Instituto Politécnico Nacional

Escuela Superior de Cómputo

Bioinformatics

Practice 2 - VMD's tutorial

Group: 3CV9

Student:
Ramos Diaz Enrique

Professor:
Rosas Trigueros Jorge Luis

Development Date:
October 7th 2020

Due Date:
October 14th 2020

1 Theoretical Framework

The **Protein Data Bank (PDB)** was established as the 1st open access digital data resource in all of biology and medicine. Through an internet information portal and downloadable data archive, the PDB provides access to 3D structure data for large biological molecules (proteins, DNA, and RNA). These are the molecules of life, found in all organisms on the planet [1].

Knowing the 3D structure of a biological macromolecule is essential for understanding its role in human and animal health and disease, its function in plants and food and energy production, and its importance to other topics related to global prosperity and sustainability [1].

RCSB PDB (Research Collaboratory for Structural Bioinformatics PDB) operates the US data center for the global PDB archive, and makes PDB data available at no charge to all data consumers without limitations on usage [1].

The Vision of the RCSB PDB is to enable open access to the accumulating knowledge of 3D structure, function, and evolution of biological macromolecules, expanding the frontiers of fundamental biology, biomedicine, and biotechnology [1].

On the other hand, **VMD (Visual Molecular Dynamics)** is a molecular visualization and analysis program designed for biological systems such as proteins, nucleic acids, lipid bi-layer assemblies, etc. The program is developed by the Theoretical and Computational Biophysics Group at the University of Illinois at Urbana-Champaign [2].

Among molecular graphics programs, VMD is unique in its ability to efficiently operate on large bio-molecular complexes and long-timescale molecular dynamics trajectories, its interoperability with a large number of molecular dynamics simulation tools, its integration of structure and sequence information, and its built-in support for advanced image rendering and movie making. Figure 1 shows what can be done with this software tool [2].

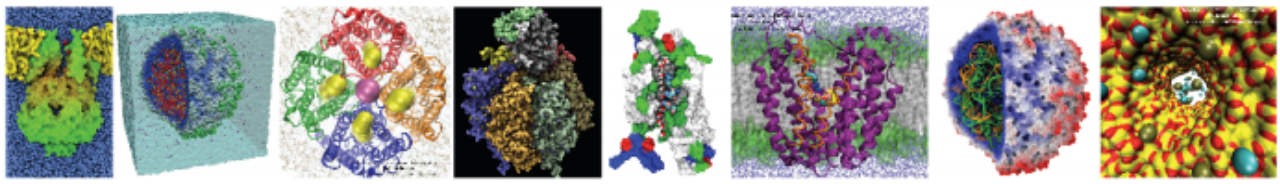


Figure 1: Example VMD Renderings. Image taken from [2].

Key features of VMD include general 3-D molecular visualization with extensive drawing and coloring methods, visualization of molecular dynamics, support for all major molecular data file formats, extensive atom selection syntax for choosing subsets of atoms, rendering of high-resolution, publication-quality molecule image, and interactive molecular dynamics simulations [2].

2 Material and Equipment

- RCSB PDB: Homepage - PDB files [1].
- VMD - Visual Molecular Dynamics Software (version 1.9.3 - using OpenGL) [3].
- "Using VMD tutorial" PDF from the VMD's official web page [2].
- (Optional) NVIDIA CUDA - only for computers with a GPU

3 Practice Development

3.1 Downloading PDB protein models

Go the RCSB PDB: Homepage through the next link: https://www.rcsb.org/. For this practice, the chosen protein models are:

1. **1UBQ - Ubiquitin** [4]
2. **6MWM - Bat Coronavirus** [5]
3. **2M1E - Human Insulin** [6]

Search for these proteins as shown in Figure 2, click on them and then its very important to download the models in **PDB Format** as shown in Figure 3.

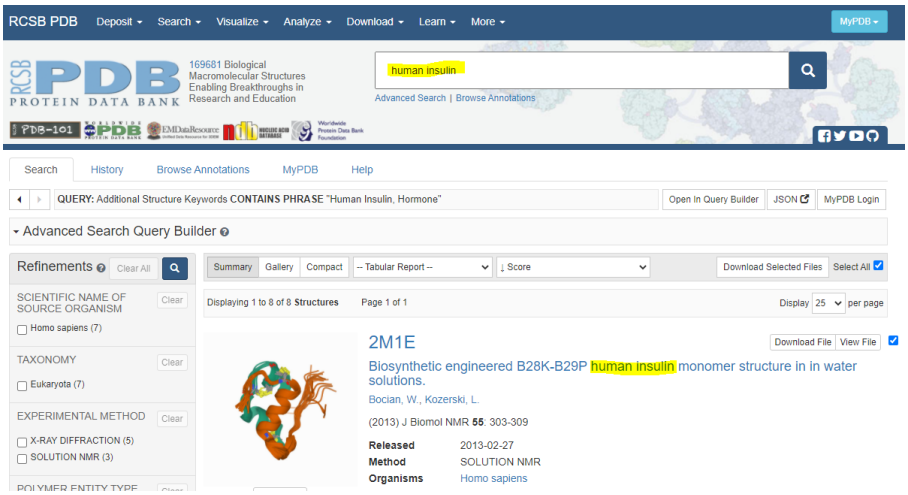


Figure 2: Search results in the RCSB web page for "human insulin".

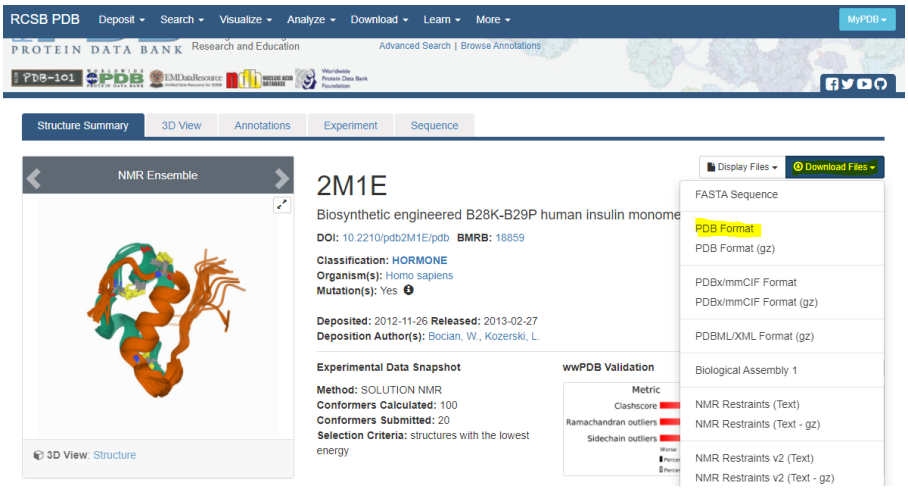


Figure 3: Downloading example for the model of the 2M1E protein. Image taken from [6]

3.2 Downloading and installing VMD

Go to the VMD software's download web page [3] and pick the correct version according to the installed operating system, as show in Figure 4. If your computer has a GPU, its highly recommended to install CUDA and after that, chose the OpenGL and CUDA version.

For this practice and the next ones, the selected version is the Windows 1.9.3 OpenGL + CUDA. After the download, launch the installer and follow the instructions on them.

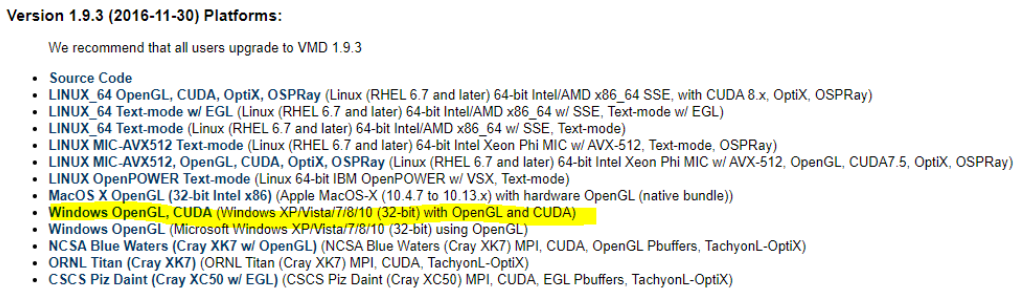
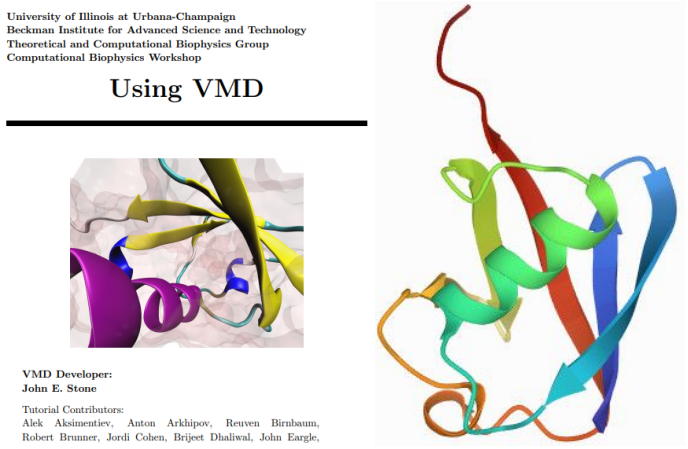


Figure 4: All VMD software's version available in its official web page. Image taken from [3].

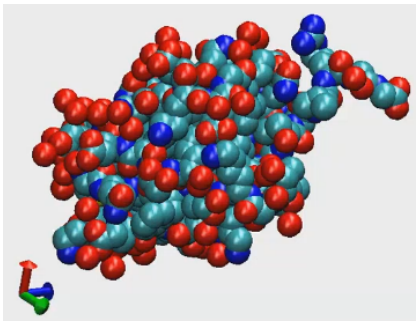
3.3 Using VMD tutorial with the 1UBQ model

In the lab session, the VMD's tutorial [2], shown in Figure 5a, was followed up to section 1.5 using the 1UBQ - Ubiquitin protein [4] as the main model example, as can be seen in Figure 5b.

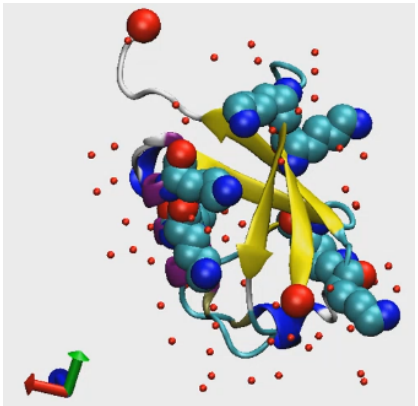


(a) Using VMD tutorial's PDF [2]. (b) 1UBQ model.

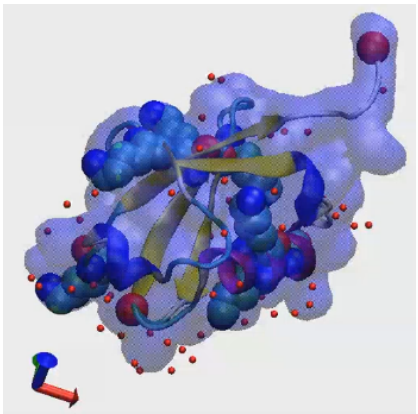
Figure 5: VMD's tutorial and the 1UBQ model.



(a) VDW drawing method.



(b) Four different selections with their respective styles and colors: protein (NewCartoon - Structure), resname LYS (VDW - Name), water (CPK - Name), CA in residues 1 and 76 (VDW - ColorID 1: red).



(c) Additional selection for the Figure 6b: protein with QuickSurf style, Molecule color and Transparent material. The previous internal variations can be seen covered by this surface.

Figure 6: 1UBQ model variations resulted from the lab session 2, following the VMD's tutorial.

Some of the resulting variations are shown in Figure 6. Because of this, the steps of the tutorial for this specific model protein won't be reported. Instead of that, the other two remaining protein models (6MWM and 2ME1) will be shown and analyzed in this paper.

3.4 6MWM - Bat Coronavirus Protein Model

In this section the VMD's tutorial [2] is followed up to section 1.4 using the 6MWM Bat coronavirus HKU4 SUD-C viral protein [5].

3.4.1 Loading and Displaying a Molecule

Start a VMD session. In the VMD Main window, choose File > New Molecule and load the PDB file through the browser window as shown in Figure 7 [2]. The Figure 8 shows how the protein model looks displayed with VMD's default settings.

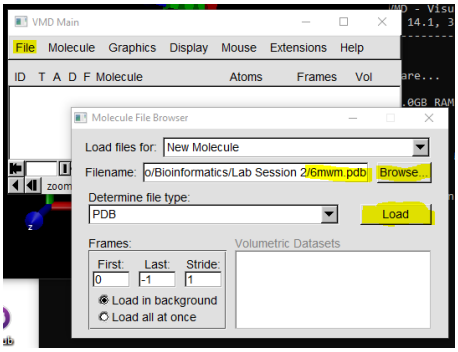


Figure 7: Loading a Molecule.

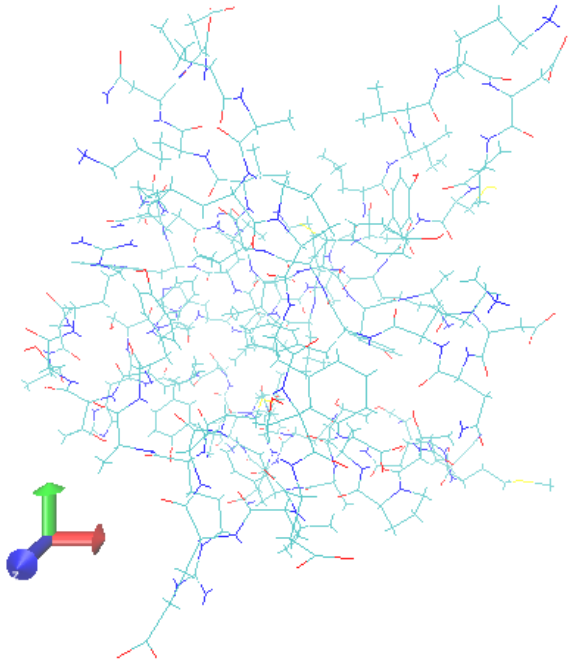


Figure 8: 6MWM Bat coronavirus model with a little zoom applied.

3.4.2 Exploring different drawing styles and coloring methods

Some of the different drawing styles available in this software are applied to the molecule through the tab Graphics > Representations in the VMD Main window, which results in the pop up of the Graphical Representations window (Figure 9), where these can be explored through the Drawing

Method list options in the Draw Style Tab [2]. Figure 10 shows three different drawing methods for the 6MWM protein: Licorice, Tube and NewCartoon.

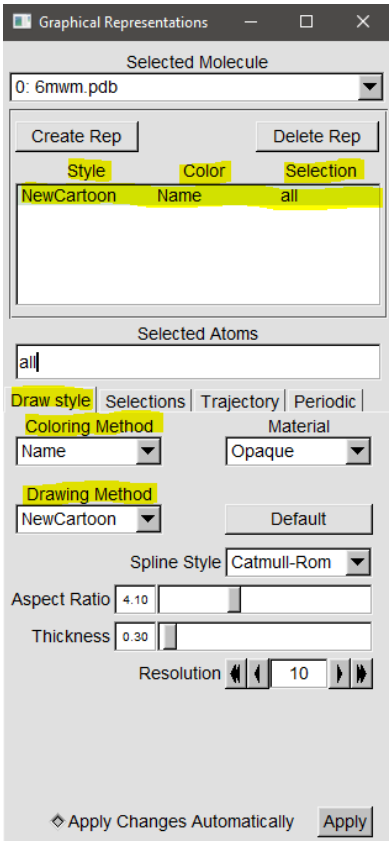
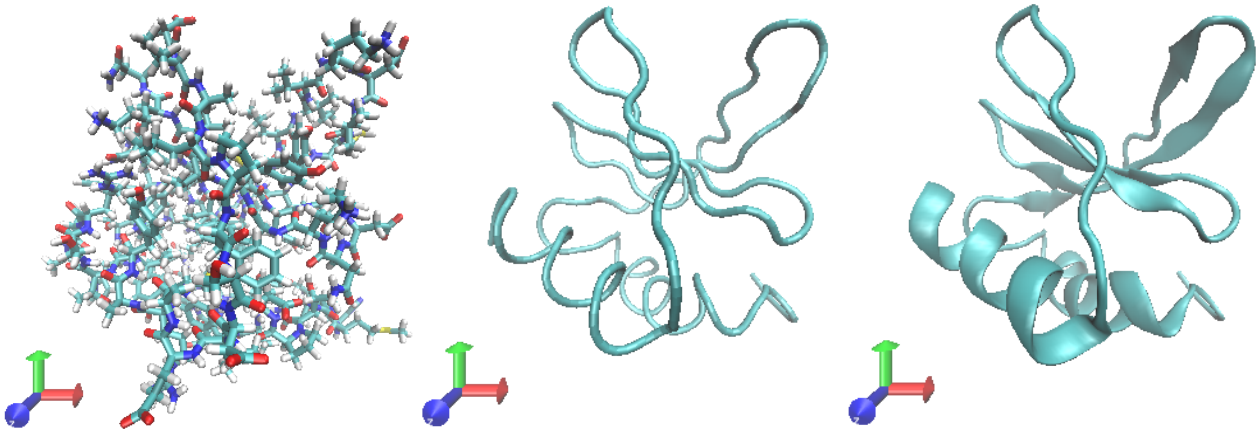


Figure 9: The Graphical Representations window.

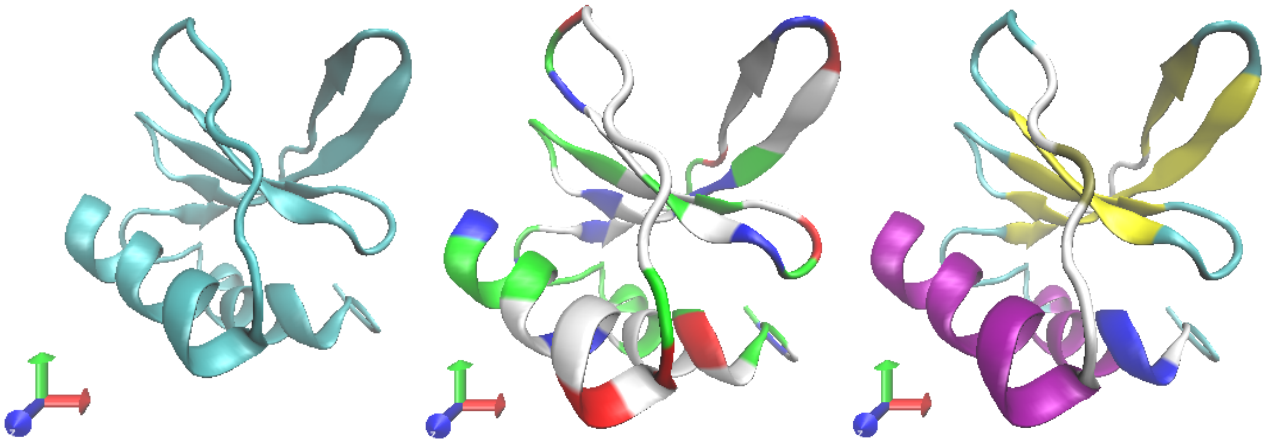


(a) Licorice drawing method. (b) Tube drawing method. (c) NewCartoon drawing method.

Figure 10: Representations of 6MWM Bat Coronavirus.

Additionally, color variations can be applied to these representations. In the same Graphical Representations window (Figure 9), these color styles can be explored through the Coloring

Method list options in the Draw Style Tab [2]. Using the drawing method NewCartoon, in Figure 11 are shown three different color variations: Name, ResType and Secondary Structure.



(a) Name coloring method. (b) ResType coloring method. (c) Secondary Structure coloring method.

Figure 11: Coloring variations for NewCartoon 6MWM style.

3.4.3 Creating multiple representations

There is a button on the Graphical Representations window (Figure 9) called "Create Rep", that allows to create multiple representations of the same molecule. There can be a mixture of different selections, each one with its drawing and coloring methods, all displayed at the same time [2]. Figure 12 shows some multiple representations as example created in the Graphical Representations window, and Figure 13 shows them displayed one by one and all at the same time with their respective description of drawing, coloring and selection styles.

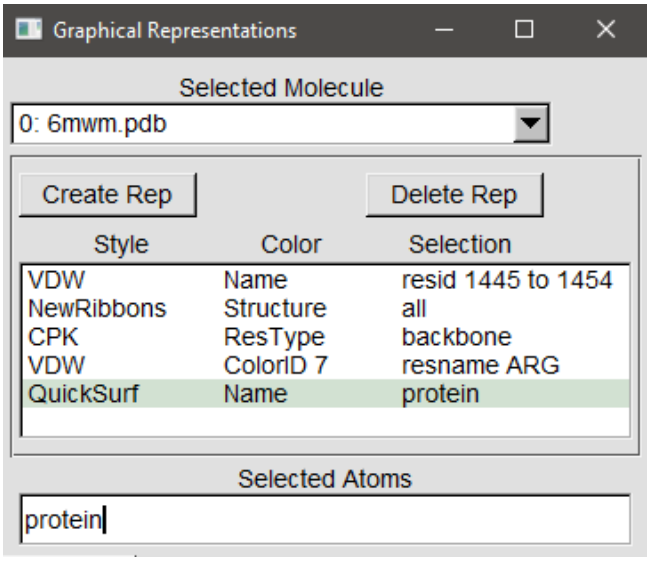
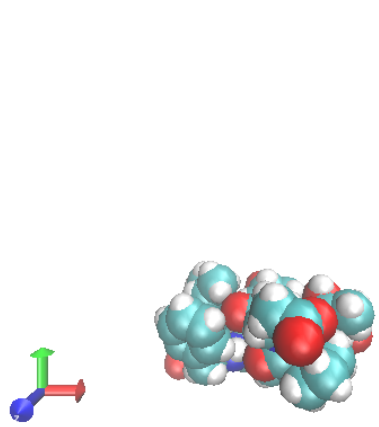
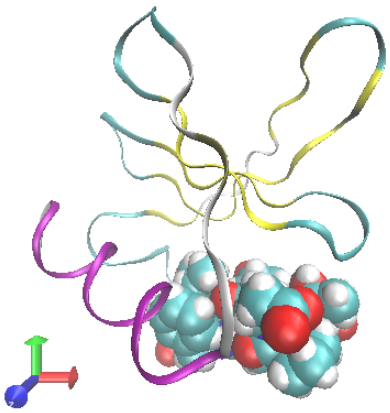


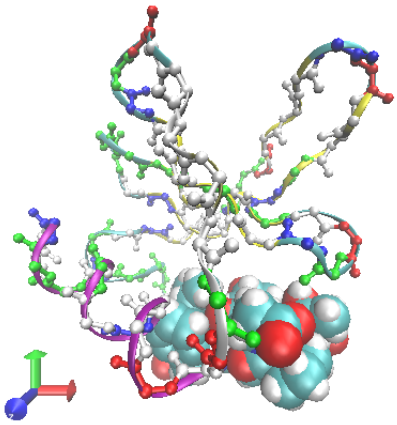
Figure 12: Different selections with different representations for the 6MWM molecule.



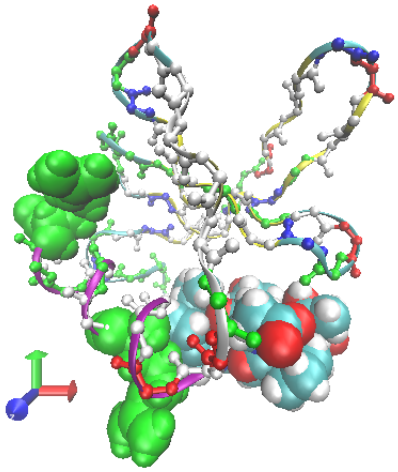
(a) 1st selection: Residues 1445 to 1454 with VDW style and Name coloring.



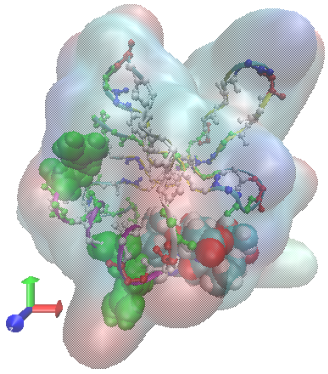
(b) 2nd selection: All molecule with NewRibbons style and Secondary Structure coloring.



(c) 3rd selection: Backbone with CPK style and ResType coloring.



(d) 4th selection: Resname ARG with VDW style and ColorID 7 (green) coloring.



(e) 5th selection: Protein with transparent QuickSurf style and Name coloring.

Figure 13: Multiple selections and representations displays for the 6MWM molecule.

3.4.4 Sequence Viewer Extension

The sequence viewer extension allows to view the protein sequence, as well as picking and displaying one or more residues. In the VMD Main window, choose Extensions > Analysis > Sequence Viewer, and the VMD sequence window for the molecule will show [2]. The protein sequence for the 6MWM molecule is shown in Figure 14.

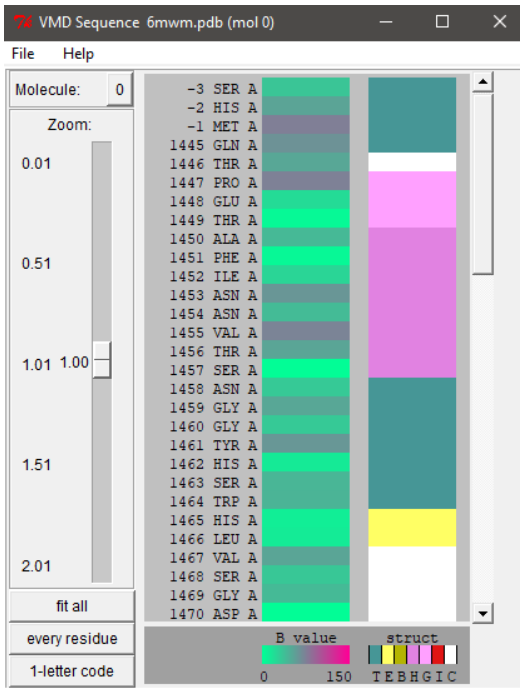


Figure 14: VMD Sequence window with the list of the amino acids of the 6MWM molecule.

3.5 2M1E - Human Insulin Hormone Model

In this section the VMD's tutorial [2] is followed up to section 1.4 using the 2M1E Biosynthetic engineered B28K-B29P human insulin monomer structure in water solutions [6].

The specific steps of the tutorial won't be reported here in order to avoid repetition and redundancy. Instead of that, only the resulted displays of the different styles, colors, selections and representations of the 2M1E molecule's model will be shown.

First, the PDB file is load in VMD (see subsection Loading and Displaying a Molecule). The Figure 15 shows how the protein model looks displayed with VMD's default settings.

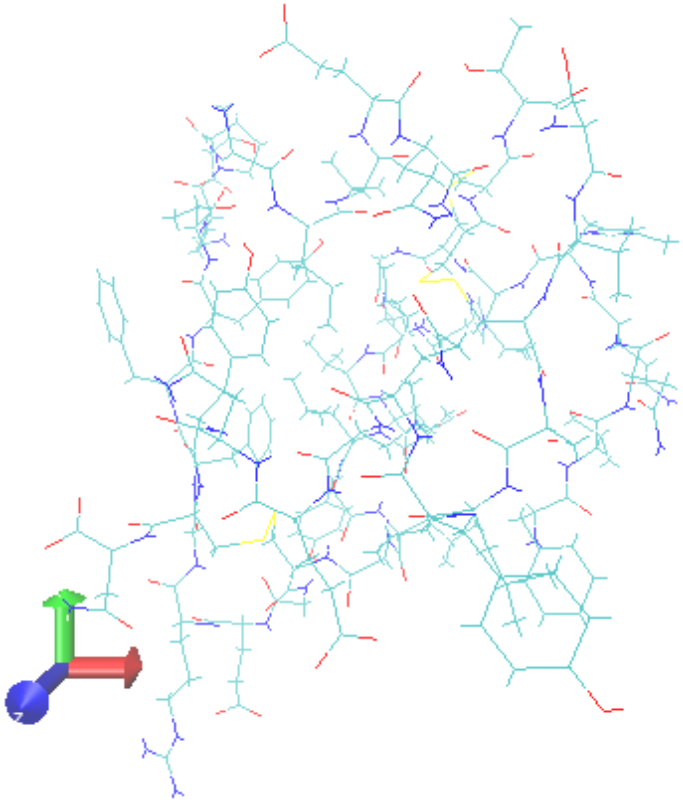
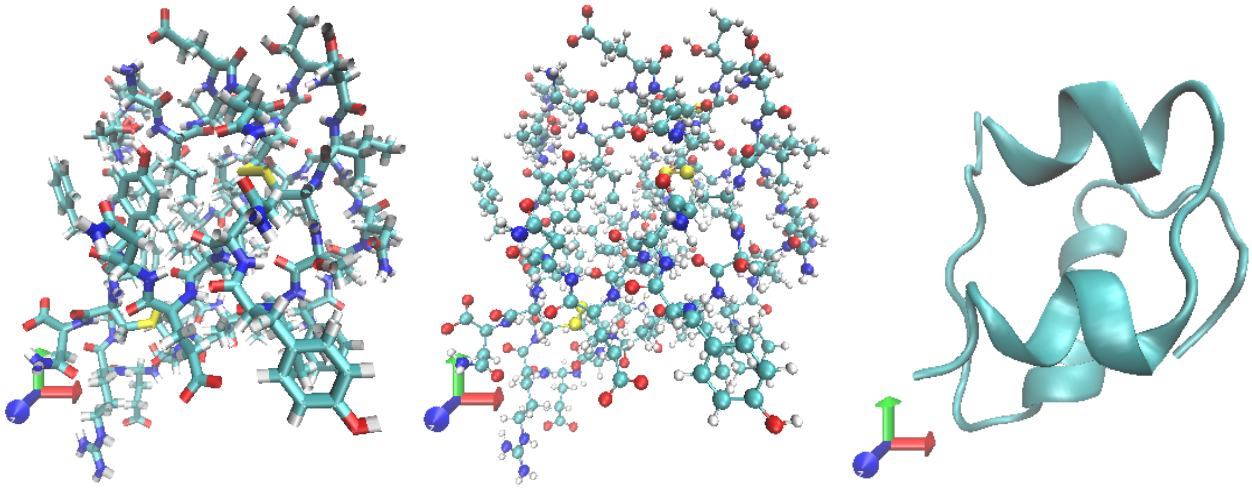


Figure 15: 2M1E human insulin model with a little zoom applied.

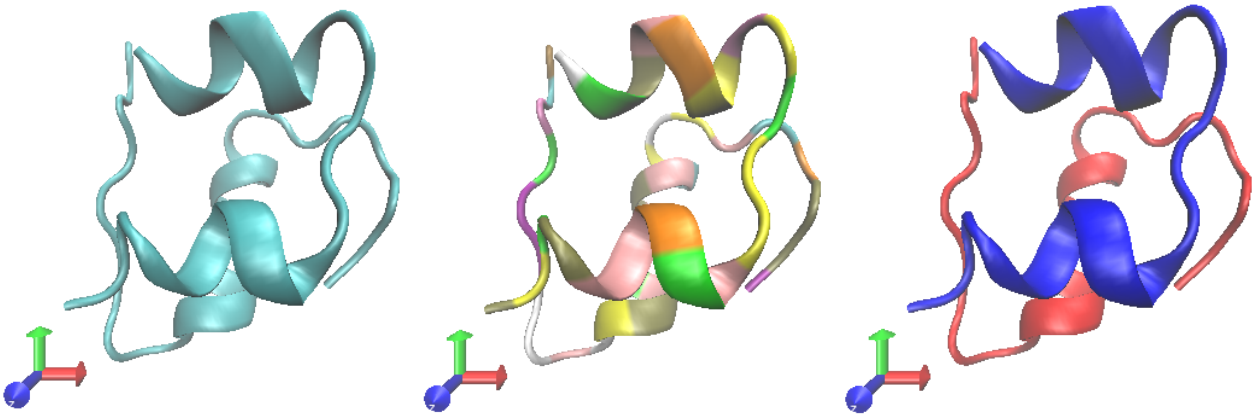
Second, three different drawing styles are applied to the 2M1E hormone model: Bonds, CPK and NewCartoon, through the Graphical Representations window (see subsection Exploring different drawing styles and coloring methods), which can be seen in Figure 16.



(a) Bonds drawing method. (b) CPK drawing method. (c) NewCartoon drawing method.

Figure 16: Representations of 2M1E human insulin.

Third, color variations are applied to these representations (see subsection Exploring different drawing styles and coloring methods). Using the drawing method NewCartoon, in Figure 17 are shown three different color variations: Name, ResName and Chain.



(a) Name coloring method. (b) ResName coloring method. (c) Chain coloring method.

Figure 17: Coloring variations for NewCartoon 2M1E style.

Fourth, Figure 18 shows some multiple representations as example created in the Graphical Representations window (see subsection Creating multiple representations), and Figure 19 shows them displayed one by one and all at the same time with their respective description of drawing, coloring and selection styles.

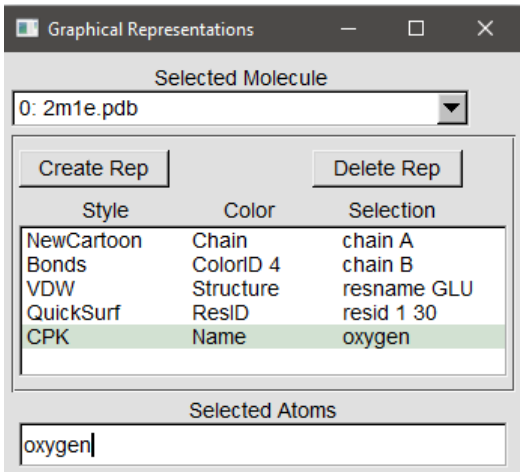
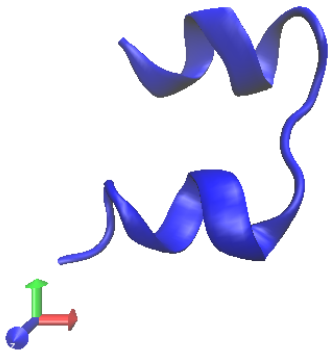
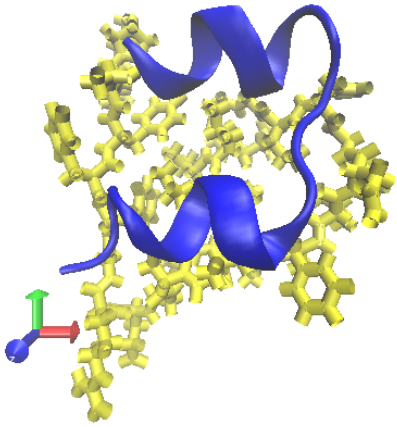


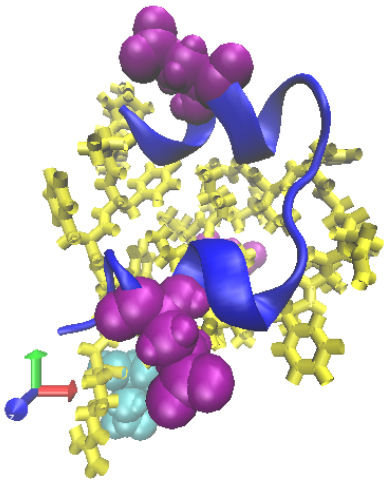
Figure 18: Different selections with different representations for the 2M1E molecule.



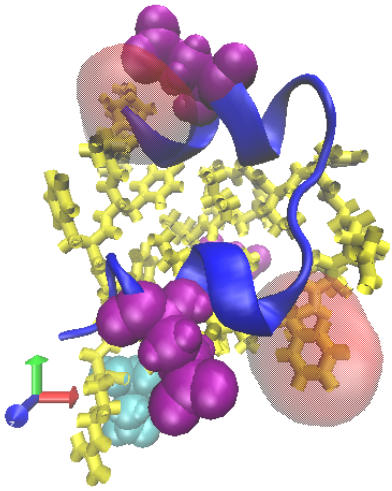
(a) 1st selection: Chain A with NewCartoon style and Chain coloring.



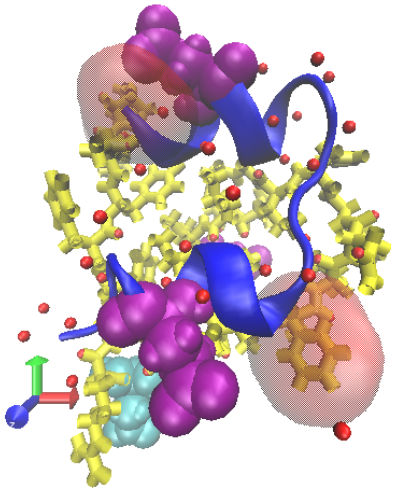
(b) 2nd selection: Chain B with Bonds style and ColorID 4 (yellow) coloring.



(c) 3rd selection: ResName GLU with VDW style and Structure coloring.



(d) 4th selection: Residues 1 and 30 with transparent QuickSurf style and ResID coloring.



(e) 5th selection: Oxygens with CPK style and Name coloring.

Figure 19: Multiple selections and representations displays for the 2M1E molecule.

Finally, the protein sequence (see subsection Sequence Viewer Extension) for the 2M1E molecule is shown in Figure 20.

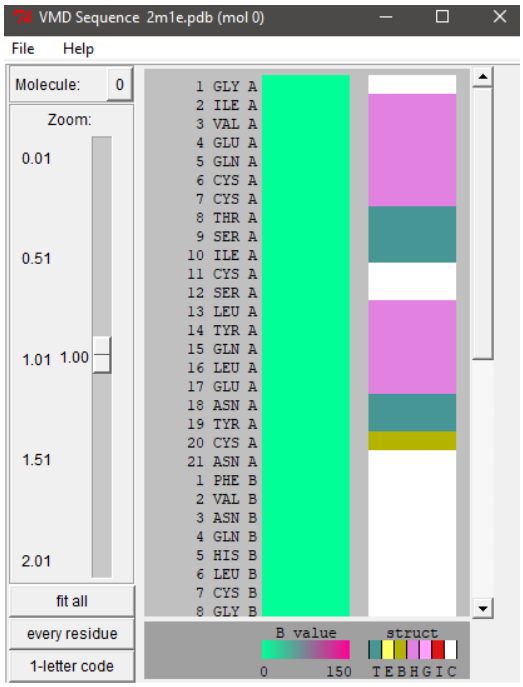


Figure 20: VMD Sequence window with the list of the amino acids of the 2M1E molecule.

4 Conclusions and recommendations

Its highly recommended to switch the color of the background in VMD from black to white, for better detailed displays on the molecule models.

There are several structure-prediction tools, easy to use and readily available via internet-based servers, and they are an important contribution of computational biology to the development of the life sciences. All available prediction methods have limitations, however, dictated by the hypotheses on which they rely, and this determines very precisely the field of application of the models produced.

5 References

[1] R. PDB, "About RCSB PDB: Enabling Breakthroughs in Scientific and Biomedical Research and Education," https://www.rcsb.org/pages/about-us/index, [Online; last access October 13, 2020].

[2] J. E. Stone, "Using VMD," https://www.ks.uiuc.edu/Training/Tutorials/vmd/vmd-tutorial.pdf, [Online; last access October 13, 2020].

[3] "Download VMD," https://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD, [Online; last access October 13, 2020].

[4] R. PDB, "1UBQ STRUCTURE OF UBIQUITIN REFINED AT 1.8 ANGSTROMS RESOLUTION." https://www.rcsb.org/structure/1UBQ, [Online; last access October 13, 2020].

[5] ——, "6MWM Bat coronavirus HKU4 SUD-C." https://www.rcsb.org/structure/6MWM, [Online; last access October 13, 2020].

[6] ——, "2M1E Biosynthetic engineered B28K-B29P human insulin monomer structure in in water solutions." https://www.rcsb.org/structure/2M1E, [Online; last access October 13, 2020].