Bioscience Discovery, 9(1): 53-58, Jan - 2018

Print & Online, Open Access, Research Journal Available on http://jbsd.in
ISSN: 2229-3469 (Print); ISSN: 2231-024X (Online)
Research Article

A comparative study of different viewing and selection options provided by protein structure visualization tools

Shaheda N. Ansari

Department of Computer Science
AKIS, Poona College of Arts, Science and Commerce, Camp, Pune-411001(MS)
Email: shad_cs@yahoo.com

Article Info

Received: 17-07-2017,
Revised: 10-10-2017,
Accepted: 06-11-2017

Keywords:
Amino Acids, R group, Side Chain, Fold etc.

Abstract

There are many different types of proteins, each with a particular shape and function, and those shapes and functions are linked. Proteins can be made up of one or more chains of amino acids, and the order of the amino acids determines the function of the protein. The protein then forms a particular structure, depending on the amino acids present. This structure is held together by a variety of different possible bonds, including covalent disulphide bonds, electrostatic interactions, and hydrogen bonds. A molecular graphics visualization tool is required to view the structure encoded by PDB files and to be able to manipulate the images to view the molecule from various perspectives. PDB file lists each atom and its numerical coordinates in 3-D space. Thus tools are required that are capable of loading and displaying huge amount of data due to the difficulties inherent in understanding raw numbers. Many tools have been developed to visualize a protein whose structure has been known. During present course of investigation a comparative study of seven commonly used freely available protein structure visualization tools viz. RasMol, Chime, Protein Explorer, Swiss-pdb Viewer, WebMol, MoLMOL and Cn3D were made based on different 'viewing' and 'highlighting part of the protein in the structure based on given selection' options, that will help the researchers to select the appropriate tool in their study.

INTRODUCTION

Proteins are macromolecules (heteropolymers) made up from 20 different amino acids, also referred to as residues (Rastogi, 2006). The arrangement of amino acids along the chain determines the structure and chemical properties of the protein. The R-group determines the identity, structure, and function of the amino acid. The structural and chemical relatedness of the R-groups allows classification of the twenty amino acids into chemical groups. Amino acids can be classified according to optical activity (the ability to polarize light), acidity and basicity, polarity and nonpolarity, or hydrophilicity (water-loving) and hydrophobicity (water-fearing). These categories offer clues to the function and reactivity of the amino acids in proteins. The biochemical properties of amino acids determine the role and function of protein in the human body (Radecki and Kim, 2007). There appear to be many different classification systems based on the chemical and/or structural properties of their side chains. A more meaningful classification of amino acids is based on the polarity of R group present in their molecules, i.e., their tendency to interact with water at biological pH (near pH 7.0). The R group of the amino acids varies widely with respect to their polarity from totally nonpolar or hydrophobic (water-hating) R

group to highly polar or hydrophilic (water-loving) R groups. This classification of amino acids emphasizes the possible functional roles which they perform in proteins (Jain *et al.*,2006).

The unique structure and chemical composition of each protein is important for its function. Differences in protein properties originate from differences in the chemical structure of the amino acids that make up the protein. Proteins are typically characterized by their size (molecular weight) and shape, amino acid composition and sequence, isolelectric point (pI), hydrophobicity, and biological affinity. Differences in these properties can be used as the basis for separation methods in a purification strategy. The chemical composition of the unique R groups is responsible for the important characteristics of amino acids, chemical reactivity, ionic charge and relative hydrophobicity. Therefore protein properties relate back to number and type of amino acids that make up the protein (Seidman and Mowery, 2006).

To be able to perform their biological functions proteins fold into one, or more, specific spatial conformations, driven by a number of noncovalent interactions such as hydrogen bonding, ionic interactions, Van der Waals' forces and hydrophobic packing. The various types of secondary structure are defined by their patterns of hydrogen bonds between the main-chain peptide groups. There are three common secondary structures in proteins, namely alpha helices, beta sheets and turns.

The most important factor governing the folding of a protein into 3D structure is the distribution of polar and non-polar side chains (Cordes *et al.*, 1996). The hydrophobic effect (Silverman, 2001) plays a crucial role in protein stability and folding. The interaction of proteins with water plays a large and essential role in almost all aspects of protein folding and function (Mattos, 2002). Heat, acid, or other conditions can disturb proteins, causing them to uncoil or lose their shape and impairing their ability to function. This is referred to as denaturation (Himburg, 2007). Although the primary amino acid sequence determines how the protein folds, this process is not completely understood. Although certain amino acid sequences can be identified as more likely to form a particular conformation, it is still not possible to completely predict how a protein will fold based on its amino acid sequence alone, and this is an active area of biochemical research (Seidman and Mowery, 2006).

Solved structures are usually deposited in the Protein Data Bank (PDB) (Berman *et al.*, 2000), a freely available resource from which structural data about thousands of proteins can be obtained in the form of Cartesian coordinates for each atom in the protein. As the genome sequencing projects proceed, scientists have gained access to tremendous amounts of biological information. Due to the difficulties inherent in understanding large quantities of data, information visualization techniques have become an attractive option for the field of bioinformatics. Using information visualization, researchers can see experimental results more clearly than by simply viewing raw numbers (Pang *et al.*, 1999). There are many well established ways of visualizing the 3D protein structures. Each way of visualization highlights a different aspect of the protein molecule (Shirky, 2000). The ability to visualize the 3D structure is of proteins is critical in many areas like, drug design, protein modeling. This is because that the 3D structure of a protein determines its interaction with other molecules, hence its function, and the relation of the protein to other known proteins. Moreover, studying the interaction between protein molecules may also require visualizing huge number of atoms, thus researchers also need tools that are capable of loading and displaying this huge amount of data (Can *et al.*, 2003). Many tools have been developed to visualize a protein whose structure has been known. Some of these tools are: RasMol (Sayle and Milner-White, 1995), Chime (MDL Inftormation Systems, Inc.), Protein Explorer (Martz, 2002), Swiss-PDB viewer (Kaplan and Littlejohn, 2001), WebMol (Walther, 1997), MOLMOL (Koradi *et al.*, 1996), and Cn3D (Wang *et al.*, 2000). The study was performed on various display capabilities of these softwares (Ansari and Sayyed, 2011). During present course of investigation the study is performed on these softwares for different viewing and highlighting part of the protein structure based on given selection option.

MATERIALS AND METHODS

For visualizing the 3-D structure of proteins seven freely available structure visualization tools were downloaded:

RasMol (http://www.umass.edu/microbio/Rasmol/),
Chime (http://www.mdl.com/products/framework/chime/)
Protein Explorer (http://www.proteinexplorer.org/)

Swiss-PDB viewer (http://www.expasy.org/spdbv/)
WebMol
(http://www.cmpharm.ucsf.edu/~walther/webmol/)
Cn3D
(http://www.ncbi.nlm.nih.gov/Structure/CN3D/)

Structure data files were downloaded from the Protein Data Bank (http://www.rcsb.org/pdb/) and NCBI
(http://www.ncbi.nlm.nih.gov/Structure/MMDB/mmdb.shtml).

One of the download protein structure file (PDB ID: 1CRN) was loaded in each of the seven softwares one by one and softwares were observed for the following properties:

Observations were made to see whether the softwares provides the options for - to rotate, translate and zoom the structure; stereo view; slab; and spin.

The study was performs to see the selection options provided by the softwares that highlights the selected portion within the structure viz. selection based on amino acids properties like properties like acidic, basic, neutral, aromatic, aliphatic, charged, polar, hydrophobic, buried, surface and forming part of helices, sheet and turn. Selection options like selecting all groups, deselecting all the groups, inverse of current selection, selecting range of residues, selecting residue based on clicked atom in the structure, selection of groups within specified distance of the clicked atom in the structure, selection of all protein chains, specific chain, selection by atom type, by residue type, by hetero atoms were also observed.

MOLMOL
(http://www.mol.biol.ethz.ch/wuthrich/software/molmol/)

RESULTS AND DISCUSSION

The results summarized in Table 1 shows that all the seven softwares have the facilities to rotate, translate and zoom the structure. Stereo view can not be seen in Cn3D but can be seen in other six softwares. Slab option in not provided in MOLMOL and Cn3D but has in other five softwares. RasMol can not spin the structure where as other six softwares can spin the structure.

Table 1 : Different viewing options provided by protein structure visualization tools.

Software	Rotate, Translate, Zoom	Stereo View	Slab	Spin
RasMol	+	+	+	-
Chime	+	+	+	+
Protein Explorer	+	+	+	+
Swiss-Pdb Viewer	+	+	+	+
WebMol	+	+	+	+
MOLMOL	+	+	-	+
Cn3D	+	-	-	+

+ = YES, - = NO

The selection based on amino acids properties studied in all the softwares, summarized in Table 2 shows the RasMol, Chime and Protein Explorer can select amino acids based on acidic, basic, neutral, aromatic, aliphatic, charged, polar, hydrophobic, buried, surface, forming helices, sheets and turns and highlight in the structure. Swiss-PDB viewer can select only acidic, basic, polar, hydrophobic, forming helices, sheet and turns amino acids. WebMol can only select buried and surface amino acids. MOLMOL and Cn3D provide no options for selection based on amino acids properties.

Table 2 : Selection based on Amino Acid Properties provided by protein structure visualization tools.

Software	Acidic	Basic	Neutral	Aromatic	Aliphatic	Charged	Polar	Hydrophobic	Buried	Surface	Helices	Sheet	Turn
RasMol	+	+	+	+	+	+	+	+	+	+	+	+	+
Chime	+	+	+	+	+	+	+	+	+	+	+	+	+
Protein Explorer	+	+	+	+	+	+	+	+	+	+	+	+	+
Swiss-Pdb Viewer	+	+	-	-	-	-	+	+	-	-	+	+	+
WebMol	-	-	-	-	-	-	-	-	+	+	-	-	-
MOLMOL	-	-	-	-	-	-	-	-	-	-	-	-	-
Cn3D	-	-	-	-	-	-	-	-	-	-	-	-	-

+ = YES, - = NO

Selection options provided by softwares summarized in Table 3 shows that all the softwares provide selection of all the groups, deselecting all the groups and selection by residue type. Inverse of current selection is possible in RasMol, Chime, Protein Explorer, and Swiss-PDB Viewer but not in WebMol, MOLMOL and Cn3D. Selection of range of residues is not possible in Chime, where as possible in all the other softwares. Selecting residue by clicking on structure is not possible only in WebMol but possible in other softwares. Selecting groups within specified distance of the clicked atom is not provided in Chime and MOLMOL, where as provided in other softwares. Protein chains can be selected in RasMol, Chime and Protein Explorer but not in other softwares. Selection by chain is not supported by WebMol and Cn3D but other softwares supports. Selection by atom type is not provided by Swiss-PDB Viewer, WebMol and Cn3D, where as other four softwares provides. Selection by hetero atoms is possible using RasMol, Chime, Protein Explorer and Swiss-PDB Viewer, but not by WebMol, MOLMOL and Cn3D.

This comparative study will help the researchers to choose the appropriate tool in their study.

CONCLUSION

A molecular graphics visualization tool views the molecule from various perspectives and highlights the different aspect of the protein molecule. A tool should have rotate, translate and zoom options to see the molecule from different angles; slab option to see inside the molecule; and spin option that rotate the view automatically which helps to show 3-D aspects of the structure and to appreciate the 3-D structure. It should have selection option based on amino acid properties viz. Acidic, Basic, Neutral, Aromatic, Aliphatic, Charged, Polar, Hydrophobic, Buried, Surface, and forming part of secondary structure Helix, Sheet and Turn; that highlights only those amino acids in the structure that falls under the selected category. It should have selection options like ‘All’ to select all groups i.e. whole molecule, ‘None’ to deselect all groups, ‘Inverse’ to select the inverse of a current selection, ‘Range’ to selects a range of residues, ‘Clicked on Structure’ to select groups by clicking them on the structure, ‘Groups within a specified distance of clicked atom’ to select all groups within the specified distance of the clicked atom in the structure, ‘Protein’ to select all protein chains, ‘By Chain’ to select all atoms in a selected chain, ‘By Residue type’ to select all residues of the chosen type, ‘By Atom type’ to select all atoms of chosen type, ‘By Hetero atoms’ to selects hetero atoms i.e. ligands or solvents.

Table 3 : Selection options provided by protein structure visualization tools.
+ = YES, - = NO

Software	All	None	Inverse	Range	Clicked on Structure	Groups within a specified distance of clicked atom	Protein	By Chain	By Residue type	By Atom type	By Hetero Atoms
RasMol	+	+	+	+	+	+	+	+	+	+	+
Chime	+	+	+	-	+	-	+	+	+	+	+
Protein Explorer	+	+	+	+	+	+	+	+	+	+	+
Swiss-Pdb Viewer	+	+	+	+	+	+	-	+	+	-	+
WebMol	+	+	-	+	-	+	-	-	+	-	-
MOLMOL	+	+	-	+	+	-	-	+	+	+	-
Cn3D	+	+	-	+	+	+	-	-	+	-	-

REFERENCES

Ansari SN and Sayyed Iliyas, 2011. A Comparative Study of Protein Structure Visualization Tools for Various Display Capabilities. *Bioscience Discovery*, **02**(2):222-226.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN and Bourne PE, 2000. The Protein Data Bank. *Nucleic Acids Res*., **28**:235–242.

Can T, Wang Y, Wang YF and Su J, 2003. FPV: fast protein visualization using Java 3D. *Bioinformatics*, **19**(8):913-922.

Cordes MH, Davidson AR and Sauer RT, 1996. Sequence space, folding and protein design. *Curr Opin Struct Biol*., **6**:3-10.

Himburg SP, 2007. *Amino Acids* [online], Nutrition and Well-Being A to Z. Available from: http://www.faqs.org/nutrition/A-Ap/Amino-Acids.html

Jain JL, Jain S and Jain N, 2005. *Fundomentols of Biochemistry*, New Delhi, S. Chand & Company Ltd.

Kaplan W and Littlejohn TG, 2001. Swiss-PDB Viewer (Deep View). *Briefings in Bioinformatics*, **2**:195–197.

Koradi R, Billeter M and Wuthrich K, 1996. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graphics*, **14**:51-55.

Mattos C, 2002. Protein-water interactions in a dynamic world. *Trends Biochem. Sci.*, **27**:203-208.

Pang A, Meads D and Hansen M, 1999. *Visualizing BioInformatics Data*, SLVG. Available from: http://www.soe.ucsc.edu/research/slvg/bio.html

Radecki J, Kim S, 2007. *Protein* [online], Nutrition and Well-Being A to Z. Available from: http://www.faqs.org/nutrition/Pre-Sma/Protein.html

Rastogi SC, Mendiratta N and Rastogi P,2005. *Bioinformatics Methods and Applications Genomics, Proteomics and Drug Discovery*, Second Edition, New Delhi, Prentice Hall of lndia private Limited.

Sayle RA and Milner-White EJ, 1995. RASMOL: biomolecular graphics for all. *Trends Biochem Sci*., **20**:374-376.

Seidman L and Mowery J, 2006. *Chapter 2: Protein Structure* [online], Biotechnology project. Available from: http://matcmadison.edu/biotech/resources/proteins/labManual/chapter_2.html

Shirky C, 2000. SevenWays of Looking at a Protein. *Feed Magazin*.

Silverman BD, 2001. Hydrophobic moments of protein structures: spatially profiling the distribution. *Proc Natl Acad Sci U S A,* **98**: 4996-5001.

Walther D, 1997. WebMol – a Java-based PDB viewer. *Trends Biochem Sci*., **22**:274–275.

Wang Y, Geer LY, Chappey C, Kans JA and Bryant SH, 2000. 'Cn3D: sequence and structure views for Entrez', *Trends Biochem. Sci*., **25**:300–302.

How to cite this article

Shaheda N. Ansari, **2018**. A comparative study of different viewing and selection options provided by protein structure visualization tools. *Bioscience Discovery*, **9**(1): 53-58.