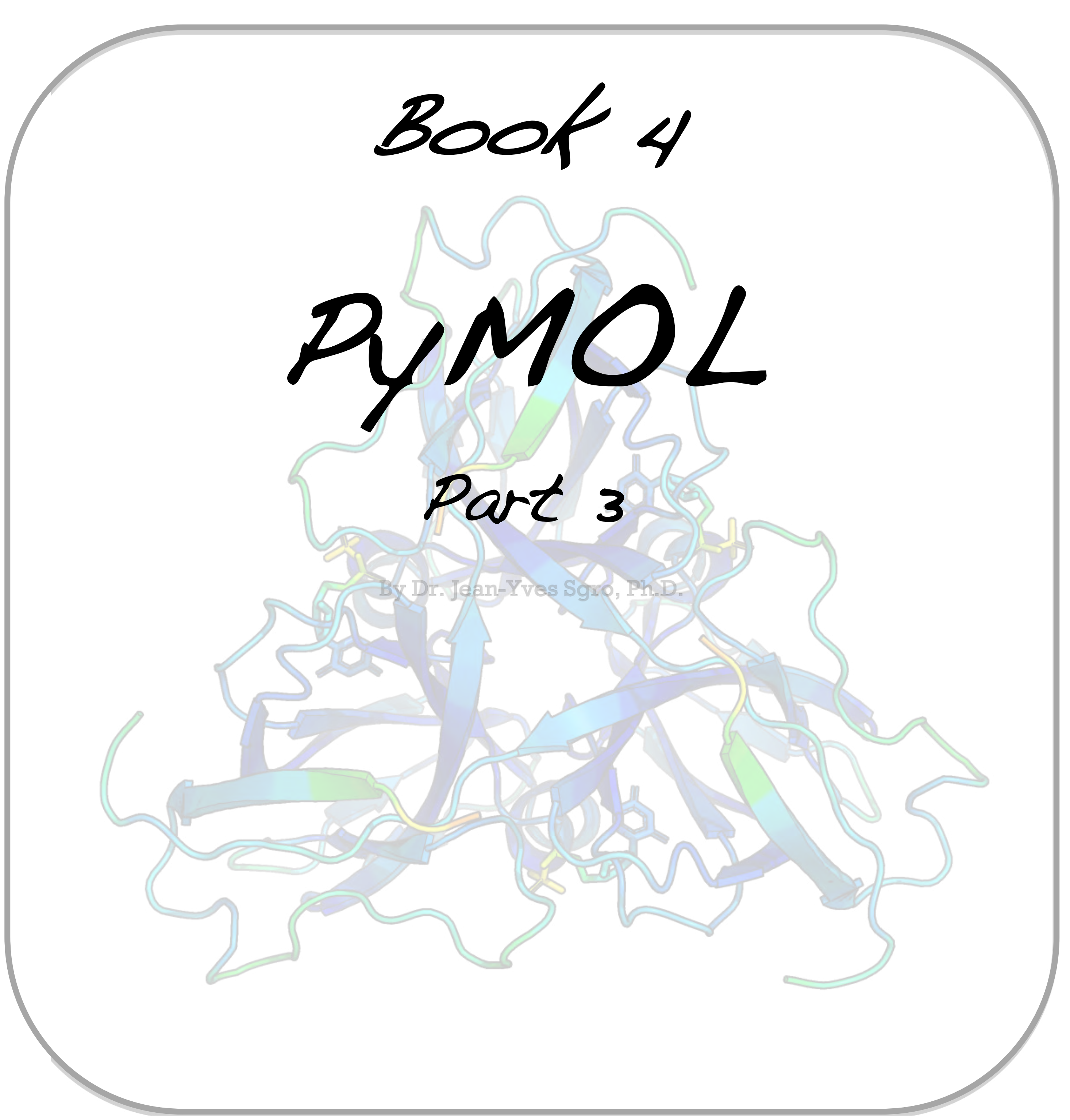


Book 4

PyMOL

Part 3

By Dr. Jean-Yves Sgro, Ph.D.

Original © 1997-2019 Dr. Jean-Yves Sgro, Ph.D.

Permission is granted to make and distribute copies, either in part or in full and in any language, of this document on any support provided the above copyright notice is included in all copies. Permission is granted to translate this document, either in part or in full, in any language provided the above copyright notice is included. The sale of this, or any derivative work on physical media cannot exceed the actual price of support media.

This work is released under the Creative Commons license
Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)

https://creativecommons.org/licenses/by-nc-sa/4.0/

You are free to:

Share — copy and redistribute the material in any medium or format **Adapt** — remix, transform, and build upon the material

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes.

ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Cover: PyMOL rendering of PDB 1DUD (*Crystal structure of the Escherichia coli dUTPase in complex with a substrate analogue (dUDP).* Larsson, G., Svensson, L.A., Nyman, P.O. (1996) Nat.Struct.Mol.Biol. **3**: 532-538)

Preamble

For many years I participated in the teaching of Prof. Ann Palmenberg classes, in particular teaching about molecular graphics on the Desktop at a time when computer use in the classroom was not yet preeminent. This class was a useful complement to the main topic of Sequence Analysis and Evolution also using computers.

I personally used what was called "Unix Workstations" (Silicon Graphics, sgi) which had powerful graphics and rather beautiful "photorealistic" renderings. For teaching molecular graphics on the Desktop, I first used Rasmol, which was an amazing software fitting in about half of a 3.5" floppy disk or about 500 kilobytes. Rasmol included a sophisticated line command language but lacked the beautiful "photorealistic" renderings of the workstations. This was remedied by adding two other software in the course, one for the "publication quality renderings" (VMD) and the other for the modeling abilities of side-chain mutations and automated 3D superimposition of structures (Swiss PDB viewer later called DeepView.)

When PyMOL was still in preliminary development at version 0.99 I spent one intense week porting all the class material to PyMOL. Now, rather than using three different software, all was possible with only PyMOL. Over the years I extended and updated the PyMOL course material.

The UW-Madison Biochemistry students were the primary audience for these classes in courses Biochem 660 and 712, and occasionally in Biochem 511. I offered the PyMOL class in Biochem 660 for over a decade. The PyMOL tutorial and preliminary molecular graphics and file format introduction were part of a very large, made-to-order physical copy of a class book of about 500 pages that also contained tutorials on using other software. The PyMOL section was about 200 pages.

2017 was the last year that Biochem 660 was offered. I have therefore decided to release the complete PyMOL tutorial which you will find split in multiple PDF files. In this final revision, I had updated all web pages, and added links to archived pages when web site were defunct to keep the text as relevant as possible for future use.

I hope that it will be useful to you in accomplishing your molecular graphics goals.

Sincerely,

Jean-Yves Sgro, Ph.D.
Distinguished Scientist | Senior Scientist
Biotechnology Center | Biochemistry Department
University of Wisconsin-Madison - USA
Email: jsgro@wisc.edu
Formerly at the Institute for Molecular Virology and VirusWorld web site creator.

JYS

Table of Contents

PyMOL (3) – Electron Density Map – (Optional) 3

1. Download an electron density map 4

1.1 Go to Electron Density Server page 4

1.2 Download electron density map in CCP4 format 7

2. Display electron density map 8

2.1 Open PyMOL and Map 8

2.2 Change electron density levels 10

3. Add PDB and reduce displayed area 10

3.1 Display electron density for the complete protein 10

3.2 Illustrate electron density resolution via map levels 11

4. Display a local area only 11

Optional Tutorials 15

1 Rasmol 15

2 VMD 15

3 DeepView 15

4 QuickTime Pro 15

“*My Molecule*” – Student Molecular Graphcics Presentations 17

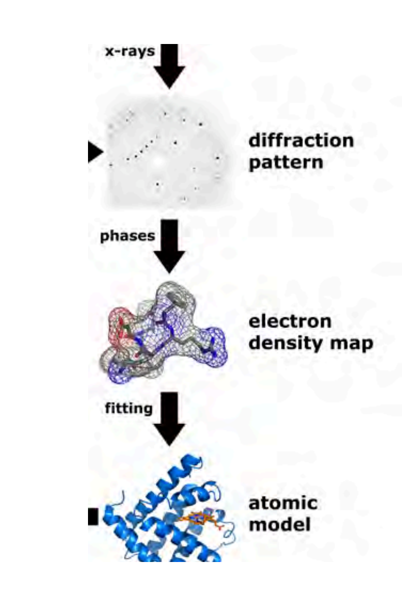
1. HOMEWORK EXERCISE (in-class presentation) 17

1.1 What you have to do: 17

1.2 Where can you work together with your partner? 20

PyMOL (3) – Electron Density Map – (Optional)

✔ **READ**: Electron density is the measure of the probability of an electron being present at a specific location. In protein crystallography, an electron density map averaging all the molecules within the crystal allows a crystallographer to build a model of the molecule.



Workflow for solving the structure of a molecule by X-ray crystallography. (source: wikipedia)

Calculation of the electron density map[1]: The data collected from a diffraction experiment is a reciprocal space representation of the crystal lattice. The position of each diffraction 'spot' is governed by the size and shape of the unit cell, and the inherent symmetry within the crystal.

The intensity of each diffraction 'spot' is recorded, and this intensity is proportional to the square of the structure factor amplitude. The structure factor is a complex number containing information relating to both the amplitude and phase of a wave and is a mathematical description of how a material scatters incident radiation.

In order to obtain an interpretable electron density map, both amplitude and phase must be known (an electron density map allows a crystallographer to build a starting model of the molecule). The phase cannot be directly recorded during a diffraction experiment: this is known as

[1] http://en.wikipedia.org/wiki/X-ray_crystallography

	the phase problem.

Initial phase estimates can be obtained in a variety of ways including molecular replacement and heavy atom derivatives.

The final atomic model is deposited after a series of refinements. Some authors choose to deposit the electron density map or a structure factor file that can be used to calculate an adequate electron density map.

For further information:

`http://pymolwiki.org/index.php/Display_CCP4_Maps`

The Protein Data Bank is the primary source of atomic models. However it now links to the Electron Density Server (EDS)[2] when a map is available.

"The Uppsala Electron Density Server (EDS; **http://eds.bmc.uu.se/**) is a web-based facility that provides access to electron-density maps and statistics concerning the fit of crystal structures and their maps."

1. Download an electron density map

We will use the now familiar 2BIW entry, biological unit 2 that was the basis of multiple exercises in previous sections. The file `2biw.pdb2` should be on your desktop or you can download it again from the web browser link below. Alternatively you can use the **`fetch`** command as before to load that single protein chain.

The electron density map can be obtained directly from the EDS server or via a link to the EDS server from the PDB page. In both cases the user will land on the page on the EDS server.

1.1 Go to Electron Density Server page

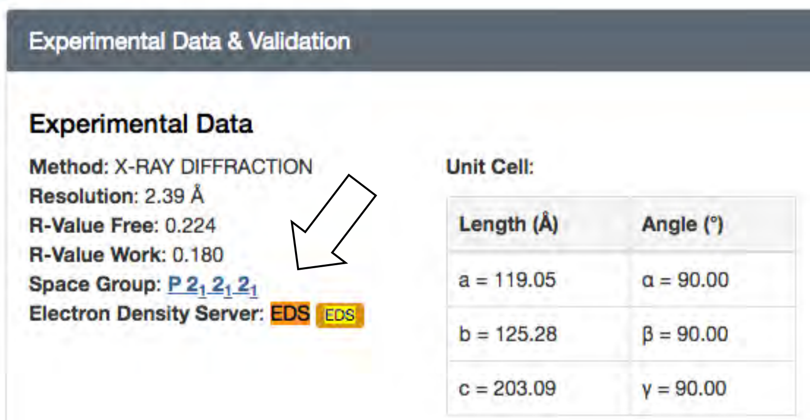
- **Open a web browser** such as Safari or Firefox.
- Point your web browser to **`www.rcsb.org`**

[2] The Uppsala Electron-Density Server. G. J. Kleywegt, M. R. Harris, J. Zou, T. C. Taylor, A. Wählby and T. A. Jones. *Acta Cryst*. (2004). D**60**, 2240-2249 doi:10.1107/S0907444904013253

TASK

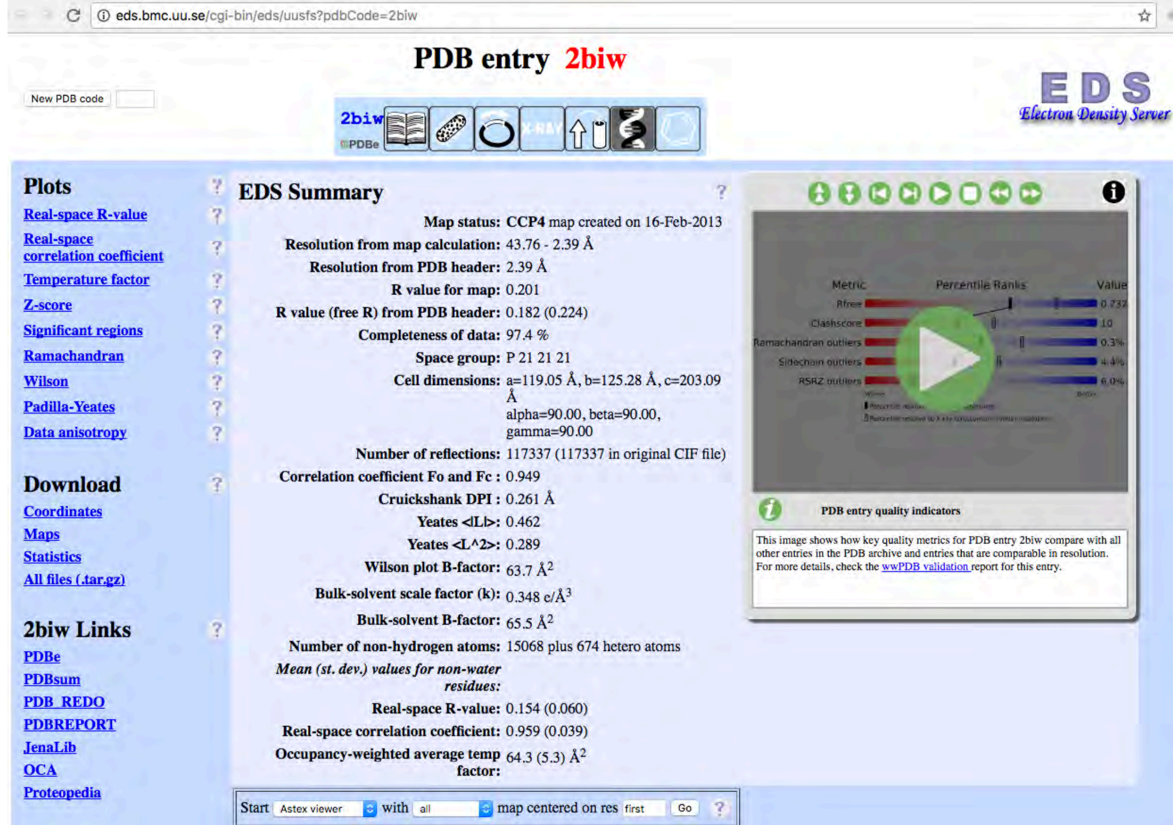
- In the Search box **type** the following ID: **2biw**
- Then **click Search** button

If the PDB entry has an EDS file, the relevant link will be provided within the "Experimental Details" panel on the bottom right of the entry page.



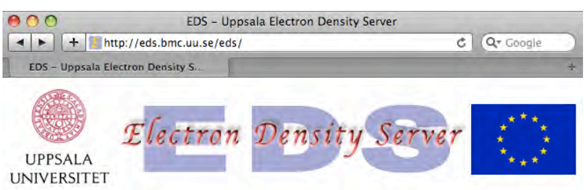
TASK

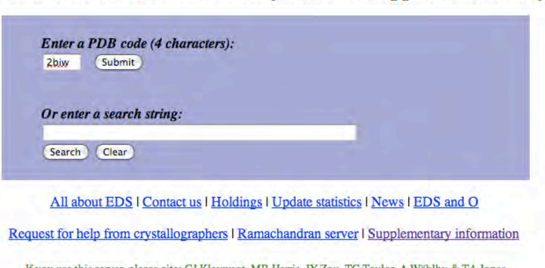
At the bottom right of the entry page **click on the EDS link** (shown by arrow.) This will open a new page displaying the information from that entry directly.



Alternate method: go to the EDS server (`http://eds.bmc.uu.se/eds/`) and enter **`2biw`** in the PDB code box and click submit.

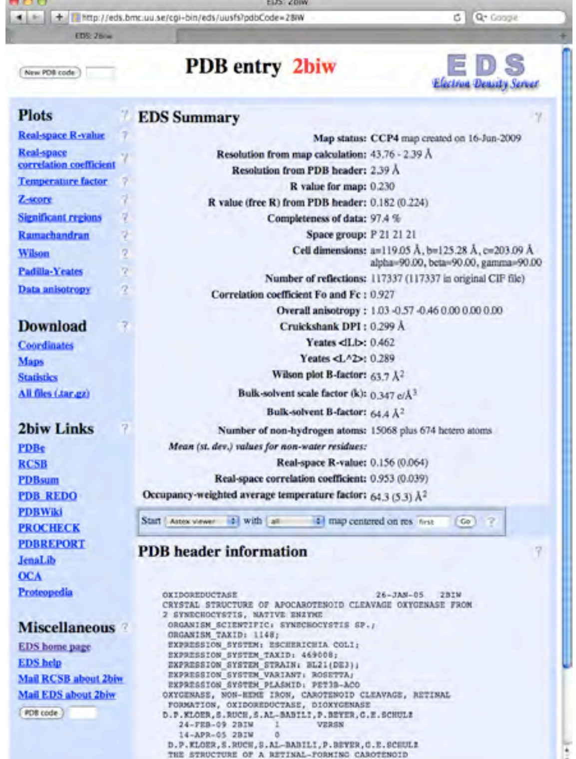
Regardless or the method, the end page will be the same.





The EDS summary page lists most parameters related to crystallography.

The bottom part of the page echoes the header portion of the actual PDB data file.

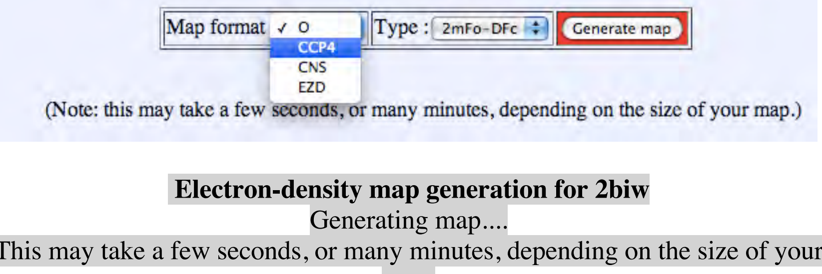


1.2 Download electron density map in CCP4 format

Just like everything else for computer files, there are various formats available. CCP4[3] is a format PyMOL understands.

TASK

- On the left panel under **Download, Click on Maps**



- **Select CCP4** as the map format
- **Keep Type** the same

- **Press "Generate map" button**

Electron-density map generation for 2biw
Generating map....
This may take a few seconds, or many minutes, depending on the size of your map.

Here is your gzipped map:2biw.ccp4.gz

- **Download** the proposed file
- **Rename the uncompressed** file: 2biw.map.ccp4

Alternate method: use the **fetch** command to download an electron density map called 2fofc (for **2**-sigma **o**bserved / **c**alculated in the crystallographers' jargon) but note that the download of the map does require some time before it is updated within PyMOL.
See the following example

```
fetch 2biw, type=pdb2
fetch 2biw, type=2fofc
isomesh 2biw_map, 2biw_2fofc, 1.06, poly, carve=2.5
```

[3] The Collaborative Computational Project Number 4 in protein crystallography or (CCP4) was set up in 1979 in the United Kingdom to support collaboration between researchers working in software development & assemble a comprehensive collection of software for structural biology. 1979. http://en.wikipedia.org/wiki/Collaborative_Computational_Project_Number_4

Alternate method 2: The new menu **File > GetPDB** can now be used to automatically download both the PDB and the map: simply click the **2FoFc Map** button.

Reminder: we are here calling for the 2nd biological monomer (Assembly) here set to 2 and it is also necessary to specify the chain ID (here B which matches the 2nd monomer.)

Get PDB File

Note: Downloading will save the files in the directory defined by the "fetch_path" setting.

PDB ID: 2BIW

☑ PDB Structure — Object name (optional)

☑ 2FoFc Map — Object name (optional)

☐ FoFc Map — Object name (optional)

PDB Structure Options

Chain name (optiona: B

Assembly (optional): 2

This will run the following command

```
set assembly, "2"
fetch 2BIWB
fetch 2BIW, type=2fofc
```

Download

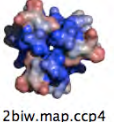
2. Display electron density map

(Assumes you completed previous exercise and downloaded map)

Note: the newest version of PyMOL will embed the map from the above command. Simply follow the "mesh" command below to visualize. However, for the command to work change the word **`2biw.map`** to the name given automatically: **`2biw_2fofc`**

2.1 Open PyMOL and Map

If the file has a PyMOL icon, simply **double-click the icon** to open a new PyMOL session. Or simply launch PyMOL and use the **File > Open** menu.



The top section of PyMOL will echo information about the opened file:

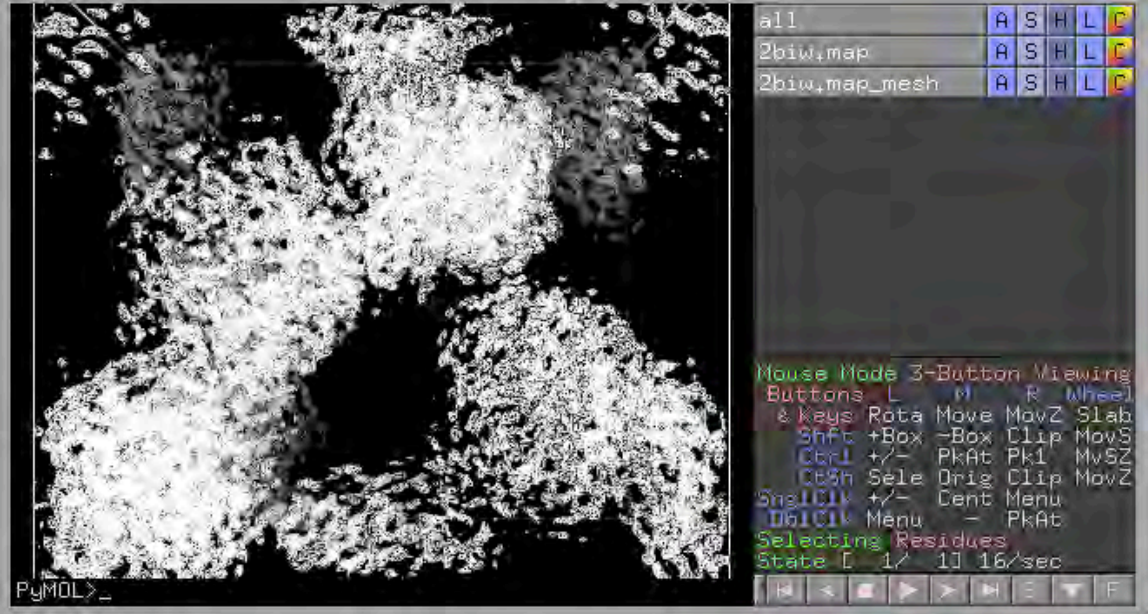
```
ObjectMapCCP4: Map Size 160 x 145 x 155
ObjectMapCCP4: Normalizing with mean = -0.008028 and stdev = 0.211077.
ObjectMap: Map read.  Range: -2.930 to 8.486
Crystal: Unit Cell          119.050  125.275  203.087
Crystal: Alpha Beta Gamma    90.000   90.000   90.000
Crystal: RealToFrac Matrix
Crystal:     0.0084   -0.0000   -0.0000
Crystal:     0.0000    0.0080   -0.0000
Crystal:     0.0000    0.0000    0.0049
Crystal: FracToReal Matrix
Crystal:   119.0500    0.0000    0.0000
Crystal:     0.0000  125.2750    0.0000
Crystal:     0.0000    0.0000  203.0875
Crystal: Unit Cell Volume  3028845.
ExecutiveLoad: "/Users/dmc/Desktop/2biw.map.ccp4" loaded as "2biw.map", through state 1.
```

Do not panic if your screen still appears black at this point!
This is normal.

In the Names panel **click** on **2biw.map** to let the volume of the map appear as an empty cube.

Then:
Click on Action (**A**) and select **mesh @ level 3.0** to show the corresponding electron density.

This will create a new entry in the Names Panel called 2biw.map.mesh



The electron density map for the complete section of space is presented at the level 3.0. Of course that is too much information and we will reduce the section of the map that is shown and play with the density level as well.

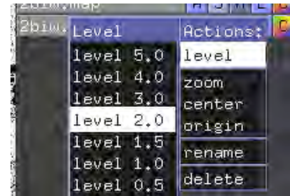
Note: we selected the “mesh” option within the Action menu. There is also a “surface” option that could be useful in some cases, perhaps with added transparency.

2.2 Change electron density levels

Typically, contour level values are reported in the absolute value of electrons/Å^3.

Click Action button for name 2biw.map.mesh and select **level 2.0**

This will change the displayed density in the viewer.



You can check the effect for other values, and then come back to a value of 2.0

3. Add PDB and reduce displayed area

The electron density map should be displayed at level 2.0 from the previous section.

Load the PDB data within PyMOL with the line commands:

```
fetch 2biw, type=pdb2
```

When the PDB file is loaded, PyMOL automatically zooms on the location of the atom coordinates. The familiar bonds between atoms are masked by the electron density map. The many other sections of the electron density map representing crystallographic symmetry can be later removed for clarity.

3.1 Display electron density for the complete protein

The following commands select which region of the map to display (variable name `site` chosen for the complete protein) and the next command creates a new map around the selected site:

✓ TASK

```
select site, 2biw
isomesh map, 2biw.map, 2.0, site, carve=1.6
```

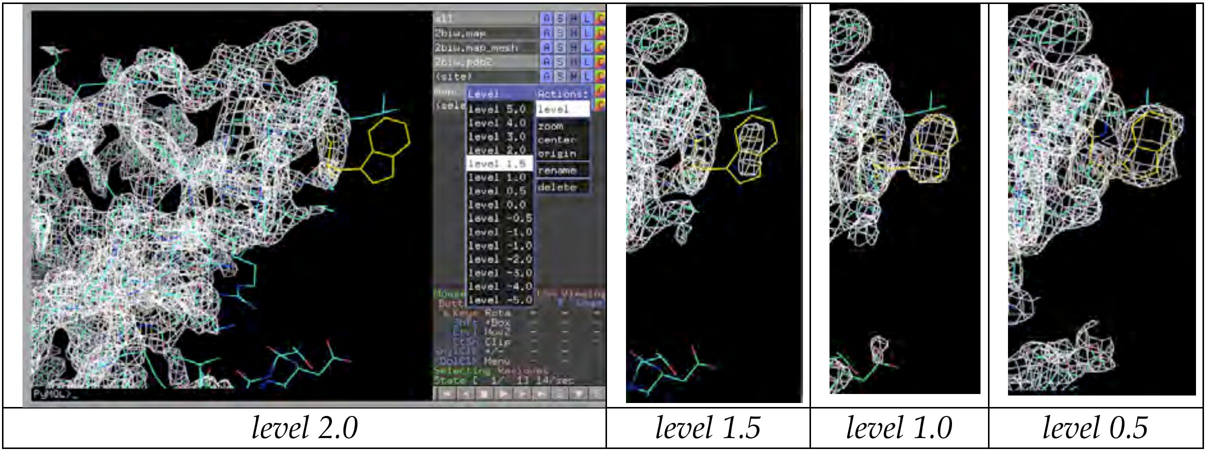
This will create a local map under the name "map" within the Names Panel.

In the Names Panel **click on `2biw.map.mesh`** to hide the general mesh map. This will reveal the local map around the single protein.

3.2 Illustrate electron density resolution via map levels

Rotate the structure to better see **Tryptophane 121**
To help you can use line command: **`orient resi 121`**
that can also be written as: **`orient i. 121`**

Change map level to 1.5, 1.0 and then to 0.5 (**map > A > level > 1.0**) to illustrate the increase of electron density map drawn at this location. It is due to the lower resolution of the map at this site: most W121 atoms would have a higher temperature factor within the PDB file.



level 2.0	*level 1.5*	*level 1.0*	*level 0.5*

Note that ASP 331 at the bottom of the page follows the same pattern of electron density map plots as the levels vary.

4. Display a local area only

Starting point: previous exercise showing map at level 0.5.

We will select a local segment which is in alpha-helix configuration and display the electron density of this local area.

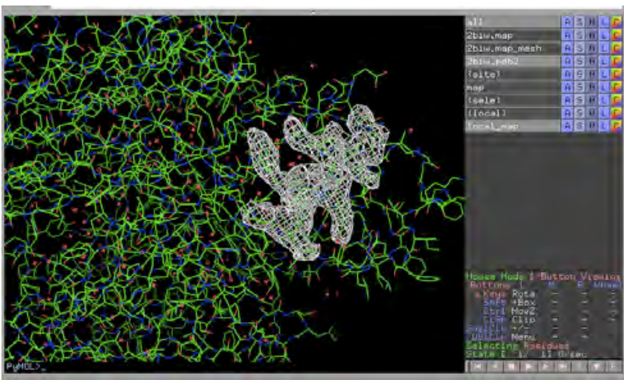
TASK

```
select local, resi 95-105
isomesh local_map, 2biw.map, 1.0, local, carve=1.6
```

Note: the words "local" and "local_map" are created by the user and are not PyMOL commands reserved words.

On the Names Panel **Click on "map"** to remove it from view.

Only the wireframe of the protein should remain and the local electron density plot we just created.



TASK - With the mouse do the following:

- Hide the protein wireframe: **2biw.pdb2 > H > everything**
- Show the local helix in wireframe: **local > S > sticks**
- **Zoom** on the local area with the mouse.

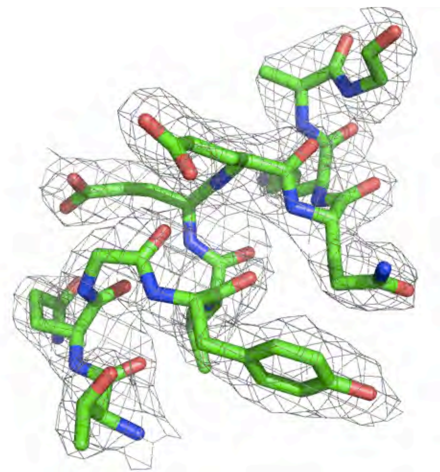
Using line commands:

```
center local    # change center of rotation to local area
color gray50, local_map  # set color to 50% gray
bg_color white           # change background to white
set mesh_width, 0.5     # makes meshes thinner, for raytrace
```

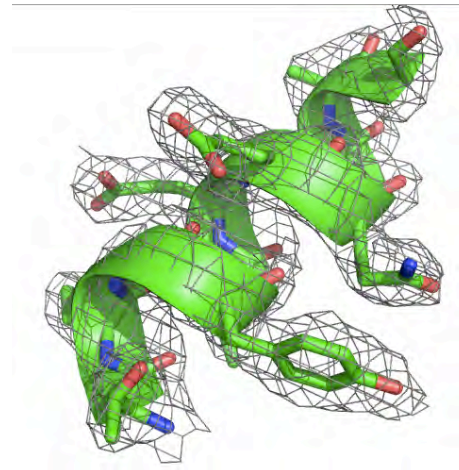
The following are optional settings: remove # to activate.

```
# set ray_trace_fog, 0 # turn off raytrace fog;
# set depth_cue, 0     # turn off depth cueing
# set ray_shadows, off # turn off ray-tracing shadows
```

Press the "**ray**" button or **type ray** within the line-command are to finish.

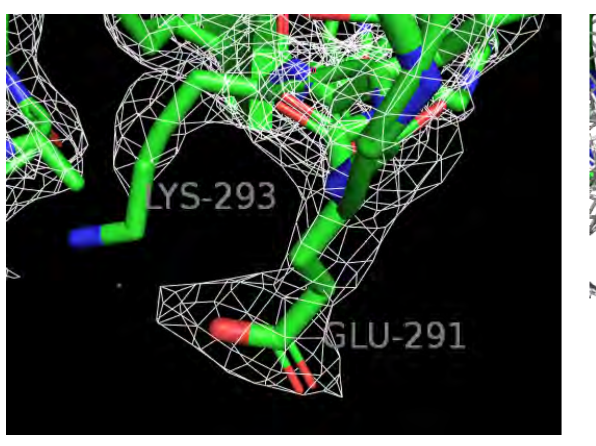


Ray-traced image from above settings

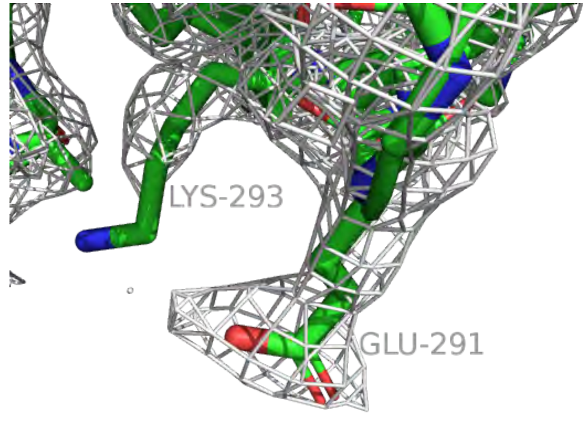


with additional cartoon & thicker mesh width (1.0)

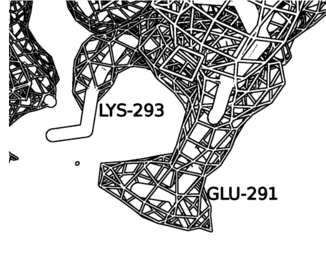
Another example for 2 amino acids clise up:



On the display

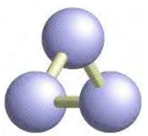
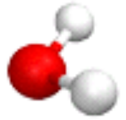


Ray traced with default parameters



Optional Tutorials

These tutorials are no longer printed in the standard book but are available as **PDF downloads** taken from previous book versions but are no longer updated.

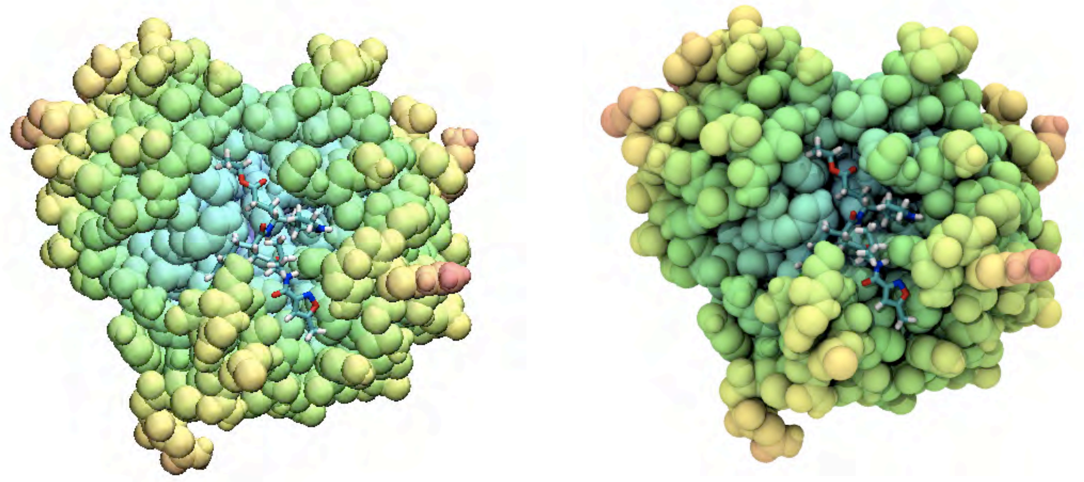
Tutorial	Icon	Description
1 Rasmol		Various versions of Rasmol exist. The tutorial is for the X11 version. All commands runs the same in all versions, only the launching is different.
2 VMD		VMD requires UNIX end-of-line. Useful option to create lipid bilayer and ambient occlusion depth-cueing options. (see next page!)
3 DeepView (Formerly Swiss PDB Viewer)		Specially useful for side-chain mutations and for 3D structure superimpositions. But now this is done with PyMOL.
4 QuickTime Pro		QuickTime is useful to assemble images into movies or re-export existing movies. Not necessary when using MacPyMOL.

Where to download: `http://virology.wisc.edu/acp`

Click on "**Class Tutorials**"

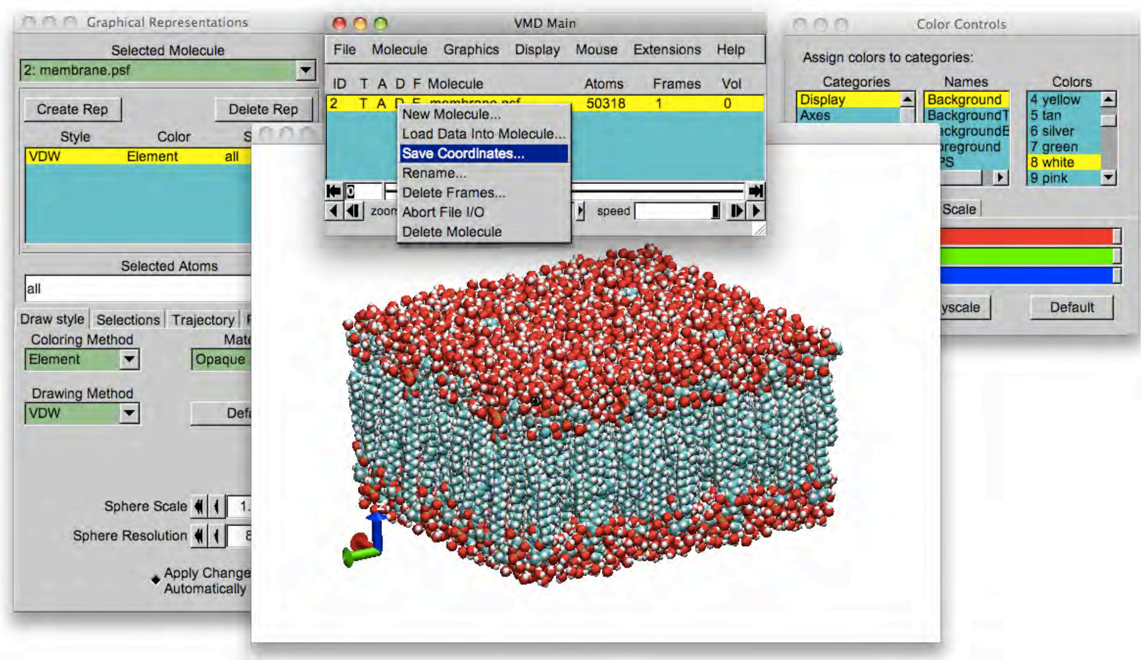
Then under "Our Tutorials" use the pull-down menu to select the PDF you want to download. It will open automatically in your browser. To save simply use the "**Save As**" command from the "**File**" menu in the browser.

VMD: Ambient Occlusion



VMD OpenGL Dispaly *TGA rendered file with Ambient Occlusion*

VMD: Membrane Builder (VMD Main: Extensions > Modelling > Membrane Builder)



"*My Molecule*" – Student Molecular Graphcics Presentations

1. HOMEWORK EXERCISE (in-class presentation)

My Molecule Homework Project: to be presented by teams in laboratory.

Now that you (will shortly) know all there is to know about structure display from your exercises with PyMOL & other optional tutorials, and you also know a bit about presentations and movies in PowerPoint, it's time to assimilate these skills and use them to illustrate some real research points.

1.1 What you have to do:

1. **Pick a project partner**. The class will be divided into groups of 2 or 3 students each. Partner selections will be finalized during the second Graphics lab (L02) segment.

2. You and your partner(s) should go online to the pdb database (http://www.rcsb.org/pdb/) and pick one (or more) crystal structure files (or an NMR structure set) that is of interest to you. It could perhaps be an enzyme related to your own research, a protease, an RNA (*e.g.* tRNA), a protein /nucleic acid complex, a set of amino acids, bases or whatever... just so long as there is a pdb file encoding it (or them).

3. Read a few of the research papers associated with this (these) structures. Remember, pdb files have references listed at the top of each file within the header just so you can learn the salient features of the molecule and why people consider it biologically interesting. You DON'T need to go into great depth here, the idea is to gather some interesting information about mechanisms, active sites, substrates etc., so you will have some concepts to illustrate.

4. Using PyMOL, VMD, Rasmol, DeepView or any other molecular display programs available to you (or even other stuff you might find on the web), create some nice images of your molecule(s) that illustrate key features of its molecular biology. For example, if you were working with HIV protease, you might be interested in highlighting the active site residues or the substrate binding pocket, among other features. On a tRNA, the anticodon loop or the bases that make contact with the synthetase might be of interest. In an NMR structure, the differences among the various structures might be of interest (where do they wiggle?) etc. Get creative with several different views, using surfaces or electrostatic potential, or whatever, to highlight and illustrate any interesting points. You might wish to compare or contrast elements from related structures. (Where is the HIV protease different from the SIV protease, for example, and why might these differences be related to their activities?) Or perhaps you may wish to show how a substrate/inhibitor binds in an active site. **Use your imagination! Use humor! No holds barred!**

5. Capture your images as graphics files (PNG, GIF etc.), import them into PowerPoint, and create a short presentation (8-10 slides) **telling a story about your molecule**. Add text, builds, short movies (if you make them), references or any of the other features of PowerPoint to your file that you think might be helpful and attractive.

6. During the laboratory on the week of these presentation (L06), you and your partner(s) should bring to lab, a (PC or Mac formatted) CD disc or a USB thumb/flash drive containing your "My Molecule" presentation**. Be sure to include on this disc: any QuickTime (MOV, MPEG or AVI) movies that are referenced by your file or they won't play when you think they would. Each partner in the group should plan to participate in a **10 minutes (total/group) oral presentation** with another **5 min for questions** about graphic methods or topic content.

Instructors will provide a Macintosh computer (Windows MIGHT be an option too) with CD drive and USB port, hooked up to an XGA-level projector. These presentations will be graded (for all students in the group by all students in the class), on the basis of originality, clarity, and creative use of program tools to illustrate some salient aspect of protein/nucleic acid biochemistry. The idea here is to demonstrate that you can use the molecular graphics tools available to you, to develop illustrations or animations that are clear and informative to an audience.

7. **BOTH/ALL partners in each group must participate equally** in the planning, execution and presentation phases of this project. Each will be assigned the same grade for the exercise according to feedback sheets filled out by your classmates and instructors (your audience) after each presentation.

8. We are not so much concerned with presentation content (which molecule you choose, or what you learned about it), but rather, how you choose to illustrate your points with the creative and clear use of graphics techniques. Have fun with this learning opportunity and your audience will have fun too!

** The instructor's Macintosh in the DMC is equipped with a CD/DVD reader and has a USB port for flash drives. It also runs QuickTime and the latest version of PowerPoint. HOWEVER, to ensure that YOUR file made on YOUR computer (or your lab's computer) runs smoothly on this machine, it is a good idea to adhere to these simple formatting caveats:

a. When creating your presentation in PPT, make sure in EDIT > PAGE SETUP you have configured your file for *OnScreenShow* or it may have funny dimensions when you go to display it.

b. It is a good idea to keep the size of your PPT file under 30 Mb (you may need to trim your picture file resolution if they are too big). Be forewarned that files bigger than this, or even movies bigger than 7-8 Mb may have trouble playing on another machine. Be careful, because you don't want to spend your 10 min presentation, fixing your file. Test it ahead of time *e.g.* the week before.

c. If you want to test a practice file or figures for preview, please be sure to bring a test disc to the lab before MyMolecule day. You will not have access to the DMC lab outside of regular class hours.

d. If you use a PC to create your presentation, do NOT use keyboard shortcuts to insert symbols, Greek letters or bullets. Only fonts selected from the standard PPT font list, and which are also common to the DMC Macintosh will display without error.

For safety, stick to simple fonts like: Arial, Times, Σψμβολ (Symbol,) Courier, Helvetica, etc.

1.2 Where can you work together with your partner?

You can make use of the 3 iMacs terminals in Biochem 301 Biochemistry Laboratories Addition (see page 5 of the book for details.) They are programmed with all course-supported software that function in Macintosh OSX: DeepView, RasMol, VMD, PyMOL, Chimera, QTpro etc. are all installed. Non-Biochem student will be given a personal username:password that will allow them to logon to these terminals and save their work on Biochem server space. Biochem students use their existing account. **Biochem 660 students should already have a username.**

Or, you can work at home from your own personal computers. To store your files between sessions, or while you are working with your partner you can use the **Enterprise Box Service** offered by UW-Madison to all students, faculty and staff **`uwmadison.box.com`**

– END –

Class notes