Content-guided Navigation in Multimeric Molecular Complexes

Mikael Trellet[1], Nicolas Ferey[1], Marc Baaden[2] and Patrick Bourdot[1]
[1]*VENISE group, LIMSI (CNRS), University Paris-Sud, Rue John Von Neumann, Orsay, France*
[2]*LBT-IBPC (CNRS), University Paris Diderot, Paris, France*

Keywords: Molecular Visualization, Constrained Navigation, Structural Biology.

Abstract: In the field of structural biology, molecular visualization is a critical step to study and understand 3D structures generated by experimental and theoretical technics. Numerous programs are dedicated to the exploration and analysis of structures in 3 dimensions. However, very few of them offer navigation algorithms that deal in an intelligent way with the content they display and the task to perform. This observation is emphasized in the case of navigation in immersive environments where users are immersed in their molecular systems, without any spatial landmark to guide their exploration. Starting from this observation, we propose to take into account some geometrical features found in multimeric molecular complexes to provide navigation guides to the users during the exploration process. It is possible thanks to the common symmetrical layout molecular complexes present. Beyond the biological meaning and importance that symmetric layouts have in proteins, they allow us to orient and guide the exploration of molecular complexes in an intelligent and meaningful manner. We present then a current work on the design of navigation paradigms based on the content and the task for the molecular visualization.

1 INTRODUCTION

Structural biology can be split into several indispensable steps that lead to the deciphering of molecular complexes. One of these steps is the visualization and exploration of molecular structures. Several technics are able to generate accurate 3-dimensional structures of molecules. One can cite NMR and X-ray as well as molecular modelling programs to be part of the top structure generators. To fully understand the function and the role that structure plays in molecular functions, no one can pass by the visualization step. Broadly used programs of molecular visualization display 3D structures of molecules in a very efficient way. However, only poor efforts were put in the navigation paradigms use to explore such 3D models. Two things can explain this: Firstly, until a very recent period, molecular systems were quite reduced in terms of size and their exploration did not require any complicated algorithms. Secondly, visualization of molecular systems in immersive environments or with stereoscopic solutions was not possible or very limited. However, these two points are now outdated. It is now possible to generate accurate models of molecular systems with more than several hundreds of thousands of atoms. Moreover, some laboratories investigate other ways to render and interact with their 3D models. They aim to put the scientist in the center of its exploration, bringing the ensemble of his senses focus on his task using multimodal rendering (Férey et al., 2009). Stereoscopic solutions as well as intuitive and natural ways to interact with a molecular system are used in this purpose. Several works on the intake of such immersive environments for the scientific visualization have been made (Van Dam et al., 2000). Numerous issues linked to the navigation in immersive environments were addressed along the last years (Christie et al., 2008) but rarely by taking into account both the content of the scene and the tasks that will be performed in it. This immersion goes together with perception issues. Several studies have been made to evaluate the impact of immersive technics during the exploration of virtual scenes. When dealing with oriented and realistic environments, user comfort can be quite well managed. However, the exploration of abstract scientific data, where no implicit or explicit orientations exist, raises several issues when coming to the user experience and efficiency. Simple navigation tasks can be perturbed

Trellet M., Ferey N., Baaden M. and Bourdot P..
Content-guided Navigation in Multimeric Molecular Complexes.
DOI: 10.5220/0004914300760081
In *Proceedings of the International Conference on Bioimaging* (BIOIMAGING-2014), pages 76-81
ISBN: 978-989-758-014-7

by a phenomenon called “cybersickness” (LaViola Jr 2000), reducing significantly the user comfort and his efficiency to perform a specific task in a virtual environment. This observation is emphasized in the context of interactive molecular simulations where users manipulate the 3D structures during a simulation in progress (Dreher et al., 2013). Consequently, the need for alternative ways to navigate in virtual environments has been emphasized by Hanson et al. (Hanson et al., 1997) who cited several families of navigation paradigms increasing significantly the experience of the users. Based on these different observations, we propose in this study to bring new navigation paradigms aiming to fill the gap between virtual and immersive environments and molecular visualization. To do so, we took into account the importance of geometry in structural biology. Indeed, it has already been shown that most of large 3D structures were constructed around specific symmetry layouts (Goodsell and Olson, 2000). Symmetry is not only a spatial feature of these molecular complexes but also plays a significant role on the stability and then the function of such complexes. Our algorithm take as input the symmetry of a multimeric complex and the geometrical transformation between each monomers. This symmetry can be manually set up or automatically approximated using principal component analysis to detect the main orientations of monomers or by specific software dedicated to this analysis (Kim et al., 2010). Such information are good enough to provide a spatial basis to orient the virtual world of exploration. Through the symmetric nature of molecular complexes and by identifying the tasks a scientist would have to perform in molecular exploration, we will setup several navigation paradigms aiming to improve the user experience and efficiency.

2 CONTENT FORMALISATION IN STRUCTURAL BIOLOGY

Development of camera models has to take into account both the constraints linked to the visualization in immersive conditions and the navigation possibilities extractable from the molecular complexes of interest.

The notion of symmetry is of first importance in structural biology. As soon as a complex is formed of more than two identical chains, a symmetric configuration has significant chances to intervene in the structuration of the monomers. These particular layouts, found in large molecular complexes, play a crucial role in most of the biological functions accomplished by these structures. They are involved in complexes stability but sometimes also in the primary function of the complex itself. Many transmembrane proteins present a pore where ions will pass through to reach the inner or outer part of a cell. Most of the times, this pore will be also the exact symmetry axis of the protein. Each monomer of the complex will present a symmetric transformation to fit another monomer. Several types of symmetries can be found among molecular complexes and we listed some of them in the Figure 1. Each of them can be used as a base for a navigation system since they provide a first step to orient the complex. Any orientation of an abstract element in a virtual scene brings a new spatial landmark for the user and allows reducing the discomfort that could appear when dealing with such large and non-oriented structures.

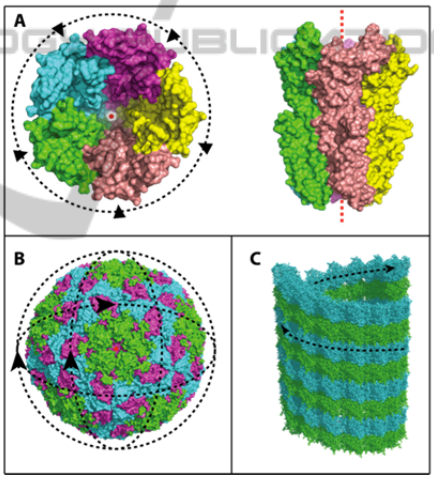

Figure 1: Several types of symmetries found in molecular complexes. (A) GLIC, transmembrane protein composed of 5 components with a single symmetry axis. On the left, top view of the protein, on the right, transversal view. (B) Virus capsid with two types of components presenting a symmetry center overlaying the gravity center of the virus. (C) Tubulins tangle shaping a microtubule with a screw-like symmetry.

We already cited the axial symmetry and it is one of the most represented symmetry among large molecular complexes. The ensemble of monomers is often organized in a circle manner around the symmetry axis that is coincident with the pore of the protein (Figure 1A).

Viruses present an important number of different symmetries, most of them central but for which angles and distances of rotation are very variable

and bring an important list of possible structures (Figure 1B).

Molecular complexes presenting important structural roles in the cell often adopt symmetries called "helical" (Figure 1C). The transformation to pass from one monomer can be associated to a screw-like movement. We can take the example of microtubules that are constituted of hundreds of subunits called tubulin and arranged with each other in a helical way. Microtubules are part of the cell skeleton and are involved in maintaining structure of the cell.

Despite their importance in the cell and the numerous functions that we just cited in a non-exhaustive manner, very few algorithms are dedicated to the identification of symmetry axes or centre in molecular complexes. Among them, only SymD (Kim et al., 2010) is available online and can be applied on a structural complex coming from a PDB structure. SymD is based on a fitting and permutation of monomers algorithm to identify the symmetries. In spite of its relative efficiency, SymD cannot deal with structures of more than 10 000 atoms and then misses most of the large complexes with the most remarkable symmetries. Moreover, the program outputs only symmetry axes when we saw that several structures present central ones. Symmetries being the starting point for the conception of navigation guides that will be described, the automatic detection of these geometric features have been developed by the team but will not enter in the scope of this study.

3 USERS NEEDS IN STRUCTURAL BIOLOGY

3.1 Context Support

Visualisation of molecular complexes has been brought to everyone's hands thanks to numerous programs like PyMol (DeLano, 2002) or VMD (Humphrey et al. 1996). In order to assess our navigation paradigms we chose to develop, in a first step, a PyMol plugin based on python language and using the API available for this program. Our choice to use PyMol comes from the fact that PyMol is recognized to be broadly used software in the structural biology community and could be then well evaluated by scientists of the field. Our developments are, however, completely independent of the visualization software used and we expect to port the paradigm on a large-scale visualization platform very soon. A short-term use case will see the integration of our paradigm in the specific and emergent case of visual analytics. This development will take place in an immersive context represented by a CAVE-like system. Together, the immersion and visual analytics bring in the limelight the need for proper algorithms addressing specific navigations tasks that we listed in a non-exhaustive way in the following section.

3.2 Navigation Tasks

The first steps of complex visualisation take place through the external exploration of the overall shape. These steps require a distant position of the camera as well as a specific orientation of the complex for the user. This orientation will be fixed during the whole exploring process and any movements of the camera will be made to avoid to the user to lose this orientation setup with respect to the complex.

It is really difficult to visualise a specific region, repeated along the different monomers of a complex. Navigation without constraints makes the quick and uniform observation of these regions almost impossible. The movements to reach these regions involve quick and precise motions to be performed in order to simultaneously keep some spatial landmarks and find the good point of view in the middle of thousands of particles. The feeling to jump from one monomer to the other in a smooth and quite instantaneous way is particularly interesting in the aim to compare binding sites or dynamic structures involved in complexes stability. Selection of similar and repeated regions of complex monomers allows the extraction of atoms as graphic objects used to perform certain standard geometrical analyses like Root Mean Square Deviation (RMSD) calculation. Several other comparisons like the number of bonds among a region or the size of binding pockets accessible for external links are also elements used to study a complex. In the same state of mind, distance between atoms or group of atoms is indispensable when coming to protein studies.

On top of the sequential navigation along the monomer, we cannot dissociate the importance of the point of view a user will have to analyse the region of interest. This is particularly true in the case of regions completely embedded and hardly accessible in the structure. We can then add an algorithm to compute the best point of view from a user position toward an interesting region.

As we already evoked, few molecular structures present unique internal arrangements that participate to their function. These functions are mainly

functions of passage where molecules of water or ions are exchanged between the inside and the outside of a cell membrane. This is made possible by the presence of a pore that also plays the role of the symmetry axis around which the different monomers constituting the transmembrane protein stand. Pores are then particular structural arrangements where many biological events can take place and of a first importance for scientists. They are really interesting to visualize but not so easy to reach and follow because of their burying state. We setup the possibility to easily anchor the camera to the entrance or exit of these biological "corridors" to let the user explore them via specific navigation paths.

Proteins are structured in tertiary structures that obey to specific rules, well known and identified for several years. It is sometimes useful to explore these geometrical patterns to value the shape and stability of the proteins. We chose to use the metaphor of the rollercoaster to minimize the interaction with the user. In this setup, the user only controls the movement speed and direction (forward or backward) but neither the path nor the orientation.

4 CAMERA MODEL

Our model of camera must address the different needs listed above in order to be pertinent for the scientists. It will be done through but putatively coupled navigation modes that we will describe now.

4.1 Navigation Modes

When a centre or axis of symmetry is detected, the molecular system is re-oriented to have its symmetry axis merged with the axis {0, 1, 0} of the {x, y, z} orthonormal reference. A translation is also applied to merge the basis of the complex with the origin of the coordinates system.

In a first mode of navigation that we could call "simple exploration", the camera is oriented in a way that it always faces either the centre of symmetry or the centre of the symmetry axis when it is located at the complex level (Figure 2, area B). In the external areas (Figure 2, areas A), camera *up* follows the global vector model (Khan et al., 2005) that keeps it fixed parallel to the symmetry axis if existing or fixed in a way to keep the user head at the same orientation with respect to the molecular system. Movements along the 3 dimensions are made in the same way, via the interaction with a unique button associated to each direction {x, y, z}. A translation along the camera *forward* will have a "zoom-in"/"zoom-out" effect.

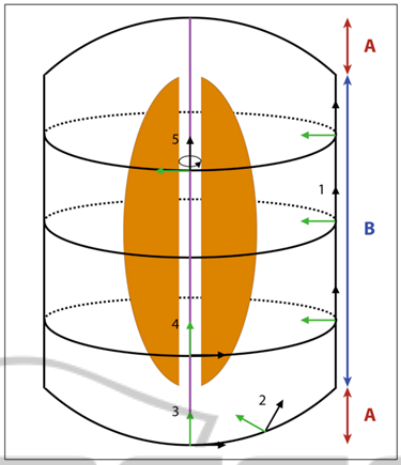

Figure 2: Schematic view of navigation paths and camera orientations with a symmetry axis (purple). The molecular complex is coloured in orange, areas A are outside the structure whereas area B encompasses the ensemble of the structure along the *y* axis. Each example of a camera position is associated to a number from 1 to 5. The camera *forward* is in green and the camera *up* in black.

A right translation will perform a turn around the structure while keeping the same distance between the camera and the symmetry centre. Finally the user will have the possibility to go up and down along an axis parallel to the symmetry axis in area B (Figure 2, position 1) but will fly over the top or base of the complex when in areas A (Figure 2, position 2). During the hovering of one of the complex extremity, the camera *forward* will be focus on the closest complex extremity.

At the intersection between the hovering trajectory and a symmetry axis, a new mode of navigation can be activated and will give the possibility to navigation along the symmetry axis (Figure 2, position 3). In this mode, the camera *up* will be kept in the same direction that it was when the user was flying over the structure to avoid any navigation issues in terms of orientation losses.

When the user is inside the complex, along the symmetry axis, we let the possibility to change the navigation mode for a free 360-turn around a specific position (Figure 2, position 5). In this mode, the camera *forward* will not be confounded with the symmetry axis anymore but will be perpendicular to it. The camera *up* will be along the axis and the user will have the possibility to look in every direction from his position. The movement might be associated to the one of a panoramic elevator

allowing the exploration of the internal structure of a protein.

To answer the basic needs of the field experts, the possibility to target a specific region of a monomer can be split into 4 successive steps:

1. Constrained navigation as illustrated in Figure 2 to face the interesting region.
2. Selection of the interesting area and best point of view computing.
3. Memorisation of the point-of-view information in terms of camera parameters.
4. Jump to the following monomer thanks to the symmetry information.

4.2 Accessibility Computing

The previous navigation mode includes a research of an optimal point-of-view relative to an interesting area pointed by the user. This algorithm operates in the following way: Starting with the center of an area identified by the user, a sphere is created around this point with a radius matching a configurable distance between the camera and the target. From each element of the sphere surface, a ray is cast toward the target. If a ray crosses a cell of the grid with an atom in it, the grid element corresponding to a surface starting point is considered as occulted.

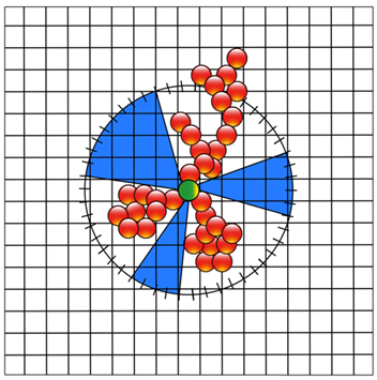

Figure 3: Grid splitting of the region surrounding an atom of interest (green). Each neighbouring atom (red) constitutes an occlusion for the camera view. The optimal point-of-view cones are represented in blue.

At the end, each element of the surface is associated to a 2D table and we extract the area with less occulted neighbors (Figure 3). The centers of these areas, once re-transformed into 3D coordinates, yield the observation points with the least occlusions to observe the target. The user is then able to choose among the optimal points of view calculated by modifying the position and the orientation of the camera.

This accessibility grid can also be used to follow the shape of protein surfaces or membranes. In this mode, the camera *up* is aligned on the surface normal and any movement can be associated to navigation on a ground where the user is limited to 2 dimensions control with a freedom degree of rotation around the camera *up* axis.

4.3 Structures Manipulation

The algorithm of optimal point-of-view reaches some limits when we come to protein areas completely buried into the structure or at the interface of two monomers. In these situations, accessibility is null and there is no simple algorithm to visualize these regions. Based on this observation, we went further in the use of the symmetric layout found in molecular structures. In order to put into light hardly accessible regions, we decided to change the 3D structure of our proteins. The user is able to control the 3D modification to fit his wishes. Some translational movements are applied to each monomer in order to further them from each other and then free several regions that were previously hidden in the structure (Figure 4). By a simple interaction with a button or by letting the program computes the movement automatically according to the camera position we can move the monomers away along an axis starting from the symmetry centre/axis to the monomer centre of mass.

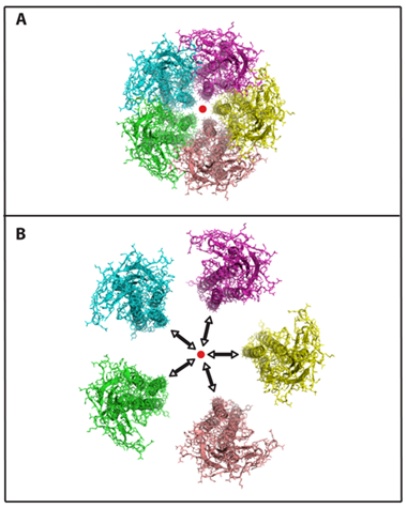

Figure 4: Top view of GLIC protein. (A) Default layout of the different monomers. (B) Spreading of the different monomers by translation with respect to a vector linking the symmetry axis centre to each monomer centre of mass.

5 CONCLUSIONS

Immersion for molecular exploration brings an interesting dimension for the study of certain biological events. However, it also brings some navigation issues that can be easily extended to desktop environments. Without any specific navigation paradigm, the exploration of abstract and scientific data can easily lead to significant spatial landmarks losses and then decrease the user experience and efficiency. To tackle this problem, navigation guides can be quickly setup. They are based on the most common feature that most of molecular complexes of a certain size share, a symmetric layout geometrically connecting the monomers between them. But beyond the simple comfort of the user, if correctly implemented, the constraints extracted from any centre or axis of symmetry can also become useful helps for the execution of daily standard tasks in structural biology.

Next steps will be the evaluation of our paradigm by structural biology experts. We will base our assessments on the broadly accepted evaluation benchmark proposed by Bowman et al (Bowman et al., 1997). The criteria used to assess the navigation paradigms group, in a non-exhaustive way: the task execution speed, its accuracy, the user spatial awareness in the virtual scene during and after the experience, the comfort of the user, etc…

It is important to notice that our paradigms are completely independent of the navigation technic or the visualisation system used. Indeed, they can be applied from immersive environments like CAVE systems to desktop ones and with mouse/keyboard as well as 3D mice or tracking solutions (Chen et al., 2013). A navigation control can easily be substituted to one other while keeping a behaviour consistency of the camera defined in function of the scene content.

REFERENCES

Bowman, D. A., Koller, D. & Hodges, L. F., 1997. Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Virtual Reality Annual International Symposium, 1997., IEEE 1997*. pp. 45–52.

Chen, W. et al., 2013. 6DoF navigation in virtual worlds: comparison of joystick-based and head-controlled paradigms. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*. VRST '13. New York, NY, USA: ACM, pp. 111–114.

Christie, M., Olivier, P. & Normand, J.-M., 2008. Camera Control in Computer Graphics. *Computer Graphics Forum*, 27(8), pp.2197–2218.

DeLano, W. L., 2002. The PyMOL molecular graphics system.

Dreher, M. et al., 2013. Interactive Molecular Dynamics: Scaling up to Large Systems. *Procedia Computer Science*, 18, pp.20–29.

Goodsell, D. S. & Olson, A. J., 2000. Structural symmetry and protein function. *Annual review of biophysics and biomolecular structure*, 29(1), pp.105–153.

Férey, N. et al., 2008. User needs analysis to design a 3D multimodal protein-docking interface. *In Proceedings of the IEEE Symposium on 3D User Interface*, 3DUI'08, pp.125–132.

Hanson, A. J., Wernert, E. A. & Hughes, S. B., 1997. Constrained navigation environments. In *Scientific Visualization Conference, 1997*. pp. 95–95.

Humphrey, W., Dalke, A. & Schulten, K., 1996. VMD: visual molecular dynamics. *Journal of molecular graphics*, 14(1), pp.33–38.

Khan, A. et al., 2005. HoverCam: interactive 3D navigation for proximal object inspection. In *Proceedings of the 2005 symposium on Interactive 3D graphics and games*. pp. 73–80.

Kim, C., Basner, J. & Lee, B., 2010. Detecting internally symmetric protein structures. *BMC bioinformatics*, 11(1), p.303.

LaViola Jr, J. J., 2000. A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, 32(1), pp.47–56.

Van Dam, A. et al., 2000. Immersive VR for scientific visualization: A progress report. *Computer Graphics and Applications, IEEE*, 20(6), pp.26–52.