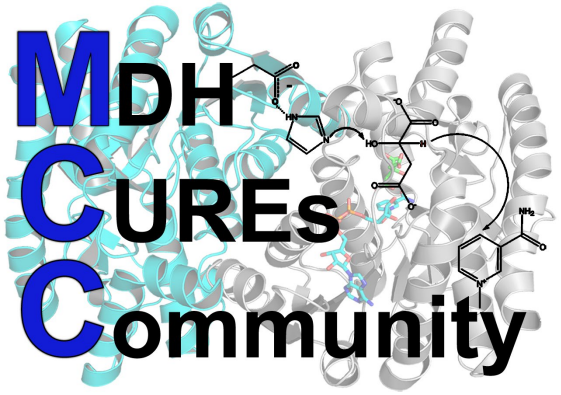


Resources for Virtual Research Projects

Developing Project Ideas and Background Molecular Visualization

A Tutorial to accompany "Teaching Virtual Protein-Centric CUREs and UREs Using Computational Tools"

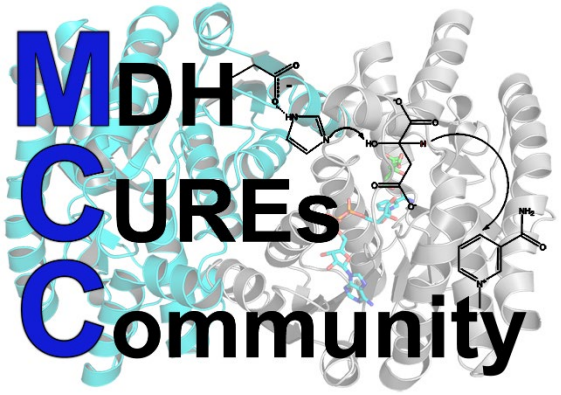
Anthony Bell[1], Laura Christian[2], David Hecht[3], Kathryn Huisinga[4] John Rakus[5] & Ellis Bell[1]

1, University of San Diego, 2, RPI, 3, SWCC, 4, Malone University, 5, Marshall University

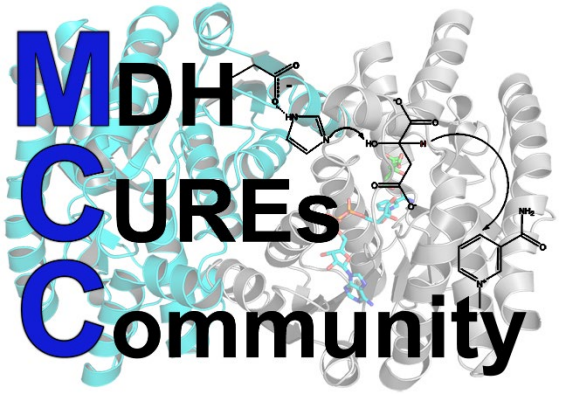
Web Sites: https://mdh-cures-community.squarespace.com/virtual-cures-and-ures

MCC: https://mdh-cures-community.squarespace.com/

Bell Labs: https://www.molecularlifesciences.org



Obtaining a pdb file from the Protein Data Base



https://www.rcsb.org/

Nucleic Acids Research, 2000, Vol. 28, No. 1 235–242

The Protein Data Bank

Helen M. Berman[1,2,*], John Westbrook[1,2], Zukang Feng[1,2], Gary Gilliland[1,3], T. N. Bhat[1,3], Helge Weissig[1,4], Ilya N. Shindyalov[4] and Philip E. Bourne[1,4,5,6]

[1]Research Collaboratory for Structural Bioinformatics (RCSB), [2]Department of Chemistry, Rutgers University, 610 Taylor Road, Piscataway, NJ 08854-8087, USA, [3]National Institute of Standards and Technology, Route 270, Quince Orchard Road, Gaithersburg, MD 20899, USA, [4]San Diego Supercomputer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0505, USA, [5]Department of Pharmacology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0500, USA and [6]The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

Received September 20, 1999; Revised and Accepted October 17, 1999

ABSTRACT

The Protein Data Bank (PDB; http://www.rcsb.org/pdb/) is the single worldwide archive of structural data of biological macromolecules. This paper describes the goals of the PDB, the systems in place for data deposition and access, how to obtain further information, and near-term plans for the future development of the resource.

RCSB PDB | Deposit | Search | Visualize | Analyze | Download | Learn | More | MyPDB

RCSB PDB PROTEIN DATA BANK

163949 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

Enter search term(s)

Advanced Search | Browse Annotations

PDB-101 | wwPDB | EMDataResource | NDB Nucleic Acid Database | Worldwide Protein Data Bank Foundation

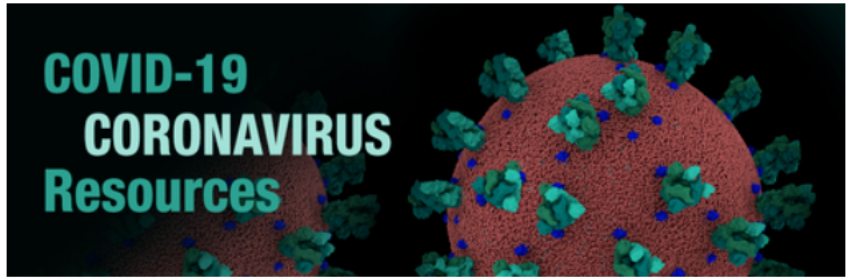
- Welcome
- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

A Structural View of Biology

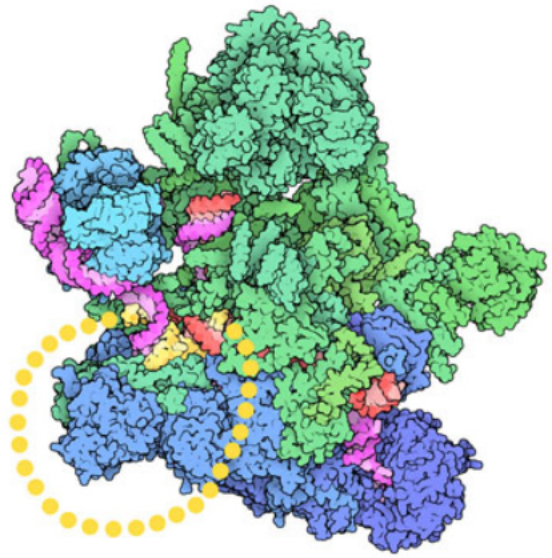
This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.



May Molecule of the Month



Spliceosomes

Contact Us

Latest Entries — As of Tue May 12 2020

Features & Highlights

News | Publications

QUERY: Full Text = "Malate Dehydrogenase"

Open In Query Builder | MyPDB Login

Advanced Search Query Builder

Refinements

Clear All

SCIENTIFIC NAME OF SOURCE ORGANISM Clear

- Homo sapiens (47321)
- Mus musculus (6816)
- Escherichia coli (5856)
- synthetic construct (4222)
- Escherichia coli K-12 (3427)
- Rattus norvegicus (3152)
- Bos taurus (2973)
- Saccharomyces cerevisiae (2729)
- Gallus gallus (1811)
- Saccharomyces cerevisiae S288C (1707)

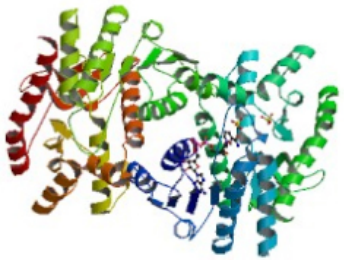
More...

TAXONOMY Clear

- Eukaryota (87747)
- Bacteria (57366)

Summary | Gallery | Compact | -- Tabular Report -- | ↓ Score | Download Selected Files | Select All

Displaying 1 to 25 of 160439 **Structures** Page 1 of 6418 ← Previous | Next → Display 25 per page



3D View

1IB6

Download File | View File

CRYSTAL STRUCTURE OF R153C E. COLI MALATE DEHYDROGENASE

Bell, J.K., Yennawar, H.P., Wright, S.K., Thompson, J.R., Viola, R.E., Banaszak, L.J.

(2001) J Biol Chem **276**: 31156-31162

Released	2001-09-19
Method	X-RAY DIFFRACTION 2.1 Å
Organisms	Escherichia coli
Macromolecule	MALATE DEHYDROGENASE (protein)
Unique Ligands	NAD, SO4

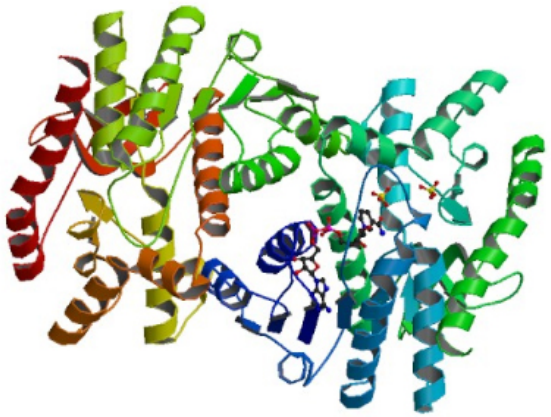
1MLD

Download File | View File

REFINED STRUCTURE OF MITOCHONDRIAL MALATE DEHYDROGENASE

Contact Us

Biological Assembly 1



3D View: Structure | Electron Density | Ligand Interaction

Standalone Viewers (To be retired June 1st)

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)
- fo-fc Map (DSN6)
- 2fo-fc Map (DSN6)

1IB6

CRYSTAL STRUCTURE OF R153C E. COLI MALATE DEHY

DOI: 10.2210/pdb1IB6/pdb

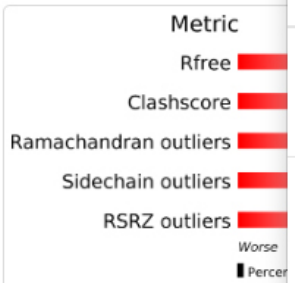
Classification: OXIDOREDUCTASE
Organism(s): Escherichia coli
Expression System: Escherichia coli
Mutation(s): Yes

Deposited: 2001-03-27 **Released:** 2001-09-19
Deposition Author(s): Bell, J.K., Yennawar, H.P., Wright, S.K., Thompson, J.

Experimental Data Snapshot

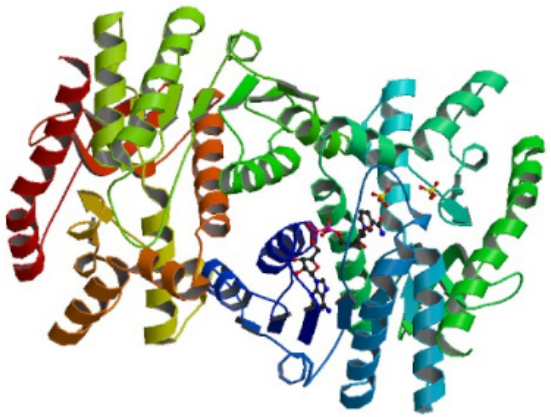
Method: X-RAY DIFFRACTION
Resolution: 2.10 Å
R-Value Free: 0.244
R-Value Work: 0.192
R-Value Observed: 0.192

wwPDB Validation



Contact Us

Biological Assembly 1



3D View: Structure | Electron Density | Ligand Interaction

Standalone Viewers (To be retired June 1st) Protein Workshop | Ligand Explorer

Global Symmetry: Cyclic - C2 (3D View)
Global Stoichiometry: Homo 2-mer - A2

Display Files | Download Files

1IB6

CRYSTAL STRUCTURE OF R153C E. COLI MALATE DEHYDROGENASE

DOI: 10.2210/pdb1IB6/pdb

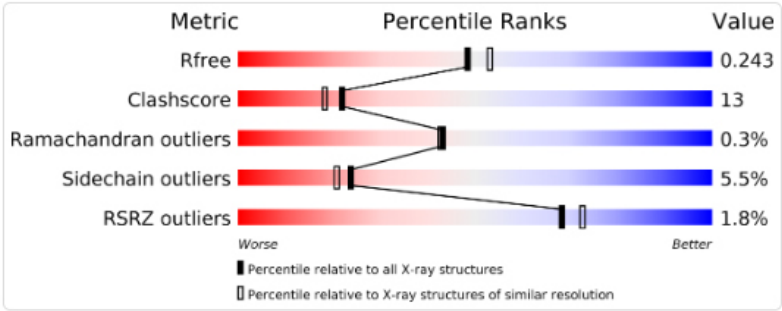
Classification: OXIDOREDUCTASE
Organism(s): Escherichia coli
Expression System: Escherichia coli
Mutation(s): Yes

Deposited: 2001-03-27 **Released:** 2001-09-19
Deposition Author(s): Bell, J.K., Yennawar, H.P., Wright, S.K., Thompson, J.R., Viola, R.E., Banaszak, L.J.

Experimental Data Snapshot

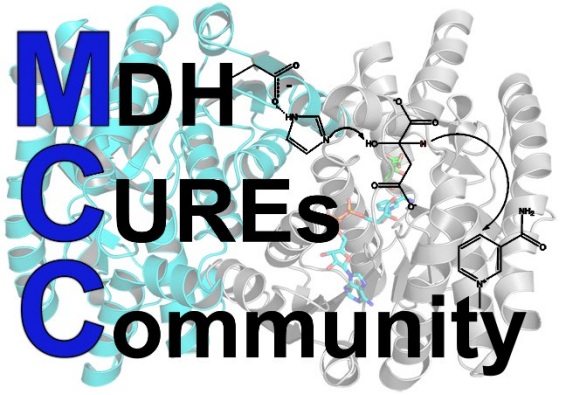
Method: X-RAY DIFFRACTION
Resolution: 2.10 Å
R-Value Free: 0.244
R-Value Work: 0.192
R-Value Observed: 0.192

wwPDB Validation 3D Report | Full Report



This is version 1.2 of the entry. See complete history.

Contact Us



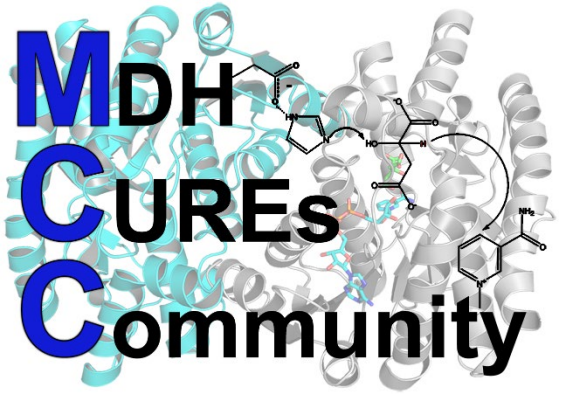
What does a pdb file look like, and what can you do with it?

The structure of a pdb file

The pdb file for the E Coli enzyme begins:

```
HEADER    OXIDOREDUCTASE(NAD(A)-CHOH(D)          25-MAR-93   1EMD       1EMD   2
COMPND    MALATE DEHYDROGENASE (E.C.1.1.1.37)                           1EMD   3
SOURCE    ESCHERICHIA COLI                                              1EMD   4
AUTHOR    M.D. HALL, L.J. BANASZAK                                      1EMD   5
REVDAT   1   31-OCT-93 1EMD    0                                        1EMD   6
JRNL        AUTH   M.D. HALL, L.J. BANASZAK                           1EMD   7
JRNL        TITL   CRYSTAL STRUCTURE OF A TERNARY COMPLEX OF          1EMD   8
JRNL        TITL 2 ESCHERICHIA $COLI MALATE DEHYDROGENASE, CITRATE     1EMD   9
JRNL        TITL 3 AND /NAD$ AT 1.9 ANGSTROMS RESOLUTION               1EMD  10
JRNL        REF    TO BE PUBLISHED                                     1EMD  11
JRNL        REFN                                          353          1EMD  12
REMARK   1                                                            1EMD  13
REMARK   2                                                            1EMD  14
REMARK   2 RESOLUTION. 1.9 ANGSTROMS.                                 1EMD  15
REMARK   3                                                            1EMD  16
REMARK   3 REFINEMENT.                                                1EMD  17
REMARK   3   PROGRAM                 X-PLOR                           1EMD  18
REMARK   3   AUTHORS                 BRUNGER                          1EMD  19
REMARK   3   R VALUE                 0.195                            1EMD  20
REMARK   3   RMSD BOND DISTANCES     0.012 ANGSTROMS                  1EMD  21
REMARK   3   RMSD BOND ANGLES        1.65  DEGREES                    1EMD  22
```

The "header" simply gives information about the general class of enzyme, the date of the file and the file name.
The final columns are the file name and line number which runs throughout the file.
The "Compound" line gives the name of the enzyme and the Enzyme Commision number.
"Source" indicates the organism that the protein was obtained from, in this case E Coli.
"Author" is the person or people who published the structure
"RevDat", for Revision Date is to indicate when revisions to the file were received.
"JRNL" is the citation to the relevant publication
"Remark" lines are for commentary about the structure and usually indicate the resolution, the program used for the refinement of the structure, the R factor value, which indicates how good the data is and is defined by:
$R = S|Fo-Fc|/SFo$
Where Fo is the actual data point and Fc is the modeled parameter.
and RMSD [root mean square deviations] for the bond distances and bond angles in the structure.



Next in the pdb file comes the 'SEQRES" section which lists the amino acid sequence of the protein with appropriate "FTNOTE" lines-in this case indicating that residue 120 is a cis-proline.
Next comes a listing of lines for "HET" which indicates whether any other molecules are in the structure-this is often the substrate, analog or inhibitor etc. Followed by the Formula of the HET molecules and a line for water molecules in the structure.
This is followed in turn by listings of structure, first "HELIX", then "SHEET" and finally "TURN" lines

The section:¶

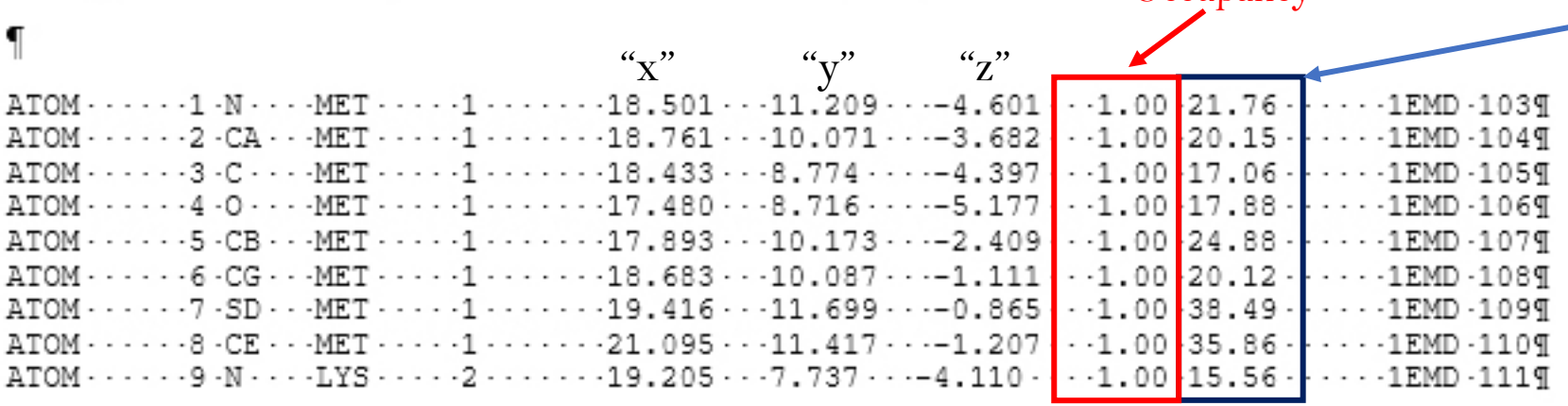
```
SITE·····1·ACT··5·ARG····81···ARG····87···ASP···150···ARG···153·········1EMD··94¶
SITE·····2·ACT··5·HIS···177·····································1EMD··95¶
CRYST1···116.800···43.050···83.740···90.00·130.10···90.00·C·2·····4···1EMD··96¶
ORIGX1·····1.000000···0.000000···0.000000·····0.00000·····1EMD··97¶
ORIGX2·····0.000000···1.000000···0.000000·····0.00000·····1EMD··98¶
ORIGX3·····0.000000···0.000000···1.000000·····0.00000·····1EMD··99¶
SCALE1·····0.008562···0.000000···0.007210·····0.00000·····1EMD·100¶
SCALE2·····0.000000···0.023229···0.000000·····0.00000·····1EMD·101¶
SCALE3·····0.000000···0.000000···0.015612·····0.00000·····1EMD·102¶
¶
```

The "Site" lines indicate in this case that 5 residues that are part of the active site have been identified as R81, R87, D150, R153 and H177.¶

¶

"CRYST!" indicates the unit cell parameters, the space group, in this case c2 and the z score: the number of asymmetric units per unit cell, in this case 4.¶

Finally, the actual three-dimensional coordinates begin:




```
                                                "x"       "y"       "z"
ATOM      1  N   MET     1      18.501   11.209   -4.601  1.00 21.76      1EMD 103
ATOM      2  CA  MET     1      18.761   10.071   -3.682  1.00 20.15      1EMD 104
ATOM      3  C   MET     1      18.433   8.774    -4.397  1.00 17.06      1EMD 105
ATOM      4  O   MET     1      17.480   8.716    -5.177  1.00 17.88      1EMD 106
ATOM      5  CB  MET     1      17.893   10.173   -2.409  1.00 24.88      1EMD 107
ATOM      6  CG  MET     1      18.683   10.087   -1.111  1.00 20.12      1EMD 108
ATOM      7  SD  MET     1      19.416   11.699   -0.865  1.00 38.49      1EMD 109
ATOM      8  CE  MET     1      21.095   11.417   -1.207  1.00 35.86      1EMD 110
ATOM      9  N   LYS     2      19.205   7.737    -4.110  1.00 15.56      1EMD 111
```

In this section, each atom is numbered in sequence and the element type and some additional information given:

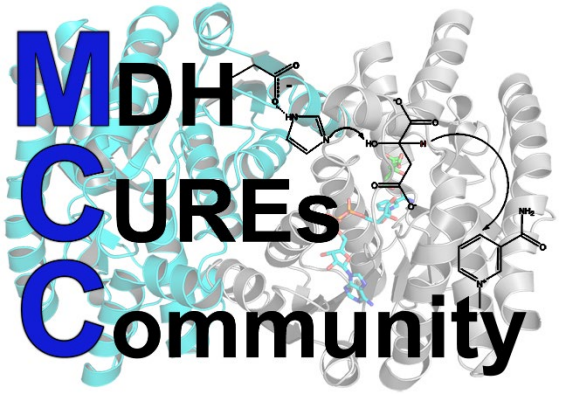
For example, CA is the alpha-carbon, CB is the beta carbon etc before the residue type and number given. After the residue number the next three numbers are the three-dimensional, Cartesian coordinates of the atom, followed by the "occupancy" of the electron density for that atom: usually 1.0. The next number, the so-called B factor or temperature factor gives an indication of the local motion of the atom: a low number indicates little motion while a high number indicates significant motion of the atom. While it is quite usual for exposed side chains such as the charged or hydrophilic residues to have relatively high temperature factors [up to 30-50] the backbone atoms often have single digit temperature factors unless significant motion is observed.

The final two columns are simply the file name and the file line number.

The end of the protein sequence [remember there may be more than one polypeptide chain: usually indicated by 1A, 1B, 1C…etc. comes the TER statement:

```
TER    2279      LYS   312                                                1EMD2381
```

Indicating the end of the protein.



This is followed by the coordinates of any heteromolecules such as ligands or water: in this case the ter statement is followed by:¶

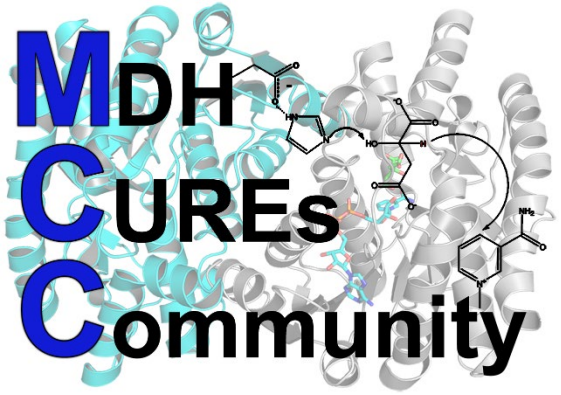
```
HETATM 2280  C1  CIT   313       5.426  -7.608  15.720  1.00 25.43      1EMD2382¶
HETATM 2281  O1  CIT   313       4.760  -7.292  16.696  1.00 22.77      1EMD2383¶
HETATM 2282  O2  CIT   313       5.710  -8.783  15.470  1.00 23.22      1EMD2384¶
HETATM 2283  C2  CIT   313       5.220  -6.691  14.572  1.00 22.61      1EMD2385¶
HETATM 2284  C3  CIT   313       5.865  -5.268  14.934  1.00 23.43      1EMD2386¶
HETATM 2285  O7  CIT   313       7.150  -5.379  15.668  1.00 20.35      1EMD2387¶
HETATM 2286  C4  CIT   313       6.220  -4.401  13.666  1.00 23.09      1EMD2388¶
HETATM 2287  C5  CIT   313       6.872  -5.350  12.811  1.00 31.49      1EMD2389¶
HETATM 2288  O3  CIT   313       7.741  -6.011  13.344  1.00 34.43      1EMD2390¶
HETATM 2289  O4  CIT   313       6.506  -5.733  11.536  1.00 32.18      1EMD2391¶
HETATM 2290  C6  CIT   313       4.833  -4.564  15.697  1.00 21.46      1EMD2392¶
HETATM 2291  O5  CIT   313       5.262  -3.941  16.722  1.00 17.15      1EMD2393¶
HETATM 2292  O6  CIT   313       3.669  -4.640  15.174  1.00 19.96      1EMD2394¶
HETATM 2293  AP  NAD   314       6.405  -11.130  4.954  0.69 26.19      1EMD2395¶
HETATM 2294  AO1 NAD   314       6.596  -12.584  5.180  0.69 27.85      1EMD2396¶
HETATM 2295  AO2 NAD   314       5.164  -10.631  4.350  0.69 19.49      1EMD2397¶
```

Note in this case the occupancy for the ligand Citrate is 1.0 while that for the NAD is 0.69

```
HETATM 2337  O   HOH   315       1.022  8.231  -1.217  1.00 23.42      1EMD2439¶
HETATM 2338  O   HOH   316       8.525  -7.240  3.696  1.00 16.79      1EMD2440¶
HETATM 2339  O   HOH   317      20.324  -12.888 -0.305  1.00 30.95      1EMD2441¶
```

[Note not all of the waters are shown here]¶

The B factors of so called “Crystallographic” waters in a structure are governed by the number of polar contacts made to the protein-ligand complex



One of the nice things about pdb files is that you can easily copy sections of the file and use just those sections in a viewer of some type. For example if you have two subunits and several ligands it is often convenient [see why later] to make separate files of each subunit and each bound ligand. You can do this for the various molecular components in the pdb file and save each set of coordinates in plain text format. While plain text format is read as pdb format by most molecular visualization programs, it is easy to simply change the extension from .txt to .pdb.

MDH CUREs Community

Getting & Using PyMol

https://pymol.org/2/

https://pymol.org/edu/?q=educational

SCHRÖDINGER.

Registration For Educational-Use-Only PyMOL Builds

Schrödinger offers **Educational-use-only** PyMOL builds available at no cost to **teachers and high school and college students** (including online courses, homeschooling, etc.) for classroom instruction, homework assignments, and to provide a means for creating high quality figures. Please note that it is not provided for the purposes of academic research or publication.

-> FAQ (Frequently Asked Questions)

The Educational-use-only PyMOL builds are provided "AS IS" with no obligation to grant download access, fix bugs, furnish updates, provide documentation, or meet any other need related to the educational-use PyMOL builds.

If you intend to use PyMOL products for academic research or publication, please purchase an Academic PyMOL subscription, which includes access to technical support, screencasts, and additional resources. See http://pymol.org/academic.

I am a: ▾

Your First Name:

Your Last Name:

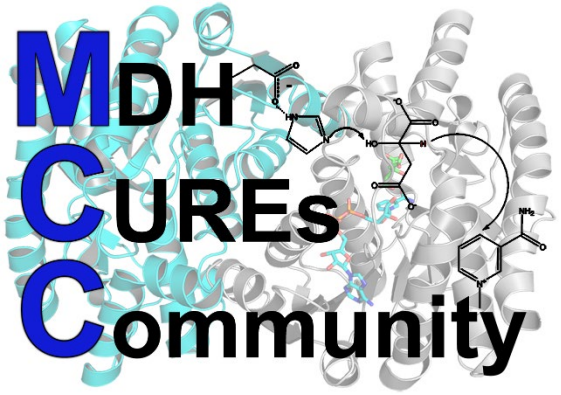
Your Email Address:

Your Telephone Number:

Institution:

Comments (optional):

Continue



As necessary there are a number of useful Pymol video tutorials you can view that will help your progress through the worksheet.

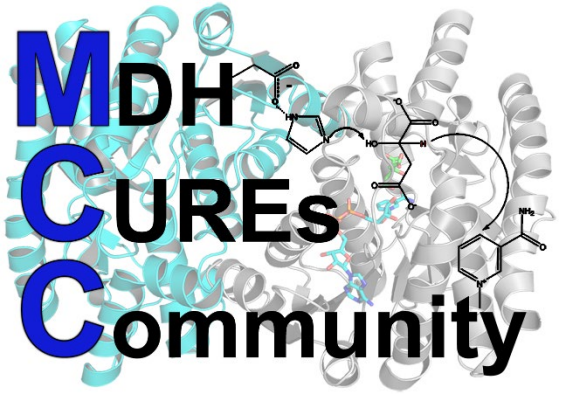
PyMOL Tutorial Interface Part 1 https://vimeo.com/44836592 22:27 min

PyMOL Tutorial Animation Part 2 https://vimeo.com/44801178 12:07 min

Basics of Pymol Part 1 http://www.youtube.com/watch?v=ai7p9Neguks 13:14 min
(downloading, color coding, saving)

Basics of Pymol Part 2 http://www.youtube.com/watch?v=uxa-9UYnIAw 13:55 min
(measuring tool, polar contacts, mutagenesis)

NOTE: In the following, some of the things that you can do to create and capture protein structure images are illustrated. To export images from PyMOL, use File → Save Image As → PNG. These .png picture files can be directly inserted into a document. All images should annotated with a figure title and figure legend describing the image composition.



The PyMol Worksheet:

Using PyMol, you can:

1) Create a complete cartoon image of the protein and ligand as shown in the basic commands and rendering tutorial videos (4 & 5 below). Insert the image into a word document and annotate to explain your picture. This should then be uploaded to the appropriate part of your eLN.

2) Zoom in on and display the ligand binding site (using the following hint – With the ligand showing, shift and drag a box to select the protein region all around the ligand. For the selected region, then under A (action), choose preset → ligand sites, and then choose how to display). Display to highlight all the binding/interacting/interesting amino acid residues on the screen. The image can then be captured, inserted into a document, and annotated appropriately

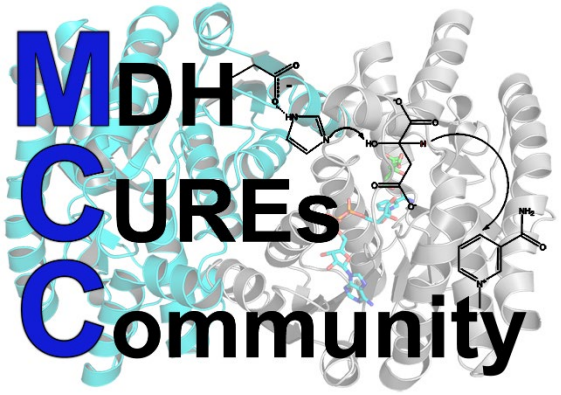
3) Measure the distance from 4 or 5 of the binding residues in your protein to the bound ligand/small molecule using the Measurement Wizard. Capture, insert, and annotate the image as before.

4) Use the Mutagenesis Wizard, mutate your assigned amino acid to: a) a conserved amino acid, b) an amino acid with the opposite chemical characteristics, and c) the designated mutation. Describe the changes in structure when you perform each mutation. Capture the more dramatic instance, then insert and annotate the image as before.

5) You can use a homologous protein, for example pig mitochondrial MDH or E coli MDH, to create an overlay of both structures. For your project you should overlay subunit C with subunit D, corresponding to the closed (citrate bound) and open forms of the loop respectively. Capture, insert, and annotate the image as before.

6) Create a publication quality image. . Use a white background and have some fun with this image. Capture, insert, and annotate the image as before.

7) Create a movie in PyMOL with your protein – see video tutorial below



- Advanced PyMol Features and Tutorials
- – Tutorial 1: Scene-based movies
- – Tutorial 2: Advanced Movies with Morphing
- – Tutorial 3: Aligning Structures and Calculating Poisson Boltzmann Electrostatics
- ````````Tutorial 4: Advanced Analyses (Ligand Binding Site Analysis, Distance Measurements, Mutagenesis, and Dihedral Angle Changes)

MDH CUREs Community

Chimera is a molecular visualization program somewhat similar to PyMol developed by UCSF
Certain applications in SwissDock work better with Chimera that PyMol
https://www.cgl.ucsf.edu/chimera/

RBVI

about projects people publications resources visit us search

UCSF CHIMERA

an Extensible Molecular Modeling System

UCSF Chimera is a program for the interactive visualization and analysis of molecular structures and related data, including density maps, trajectories, and sequence alignments. High-quality images and animations can be generated. Chimera includes complete documentation and is free of charge for academic, government, nonprofit, and personal use. Chimera development was supported by the National Institutes of Health (P41-GM103311).

UCSF ChimeraX is the next-generation molecular visualization program from the RBVI, following UCSF Chimera. We encourage Chimera users to try ChimeraX for much better performance with large structures, as well as other major advantages. ChimeraX replaces a significant subset of Chimera features, includes several completely new features, and is under active development. Users may certainly choose to use both programs, and it is fine to have both installed.

Quick Links

Documentation
- Getting Started
- User's Guide
- Command Index
- Tutorials and Videos
- Guide to Volume Data
- Release Notes

Download
- What's New in Daily Builds
- Map of Download Locations

Galleries
- Image Gallery
- Animation Gallery

Publications and Talks
Related Databases and Software
Citing Chimera
Contact Us

Recent Citations

Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. Dodonova SO, Zhu F *et al. Nature.* 2020 Apr 30;580(7805):669-672.

Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. Zhang L, Lin D *et al. Science.* 2020 Apr 24;368(6489):409-412.

Determination of the melanocortin-4 receptor structure identifies Ca2+ as a cofactor for ligand binding. Yu J, Gimenez LE *et al. Science.* 2020 Apr 24;368(6489):428-433.

The ABC exporter IrtAB imports and reduces mycobacterial siderophores. Arnold FM, Weber MS *et al. Nature.* 2020 Apr 16;580(7803):413-417.

Integrative modeling of a Sin3/HDAC complex sub-structure. Banks CAS, Zhang Y *et al. Cell Rep.* 2020 Apr 14;31(2):107516.

(Previously featured citations...)

Feature Highlight

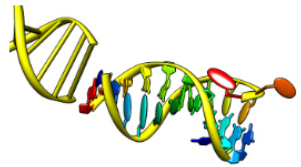
Nucleotides

Special representations of DNA and RNA can be displayed with the Nucleotides tool or the command nucleotides. Different levels of abstraction are available. The figure shows a ribbon backbone combined with the following sidechain (sugar/base) options:

- ladder rungs
- filled-ring atomic representations
- "lollipops" in which bases are shown as ellipsoids and sugars as tubes

Bases can also be displayed as boxes or elliptical tubes, with or without bumps to indicate orientation. The colors of the special representations will update automatically to match the corresponding atoms.

(More features...)



Gallery Sample

Peroxiredoxin Wreath

Peroxiredoxins are enzymes that help cells cope with stressors such as high levels of reactive oxygen species. The image shows a decameric peroxiredoxin from human red blood cells (Protein Data Bank entry 1qmv), styled as a holiday wreath.

See also the RBVI holiday card gallery.

(More samples...)

Chimera Search

Go
Google™ Search

News

November 13, 2019

Chimera production release 1.14 is now available. See the release notes for what's new.

September 21, 2019

A production release candidate (version 1.14) is available; please try it and report any problems. See the release notes for what's new.

November 17, 2018

Chimera production release 1.13.1 is now available; see the release notes for what's new. The Mac version requires OS 10.10 or later.

(Previous news...)

Upcoming Events

MDH CUREs Community

Chimera Tutorials & Videos:

https://www.cgl.ucsf.edu/chimera/tutorials.html
https://www.cgl.ucsf.edu/chimera/videodoc/videodoc.html