Structural Biology
Prof. Saugata Hazara
Department of Biotechnology
Indian Institute of Technology - Roorkee

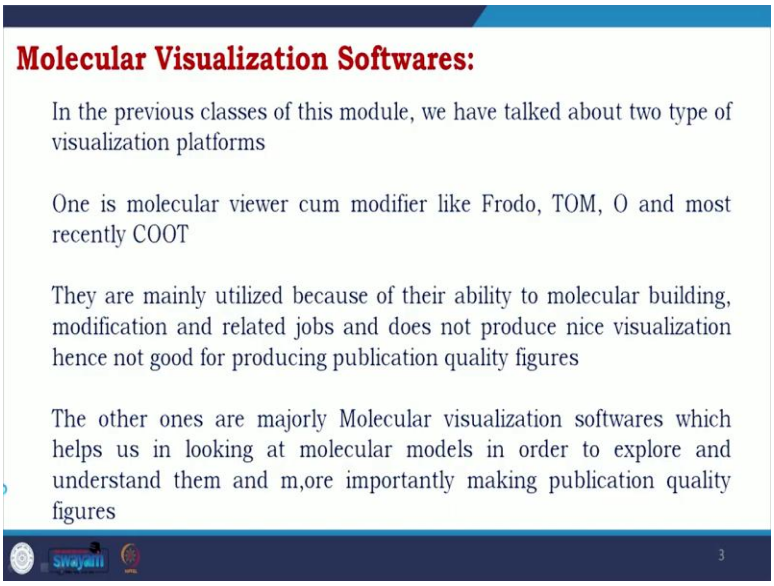
Lecture – 44
3D Visualization Using Pymol

Hi everyone, welcome again to the course of structural biology. We are continuing with the module visualization. I have talked about the history of visualization from real-life modeling with metal rods, wood with glass, and all those sculpturings. We have traveled to computer modeling, and computer models have significantly improved. We have talked about all of them. There we talked about the files which are needed for visualization.

We have talked about dot PDB which is the most accepted file in this world, and not only that, PDB is also a name of a database protein data bank wwwpdb is now the only database that has to accommodate all possible 3D structures which are solved by any experimental methods of the biological macromolecule. Then we talked about the software called Coot, what it does how the software's developed.

Today we would start with the software mainly focused on visualization. Coot is also a visualization software, but people do more functional work like modification. So, let us come to 3D visualization with the visualization platforms, and I would focus on a software platform called Pymol.

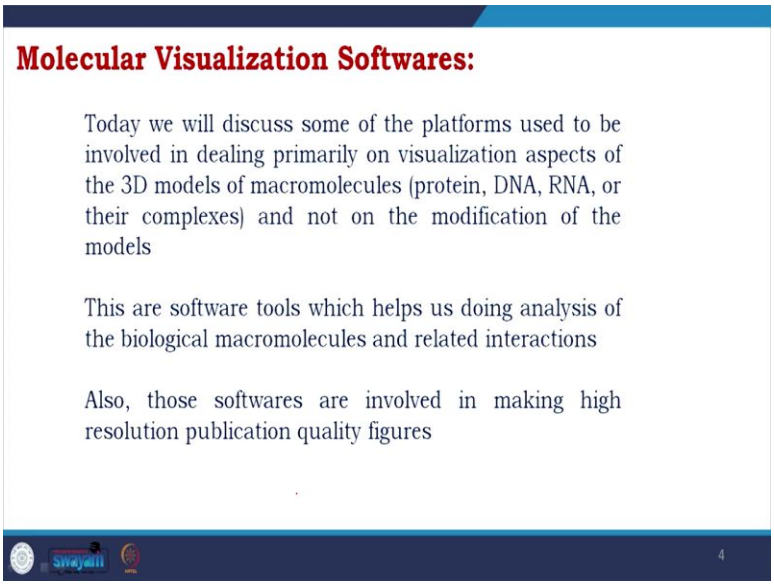
(Refer Slide Time: 02:30)



So, if we talk about molecular visualization software, as I told you in the previous classes of this module, we have talked continuously about two types of visualization platforms; one is molecular viewer comes with a modifier like Frodo, Tom, O, and most recently Coot which I talked in details which I demonstrate but there are these software's are mainly utilized because of their ability to molecular building, modification and related jobs and does not produce excellent visualization hence not good for creating publication quality figures.

So, yes, they are visualizing platforms, but they are much more. They are majorly working on a model building; like in crystallography, you have electron density, you come to Coot, and you develop the model from the electron density. So the other one is molecular visualization software, which helps us look at molecular models to understand them for analysis and, more importantly, make publication-quality figures.

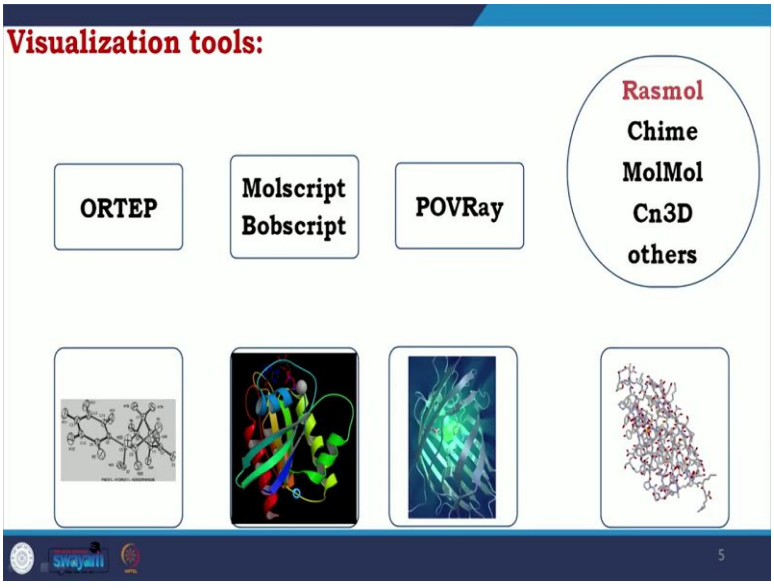
(Refer Slide Time: 03:53)



Today we will discuss some of the platforms used to deal primarily with visualization aspects of the 3D model of macromolecules protein, DNA, RNA, their complexes protein-small molecule, DNA small molecule, protein-protein, protein-DNA and all sorts of them and not on the modification of them. These software tools help us analyze biological macromolecules and related interactions. Also, the software's involved in making high-resolution publication-quality figures.

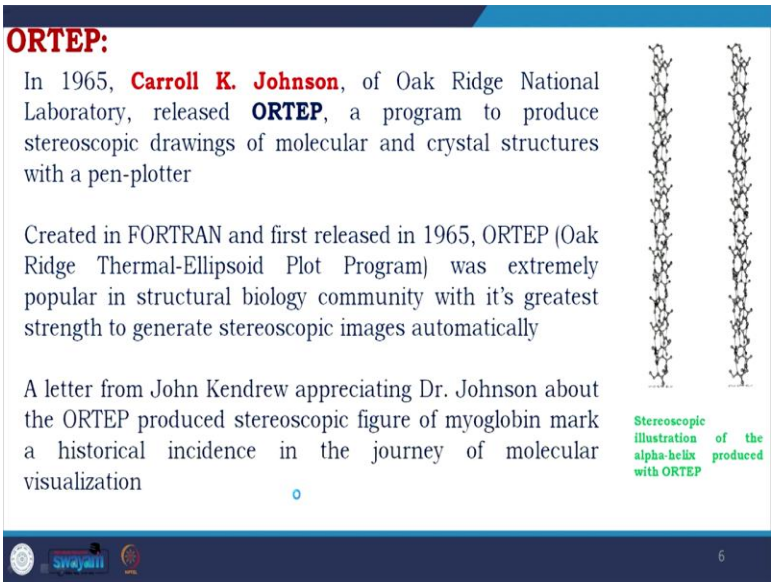
So, you do your modification, do your analysis, and then come to these platforms and produce high-resolution figures that would be accepted for publication.

(Refer Slide Time: 04:39)



And when we talk about them, we have a flow starting from ORTEP. ORTEP is the first kind of computational platform, and then it comes to script writing sMolScript, and Pov script. There are, you know, Povray, and Raster 3d, which help improve the quality of the figure, and last but not least, the visualizing platforms like Rasmol, Chime, MolMol, Cn3D, and many others.

(Refer Slide Time: 05:24)



We will try to have a look in brief about those. ORTEP we have already talked about; in 1965, Carroll K. Johnson of Oak Ridge National Laboratory released ORTEP, a program to produce stereoscopic drawings of molecular and crystal structures with a pen plotter. So, they use a pen plot and correlate with the computer, and according to the change in the structure, the pen here will plot, and you will see this is a stereoscopic illustration of the alpha helix produced with ORTEP.

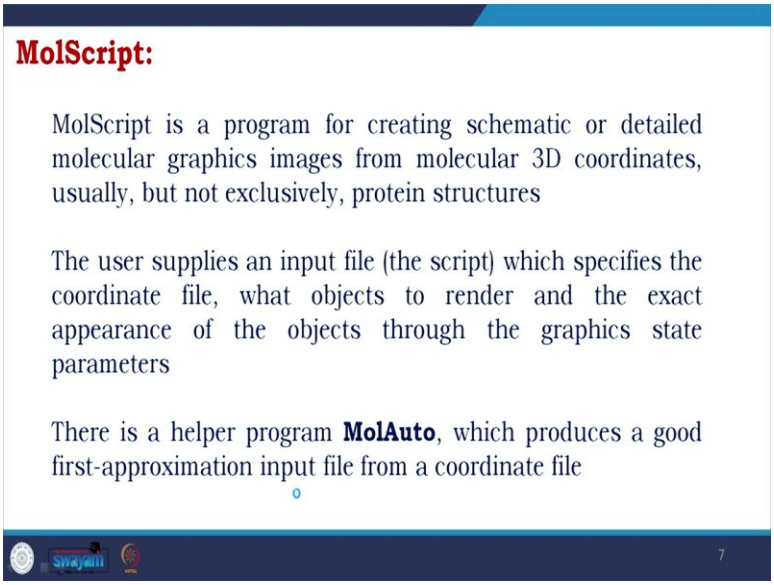
It was created in with Fortran language and first released in 1965; as I told ORTEP is named because it is Oak Ridge Thermal Ellipsoid Plot Program; that is why ORTEP was extremely popular in the structural biology community of that time because this was the only software which has helped them produce publication quality figures. And it had greater strength to generate stereoscopic images automatically.

How popular is one of the historical evidence comes from the fact a letter which I also talked about from John Kendrew appreciating Dr. Johnson, Dr. Carroll K Johnson was the inventor of

ORTEP, ORTEP produced the stereoscopic figure of myoglobin, which can make Kendrew very happy and this is a historic incident in the journey of molecular visualization. This is also proof of how ORTEP was popular at that time.

So, John Kendrew, the novel laureate and the first one to solve the structure of the protein myoglobin, has appreciated the work of Carroll K Johnson.

(Refer Slide Time: 07:30)

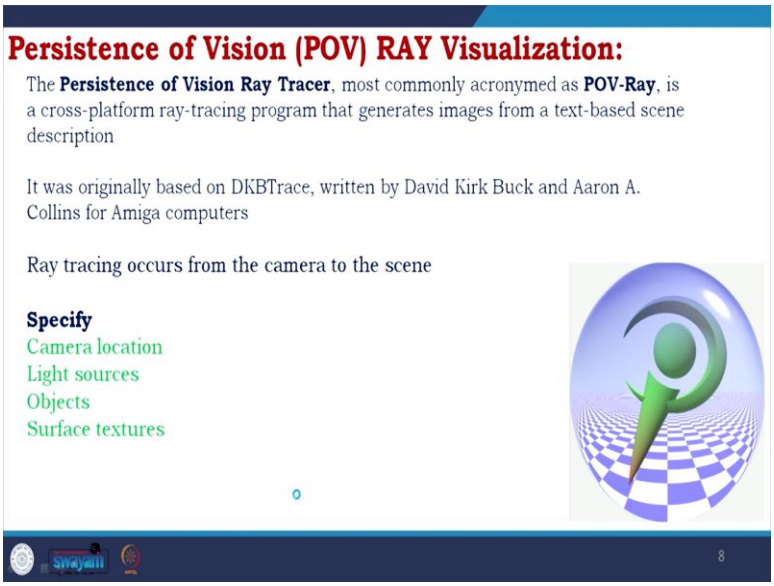


Now, coming to scripting MolScript: MolScript; is a program for creating schematic or detailed molecular graphics images from molecular 3D coordinates, usually but not exclusively for protein structure. It can be used to make an image of many other things, like in different fields, but MolScript has taken a critical role in bringing high-resolution images. We will see them. The user supplies an input file.

When I started my carrier as a Ph.D. student in the initial years, the Pymol had already appeared, but we used to make figures using MolScript. It is not that we had to make the entire script; there are a lot of scripts there. We take that script and change them according to the requirement of our structure. So, the user supplies an input file, the script which specifies the coordinate file what objects to render and the exact appearance of the objects through the graphic state parameters.

If you still go back and prepare figures with MolScript, you will see that, generally, the figures are very high resolution, which is my personal experience. There is a helper program MolAuto which produces an excellent first approximation input file from the coordinate file, and then you go through and make changes if you think so.

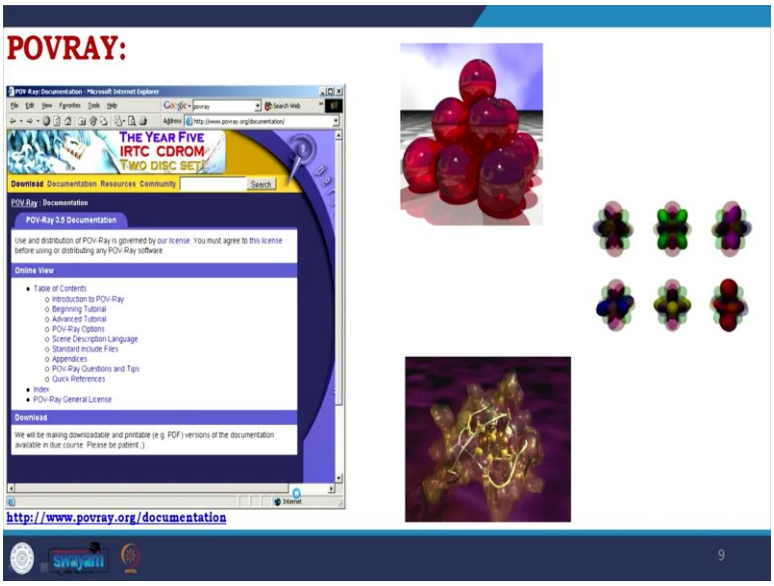
(Refer Slide Time: 09:21)



Coming to the persistence of vision, POV, or Pub Ray Visualization: You could see the image to understand proper visualization. The persistence of fission rate pressure, commonly acronym as POV-Ray, is a cross-platform ray tracing program that generates images from a text-based scene description. So, it does not apply exclusively to biological macromolecules but to any image formation where retracing makes it high-resolution.

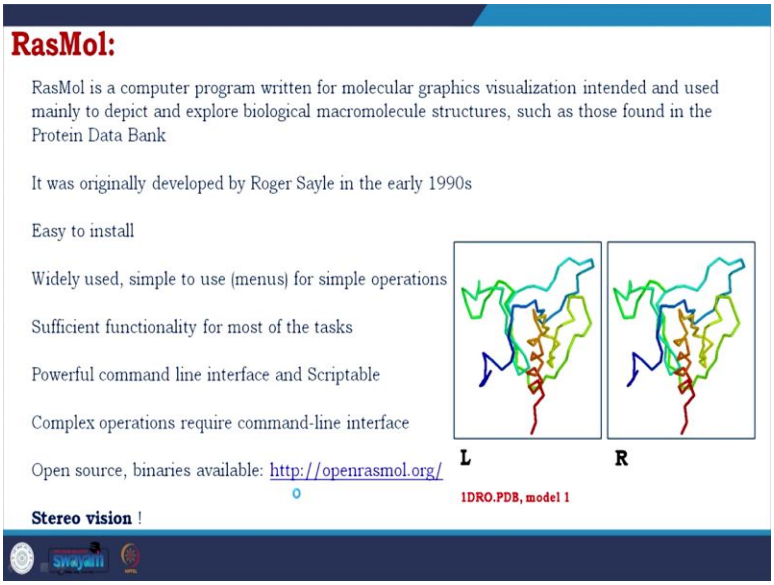
It was originally based on DKBtrace, written by David Kirkbach and Aaron A Collins of Amiga Computers. Ray tracing occurs from the camera to the scene, and the user must specify the camera's location, the light's source, and the object on which you need the retracing and image formation that surface structures. Atmospheric media like fog hedge or fire.

(Refer Slide Time: 10:38)



This is the page of POV-Ray you will see the table of contents documentation, which are high-quality figures made from POV-Ray. This is a figure of balls that shows like general transparent balls. These are figures of orbital, and these are figures for protein. If you are interested, the links are given in for all of them; you will get them here.

(Refer Slide Time: 11:17)



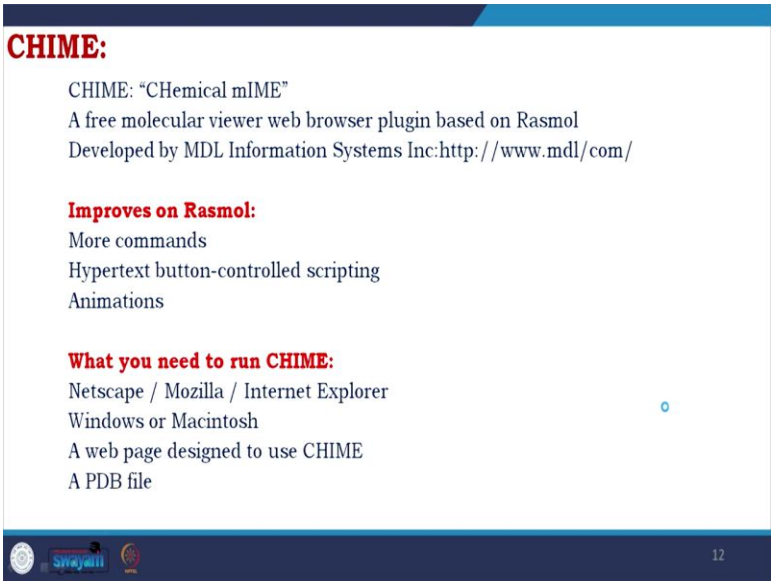
Coming to molecular visualization platforms, as I talked about Rosmol, I have already talked about Rasmol in the history part Rasmol is a part of history because the first time people started making beautiful figures as it comes to common people through the innovation of Rasmol Roger SL. So, Rasmol is a computer program written for molecular graphics visualization intended and

used mainly to depict and explore biological macromolecular structures such as those found in the protein data bank PDB.

Rogers Sayle originally developed it in the early 1990s. Rasmol got popular because it is easy to install it is widely used simple, and widely used; more than one million people at the then time used Rasmol for structure presentation and publication widely used, simple to utilize for simple operations, sufficient functionality for most tasks about representing the 3D molecule, powerful command-line interface.

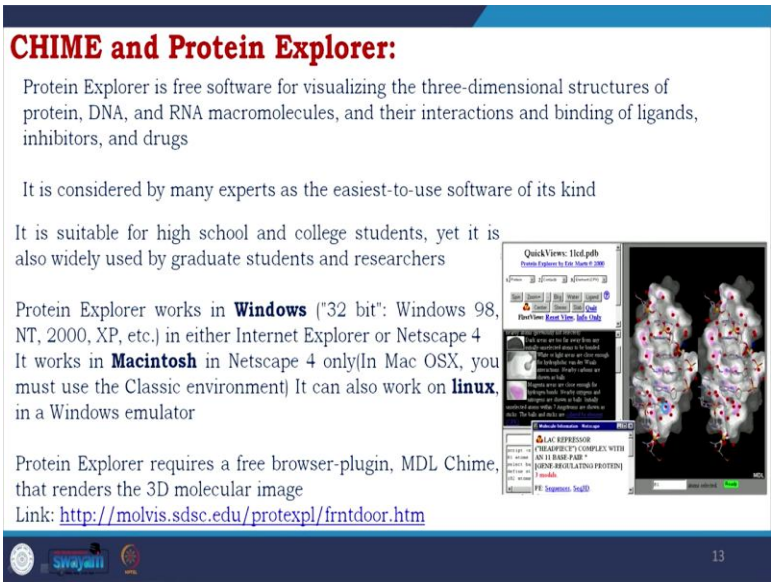
And it was a complex scriptable operation requiring a command line interface and open source freely available; you get the binaries link. Also, one of the strongest things is having a stereo vision. This is the presentation you see of the PDB R1DRO, and the presentations have L and R, so stereoscopic presentation.

(Refer Slide Time: 13:20)



Another software is CHIME. CHIME is named CHemical mIME, a free molecular viewer web browser plugin based on RasMol. It was developed by mdl information system; the link is provided here. It improves on RasMol; having more commons means more functionality, hypertext button control scripting, and animations. And for running CHIME, what you need is Netscape, Mozilla, Internet Explorer, windows, or Macintosh. A web page designed to use a chime and a PDB file.

(Refer Slide Time: 14:12)



Protein Explorer is free software that uses the background of CHIME. So, protein explorer is a free software for visualizing the three-dimensional structure of protein DNA and RNA macromolecules, and their interactions and binding of ligands inhibited on drugs. Many experts consider it the easiest-to-use software of its kind. It is suitable for high school and college students, yet it is also widely used by graduate students and researchers for their real research.

To work, it needs Windows with 32-bit Windows 98, NT, 2000, XP, etcetera, in either Internet Explorer or Netscape 4. It works in Macintosh, in Netscape 4 only. In the Mac operating system, you must use the classic environment; it can also work in Linux in a Windows emulator. Protein Explorer requires a free browser plugin mdl CHIME that renders the 3D molecular image again. The link is provided here.

This is the page of protein explorer where you see the Lac Repressor complex with an 11-base pair gene-regulating protein. So, you see the Lac Repressor and its DNA bound to it. So, this is the presentation platform of protein explorer.

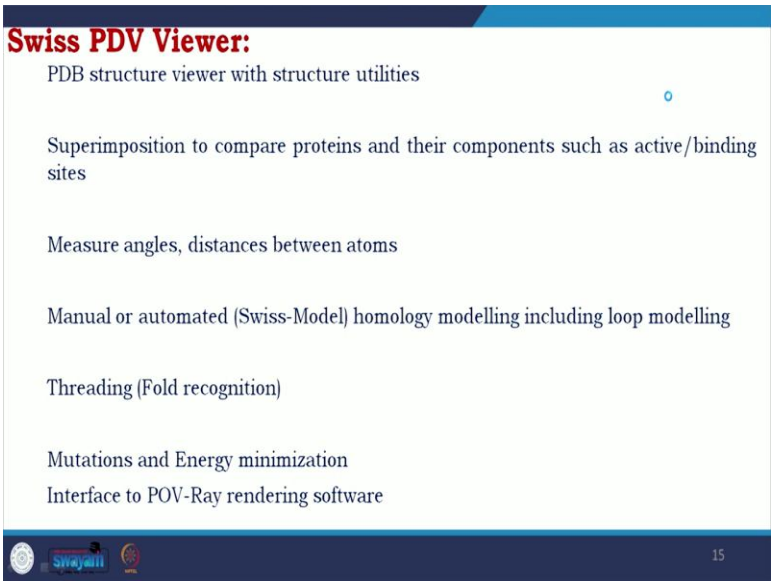
(Refer Slide Time: 15:51)

Coming to Cn3D is developed by NCBI again, open source; the binaries are available the link is provided here first OpenGL graphics annotation engine is provided can fetch a structure over the internet is the first of a kind where now you will see this feature in most of the softwares. If you have internet, you could directly fetch the 3D structure by providing the PDB id. I am sure that you guys now remember what the PDB id is. And what is a UNIPROT id?

The PDB id is unique for a structure the UniProt id is unique for a protein. Another feature it brings could display protein movies for the NMR ensembles. Remember, we do not get a single high-resolution structure in NMR as in X-ray; instead, we get tiny models in an ensemble. The ensemble is represented as a movie in Cn3D; protein folding trajectories are also expressed in movies.

To get the Cn3D structure files, you must use MMDB files to retrieve those MMDB files using MMDB identifier PDB identifier, concept domain database blast search, PubMed query, and text search. This represents Cn3D, the Hen Egg White lysozyme structure complex with tri N acetyl maltotriose. It is a small molecule that is complex with Egg White lysozyme.

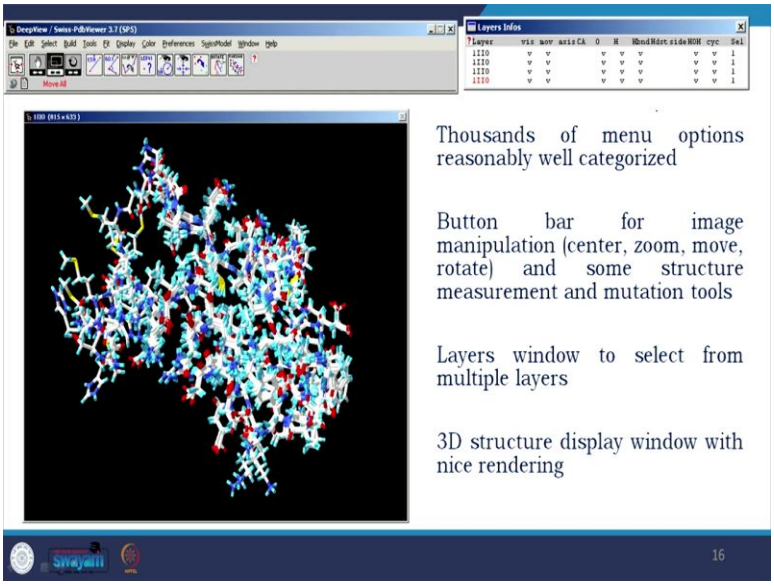
(Refer Slide Time: 17:51)



Swiss PDV viewer: The PDB structure viewer with structure utilities; though I told you most of them are molecular viewers, the Swiss PDV viewer also has a modification facility. Super positions to compare proteins and their components, such as active binding sites, are available. They could measure angle distance between atoms, manual or automated homology modeling, including loop modeling as a specialty threading or fold recognition.

So, it could build a model in different ways by homology by threading and performing mutation and energy minimization. As I talked about POV-Ray, which does the rendering, it also takes the interface of POV-Ray to improve the picture quality.

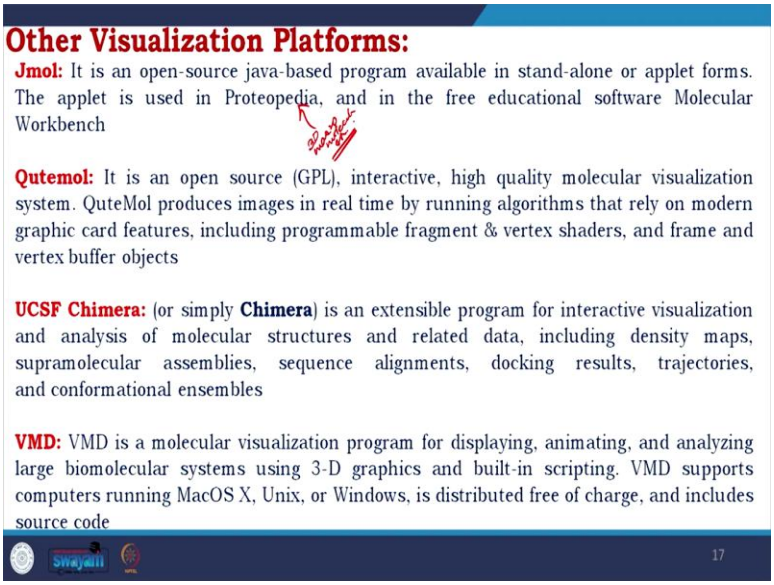
(Refer Slide Time: 18:56)



This is the platform where the protein DNA and other macromolecules are displayed. So, it contains thousands of menu options that are reasonably well categorized; you can see the window here; all the options are there; you click, and you get many options for each of the components here. File, edit, select, build tools, and feed any of them. These are also options where you could get distance; you could define the angles and all those you know facilities.

The button bars for image manipulation center, zoom, move, rotate, and some structure measurement and mutation tools, which you can see here. Also, there are layers windows to select from multiple layers. You could see a 3D structure display window with an excellent rendering. So, this is the layer information here; different layers have different macromolecules and structures. And this is as I talked about the main display where the 3d structure is displayed.

(Refer Slide Time: 20:15)



There are many other visualization platforms because of the time constant. I could not talk about all of them, but to mention about few of them again in very brief Jmol. Jmol is an open source java based program available in standalone or applet forms. The applet is used in Protopedia if you did not hear about Protopedia. This is a very interesting collection of 3D macromolecular structures in the free educational software called molecular workbench.

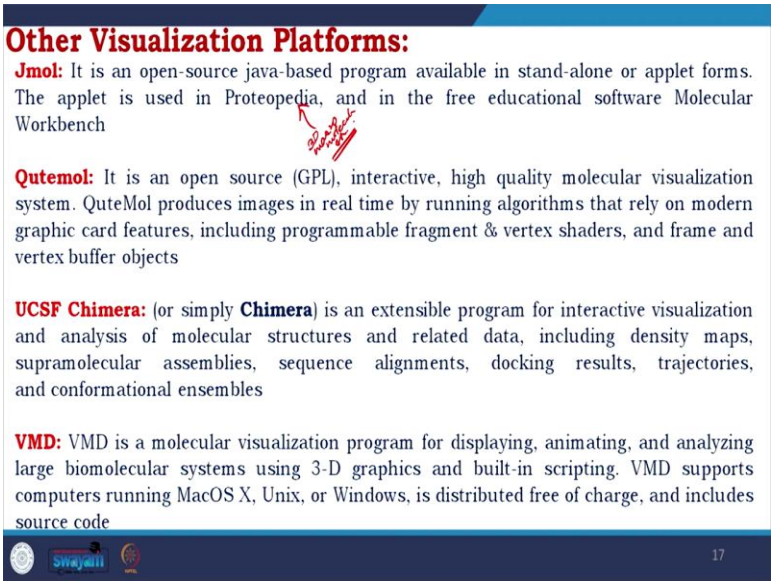
Qutemol is an open-source GPL interactive high-quality molecular visualization system. Qutemol produces real-time images by running algorithms that rely on modern graphic card

features, including programmable fragments, vertex shaders, and frame and vertex buffer objects. UCSF Chimera, or simply Chimera, is an extensible program for interactive visualization and analysis of molecular structures and related data, including density maps, supramolecular assemblies, sequence alignments, docking results trajectories, and conformational ensembles.

It is a very widely used program and a really good platform. VMD: VMD is a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3D graphics and built-in scripting. VMD; name for visual molecular dynamics by UIUC biophysics group University of Illinois Urbana. VMD supports computers running the Mac operating system, Unix, and Windows, is distributed free of charge, and includes source code.

There are many, but this briefly describes a few popular visualization software's. But as I discussed initially, my focus would be on Pymol.

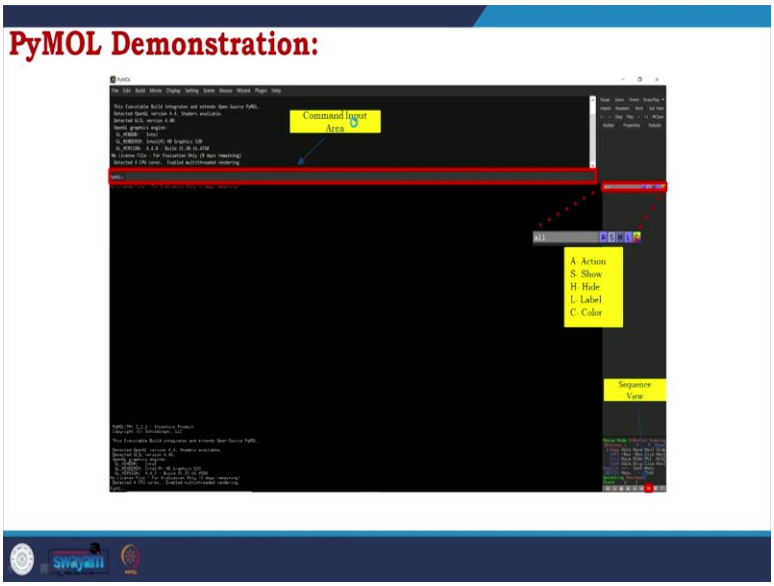
(Refer Slide Time: 22:47)



So, Pymol is a user-sponsored molecular visualization system on an open-source foundation maintained and distributed by Schrodinger. Pymol is a molecular visualization program that is written in the language Python. It is a molecular graphics program that can be downloaded and installed on Windows or Linux PCs and Apple computers running the Mac operating system and other Unix-based systems.

These are the links you can download for free. Standalone Pymol binaries for Windows provide the easiest installation; you have to download the Pymol installer file, install pymol by clicking on setup dot exe, and then Pymol can now be launched from the start menu of the computer.

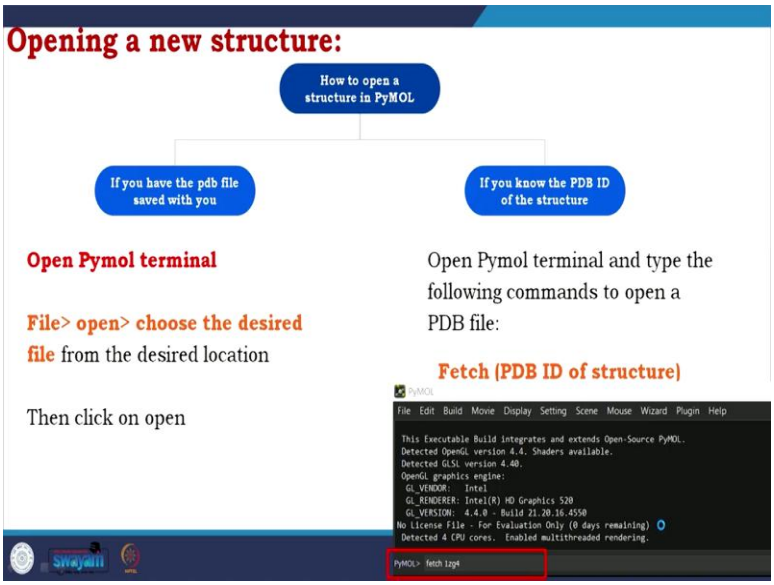
(Refer Slide Time: 23:46)



This is the first page of Pymol. I will show you through a demonstration; here, there are command input areas where you could provide commands, but you have a lot of menus. If you see here, File, Edit, Build, Movie, Display, Settings, Scene, Mouse Wizard, Plugin, help. But there is another window; these windows have other options. I will talk about all these, and there is another window connected with the display area again few menus action A S, so, H hide, L label, C color. So, you could do many things; here, you could also have different options; you could do sequence view, you could control the mouse, and all of this you could do using them.

So, if you remember, while I was talking about Coot, I talked about beta-lactamase as a protein and described it today. I will talk about a few proteins. I will show you how they could be displayed and how they could be studied.

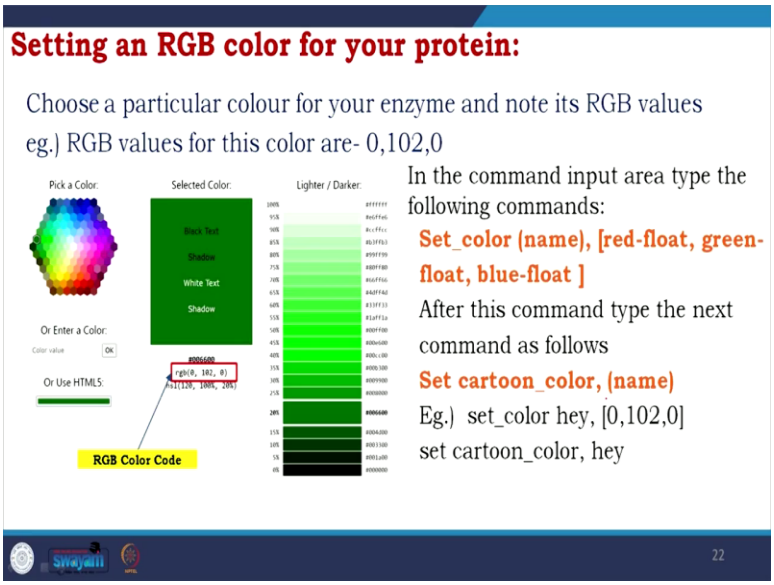
(Refer Slide Time: 25:16)



So, to go there first, we will go for opening a new structure. So, how to open a structure in Pymol if you have the PDB files saved with you, you have to go in one way; if you know the PDB id of the structure, then there is another option. So, if you have the PDB file saved with you, you go to the Pymol terminal, you open the Pymol terminal, you go to the file open, and choose the desired file from the desired location.

Then click on open, and you get the PDB structure displayed on your display window, which I showed you. If you know the PDB idea of the structure, open the Pymol terminal and type the following common to open the PDB file fetch; then, you put the PDB id of the structure you want to see. As I told you, in this common information, you could put fetch and then PDB id here; the PDB id given is 1ZG4.

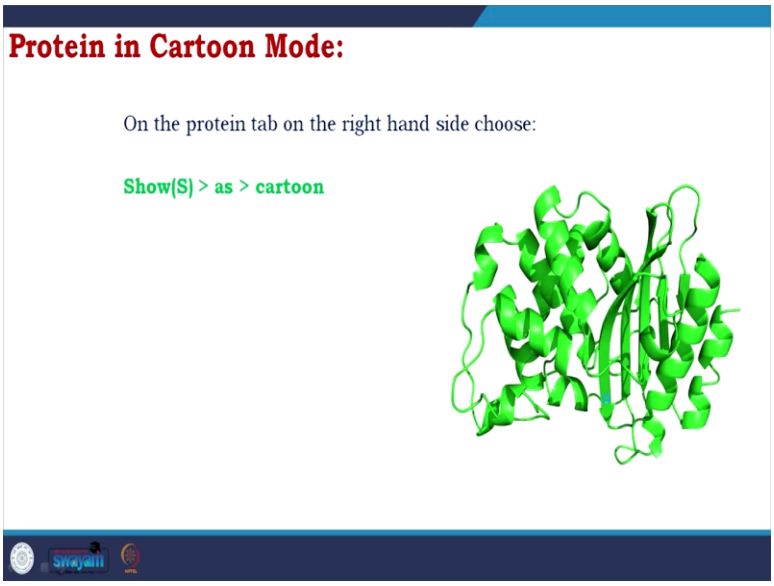
(Refer Slide Time: 28:38)



You could provide color to the protein, you could choose a particular color for your enzyme, and you could note the RGB values. RGB values are an excellent representation of the color R is red, G is green, B is blue red, and green-blue is RGB. So, providing the value of all three of them, the mixing gives you a new color. Here you could see you could pick any color you could, select the color, and you get the RGB value; this is the RGB color code, and for different of them, you see you could travel from lighter to darker.

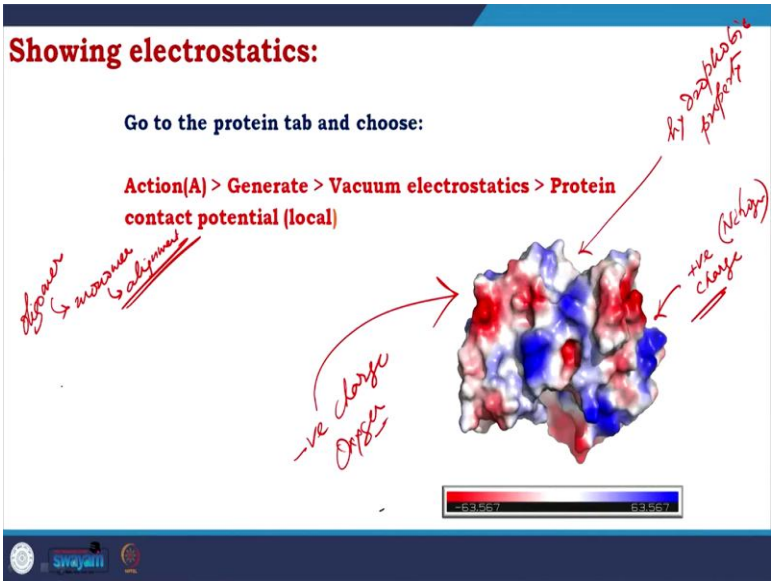
In the common input area, type the following command set underscore color name red float green float. So, float means you must put the value after this common type. The next common would be a set cartoon underscore color name. So, if you put 01020 and which is the "color hey set cartoon underscore color comma, you will get to see the color".

(Refer Slide Time: 28:09)



You can see the protein in cartoon mode. I will show you the presentation on the protein tab on the right-hand side; you have to choose show as cartoon. So, you will see this representation where this representation is called a cartoon.

(Refer Slide Time: 28:27)



Electrostatics: Electrostatics is a very critical representation; it would not only represent the protein surface but also tell you about the electrostatic property of the protein, which, in many cases, has functional significance. So, you go to the tab protein tab and choose action to generate action, which means the A I showed you generate vacuum electrostatic of the protein contact potential, and you will get here, see there.

You have seen the white color portion. The white color portions represent the hydrophobicity or hydrophobic property of the surface. The blue one not only presents the hydrophilic part but also presents the presence of a positive charge, or you could say nitrogen. The red part gives a negative charge; you could say oxygen. So, how beautifully the three components of the protein the three elements majorly develop the proteins are presented here.

White is the color of carbon, blue is the color of nitrogen, and red is the color of oxygen; I will show you the demonstration of different aspects of Pymol.

(Video Start Time: 30:13)

So, now you see, this is the software Pymol I was talking about. So, this is the educational version. So, it is for educational use only. So, as I told you here, you could write comments first; let us load a biological macromolecule; you have to click a file, then you have go for open, and then you have to go to the desired folder, and then you click a PDB file. So, this is a PDB file that represents DNA now; you could move the DNA could rotate the DNA using your mouse.

Also, what you could do if you do not like the background, you could change the background; you also could have to see the sequence. So, you know the sequence; this is one strand and the other strand. So, you could have selected one strand, which comes here; remember these are the five A for action S for a show, H for hide, L for a label to do the labeling and C for color. So, you go to action and do the rename because if you do not rename, there might be a problem later when you select the other one.

So, you could say chain A DNA. So, you save here the beautiful option Pymol is providing you: you could go to file and look at the options where you could do the export molecule. So, when you look at explore molecule, you see your selected chain A underscore DNA is also there, and you save it. So, you name it chain A DNA, and you save it. Now you have nothing here; go there, open, get the chain A DNA, and see that the single-strand DNA is present.

So, in that way, you could save a part of the 3d structure, which enables you to do a lot of differential figure development and presentation according to the science you want to present.

So, that is one thing. Then if you want to present, what type of presence could you make? So, if you go to the wizard and go to the demo section and you do the representations, you see you could have a lot of different representations.

So, we have to make a single chain of the DNA. Now we will also bring DNA protein. So, every time you go to file, you have to open and then go to DNA protein. So, you see that the protein is interacting if you see a certain part of the protein is still there. They form a dimer. I have picked out a monomer here. You will see that for DNA, there is a major group and a minor group.

The interaction of a protein with the DNA happened mainly in the major group, and as you see here, you will see the helix, which directly interacts with the significant group; there is a loop and helix. So, this is a typical super secondary structure motif you remember when we are going through the protein structure protein domain I talked about; there is an awesome secondary structural domain that helps you to understand the function directly.

You do not even know about the protein, but when you see a helix loop helix-helix turn helix, you know that this protein is involved in protein DNA interaction, which you can see here. Now if you remember, I talked about display; when you represent the protein, you could now select the protein as I talked, you get to get a sele, and again you could rename it to prot. Now, if you come to the plot and make different representations, you could have seen those changes in the representation.

So, if you hide everything you do not get, you could have come up with different changes, like a cartoon presented as usual. Now if you again hide, you could show in ribbons that is a different representation. It was a cartoon; now it is a ribbon you want to show other ones so, you so, sticks, and you will see that this is presented at sticks. So, these types of protein representation will help you understand them.

So, those are the things we could do with the help of Pymol. As I told if you see, there are. So, many options in each of the menus; you could open a new Pymol window, you could open a file anything you could get a recent file that you have used, you could get a PDB by giving the PDB

id if you are connected to the internet also there is a very interesting option available in Pymol which is called saving the session.

When I saved, if you remember the PDB, the part of the DNA it was saved as a PDB file dot PDB file, it generates the coordinates again as a part of the two-strand or double-strand DNA. But if you want to save the changes you made in the session you are working you will do substation, and if you see the format is dot PSE or dot PZE, save it like a test. If you save it, you could directly open the session I would show you.

Let us show if we cut; we will go to the file where it was there; you click and save it. So, once you want to see the fonts, if it is not good enough to see, you could go to setting, edit all you can bring to display, and enhance the scale; if you enhance the scale, it will be very big. So, 2, I think, is a good number. So, you could see that you returned to the same confirmation where you ended.

Now if you want to see protein singly, you open a protein; let us open 1ake dot pdb. So if you know, you will see that this is a dimer. How to make it a monomer? First, you go to see the color and color by chain. When you do that, as you have correctly predicted, this is the dimer they produce. So, a dimer is an oligomer with two chains. So, when your color changes differently, you get sky blue on one chain and green on the other.

Now if you need the monomer to see, you could save it; if you want to save the waters, you could save here the complete if not, then you could save up to this. So, you get sele and rename it; this is 1ake. So, you make 1ake underscore mon. So, you could export the molecule, get 1ake underscores mon, and save it 1 ake underscore mon. Now if we see any other file, we see 4ake dot pdb.

So, this is 4ake dot PDB if. We color them by chain; you will get two chains. So, you know this is a dimer also. Now what you could do you could again is make a monomer out of that; you could save, and you could rename it as you did the first one, 4ake underscore mon, and then you

could save this by export molecule where you get 4ake underscore mon and save it as 4ake underscore mon. Now they are saved as monomers, but they also exist as dimers.

So, is it possible to align them? If you try to align 1ake with the 4ake, you see that you did an alignment. Is it a good alignment? How could you decide the alignment to be a good or bad alignment? The decision comes through a statistical technique called root mean square deviation or RMSD. Here you see the RMSD value is calculated. And the RMSD's value is calculated for 3370 atoms. The value is 18; what is the significance of this RMSD value?

If your RMSD comes between three, you declare the two structures have similar folds. So, these conclude that these two structures do not have a similar fold; you agree with me; if you agree with me, then you come and make a fool of yourself. What I mean now is you compare between 1ake and 4ake. In 1ake, the first residue is m the first residue of 4ake is m r r i i i i l l l l g g; what is going on put some time, and you will realize that 1ake and 4ake are representing the same molecule.

So, why do you get so high RMSD? Because we did not take care of a fundamental aspect, I made this mistake. So, that you all remember a critical thing about the structures, what is that critical thing? The critical thing is symmetry. Remember I talked about the symmetries; whatever the symmetries are, suppose this is a monomer, and this is a monomer. It depends on the building up of this symmetry; those two proteins would be coming like this or these or that.

So, they could be rotating in this screw axis. So, you should never remember very, very correctly. This is one of the most fundamental relations to use molecular visualization or to deal with the structure; you never align two structures in their oligomer. Whenever you need to align or compare two structures, you always make monomers out of them. However, you already make monomers. So, let us see if. We take the monomers and what will happen.

So, we take them out again when a new screen appears, and we open 1aka underscore mon and open 4ake underscore mon. They are now doing the alignment. So, align this as I told the command input area. So, 1ake underscore mon 4ake underscore mon. So, we have two, so,

better. We could have these two structures, and when so, 4ake underscore mon, and you see that they perfectly align with the value of 2.069.

Now you understand. So, the RMSD value of 18.6 is reduced to 2.069. So, I hope you have understood, and you have to take a very important relation from the current study that whenever you want to compare two structures, you will never do that in their dimer, trimer, or tetramer, which means oligomers you always do the first initial step you convert the oligomer to monomer and then compare the two monomers.

Now some smart people between you will ask ok, sir, we understand, but you have shown us that these two proteins are the same. Then why even the difference of 2.069? 2.069 only tells us they have similar folds, but we see they are the same. So, they should be aligned totally with an RMSD value of 0, and if you think you are thinking properly, but there is a trick, what is the trick?

This is the enzyme called adenylate kinase. Adenylate kinase has a beautiful history in that it could be thermodynamically stable in two conformations. So, if you look it very carefully, you will see that among the two structures, if I take this right, the blue structure has ap5, and if I take it out, the reddish structure is reddish brown. You could say it is not bound to anything. In the language of biochemistry, we call the conformation adopted in the crystal structure with the pdbid 4ake.

As an open structure and this is closed so, you could have probably now understood that there is a change of confirmation this part of the structure come down. And the coming down is responsible because of the presence of ap5. Kinases, if you go and check in the literature, are enzymes that undergo a huge conformational change. They have a big loop called the p loop. And this p loop changes its conformation; it is open to close, close to open, and this is how a protein declares its state of work. It has to do with some reaction or not.

When it is in the open state, it is not working. When it is in the closed state, it brings the substrate together to start the reaction, and this is one way it reduces the activation energy. If you

remember, I told you for a reaction to happen that, del G should always be negative for a spontaneous reaction. For biological reactions, a lot of times, the reactions have to be initiated. And initiation means the achieving of the activation barrier.

In a chemical laboratory, you produce heat; you have pressure and all those external powers to start a reaction. In physiological conditions, you cannot eat some food and then sit on the burner. So the heat comes on your body through the burner. It cannot happen right; you would destroy yourself if you could not put pressure in your mouth from an external source.

So, enzymes are responsible for how the reactions are possible in the physiological conditions in the normal pH and temperature. They have to do something to overcome those barriers, which you can see here. Also, this talks about switching because you could imagine the cell's body as a factory. But in the factory, they are in charge of their decision to stop the button and close the functionality.

Who will do that in their body? This is done by switching from an open state to a closed state, which decides whether the protein is going to be functional or not going to be functional. So, we have learned all these things now. If you see the protein of our interest, you will see exciting things. So, the window is already crowded. So, I will start from the beginning with the file if you remember the dot psc file we have.

So, we have the dot psc file. I will again make the setting for the good bigger view, which is two is good enough that you can see now. Now we have protein. We hide everything and. We presented this cartoon, and then we could find the different properties. If you know the action is deleted selection, you could delete the selection could, rename it as we have used could zoom could, orient it to the center it goes to the origin you could, drag the coordinate could, clean the labels, could modify around expand, extend inward complete restrict include exclude you could do the preset you could do find.

When you go to find polar contacts, you get many options. When you go to any contacts, you get option 33.54; when you go to high interaction, it will tell you what type of pi interaction. So, all

these types of interactions you could make, you could also align from there through command. So, these are all we could do with the options, and then you could generate different selection symmetry mates.

Who are the symmetry mates here in vacuum electrostatics? When you do vacuum electrostatics, if you remember, I will talk about that. So, you get three colors one is blue color another is red, color another is white color. Did you see anything special here? If you do not, I am telling you, you see a big blue part in the small protein. Why do these proteins have this unusually big blue stretch? This is because this protein is a DNA-binding protein, and if you see the charge of the DNA surface, this is made by phosphate.

So, this is a negative charge. So, to bind to the negative charge to interact with the phosphate, the proteins have a huge positive charge zone. So, looking at the electrostatic, you could see some sciences looking at the positive charge. You understand that this protein can bind to the negatively charged molecule here, which is proven because this protein is binding to the negatively charged DNA.

So, we will get all such kinds of information through the view of Pymol. So, to finish today's session, I have talked about the basics of Pymol. You see the windows. This window has a menu with file, edit, build, movie, display settings, scene, mouse, wizard, plugin, and help. So, you could use a lot of them, and there are many options. The builder option, which is coming later, allows you to modify. Remember, I talked about Coot as a model builder software O, Coot, SPDB viewer, and all of them.

Pymol is also updated and gives you the option to develop fragments it could, add residues could, and do sculpting; a lot of you know different building options are there. There are movies where you can create movies with a click. Then there are display options. As I have shown, you could have a display sequence, which you did. Then you could go to stereo mode, different types of stereo, and then overlay. You could change the background. Generally, we keep the background white.

When you make it white, taking pictures is more accessible, which I will show you in the next part of the talk in the next class. Also, I talked about these where you have five options A, S, H, L, and C. A is action; I have shown you different options in action then. The show will again give you other options to represent your biological micro molecule in lines non, bonded as ticks as fears ribbons, cotton, dots as mace.

So, you do different see you will get other presentation. Now you will understand that the part where the DNA comes and interacts shows the positive charge. So, you could do all of them and hide all the options you have given. You could label the residues the chain segments atom name element sigma symbol residue name one letter code residue identifier a lot of them. B factor occupancy Vander Waal radius what not.

And then you could color by the element, chain, second structure, representation, lines, steaks, ribbon, cartoon, and all these things. Then you could also bring different colors. Remember what I talked about here? You get variations of red, greens, blues, yellows, magenta, scions, oranges, teals, and grays. But if you go to RGB and provide RGB code, you could even give more colors which we have already discussed.

Now here also, there are many options you could reset, zoom, orient, and one of them is restoring, which is ray rastering makes the structure finer. When using ray, you are giving your structure much better. So, you could have saved it as a figure. So, you see the changes this is the ray has performed the rastering the rendering we have learned which came through the innovation of POV-Ray.

It is now applied regularly on the biological macro molecule to develop high-resolution structures. So, in that way, we could have performed many calculations, and I will finish the first demonstration of Pymol in that stage.

(Video End Time: 1:03:45)

So, this is the slide where I left and went to the demonstration, now returning to this. So, in this journey, we have learned the basics. We talked about the visualization software and the

visualization platform. We take them into the fourth step: the first one, the ORTEP; the scripting one, MolScript, Bobscript; the rendering, Pov-Ray, Raster 3D, and all the platforms. We have talked about them. We have introduced them.

But we came to talk in detail about the software called Pymol. I talk about the basics of Pymol Pymol is written in Python. After talking about the basics, we demonstrate different aspects, first going to the window on how to open a PDB file and then how you could define various aspects of the PDB file. How could you look at the PDB file? The PDB file could be represented like a cartoon, a wire, a ribbon, and sticks in all of such representations.

You could also change the color; there are different color options in the Pymol window; you could directly click them and do them. But there are also RGB codes that give you the universal ability to put any and every possible color. Then we look at the alignment to do the structural alignment. What we have learned is we must go and make the oligomer. So, we understood that for any oligomer, we must make a monomer and then do the alignment.

This is a very fundamental relation. We have learned which is true to understand structural biology. Then we see the change of conformations in a protein, which clearly shows how the protein functions. So, in addition to the basic visualization, we go into the deep thought of function. The functional understanding comes through the comparison of the two conformers of the protein adenylate kinase.

We see the protein exist in open form when the substrate does not come and bind to it and when the substrate comes, it just comes down and makes close conformation which is the proper conformation to perform the enzymatic catalysis. Very interesting for you guys to learn. Now, if you see the residue in the protein and how they interact with the substrate, you will find the atomic-level reasoning behind this; you will get into the more profound science.

And all these are possible by learning the software Pymol. So, with that, I will finish this class. Thank you very much for listening, thank you.