ViewMotions Rainbow: A new method to illustrate molecular motions in proteins

Gregory M. Cockrell, Evan R. Kantrowitz *

Boston College, Department of Chemistry, Merkert Chemistry Center, Chestnut Hill, MA 02467, USA

ARTICLE INFO

ABSTRACT

The biological functions of many enzymes are often coupled with significant conformational changes. The end states of these conformational changes can often be determined by X-ray crystallography. These X-ray structures are snapshots of the two extreme conformations in which the macromolecule exists, but the dynamic movements between the states are not easily visualized in a two-dimensional illustration. Here we have developed a new method to visualize macromolecular motions called a ViewMotions Rainbow diagram. These diagrams show the initial and final states overlaid along with approximately 30 intermediate structures calculated by linear interpolation of the backbone coordinates of the initial and final states. This group of structures is then spectrally colored from the initial structure in blue to the final structure in red. ViewMotions Rainbow diagrams provide the reader with a much easier way to understand the macromolecular motions using a single two-dimensional illustration. Since producing these diagrams requires a number of different software packages, we have setup the ViewMotions Web Server (http://viewmotions.bc.edu) to automatically generate these diagrams from two Protein Data Bank files or from the Database of Macromolecular Movements (http://molmovdb.org).

1. Introduction

Many enzymes undergo significant conformational changes during the course of their biological function. These conformational changes may be induced in a variety of ways including the binding of substrates, allosteric effectors, or by phosphorylation. The end states of these conformational changes can often be determined by X-ray crystallography or structural NMR. While these structures are snapshots of the two extreme conformations in which the macromolecule exists, the dynamic movements between the states are not easily visualized.

Krebs and Gerstein have developed the Internet accessible Database of Macromolecular Movements [1], which includes the Morph Server [2] that can be used to generate movies depicting the movement of a macromolecule from its initial to its final state. These movies are an excellent resource to visualize macromolecular motions; however, these molecular motions are much more difficult to visualize using a single diagram. A variety of methods to illustrate conformational changes have been employed. However, each of these two-dimensional representations lacks one important advantage of a movie depiction of the conformational

change, which is to allow the reader to visualize the actual molecular motion. In order to solve this problem we have developed a new method to visualize the conformational changes of a protein in a two dimensional illustration. This new method uses a number of existing software packages to generate a series of intermediate structures, typically thirty, on the path between the initial and final conformations of the same protein. All of the structures are shown in a single image using a spectral color gradient to indicate the molecular motion. We have published a number of these ViewMotions Rainbow diagrams as an all-inclusive way to convey molecular motions [3–5]. However, the steps required to generate these diagrams are cumbersome, therefore we have developed a Web Server to allow the user to easily create ViewMotions Rainbow diagrams.

2. Computation methods

In order to generate a ViewMotions Rainbow diagram the coordinates of the initial and final positions of the macromolecule are required. These coordinates can be supplied either with two PDB files or a MorphID corresponding to a particular macromolecular movement from the Database of Macromolecular Movements [1]. A flow chart of the steps necessary to create a ViewMotions Rainbow diagram is illustrated in Fig. 1.

After data input, the two structures are then aligned and superimposed by one of four alignment algorithms: a standard least-squares procedure as described by Kabash [6] (sRMSD), the globally optimized alignment method of Needleman and Wunsch
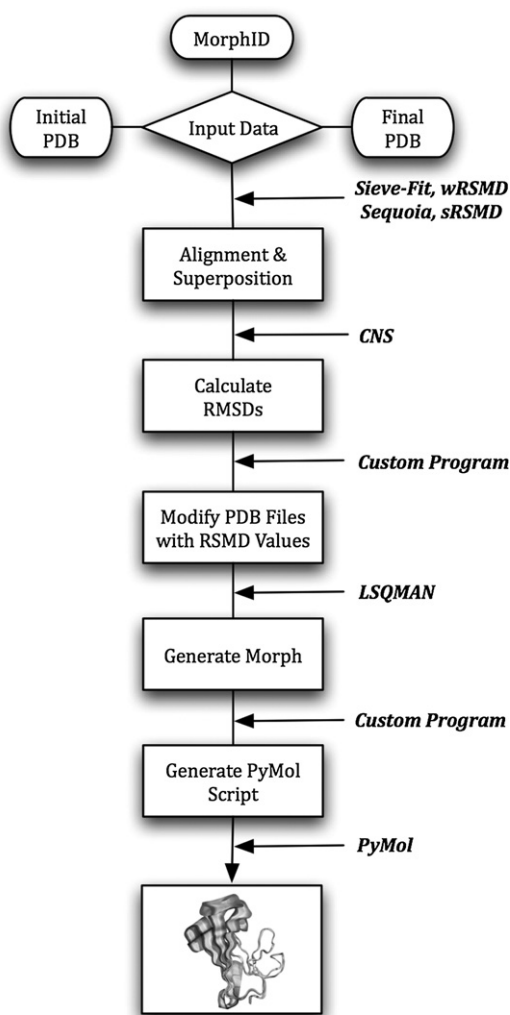
Fig. 1. Flow chart for producing a ViewMotions Rainbow diagram. The input data is either two PDB files corresponding to the initial and final coordinates of the motion or a MorphID corresponding to a macromolecular motion described by the Database of Macromolecular Movements [1]. Each of the computational steps is indicated by a box and the software used to perform the calculation is indicated.

[7] as implemented in SEQUOIA program [8], the sieve-fit algorithm [9,10], and a Gaussian-weighted algorithm [11] (wRMSD). The sieve-fit algorithm [9,10] usually yields the best alignments, particularly for domain motions.

CNS [12] is next used to calculate the backbone and side chain RMSD values between the beginning and ending structures using the rmsd.inp procedure. The rmsd.inp procedure generates a list of RMSD values for the backbone and sidechains of each residue. The backbone RMSD values are used for the coloring of the ViewMotions Rainbow diagram. In order to accomplish this a custom written C language program was written to replace the b-factor and occupancy values in the initial and final PDB files with the corresponding backbone and sidechain RMSD values, respectively, output from the rmsd.inp procedure of CNS. LSQMAN from the Uppsala Software Factory (http://xray.bmc.uu.se/usf/) is then used to generate the intermediate structures using its morph option. Usually 30 intermediate structures are generated between the initial and final structures by LSQMAN. However, if the maximal RSMD for the macromolecular movement is greater than 1.8 Å or 3.0 Å, 45 or 60 structures are generated, respectively. Finally, a custom shell script is used to generate a command file for PyMol (http://www.pymol.org) to generate a ViewMotions Rainbow

diagram. PyMol and Chimera (http://www.cgl.ucsf.edu/chimera) session files are also generated, which can be downloaded by the user.

3. Results and discussion

3.1. Commonly used methods to view macromolecular motions

The catalytic cycle of most transferases involves a domain closure, which brings two substrates in close proximity for reaction and may also change the active site environment to one more favorable for reaction. The binding of glucose to hexokinase and glucokinase are two such examples. Shown in Fig. 2(A and B) are the open and closed forms of hexokinase as well as the two forms overlaid (Fig. 2(C)). From this static representation of the conformational change in hexokinase, it is difficult to envision the actual motion. A modification that is often used to help direct the viewer to the important portion of the molecule is to use a neutral color in regions where the conformational changes are the smallest (see Fig. 2(D)).

To provide the reader with alternative methods to view molecular motions, a variety of different graphical representations have been employed. A tube representation can be used in which the width of the tube is linearly proportional to the displacement of the protein backbone atoms (Fig. 2(E and F)). A significant drawback to this representation is that the variable tube width must be applied to a single structure, in this case either the open or closed form. Neither of these individual representations clearly conveys the conformational change that the protein is undergoing. A CPK representation of each of the two states can also be used as illustrated with hexokinase in Fig. 2G (open form) and 2H (closed form). Although the domain motion is apparent, this representation has its drawbacks as small conformational changes, or conformational changes in several parts of the protein are not well represented.

Krebs and Gerstein recognized the complexities of trying to view molecular motions in proteins and therefore developed a Morph Server [2] and a database [1], as a means to create movies of molecular motions and the database to track the various types of molecular motions. Viewing a movie of a conformational change provides a much easier way to understand a molecular motion than any of the two-dimensional representations mentioned above.

In order to overcome the problem of viewing protein conformational changes in textbook and journal diagrams we have devised a new representation, which we call the ViewMotions Rainbow diagram. This new representation starts with the initial and final coordinates of the protein (see Fig. 2(I and J)) and thirty or more intermediate structures generated by linear interpolation of the protein backbone coordinates. Although, a linear interpolation of macromolecular motions does not truly represent the actual motion of the protein, the resulting structures are useful for illustrating the molecular motion. All of these structures are shown simultaneously and spectrally colored from the initial structure in blue to the final structure in red. In order to focus the viewer on the parts of the protein undergoing the conformational change, only those portions of the protein that show a RMSD of 1.3 Å or greater are colored (see Fig. 2(L)).

The Database of Macromolecular Movements [1] classifies macromolecular motions into three main categories: (I) motions of fragments smaller than domains, (II) domain motions, and (III) larger movements than domain movements involving the motion of subunits. Furthermore, each of these categories is subdivided by the type of motion (see Table 1). The first two categories involve motions of a single polypeptide chain and are those best visualized using a ViewMotions Rainbow diagram.