Visualization and modelling of protein motion. Autodesk Maya as a reliable and fast instrument for macromolecular dynamic representation

Yuri Porozov[1], Raluca Andrei[1,2] and Monica Zoppè[1]

[1] Scientific Visualization Unit, Inst. of Clinical Physiology CNR, Pisa, Italy.
[2] Lab. of Molecular Biology, Scuola Normale Superiore, Pisa, Italy.

Introduction and objectives

Proteins are in constant motion: besides moving from one place to another, they also change conformation, due to spontaneous vibrations or upon stimulation from external forces, such as binding to other cellular components, small molecules and ions or changes in the chemical conditions (pH, salt concentrations etc.). However, it is difficult to figure the dynamic events that take place in the living cell, mostly because information about structure is still.

We present a novel approach for fast modeling and visualization of interconformational transitions in proteins. This approach is based on using the 3D animation package Autodesk® Maya® as a media for visual representation and as a source of fast programmable mathematical engine. Comparison of simulated interconformational transition of Calmodulin with NMR data shows the applicability of the proposed system. At the same time, observation of the movements in the 3D space of Maya conveys a clear impression of natural, smooth motion capable of delivering movement information in a way that we can interpret as meaningful.

Methods

Modeling: as we are interested in molecules with defined (although flexible) structure, we have taken information from chemical and structural biology databases and imported spatial coordinates of atoms and their connections with a script that reads the .pdb files.
Animation: means to interpolate intermediate positions between a start and an end position or conformation. This can be accomplished through direct animation, path animation, kinematics (determined by a system of bones and joints), inverse kinematics (ik) or dynamics (force field animation). Models can also be treated with some dynamic properties: collisions can be detected (and avoided), fields can be introduced and other properties can be encoded.
Two or more stable conformations were taken from NMR data. Then they were imported into Maya using our scripts, which read a PDB file line by line, extract information about amino acids, atomic coordinates and identities, and represent atoms as particles. Many factors, considered as representations of chemical and physical properties are evaluated by Maya engine. They are: physical diameters of all atoms, stiffness, damping and momentum as attributes of springs, strengths of the goal and strength of the radial field on per-atom basis, which control sterical conflicts during motion.
For modeling of Ca-free Calmodulin we used dynamics animation because of large amount of atoms (2262 with H) in the molecule represented as particles. Chemical bonds were represented as "springs", a particular Maya object. Springs connect two or more particles and affect their spatial behavior. It is possible to correct various parameters of motion by fine tuning the parameters of springs. To obtain motions between conformations we used a "goal", a set of point to which each particle is directed. The radial field was implemented on per-particle basis and prevents collisions of particles. It also has properties as magnitude, radius and some other.
To import proteins in Autodesk® Maya® and reconstruct possible interconformational transitions with dynamics animation we have written the MEL script **protview_sp.mel**. This program reads .pdb files which contain the NMR data and creates "atoms" represented as particles, in corresponding xyz coordinates.
To reconstruct possible path of interconformational transition of protein we use data from NMR studies. The two conformations were placed into scene as key-frames and each of them were associated with a goal. The goal object attracts each particle of the other object paired with it. So the combination key-frames – points on timeline for obligate "visiting" and goals allow us to simulate a motion of protein. During the motion a protein changes it shape in order to fit into goals and so far to fit into different conformations.

Results

The major suggestion is that if more than one conformation for a molecule exist, then should exist a motion path that molecule passes through to reach each conformation. In order to evaluate that path we build some conformation as "starting point" and "end point" in Maya scene. Other conformations available in NMR are considered as steps, visited by molecule during interconformational changes of geometry.
For our investigation we choose the NMR models retrieved by Kuboniwa H. et al (PDB ID 1CFC) [1].
In order to retrieve a full range of motion we calculated RMS distances of backbone atoms between all available conformations using VMD RMSD TT [2] plug-in and selected the 2 most distant conformers in terms of Cartesian distance: conformation 7 and conformation 2, whose RMSD is 11.523 Å.
Conformation 7 was set as starting position in Maya scene and conformation 21 was set as a goal. During simulation of movement the Maya engine interpolate all atoms in the space to achieve the final goal (in another words Maya try to fit all atoms of conformation 7 into conformation 21) in the space. Model 7 fits almost exactly with the conformation 21 at 30 frames of timeline.
To evaluate the validity of our system we retrieved Cartesian coordinates of all atoms of the models interpolated by Maya. The data were collected every 5 frames during simulation. RMS distances between backbone atoms for these intermediates and all NMR models were calculated (fig. 1). for conformations within 3 Å RMSD. The graph shows, that simulated virtual molecule fits sequentially into some of known models. At frame 10 the backbone distances between chain moved by Maya and model 19 (NMR) is 1.926 Å. At frames 15 and 20 the virtual model of CaM driven by forces in Maya scene, visits sequentially NMR models 22, 13 and 3. Moreover the distance of simulated model to NMR model 3 decreases more then 3 times.
In order to control protein geometry during transitions we have checked the Ramachandran plots of intermediates between models 7 and 21. As shown at figure 2, the common shape of the CaM is stable during simulation. Number of aminoasids with unrealistic torsion angles in not increased, indicating that the protein keeps it's fold during conformational transitions.
It should be noted that for different compounds different concept (techniques) of animation and 3D simulation are implemented. However for modeling of whole protein a more sophisticated system consisting of springs and fields, decreasing dramatically amount of sterical conflicts, is preferred.

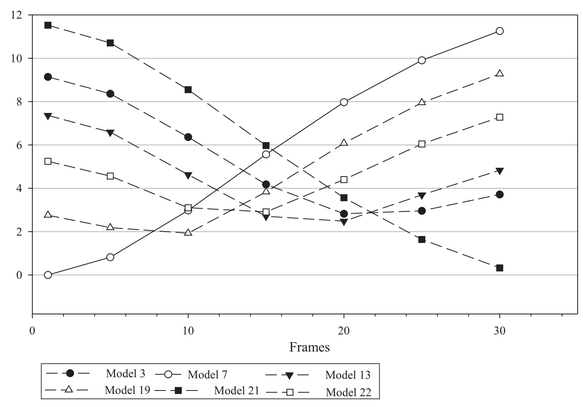

Fig.1 Dynamics of RMSD between Maya intermediates and some NMR models of Calmodulin during simulation.

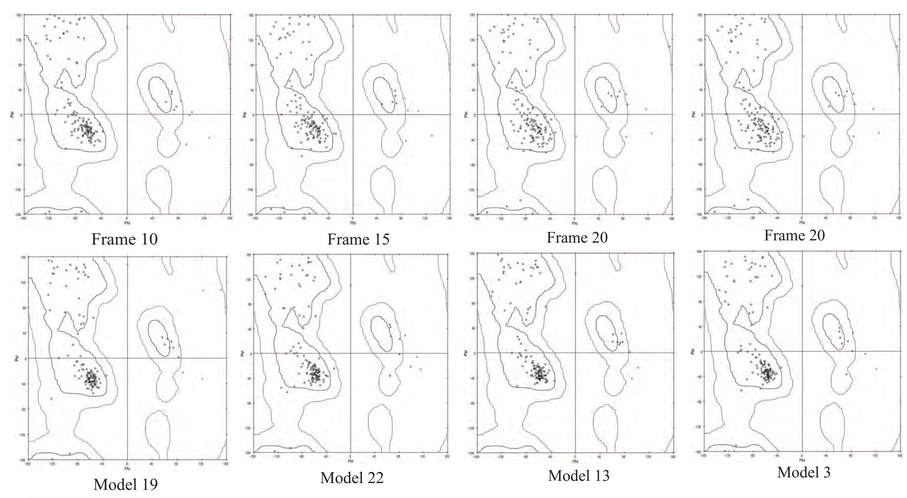

Fig.2. Ramachandran plots of Maya intermediates and corresponding NMR models of Calmodulin

Conclusions

In this report, we describe how, starting from information gathered from scientific sources, we were able to produce and analyze movement of protein in a direct, visual way as a first step towards the visualization of much larger molecules, such as big multichain proteins, protein complexes and other macromolecular assemblies, from DNA to ribosomes to membrane systems.

In case of polypeptide chain it is relatively easy and significantly faster to build a model of protein motion that is meaningful from biological point of view and is not in contradiction with basic chemical rules.

We use only geometrical parameters, which are much easier and faster to calculate (in particular relative to molecular dynamics simulations), and only concentrate in the direct trajectory between a start and an end states. However, it is possible to introduce other rules and constraints derived from chemical and physical information to make the system more reliable as a 'fast simulator' of biological processes.

When the structure of a protein is known in more than one conformational state, our system can propose a trajectory for such change, in a very fast and inexpensive way. However, in order to obtain a biologically meaningful representation of motions, all information available, derived from any experimental procedure, must be integrated.

1. Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C.B., and Bax, A. Solution structure of calcium-free calmodulin. Nat Struct Biol. 1995, 2, 768-776.
2. Humphrey, W., Dalke, A., and Schulten, K. VMD: visual molecular dynamics. J Mol Graph. 1996, 14, 33-38, 27-38.