**The DynDom3D Webserver for the Analysis of Domain Movements in Multimeric Proteins**

Christopher Girdlestone1 and Steven Hayward1,\*

1D’Arcy Thompson Centre for Computational Biology, School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, U.K.

**Running title**: DynDom3D webserver

**ABSTRACT**

DynDom3D is a program for the analysis of domain movements in multimeric proteins. Its input is two structure files that indicate a possible domain movement, but the onus has been on the user to process the files so that there is the necessary one-to-one equivalence between atoms in the two atom lists. This is often a prohibitive task to carry out manually which has limited the application of DynDom3D. Here we report on a webserver with a pre-processor that automatically creates an equivalence between atoms using sequence alignment methods. The processed structure files are passed to DynDom3D and the results are presented on a webpage that includes molecular graphics for easy visualization. The website can be found at: http://www.cmp.uea.ac.uk/dyndom/3D

**INTRODUCTION**

Domain movements in biomolecules can be intimately involved in function. Prominent examples include the domain movement of the β subunit in F1-ATPase (Boyer, 1997) and the ratchet movement in the ribosome (Frank and Agrawal, 2000). The original DynDom project (Hayward and Berendsen, 1998; Hayward and Lee, 2002; Lee, et al., 2003) focused on the analysis of domain movements within a single protein chain. However, it is apparent that the analysis of subunits in isolation is inadequate for the detection of regions involved in inter-subunit communication as in allostery, and for the analysis of mechanism in molecular machines. DynDom3D (Poornam, et al., 2009) was developed in response to this need. Both DynDom, referred to as “DynDom1D”, and DynDom3D are based on the idea that domains can be identified from their differing rotational properties in the movement between two structures representing a conformational change of a biomolecule. In DynDom3D the elements of this rotational analysis are sets of atoms within overlapping blocks placed on a grid spanning the whole molecule. The rotation of atoms from a block between the two structures is calculated with a quaternion-based least-squares best-fit routine. In these routines each atom has two sets of coordinates one from each of the two conformations. However, there are several reasons why two structures of a protein representing a possible conformational change will not have the one-to-one correspondence between atoms required for fitting routines: one or more residues are different or missing due to mutation, atoms are missing due to experimental limitations, or there are substantial differences as the structures are from different species. In practice, this means that when two such structures are read from two coordinate files one has to make the one-to-one association of coordinates by identifying equivalent atoms. If this is not done then the output from DynDom3D is usually meaningless but it can be prohibitive to carry out manually. Thus a webserver that is able to automatically perform this task, pass the results on to the DynDom3D, and present the results in a clear and informative way on a webpage would be of considerable value.

**2 METHODS**

**2.1 Basic Methodology**

According to the Protein Data Bank (PDB) file format, atoms are in an ordered list and if two files are identical in all respects apart from the coordinates of the atoms (i.e. they are different conformations of the same protein), then the *i*th atom in the list in file 1 would be equivalent to the *i*th atom in the list in file 2. If this is not the case the files need to be edited to achieve this equivalence. These differences may occur at the model level, the chain level, the residue level or the atom level. Figure 1A illustrates what happens at the model and chain level. Only chains within the same model are compared and it is presumed that corresponding chains (identified by chain terminators “TER”) are ordered equivalently in each model. The next stage utilises the dynamic-programming algorithm for pair-wise sequence alignment of amino acid residues within each corresponding chain. The amino acid sequences of each corresponding chain (extracted from the “ATOM” field) are aligned (globally using the BLOSUM64 substitution matrix with a gap penalty of -8 without penalizing overhanging ends). Inserted residues (including those in overhanging ends) in both aligned sequences are excised as shown in Figure 1B. Then, in the final stage, atoms in each aligned residue are themselves aligned (by PDB atom name using an identity matrix and gap penalty of 0) and inserted atoms are excised. This leaves only those atoms that have an exact match in atom name. It would mean, for example, that if a phenylalanine were aligned against a tyrosine, the atoms N, CA, C, O, CB, CG, CD1, CD2, CE1, CE2 and CZ would remain with only the OH of the tyrosine side-chain removed as shown in Figure 1C. This stage would also remove all hydrogen atoms in one residue and not the other, if for example one structure were solved by neutron diffraction and the other by X-ray diffraction. The method developed thus adheres to the principle of retaining as much as possible that is in common between the two structures. This procedure leaves two structure files with the desired one-to-one equivalence between atoms. These files are then passed on to the DynDom3D program.

**2.2 Implementation details**

The PDB file pre-processor described above was developed in Java and included as a library in a Java web application which acts as both interface and wrapper for the programs. The application takes advantage of web tools to provide a user interface, before utilising servlets to invoke the pre-processor and DynDom3D. Once DynDom3D has run, results are forwarded to a JSP page. The JSP page displays the PDB file (“Conformer 1” on input page) using the molecular graphics software, JSmol, (wiki.jmol.org), a JavaScript implementation of the popular Jmol that avoids security issues associated with the Jmol applet. The JSP page uses Ajax calls to obtain structured data for JSmol from the server, e.g. the domain identification colour of each atom. The server also returns a processed JavaScript file to make a JavaScript representation of the PDB objects available. This allows for manipulation of the molecular display using the JSmol console.

**3 RESULTS**

The input page allows the user to input two PDB accession codes or upload PDB formatted files, e.g. biological unit files downloaded from the PDB website. DynDom3D has a number of parameters that control its behaviour. On the input page the user is able to adjust four of these: the “Grid Size”, the “Block Factor”, the “Occupancy”, and the “Minimum Domain Size”. The meaning of these is described below. The first molecule (Conformer 1) is spanned by a cubic grid, the cell length, *g*, of which, is termed “Grid Size”. In the process of determining domains based on clustering of rotation vectors, cubic “blocks” are placed at each grid point. Given a grid, the length of the block is set equal to *b*×*g*, where *b*, an integer, is the “Block Factor”. The grid size, *g*, determines the spacing of neighbouring blocks, and *b* determines the number of grid cells neighbouring blocks share, which has a maximum value equal to *b*3-*b*2. Thus the larger *b*, the greater the overlapping of blocks and the greater the smoothing. When *b*=1 there is no overlapping. The “Occupancy”, *occ*, is a number between 0 and 1, which gives the minimum occupancy of a block. Let *N*max be the greatest number of atoms found within any block, then only blocks with at least *occ*× *N*max atoms will be passed on for calculation of their rotation vectors. Removal of low occupancy blocks helps reduce noise. The final parameter that can be specified is the “Minimum Domain Size”, which is the minimum number of atoms for it to be accepted by the program as a domain. The user is able to change these parameters, which have default values of 4Å, 2, 60% and 200 atoms, respectively. There are other parameter’s involved but the user is not free to alter these. For further details see the DynDom3D method article (Poornam, et al., 2009).

At the output page the user can download the processed PDB files, a report detailing the models, chains, residues and atoms removed, and the DynDom3D output files. In addition the webpage shows the domains and interdomain screw (hinge) axes in JSmol, coloured for identification, as well as tables detailing the relative movement of the domains giving information such as angle of rotation, translation along the axis, percentage closure, etc. Figure 2 shows the output page for the large allosteric protein, aspartate transcarbamylase.

The website also includes a help page.

**DISCUSSION**

A webserver for the analysis of domain movements in protein complexes is presented. It can be used in those cases where a domain movement is thought to exist between two structures representing a conformational change in a protein. As it makes use of least-squares best-fitting routines there is the requirement of there being a one-to-one correspondence of atoms as they appear in the PDB files of the two structures. There can be many reasons that this may not be the case and leaving the onus of the user to correct this has limited application of DynDom3D. Here a pre-processor to a webserver for running DynDom3D that remedies this has been described. The output from DynDom3D includes a JSmol molecular graphics window displaying Conformer 1 coloured according to domains with interdomain screw axes depicted as arrow molecules. In multi-subunit proteins, particularly those that are known to have allosteric mechanism, this depiction is particularly interesting as it may show how a movement is communicated to other parts. An interesting finding from the original work showed that domains do not confine themselves to subunits and often span subunit boundaries. It was proposed that interesting regions to focus on are those where there is a common subunit and domain boundary.

DynDom1D can be used on only monomers or individual subunits and is able to determine “interdomain bending regions”. These are regions of the backbone where the hinge bending takes place. In many cases the interdomain screw axis (hinge axis) passes through the interdomain bending regions. By its design, DynDom3D does not assign hinge bending regions but it is clear that any region where there is a transition between domains is a region of interest as these regions are often involved in control of the movement. It is particularly interesting to compare the global DynDom3D result on the whole oligomeric structure with the results of the application of DynDom1D (a webserver for which can be found at the same website) to each individual subunit.

It is not unusual for DynDom3D to give a null result. This means that according to the program there is no domain movement. A critical parameter for this is the ratio of interdomain displacement to intradomain displacement. This parameter is set to 1.0 and cannot be adjusted at the webserver. Empirically, the value of 1.0 has been found to work well for DynDom1D and is the value used for the DynDom1D webserver. Although DynDom3D has been shown to be fairly robust against variations in *g* and *b*, sometimes adjustment of these from their default values may be required. At present the webserver can only deal with protein molecules as the pre-processor has been developed specifically for proteins. In future work we hope to make it available for all types of biomolecules.

**ACKNOWLEDGEMENTS**

The authors are grateful to Russell Smith for technical assistance.

**AUTHOR DISCLOSURE STATEMENT**

No competing financial interests exist.

**REFERENCES**

Boyer, P.D. 1997. The ATP synthase - A splendid molecular machine, *Ann. Rev. Biochem.*, 66, 717-749.

Frank, J. and Agrawal, R.K. 2000. A ratchet-like inter-subunit reorganization of the ribosome during translocation, *Nature*, 406, 318-322.

Hayward, S. and Berendsen, H.J.C. 1998. Systematic analysis of domain motions in proteins from conformational change: New results on citrate synthase and T4 lysozyme, *Proteins*, 30, 144-154.

Hayward, S. and Lee, R.A. 2002. Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50., *J. Mol. Graph. Model.*, 21, 181-183.

Lee, R.A., Razaz, M. and Hayward, S. 2003. The DynDom database of protein domain motions, *Bioinformatics*, 19, 1290-1291.

Poornam, G.P., et al. 2009. A method for the analysis of domain movements in large biomolecular complexes, *Proteins*, 76, 201-212.

**Figure legends**

**Figure 1.  A** Models from the two PDB files A and B are compared in the order they appear and only chains within corresponding models are compared assuming corresponding chains (identified by chain terminators “TER”) are ordered equivalently in each model. **B** The alignment shows an inserted Leu in chain A from PDB file A and inserted Asp and Ala in chain A from PDB file B. Atom data corresponding to these residues are removed from the PDB files so that the sequence alignment would show no indels. Overhanging ends are also removed. **C** Atom name sequences in aligned residues are aligned and if there are inserted atom names, as here for Phe and Tyr, atom data from the inserted atoms are removed. Here atom data of the OH on Tyr would be removed from PDB file B.

**Figure 2.** Part of the DynDom3D server output page for aspartate transcarbamylase using PDB structures 1AT1 and 1RAA (biologi-cal unit files downloaded from PDB website) with grid size 7Å, block factor 4, occupancy 60% and minimum domain size 200 atoms. There are 5 domains in this 12 subunit protein coloured according to domain in the JSmol window. The arrows depict interdomain screw (hinge) axes for the relative movement of domains in contact. The tables below the JSmol window give details on the relative movement of domains. In all 24 protein atoms were removed from 1AT1 and 372 from 1RAA.