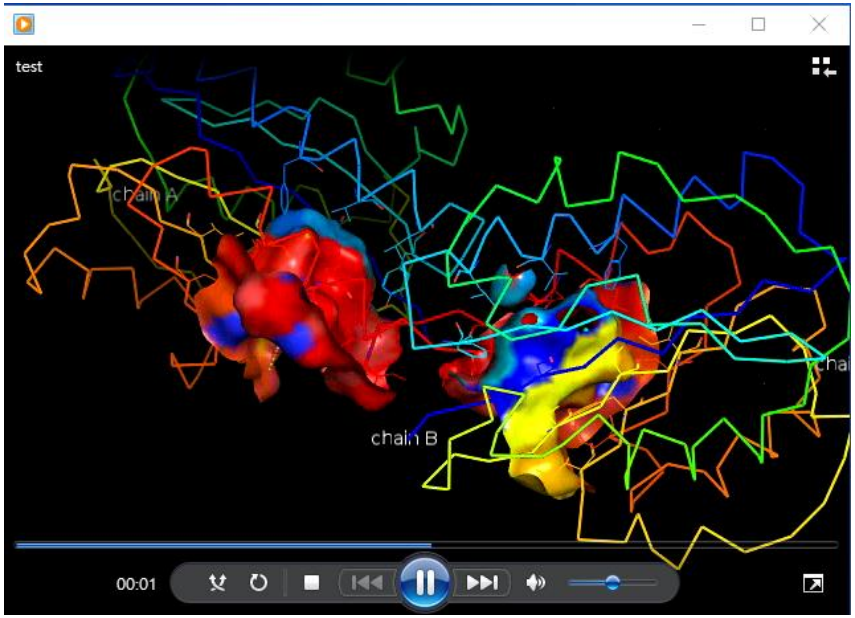
Molecular movement

- File → Open 1cqp.pdb
- select /1cqp//B
- (sele) A → remove atoms
- File → Open 3m6f.pdb
- Plugin → Alignment/Superposition → one by one
- 1cqp A → center
- 1cqp A → generate → wizard…
- Morph Wizard → sele2: **3m6f;** method: **linear** → run
- Display → Sequence
- click on Y257 of morph01
- (sele) S → sticks; C → by element → carbon atoms in yellow
- Morph Wizard → run



Movie and animation

- File → Open 1cqp.pdb
- 1cqp A → present → ligand sites → solid surface
- 1cqp L → chains
- Movie → 30 FPS; Program → Camera Loop → X-Roll → 4 seconds
- File → Export Movie As → Save Movie as
- File → Export Movie As → PNG Images

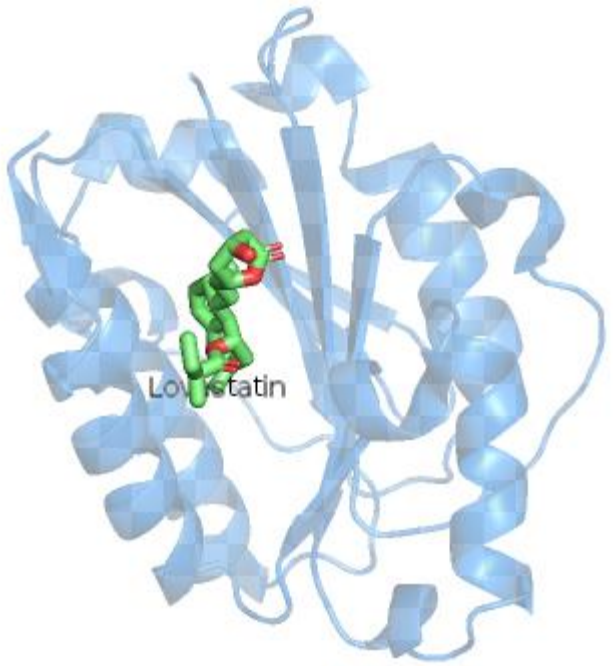


- At this step, 120 PNG images are saved. An external software is used to convert these images into a movie or an animation.

A Simple PyMOL Script

- Make a text file of commands:
  - BasicScript.pml

```
load C:\Users\USERX\Desktop\1cqp.pdb
bg_color white
select chB, chain B
remove chB
hide everything
show cartoon
set cartoon_color, marine
set cartoon_transparency, 0.5
select ligand, /1cqp//A/311
show sticks,ligand
set_bond stick_radius, 0.5
label (311/C21),"%s" % "Lovastatin"
alter 190-192/,ss="L"
alter 267-268/,ss="L"
rebuild
center all
draw
```



- Run this script file by typing in:
  - run FILEPATH\BasicScript.pml

  [or] @FILEPATH\BasicScript.pml

Resources: PyMOLWiki

Practice Makes Perfect !

Thank You