TuftsViewer: An Intuitive Protein Viewer

Caitlin Crumm*

Daniela Extrum-Fernandez[†]

Connor Gramazio[‡]

Sara Su[‡]

Lenore Cowen[‡]



Figure 1: TuftsViewer Window showing the residue pane on the left, sequence pane on the bottom, and the menu of options

Abstract

Interactive models of proteins have become popular educational and presentation tools. In this paper we present TuftsViewer, a protein visualization tool that focuses on ease of use and simple, clear rendering styles. We also introduce a new feature in protein visualization: side-by-side multiple alignment mode. We tested our program against a popular protein viewer, Jmol, in a user study. TuftsViewer demonstrated faster rendering changes than Jmol and was comparable for selecting and coloring tasks. We believe TuftsViewer can be an educational tool to introduce students to protein visualization.

Keywords: TuftsViewer, protein visualization, user interface, direct manipulation

1 Introduction

When a user is comfortable with a particular software, it is easier for them to work effectively with that software. In technology experiments, researchers find that users have improved experiences with software that focuses on intuitive social and spatial organization, including software that uses 3D visualization [Horvath and Lombard, 2010]. The newest wave of software research design focuses on developing products with this mindset and users are fascinated by manipulating data using everyday motions and techniques. Though new hardware such as the Microsoft Surface has been created to meet the need for intuitive interfaces [Shaer and et al., 2010], it is more economically feasible to improve software design in current hardware models. We present TuftsViewer, a program with an emphasis on intuitive spatial organization and simplicity. TuftsViewer takes the focus of protein visualization from maximizing scientific functionality to improving a user experience through a more hands-on feel when interacting with proteins. The hands-on feel is achieved using clear visual correlation between a user's action and its effect upon a model, which has shown to be a successful user interface [Saraiya et al.]. Our primary contributions are improved rendering styles, increased spatial manipulation capabilities and simplified user tools. TuftsViewer also supports direct selection and direct manipulation of protein residues. Though many of the basic user interface techniques are modeled after previous software, TuftsViewer uses the new Multiple Alignment with Translations and Twists algorithm [Menke et al., 2008] developed at Tufts University in a unique multiple alignment mode. This feature is an exciting new tool for scientists to interact with multiple aligned proteins while they lie in separate windows.

The TuftsViewer program was tested against a leading protein visualization software (Jmol [York]) in a pilot study with computational biologists and biologists in an attempt to gauge user response. This study took place over 10 weeks and included 17 study participants. Participants were led through a series of specific, directed tasks in Jmol and then TuftsViewer. Comparative analysis was performed on time to complete each task, choice of method to complete each task, a pre-study survey and a post-study survey.

2 Related Work

As visualization technology develops, laboratory data only decipherable by scientists can now be presented in a way that almost anyone can understand: an image or model. Introductory Biology and Biochemistry courses can incorporate complicated protein data into their curricula through visualization tools. Programs such as

^{*}University of Texas at Austin, www.eecs.tufts.edu/ ccrumm0a

[†]Mills College , www.eecs.tufts.edu/ dextru0a

[‡]Tufts University

KiNG, studied by Chen, Davis, and Richardson [2009], and Swiss-PdbViewer, explored by Guex and Pietsch [1997], have been evaluated as effective protein visualization tools. Unfortunately, despite the fact that these programs vastly simplify protein data, the large amount of options available can complicate a user interface by cluttering the workspace or option menus with possibilities for interaction. Hearst asserted in her study that inexperienced or first-time users of a system often can't choose between options when a system is complex [1999]. Indeed, the balance between complexity and usability has often been researched in an attempt to maximize both. Buja et al. concluded that it is more efficient to have multiple views that address different parts of a data set, instead of combining all options and views into a single window [1991]. This technique is known as "brushing and linking".

In the development of a protein visualization tool, there is always the possibility of building upon previous tools or using similar user interfaces. Finstad hypothesized that inexperienced users of a program would perform better with a new piece of software if that software was not modeled after the original program they had little experience with. Finstad's study revealed that this was indeed true and that when the new software had a unique appearance, an inexperienced user performed just as well as an experienced user of the original program [2008]. This study encourages program development from "scratch" rather than modeling currently existing tools in programs designed to be more friendly for novice users.

Finally, Dalgarno, Bennett, and Harper hypothesized that direct and active investigation of a 3D environment would result in greater spatial learning than viewing an equivalent animation of an area. A series of two studies showed that their hypothesis was robust-a user's control over exploration of a 3D model resulted in better spatial knowledge than an equivalent animation [2010]. In terms of educational tools, this indicates that allowing students to have direct manipulation of a model will increase spatial awareness and understanding of objects not normally visible to the naked eye.

3 Experimental Design

3.1 Users

An open email was sent out to recruit users. Potential users needed to be able to identify alpha helices and beta sheets when a protein is in the cartoon drawing style to participate in the study. Our recruited users were primarily Computational Biologists and Biologists with one Computer Scientist and one Administrative worker. There were a total of 17 users, 71% female and 29% male. Each



Figure 2: Jmol Window showing the menu of options



Figure 3: Task performed in TuftsViewer involving coloring protein chains different colors



Figure 4: Task performed in Jmol involving changing the drawing style to spacefill

user signed a consent form in accordance with IRB and was informed of audio and screen recording before the start of the study. After the study, the users received \$20 compensation.

3.2 Apparatus

The studies were conducted on either a 15.4", 1400x900 resolution MacBook Pro running the Windows 7 environment or a 15", 1280x800 resolution Dell laptop running Windows Vista, with no difference in computer setup. We installed TuftsViewer (Figure 1) and Jmol (Figure 2) software on the computers in addition to several pdb files. The only software difference between the computers is that Dell laptop could not operate the multiple alignment feature. Each study was conducted in the computer lab and only one user was evaluated at a time. The investigator giving the user instructions sat next to the user and the other investigators taking notes sat behind the user.

3.3 Survey

The users completed two surveys, pre-study and post-study surveys. The pre-study survey was designed to collect general information about the user's employment status and determine any previous knowledge or usage of protein visualization software. Users also assessed their abilities in sense of direction, spatial reasoning, visualization in 3D and acceptance of new technology.

The post-study survey was designed to gauge the user's experience with TuftsViewer. Users assessed the learnability, simplicity, visibility, user control and freedom, forgiveness of manipulation mistakes, efficiency, and graphic design of the software. Users assessed TuftsViewer's usability on a 7-point Likert scale. The survey also asked specific questions about the usability of the multiple alignment feature. Users could comment on any features they enjoyed or wished to change.



Figure 5: Task performed in TuftsViewer involving multiple alignment and chain selection



Figure 6: Task performed in TuftsViewer invloving selecting, coloring, and focusing on an alpha helix

3.4 Additional Tools

The investigators used the software Camtasia [TechSmith, 2010] to record the audio and computer screen for later analysis and AVSVideoEditor [AVS4You, 2010] was used to analyze the video.

The investigators developed the program PrettyPinkNoteTaker to assist the investigators in noting user actions and comments during the study. The program recorded typed notes and the time the note was created in hours:minutes:seconds. Notes recorded by the program could be matched to the recorded video to determine which directed task the user was performing when the note was recorded.

3.5 Directed Tasks

One of the investigators wrote the script used during the user study. Directed tasks were written such that a task performed in Jmol (Figure 4) was comparable to a task performed in TuftsViewer (Figures 3, 5, 6, and 7) and could be performed in either program. The tasks were designed to test the methods of selection, coloring, hiding of protein residues, and camera manipulation for each program and reflect real tasks performed by a scientist. There were six selecting and coloring tasks in Jmol and nine in TuftsViewer, four rendering (i.e. drawing style) in Jmol and three in TuftsViewer, and one hiding task in Jmol and four in TuftsViewer. While each task was written to be comparable to another, not all tasks were entirely comparable to other tasks i.e. users can select entire alpha helices in TuftsViewer without knowing the residue numbers but must know the residue numbers in Jmol in order to do so.



Figure 7: PrettyPinkNoteTaker for recording user actions

3.6 Protocol

Users were introduced to the investigators and directed to read and sign the consent forms (five minutes). The pre-study survey was completed after the consent forms were signed (five minutes). At the start of the study, the user was given a guided tutorial of Jmol (ten minutes) and the actions they would perform during the directed portion of the study. Users were encouraged to perform the actions presented to them and ask questions. After practicing the actions, the users were directed to perform twelve tasks on five proteins ranging in their complexity (one minute each). The guided tutorial (ten minutes) and directed tasks (eighteen tasks, one minute each, on five proteins and two aligned proteins) were repeated using TuftsViewer with the addition of the multiple alignment feature increasing the number of directed tasks. Undirected observation of protein manipulation in TuftsViewer filled additional time left during the study. At the end of the study, the user completed the post-study survey and was given compensation (ten minutes).

3.7 Analysis of Results

We used Microsoft Excel to compute the average score for each category assessed in the survey.

Directed tasks were categorized by their action performed, protein complexity and other tasks they are comparable to. As previously mentioned, AVSVideoEditor AVS4You [2010] was used to analyze the video. The time for each task to be completed and any user errors were recorded. Also, Google Docs and Microsoft Excel were used to analyze the total time for each task recorded from the videos



Figure 8: Task in TuftsViewer involving selecting a beta sheet and coloring it



Figure 9: Select and Color Comparison

and calculate p values. We compared the average total time for a task performed in Jmol to the average total time for a task of similar model manipulation and protein complexity in TuftsViewer. Data was excluded only if the investigator determined the task to be a mistrial. Mistrials were defined as a task the user did not complete as directed.

4 Results

The comparison between the average times for the selecting and coloring tasks (Figure 8) showed no significant difference between the two programs: neither program demonstrated a faster method of selecting followed by coloring. We also noted that protein complexity appears to have no correlation to the average total time for a task. Tasks on very complex proteins were completed faster than simple or somewhat complex proteins.

The comparison between the average times for the rendering tasks (Figure 9) showed a significant difference between Jmol and TuftsViewer: TuftsViewer was significantly faster for the tasks performed on simple (p value = 0.00562) and complex proteins (p value = 0.00085). There were no tasks in TuftsViewer to compare to tasks performed on the somewhat and very complex proteins in Jmol. Jmol rendering required 15.829 seconds on a simple protein and 28.900 seconds on the second protein and 12.289 seconds on the complex protein. In both cases, Jmol took more than twice as long to complete the task. We did not find any significant difference in the instructions of the tasks that may have influenced the results. The reason for the difference is likely due to the difference in adding and removing drawing styles, which we will discuss later in the paper.

The comparison between the average times for the selecting and hiding tasks (Figure 10) showed a significant difference between Jmol and TuftsViewer: TuftsViewer took longer, 43.735 seconds, than Jmol, 12.399 seconds, for a task performed on a complex protein (p value = 0.00140). We do not believe this result is due to inferior techniques used by TuftsViewer. The task performed in TuftsViewer was not comparable to the task performed in Jmol: the task in TuftsViewer required the user to perform a significantly



Figure 10: Rendering Comparison

higher number of selections than the Jmol task, greatly influencing the total time.

From the post-study survey, users found TuftsViewer to be easy to learn, simple to use, and visually pleasing (Table 1). Only error handling scored an average below 6 on the Likert-scale: users may not have been aware of or misrecollected efficient correcting methods presented in the guided tutorial. Many users found the ability to select an entire alpha helix or beta sheet to be extremely useful but suggested adding the ability to select all alpha helices or beta sheets. Also, some users, Computational Biologists and Biologists, grew comfortable with the console tool in Jmol and advocated for a console in TuftsViewer. Finally, many users grew fatigued of having to repeatedly right click to bring up the menu of actions. A tool bar or window of actions was suggested to speed up the completion of tasks.

5 Discussion

Users selected and colored in comparable time in TuftsViewer and Jmol. This was a positive result, indicating that first-time users were not overwhelmed by multiple windows resulting from the use of brushing and linking in TuftsViewer. This was also an interesting result because the method of selecting and coloring during the user study in TuftsViewer varied widely, with multiple techniques and utilization of various display windows. Selecting and coloring in Jmol tending to be limited to the console, with few users choosing to attempt either action with the right click menu. This could be because of frustrations related to the mouse precision required to navigated the multiple submenus to perform a task.

TuftsViewer was significantly faster for tasks involving changing the drawing style. This significant difference is assumed to be related to the fact that TuftsViewer applies any changes to the whole model when nothing is selected. Users in the study were often inefficient in changing the drawing style in Jmol because they had trouble mastering the concept of keeping a tally in their mind of what is currently selected (selection isnt by default visible on the model in Jmol). This resulted in users frequently changing their last selec-



Figure 11: Hiding Comparison

tion to a new drawing style, instead of the entire protein. It is also possible that the simplicity of drawing style modes in TuftsViewer (it is only possible to turn them on or off) resulted in this significant difference. Certain modes of drawing in Jmol must be layered to be turned on with the console (for example, ball and stick requires layering of wireframe and spacefill). Because users tended to prefer the console, it took extra time for many of them to learn to associate layered drawing styles with certain common names like ball and stick.

Hiding residues in TuftsViewer took significantly longer than Jmol. This could be because the specific task that defined hiding in TuftsViewer sometimes was completed using a large number of selections. Selection could have been performed using a single tripleclick, but this functionality was not mentioned in the user instructions and few users discovered it on their own. Performing the similar task in Jmol did not require a large number of selections. Several users didnt even complete the task successfully in TuftsViewer, indicating that it is possible that the directions for the task werent clear enough, or that the task required some sort of higher spatial intelligence. However, users that had difficulty with the task or didnt complete the task didnt rate themselves as below average in terms of spatial reasoning on the prescreening survey.

The survey results were positive, with only error handling rated below a 6/7. We hypothesize that this was because many users werent aware of the Ctrl-Z undo functionality of TuftsViewer, as it was not mentioned in the user instructions. The user comments on the survey indicated that some users actually liked the console in Jmol enough to want an equivalent in TuftsViewer. This idea has been taken under consideration for an advanced mode for expert users but will most likely not be implemented. Another frequent comment was that users found interaction with proteins more intuitive in TuftsViewer than Jmol. Several of the users had some free interaction time with TuftsViewer after their study, depending on how much time was left. These users were encouraged to open any fa-

 Table 1: Post-Study Survey Results

	Average Rating
Learnability	6.1
Simplicity	6.0
Visibilty	6.5
User Control & Freedom	6.2
Error Handling	5.7
Efficiency	6.0
Graphic Design	6.4

vorite proteins or choose from a list of pdb files and experiment with TuftsViewer features.

Some thought was put into whether or not a learning effect occurred in the user studies. More specifically, we considered that users might improve performance in TuftsViewer because they become more comfortable with protein interaction and interface commands while working in Jmol. It was assumed that any learning effect would be negligible due to the significant differences in user interface in each program, but alternative ideas are discussed in the conclusion that might counteract a learning effect in future studies.

6 Conclusion and Future Work

From our results, we conclude that TuftsViewer has great potential as an educational protein visualization tool. We hope that future studies will further shape TuftsViewer development to make it more efficient and easier to use. In future studies, to counteract any learning effect, we suggest that each user work with only one program- Jmol or TuftsViewer. This would remove any possibility of a learning effect. Alternatively, the researchers could alternate the order of the study, having a user work with Jmol first in one trial and a different user work with TuftsViewer in another. We suggest that future studies might investigate new features, like automatic rotation of the proteins, bond length measurement, and a protein-protein interaction mode that shows possible bonding and reactions between two protein structures.



Figure 12: Prototype of TuftsViewer Tool Bar

Since the constant action of bringing up the menu by right clicking fatigued some of the users, we have begun prototyping a tool bar to include in TuftsViewer that would allow users to change the drawing style, color of the protein, or selected portion without bringing up the menu. At the time of this writing, the prototype tool bar feature uses screen shots to simulate a change in drawing style or color of the entire protein. We hope the prototype will be used in additional user studies to assess the usefulness of adding a tool bar to TuftsViewer.

Acknowledgements

Thanks to Matt Menke for all of his work developing TuftsViewer. Thanks to Lenore and Sara for their guidance and patience during this project. Thanks to DREU for sponsoring Caitlin and Daniela and Kathryn McKinley, Stuart Reichler, Barbara Li Santi, and Susan Wang for recommending Caitlin and Daniela for the DREU program.

References

- AVS4You. Avs video editor. http://www.avs4you.com/AVS-Video-Editor.aspx, 2010.
- Vincent B. Chen, Ian W. Davis, and David C. Richardson. King (kinemage, next generation): A versatile interactive molecular and scientific visualization program. *Protein Science*, 18(11): 2403–2409, 2009.
- Barney Dalgarno, Sue Bennett, and Barry Harper. The importance of active exploration, optical flow, and task alignment for spatial learning in desktop 3d environments. *Human-Computer Interaction*, 25(1):25–66, 2010.
- Andreas Buja et al. Interactive data visualization using focusing and linking, 1991.
- Kraig Finstad. Analogical problem solving in casual and experienced users: When interface consistency leads to inappropriate transfer. *Human Computer Interaction*, 24(3):381–405, 2008.
- Nicolas Guex and Manuel Peitsch. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis*, 18(15):2714–2723, 1997.
- Marti Hearst. Modern Information Retrieval (Chapter 10: 'User Interfaces and Visualization'). Addison Wesley Longman Publishing Co. Inc., 1999.
- Karl Horvath and Matthew Lombard. Social and spatial presence: An application to optimize human-computer interaction. *Psych-Nology Journal*, 18(1):85–114, 2010.
- Matt Menke, Bonnie Berger, and Lenore Cowen. Matt: Local flexibility aids protein multiple structure alignment. *PLoS Computational Biology*, 4(1), 2008.
- Matt Menke, Sara L. Su, Connor Gramazio, Caitlin Crumm, Daniela Extrum-Fernandez, and Lenore Cowen. Tuftsviewer: An intuitive interface for viewing 3d protein structures. *In Proceedings of 3DSIG 2010: The 6th Structural Bioinformatics and Computational Biophysics Meeting*, 2010.
- Purvi Saraiya, Chris North, and Karen Duca. An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4), July .
- Orit Shaer and et al. G-nome surfer: a tabletop interface for collaborative exploration of genomic data, 2010.
- TechSmith. Camtasia studio. http://www.techsmith.com/camtasia.asp, 2010.
- Will York. Jmol: an open-source java viewer for chemical structures in 3d. http://jmol.sourceforge.net/.

Appendix A: User Study Script

Index

- Overview and Explanation
- Preparation, Consent Form, and Presurvey
- Introduction to JMol
- Introduction to Tufts Viewer
- JMol Directed Tasks
- TuftsViewer Directed Tasks
- TuftsViewer Multiple Alignment Directed Tasks
- Undirected Tasks
- Postsurvey, Debriefing, etc.

Overview and Explanation (For use of experimenter, and not research subject)

Tasks should adequately reflect real tasks used on simple proteins. These include model rotation and other camera manipulations, selection, coloring, and changing draw styles. The times listed next to each heading are approximate. Each protein should take about 3 minutes to complete, and the directed task portion of the experiment should fit under an hour. Left over time should be filled with undirected task observation.

Proteins in Part 2 that are prefaced with "For this protein do not move the camera unless prompted." are testing how a user selects residues that cannot be seen if moving the camera is not an option. The prohibition on camera movement is also used when the protein has a mouse-geared task. Because these tasks use physical description instead of residue number, the on-load image must be preserved. There are some tasks that test these two procedures near the end of a protein, so instead of using the "For this protein..." the user is told to recenter and readjust the camera before making a Mouse-Geared, or occlusion selection.

The instruction sections for TuftsViewer and JMol should be interactive -- when you are reading the instructions encourage the participant to follow along and carry out the program feature that is being described. Questions are encouraged, as long as the time limit for instructions is kept. Better they ask questions in the instruction phase before we start annotating and recording data.

Preparation, Consent Form, & Presurvey (5 minutes)

(Reset program if needed) (Refresh/Load the PDB search website) (Reset/Ready the monitoring software) (Give participant Release form)

Preface

For this experiment I will be recording audio and the screen. If you would prefer not to be recorded, accommodations can be made. Additionally, if at any point in the experiment you feel that you would like to discontinue the experiment you may do so. Revise!

Part 1: Practical Testing

Identify an alpha helix and a beta sheet (have a sheet with alpha helix and beta sheets in color that they can identify)

Part 2: JMol Experience

If you have had experience using JMol or other protein visualization software, please fill out the following questions to the best of your ability. [Give MolVis-PrevExp to the subject]

- 1. How long have you been using protein visualization software?
- 2. How frequently do you use protein visualization software?
- 3. What do you principally use protein visualization software for?
- 4. Are there any tasks that you have difficulty doing with current protein visualization software?
- 5. Do you have a favorite protein that you would like to look at?

[User to fill out pre-study survey]

Introduction to Jmol (10 minutes)

Disclaimer

If you have prior experience using JMol, please only use the following manipulation techniques unless otherwise noted.

Introduction

[Note: Adapted from 1KSA tutorial: http://bit.ly/wileytutorial] [Launch Jmol; Load 1KSA]

This is the Jmol program, and the structure we are viewing is a bacterial chlorophyll A protein. It is composed of multiple parts; there are two polypeptide chains, displayed as flat ribbons, each in a different color. Each chain also has a non-protein component, displayed in a stick model. The display shows the folding patterns of the polypeptide chains, as well as the bonds of the pigment. There are three ways to interact with structures: using the mouse, accessing the menu, or using the console.

The Mouse

To rotate the molecule, left-click and drag on it.

To translate, shift-double-click and drag on the structure--the molecule will follow the mouse. To zoom, shift-click and drag. Drag the cursor down to zoom in, up to zoom out.

To reset the camera on the main menu bar go to View -> Front.

The Menu

Let's see how to can change the display to view different aspects of the model. Access the menu by right-clicking anywhere. Choose the "Select" submenu, then click on "All". This selects all the atoms in the structure but the display will not change yet.

Access the right-click menu again. Choose "Color | Atoms | By Scheme | Secondary Structure". The atoms are now colored so that magenta-colored atoms are part of alpha helices, and yellow-colored atoms are part of beta sheets.

The Console

Jmol can also be controlled using a command language. Use the Jmol menu, choose "File | Output Console", or right click on the display frame (the giant black portion) and select "Console". Type "select all", and press the "enter" on the keyboard. Jmol writes a few things into the Console window after executing the command, including the number of atoms selected.

Now type: "color purple", and press "enter" again. Finally, type "cartoon", and press enter.

To select a residue type "select" and then the number of a residue.

To select a string of residues type something like "select [the beginning residue number] [dash] [the ending residue number]".

To select a residue on a certain chain type "select #:chain letter".

To select an entire chain type "select :[chain letter]".

To select multiple sections type "select" and then your first selection, a comma, and then your second selection. You may select individual or multiple residues, and chains this way.

To restrict the view to a residue type "restrict #-#". To restrict the view to only a given chain type "restrict :[Chain Letter]". You may also restrict the view to several different sections in the same way as you made multiple selections.

Drawing Styles

The three drawing styles, ways to view the protein model, we will be using are wire frame and spacefill (these you have already seen), and also cartoon, which displays the secondary structure of the protein. JMol presents two ways to turn drawing styles on or off -- through the console and

through the menu. To turn on a style with the mouse, first make a selection, then right click and go to "Style | Scheme". Here you can turn cartoon, ball and stick, and spacefill on or off.

To change the way the protein is displayed (drawn) using the console use the following commands:

- For cartoon type "cartoon". Now type "cartoon [space] off" to turn that view off.

- For spacefill type "spacefill" (one word). Now type "spacefill [space] off" to turn space fill off.

- For ball and stick type "wireframe" (one word) and then "spacefill [space] 100" (one word). To turn wireframe off

type "wireframe [space] off" and then "spacefill [space] off".

To turn drawing styles off with the mouse, just navigate back to "Style | Scheme" and click on the drawing style you want to turn off.

JMol Practice Run

Now I'd like you try some more complex interactions with the structure. You may use any of the visualization controls or the dropdown - direct manipulation of the structure may be required. [Load 2cx6]

First, please change the view to cartoon.Now restrict the view to show only chain A.Color chain A dark green.Refocus the camera so chain A is centered and fills the entire frame.

Questions

Before we start the JMol tasks do you have any questions?

Part 1: Directed Tasks JMol (3 minutes for each)

In this part of the study, I will ask you to make some edits to existing 3D models using JMol. I'd like you to talk me through what you are doing as you work.

Protein 1, simple: Please open protein 1r4y.

Task 1 (Selection, Coloring): Color residues 13 through 23 yellow.

Task 2 (Rendering): Change the view to spacefill. Protein 2, somewhat complex: Please open protein 1hle.

Task 1 (Rendering): Have Cartoon overlay in addition to Ball and Stick

Task 2 (Selection, Coloring): Color residues 260 through 266 and 268 through 279 red.

Protein 3, complex: Please open protein 1by7.

Task 1 (Selection, Coloring): Color residue 123 through 132 and 196 through 204 green.

Task 2 (Rendering): Change the view to Cartoon only

Protein 4, complex: For this protein please do not move the camera unless instructed to. Please open protein 20ez.

Task 1 (coloring, rotation): Color residues 108 - 125 on Chain B green. Rotate the protein model to bring the selected portion into view

Task 2 (Selection, Hiding, View Manipulation): Restrict the view to protein chain A. Recenter the camera

Task 3 (practical application of coloring each chain, and also uses selection by chain): Change the color of chain A to orange, and chain B to purple

Protein 5, very complex (3 chains): For this protein do not move the camera unless prompted. Please open protein 1a4e.

Task 1 (Selection and Color): Color chains C and D orange. Task 2 (Selection of Occluded Objects and View Manipulation): Please color residues 217 - 235 on Chain C red. Rotate the camera to bring the selected residues into view.

Task 3 (Camera and View Manipulation): Reset camera. Change view to spacefill. Manipulate camera to get fuller cross section of model.

Introduction to Tufts Viewer (10 minutes)

[Launch Tufts Viewer] [Open 1xyx]

This is another visualization program for viewing molecular structures in 3D. The structure we are viewing is a Mouse prion protein fragment. Instead of a wireframe, TuftsViewer's default drawing style is "cartoon".

As with Jmol, you can interact with structures in Tufts Viewer using the mouse or the menu. There is no console.

The Menu and What You See

Access the display menu either by right-clicking or by selecting "Display" the menu bar at the top of the window. Here you can select Cartoon, Ball and Stick, and Space Fill display methods. The "Tools" section in the menu bar allows you to create a multiple view, and also gives you access to settings where you may change the default colors of the program. Windows allows you to toggle the residue and sequence windows on and off.

The three sections of the program are the display window, the sequence pane, and the residue pane.

The display pane is where the protein is visualized.

The sequence pane is the pane on the bottom of the window. Here multiple chains can be displayed.

The residue pane is on the right hand side of the window. Only one chain is displayed at at time here. The chain may be changed by using the drop down menu at the top of the pane.

The Mouse

te the molecule, left-click and drag on it.

To zoom in and out, use the scroll wheel of the mouse. Alternatively, you may also hold both left and right mouse buttons down and move the mouse forward or backwards to zoom. To pan hold the right mouse button and drag. To reset the camera position, right-click on the display pane and select Reset Camera. You can also reset zoom, and center the camera on the right-click menu.

You may also use the mouse to gain more information about any given residue. In the bottom pain mouse over any residue letter. Note that information is displayed in the bottom bar of the program, including residue number. Residue numbers are also labeled on the right-hand column.

Selecting

There are three different ways to select one protein residue.

To select using the mouse, move the cursor over the residue you want to select, and left click on the model . You may also select using a bounding box by holding shift and then left clicking and dragging.

To select using the sequence pane click on a residue's letter. To select multiple residues at one time you may click and drag.

To select using the residue pane click on the wanted residue. To make a selection of contiguous residues, click on the first entry, then hold shift and then click on the last residue wanted. To select disjoint residues, press control. Note that, multiple sections can be selected by holding ctrl + shift.

Selections made by any of these methods update all panes. For example, a selection made in the display pane will also show highlighted residues in the sequence and residue panes.

Residue number is shown next to the residue identifier in the residue pane. For the display and sequence panes when you mouse over a residue (either on the model, or over a letter) numbers are displayed on the bottom bar of the protein. Residue number is the second number from the right.

To select a beta sheet or alpha helix, double click on a residue letter in the sequence pane, or on the portion in the model.

To select an entire chain in the display pane is by triple clicking the chain. To select using the sequence chain, left-click on the chain identifier (for example 1XYX: A).

To deselect click on any portion of the black space in the protein model frame, or an empty part of the sequence pane. Note that deselection clears all selections made.

You may also draw attention to certain areas by coloring protein residues, or restricting residues are seen in the model. To accomplish these, first a portion of the protein must be selected. To recolor, right click and select Pick Color....

To restrict the view, right click and select Hide All But Selected.

To uncolor all, right click and select "Color by Structure", or another coloring preset. To unhide all, make sure you have nothing actively selected and under Display or the right-click menu select a model display type (Cartoon, etc.).

To show color in spacefill mode, make sure that on either the right click menu or the Display menu "Use Atom Colors" is deselected. Once that is done previously colored residues will be shown, and you also have the option to color by structure or alignment.

Multiple Alignment

You also have the ability to see how proteins are aligned with the proteins overlapping, side-byside, or in separate windows. Multiple Alignment can be accessed through the main menu bar in Tools -> Alignment. When activated you will be prompted to load proteins from files, select which chains to add to the alignment, and how the different proteins and chains are laid out. Shortly you will get to see how this appears in the program.

Training Task

I'd like you to spend a few minutes playing around with it to get a feeling for how it works.

Please open protein 1yvs.

Please try zooming in and out, panning, and changing the camera angle. Using shift, click, and drag, select the alpha helix outlier. Using the bottom pane select the beta sheets. Using shift+click drag, Select the rest of the protein. Color the protein yellow. Change the view to Spacefill.

Multiple Alignment (Brief) Intro:

Now I'd like for you to get a brief introduction to the multiple alignment features.

Please open Multiple Alignment. It is found under Tools in the menu bar. Please open 1gou, and click add. Make sure "Use multiple frames in one window" is selected and then click align.

Try rotating one of the protein chains. Select a pair of aligned alpha helices using the sequence pane. (If the user needs it: aligned = overlapping alpha helices) Change the view to space fill.

Questions

Before we start the JMol tasks do you have any questions?

Part 2: Directed Tasks TuftsViewer (3 minutes for each)

In this part of the study, I will ask you to make some edits to existing 3D models. I'd like you to get enough experience with it so that you can comment if you could see yourself using something like this for your own work. As with JMol, I'd like you to talk me through what you are doing as you work.

[Compare different manipulation techniques]

Protein 1, simple: Please open protein 1e0g.

Task 1 (Selection, Coloring): Color residues 4 - 7 brown.

Task 2 (Rendering): Change the view to space fill.

Task 3 (View Manipulation): Restrict the view to residues 4-7 and 41-44.

Protein 2, somewhat complex: For this protein do not move the camera unless prompted. Please open protein 1pqy.

Task 1 (Selection by Mouse, Coloring): Color the most upper left alpha helix orange. (Residues intentionally omitted, but 110 - 118) Zoom in on the selected alpha helix.

Task 2 (View Manipulation, Camera Manipulation, Rendering): Recenter camera and reset camera position. Restrict the view to the bottom half of the protein. Change the view to display only Ball and Stick

Protein 3, complex: Please open protein 1c8z.

Task 1 (Selection by Mouse, Color): Color all beta sheets that do not exist within the central "ring" brown.

Task 2 (Occlusion selection and occlusion coloring, View Manipulation): Color residues 316 through 321 pink. Rotate the camera to bring residue into view.

Task 3 (Selection by Mouse, Hiding): Restrict the view to only the beta sheet ring.

Protein 4, very complex (2 chains): Please open protein 1p1u.

Task 1: (Mouse-Geared Selection, Coloring): Color the right-most alpha helix on purple. (For experimenter's reference: residues 22-31:B)

Task 2: (practical application of coloring each chain, and also uses selection by chain) Change the color of chain A to orange, and chain B to purple

Task 3: (Selection, Hiding, View Manipulation): Restrict the view to chain A.

Task 4: (camera manipulation, occlusion selection) FIX SO THAT SELECTION IS OCCLUDED Reset the camera position, and recenter the camera. Color residues 72-80 on chain A light blue, and then rotate the protein model to bring the selected portion into view

Protein 5, very complex (3 chains): For this protein do not move the camera unless prompted. Please open protein 2nud.

Task 1 (Selection of Occluded Objects and View Manipulation): Please select residues 145 - 163 on chain B. Rotate the camera to bring the selected residues into view.

Task 2 (Selection and Color): Color chains C and D bright green.

Task 3 (Camera, Rendering): Reset camera. Change view to spacefill. Manipulate camera to get fuller cross section of model.

Multiple Alignment: Please open in protein 2q4m and 1i7e in multiview (both are in the super family Tubby C- Terminal domain-like)

Task 1 (Selection, Coloring): Please color a pair of aligned alpha helices orange.

Task 2 (Selection, View): Hide the outlying group of beta sheets in 1i7e.

Task 3 (Rendering): Change the view for both proteins to spacefill. Color by structure, not by atoms.

Part 3: Undirected Session (10 minutes)

Now I would like you to spend 10 minutes manipulating this model as you like. There is no specific task I want you to accomplish. Just spend some time exploring the features of the software.

Nice proteins to use if user doesn't have a favorite: 2nud, 1cz8, 1a4e

Post-Survey

Now that you've had a chance to use our prototype software, I'd like to ask you few questions to wrap up.

[User to complete post-study survey]

Those are all the questions I have. Do you have any questions for me? [Thank participant and give gratuity]

After Post-Survey

[Ask informally if there is anything that is not currently implemented that you would like to see.] [As time moves onward perhaps do pen and paper models to test prototypes others have suggested] Appendix B: Informed Consent



OFFICE OF THE VICE PROVOST

Social, Behavioral, and Educational Research Institutional Review Board

FWA00002063

Re: IRB Study # 0912011 Title: Evaluation of Molecular Visualization Software PI: Lenore Cowen Co-Investigator(s): Sara Su, Robert Jacob IRB Review Date: 12/22/2009

January 6, 2010

Dear Lenore,

The Institutional Review Board (IRB) has reviewed the revisions received on 1/5/2010 for the above referenced study.

This protocol now meets the requirements set forth by the IRB and is hereby approved. Approval is valid from the date of this letter for a period of one year from the original IRB Review Date and expires on 12/21/2010.

Enclosed you will find stamped consent forms and other study materials that show the date through which these materials are valid. Only copies of these stamped consent forms and materials may be utilized for conducting your study.

This protocol has been approved for a total of 100 participants for the duration of your study.

Any changes to the protocol, consent forms or study materials must be submitted to the Office of the IRB for approval by completing the *Request for Protocol Modification* form. In addition, all Adverse Events and Unanticipated Problems must be reported to the Office of the IRB promptly, and by utilizing the appropriate reporting forms.

Investigators are required to submit a *Request for Continuing Review* or a *Request for Study Closure* six weeks prior to the expiration date of the protocol.

Please know that the PI is responsible for all information contained in both this letter and on the Investigator Responsibilities Sheet. If anything is unclear or if you have any questions, please contact the IRB office at (617) 627-3417.

Sincerely,

Yvonné Wakeford, Ph.D. IRB Administrator

IN THE KNOW

BULLETIN OF THE TUFTS SBER IRB MEDFORD CAMPUS

UPDATED IRB FORMS

All the IRB forms were updated with a view to streamlining the approval process:

- For Researchers forms are clearer and easier to complete.
- For Reviewers information is better organized and more comprehensive.

The updated forms can be found on our website: http://www.tufts.edu/central/research/IRB/Forms.htm

NOTE THAT AS OF SEPT 1, 2009 WE WILL ONLY ACCEPT THE REVISED FORMS

Please refrain from using previously saved templates of outdated forms. Please go to the website *EVERY* time you submit to ensure that you always use only the latest version of the forms.

The revisions include changes to the following sections:

- Participant enrollment information
 - The IRB recommends that researchers maintain an ongoing log of participants in each study to ensure enrollment does not exceed the number approved by the IRB
- Funding information
- Recruitment information
- Cultural details for international research

Please know that feedback and suggestions are always welcome. Please email <u>sber@tufts.edu</u>

TUFTS UNIVERSITY DEPARTMENT OF COMPUTER SCIENCE

CONSENT TO PARTICIPATE IN RESEARCH STUDY

Evaluation of Molecular Visualization Software Title of Protocol:

Principal Investigator: Lenore Cowen, Professor of Computer Science Sara Su, Visiting Assistant Professor of Computer Science Co-Investigator:

You are asked to participate in a research study conducted by Lenore Cowen and Sara Su from the Department of Computer Science at Tufts University. You were selected as a possible participant in this study because of your familiarity with computer software. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

PARTICIPATION AND WITHDRAWAL

Your participation in this study is completely voluntary. Should you decide at any time during the study that you no longer wish to participate, you may withdraw your consent and discontinue your participation without penalty or loss of benefits. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

PURPOSE OF THE STUDY

The proposed study will assess the effectiveness of software tools we have developed for molecular visualization. This work is applicable to biology, computer graphics, and user interface research. Our goal is to enhance the user experience for both experienced biologists and non-experts. You will be asked to perform a series of directed and undirected tasks using our software. Studying subjects' facility with these tasks will help us evaluate our methods against existing software and improve our design.

PROCEDURES

If you volunteer to participate in this study, we would ask you to do the following things: You will first be given a guided tutorial to using our software and an explanation of the study procedures. For the main computer-based session, you will be asked to perform a series of tasks related to manipulation of 3D molecular models. Directed tasks may include selecting a particular part of the model (e.g. changing the color of one protein residue) or performing a sequence of manipulations of the model. The undirected session will involve freely exploring the software. The directed and undirected tasks together will last approximately 45 minutes. This will be followed by a brief interview to record your subjective impressions of the software and tasks. You may choose to stop the session at any time. The complete test, including informed consent, should taken approximately 60 minutes.

APPROVED

EXPIRES

JAN 06 2010 Tufts SBER IRB

DEC 21 2010 Tutts SBER IRB

POTENTIAL RISKS AND DISCOMFORTS

You will be viewing a computer screen for up to 60 minutes. If you have a prior history of eye strain or wrist injury, you may wish to reconsider your participation in the study. There are no other foreseeable risks or discomforts associated with this study.

POTENTIAL BENEFITS

It is not expected that you will receive any direct benefits as a result of your participation in this study. This research will be used to evaluate the relative performance of our visualization software against the state of the art. The results of the study will help to contribute to the body of knowledge on biology, computer graphics, and user interfaces.

COMPENSATION

As compensation for the time and trouble associated with participating in our study, you will receive a gratuity of \$25. The investigator will explain how to collect your payment.

CONFIDENTIALITY

The results of this study may be published in a scholarly book or journal, presented at a professional conference, or used for teaching purposes. However, your name and other identifiers will not be used in any publication or teaching materials. Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission. The only identifiable information that will be included in this study are the participants' names and e-mail addresses. This information will be stored electronically and will be accessible only to the researchers who are directly involved in administering the study. Data will be electronically archived following the study. If other researchers use the data in future projects, personal identifiable information will be excluded.

REQUEST FOR MORE INFORMATION

If you have any questions or concerns, please contact one of the following investigators:

Lenore Cowen, Principal Investigator Daytime phone: 617-627-5134 E-mail address: cowen@cs.tufts.edu Sara Su, Co-Investigator Daytime phone: 617-627-2487 E-mail address: sarasu@cs.tufts.edu

In addition, you may contact Yvonne Wakeford at the Office of the Institutional Review Board at (617) 627-3417.

APPROVED

JAN 0 6 2010 Tufts SBER IRB EXPIRES

DEC 21 2010 Tutts SBER IRB

SIGNATURE OF RESEARCH SUBJECT OR LEGAL REPRESENTATIVE

I confirm that I understand the procedures described above and the purpose of the research. My questions have been answered to my satisfaction. I understand that I may ask additional questions at any time and can withdraw my participation without prejudice. I have read this consent form. My signature below indicates my willingness to participate in this study.

Printed Name of Subject

Signature of Subject

Date

SIGNATURE OF INVESTIGATOR

In my judgment the subject is voluntarily and knowingly giving informed consent and possesses the legal capacity to give informed consent to participate in this research study.

Printed Name of Investigator

Signature of Investigator

Date

APPROVED

JAN 0 6 2010 Tufts SBER IRB EXPIRES

DEC 21 2010 Tufts SBER IRB

Molecular Visualization User Study: Sample Script

Participant:

Date/Time:

Background

I am a researcher in the Tufts Department of Computer Science. Having people like you try out our software gives us important feedback to help us improve it. In the study, I'm going to ask you some questions about your familiarity with visualization software. Then I'm going to ask you to manipulate some 3D models of molecules in prototype software that the team has been working on. Finally, I'll ask you some wrap up questions to get your thoughts on the software. The total time will be about 1 hour.

[Get informed consent form]

This is not a test of you; it's a test of the software. Your honest feedback is what we are looking for. Before we get started, I'd like to ask you some questions to learn a little bit about you and what you do.

- 1. What is your age?
- 2. What is your occupation?
- 3. [If biologist]
 - a. What kinds of biological structures do you work with?
 - b. [Ask about typical workflow]
 - c. Have you used any molecular visualization software in your work?
 - d. Have you experienced any challenges or problems using [tool]? Can you give me a recent example?

Tutorial

Now we're going to start the hands-on portion of the study.

[Launch program]

This is a visualization program we have developed for viewing molecular structures in 3D. I'm going to demonstrate a few features so you can see how this works.

[Demo UI]

Now I'd like you to spend a few minutes playing around with the software to get a feeling for how it works.

[Allow subject to use program for about 3 minutes]

Now I'd like to show you a particular function for highlighting or coloring a particular piece of the model.

[Demonstrate how to select and re-color one protein residue on the model in file [XXX].]

I'd like you to spend a few minutes playing around with it to get a feeling for how it works.

APPROVED JAN 0 6 2010 Tufts SBER IRB

DEC 21 2010 Turts SBER IRB

EXPIRES

Part 1: Directed Tasks

In this part of the study, I will ask you to make some edits to existing 3D models. I'd like you to get enough experience with it so that you can comment as to whether you would see yourself using something like this for your own work. I'd like you to talk me through what you are doing as you work.

[Compare different manipulation techniques]

Select residues [##] to [###].

Color these residues purple.

Restrict the view to residues [##] to [###] and center on them.

Select the alpha helix corresponding to sequence [###].

Now, open the pair of models named [XX] and [YY]. Color only those portions that have been aligned.

Color those residues of protein [XX] that are touching those of protein [YY].

Part 2: Undirected Session

Now I would like you to spend 10 minutes manipulating this model as you like. There is no specific task I want you to accomplish. Just spend some time exploring the features of the software.

Exit Interview

Now that you've had a chance to use our prototype software, I'd like to ask you few questions to wrap up.

1. On a 7-point scale, how would you rate your experience with the software in the following areas?

Learnability:

Simplicity:

Visibility:

User control & freedom:

Error handling:

Efficiency:

Graphic design:

2. The next time you are interacting with a 3D model, is this something that you would use?

- 3. Would you recommend the [XXX] feature to your colleagues?
- 4. What would you change about the way the [XXX] feature works?
- 5. How would you compare this software to other programs you have used?

Those are all the questions I have. Do you have any questions for me?

[Thank participant and give gratuity]

APPROVED

JAN 06 2010 Tufts SBER IRB

XPIRES

DEC 2 1 2010

Subject Recruitment Letter for Tufts Social, Behavioral & Educational IRB

Title of Protocol:	Evaluation of Molecular Visualization Software
Principal Investigator: Co-Investigator:	Lenore Cowen, Professor of Computer Science Sara Su, Visiting Assistant Professor of Computer Science
Date:	December 2, 2009 (revised December 29, 2009)

Study participants will be recruited using a letter of the form below, which will be posted on the web and sent to email distribution lists.

Subject: Subjects Needed for Behavioral Study, Feb 23-26

Participants are needed for a behavioral study at Tufts University next week. The available slots are:

Mon 2/23: 11:00am, 1:30pm, 3:00pm Tue 2/24: 1:30am, 3:00pm, 4:30pm Wed 2/25: 11:00am, 1:30pm, 3:00pm Thu 2/26: 10:00am, 11:30am, 1:00pm

JAN 06 2010

APPROVED

Tufts SBER IRB

DEC **2 1 2010**

SBER IRB

You must

- * have normal or corrected-to-normal vision
- * be at least 18 years old
- * be comfortable using software in the Windows operating system

The experiment involves viewing and manipulating 3D models of molecular structures in our prototype visualization software. It takes about an hour to complete, and you will receive \$25 compensation for your time.

If you believe you meet the requirements, please reply to sarasu@cs.tufts.edu with the time slots you prefer.

We often get many replies to these messages, so if we don't write back, it means we have already filled all available slots. Thank you!

We plan to recruit from two populations: 1) general computer users and 2) biologists who have prior experience with molecular visualization software. The letter above targets the former. To recruit biologists, we will amend the letter with the following requirement:

You must be a professional in the biological sciences who has some prior experience using molecular visualization software. You do not need to be an expert user of this software.

Appendix C: Data and Graphs

			NOT ENOUGH	
GUIDE:	NOT SIGNIFICANT	SIGNIFICANT	DATA	
		somewhat		
	simple	complex	complex	very complex
			Jmol P3:T1, P4:T1,	
Select/Color	Jmol P1:T1	Jmol P2:T2	P4:T3	Jmol P5:T1
				Tufts P4:T1,
			Tufts P3:T1 P3:T2,	P4:T2, P5:T1,
	Tufts P1:T1	Tufts P2:T1	MA:T1	P5:T2
Pvalue	0.226117497	0.731503886	0.618081085	0.38632683
Render	Jmol P1:T2	Jmol P2:T1	Jmol P3:T2	Jmol P5:T3
	Tufts P1:T2		Tufts MA:T3	
Pvalue	0.005622433		0.000851364	
Hiding			Jmol P4:T2	
	Tufts P1:T3		Tufts P3:T3, MA:T2	Tufts P4:T3
Pvalue			0.001400307	

Table 1: Task type comparison between programs with p-values

	Avg Time		Avg Time
Jmol		TuftsViewer	
P1:T1	18.844	P1:T1	15.166
P1:T2	15.829	P1:T2	4.849
P2:T1	15.893	P1:T3	30.464
P2:T2	23.018	P2:T1	24.030
P3:T1	17.921	P2:T2	32.358
P3:T2	28.900	P3:T1	32.443
P4:T1	23.554	P3:T2	25.640
P4:T2	12.399	P3:T3	43.735
P4:T3	34.811	P4:T1	13.204
P5:T1	16.422	P4:T2	19.223
P5:T2	20.892	P4:T3	9.358
P5:T3	21.626	P4:T4	32.885
		P5:T1	26.050
		P5:T2	14.830
		P5:T3	13.336
		MA:T1	21.831
		MA:T2	26.651
		MA:T3	12.289

 Table 2: Average times for every task, both programs



Graph 1: Time vs. Task for Jmol



Graph 2: Time vs. Task for TuftsViewer