Protein structure, dynamics and function

Bioinformatics based approaches to understanding the relationship between a protein's structure and its dynamical behaviour. For example, enzymes exhibit conformational change upon substrate binding and product release. Detailed, atomic level information on protein internal motions usually either comes from X-ray crystallography or NMR experiments. In the former, for example, different conformations are observed when an open structure of an enzyme is found in the absence of a bound substrate, or substrate analogue, and the closed structure in the complexed state. In these cases one has detailed information about the internal rearrangements that occur upon substrate binding. More usually is not the case and the conformational states have to be predicted using computational models such as those developed by the group: DynDom.

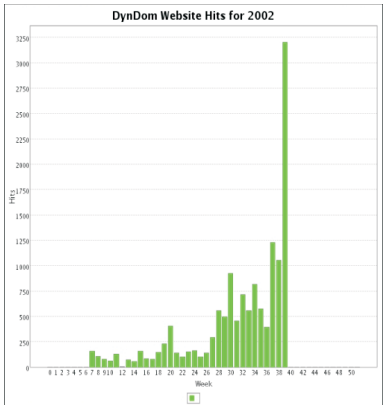

Figure 7.4: Hits on DynDom database website per week (our hits excluded).

DynDom is the most popular program for analysing domain motions in proteins and is available from Collaborative Computational Project 4, CCP4 (X-ray crystallography) website, a major resource used by structural biologists from all over the world (see http://www.ccp4.ac.uk/main.html). The DyDom database (see http://www.sys.uea.ac.uk/dyndom), is an online database of protein domain motions that have been analysed using the DynDom program. This site was set up in direct competition to a site on Macromolecular movements in Yale (see http://molmovdb.mbb.yale.edu/MolMovDB/) and is already proving itself to be popular, receiving hits from all over the world.

The DynDom Software Project

If one has two conformations of a protein then these can be analysed for domain motions. Hayward and Berendsen developed a unique and rigorous methodology based in rigid-body kinematics and implemented in the DynDom program for the specific purpose of analysing movements in proteins in terms of the relative motions of quasi-rigid parts [228, 227]. DynDom is part of the CCP4 software suite (*URL: http:// www.ccp4.ac.uk/ main.html*) and is the most popular choice for the analysis of domain motions in proteins (this has been ascertained by looking at the number of citations of the DynDom paper in comparison with the two other programs available: HingeFind [450] and DomainFinder [236]. The DynDom methodology is unique in that as part of the process of determining domains and hinge axes, it also determines interdomain bending regions. These are of great interest as it is these that collectively control the domain motion. The latest version of DynDom (version 1.50) incorporating a number of impor-

tant improvements is now available from CCP4 and article describing this new version is soon to be published [229].

Structural Genomics programmes [63] aim to cover all the variety of structures within a genome using large-scale, high-throughput X-ray crystallography and NMR. They use target lists derived from recently sequenced genomes. The implications for computational structural biology are enormous with the current database of protein structures built up over more than 20 years set to triple in two citeberman.bhat.ea:nature. Hypothesis driven structural studies aim primarily to solve structures in order to understand function. These studies, which are the antithesis of structural genomics, often result in two main structures, one where the protein is unliganded, the other where it is liganded to a substrate, or substrate analogue. These two structures can then be used to understand the mechanism of the functional movement. In the structural genomics programmes, proteins will be solved in the absence of their active ligands and consequently there will be no information on functional movements. An information gap will arise that is analogous to the information gap between sequence and structure (this information gap has led to huge efforts to predict structure from sequence as exemplified by the CASP programme: *http:// predictioncenter.llnl.gov/*).

Research Team: Dr. Steven Hayward, G.Qi. Funding: BBSRC

DyDom Database on Protein Conformational Change

We have recently developed a new database system based on the output from the DynDom program, called the ”DynDom Database” (*http:// www.sys.uea.ac.uk/ dyndom*). This database grew out of an initial study on 24 proteins [226]. To our knowledge, the only other site where X-ray derived data on functional movements in proteins has been brought together, is the ”Database of Macromolecular Movements” at Yale [200]. In the Yale database, domain motions are classified into a number of categories based on the idea that there are two main types of domain motions in proteins: “hinge” and “shear” [201]. The assignment of a particular domain motion to one of these categories appears not to have been made on the basis of a consistent and rigorous methodology, but rather on the basis of a subjective interpretation. Although there are doubts as to the scientific underpinnings of the Yale database, its high hit rate (65,000 hits per month at peak) reflects the essential importance of protein conformational change to many areas of active biological research. Although the DynDom database has only recently gone live, the steady increase in number of hits is very encourag-

ing (see Figure 7.4). An article describing the database is currently under review (Lee, submitted).

One of the most important and unique features of DynDom is its ability to accurately determine interdomain bending regions. These are the regions that collectively control the interdomain motion but have hardly been described in the literature apart from in [226, 227, 229]. The database itself focuses particularly on these regions. If the domain motion is part of function, then interdomain bending regions are functional sites, and just like binding sites are potential targets for therapeutic molecules and mutation studies. The importance of these sites for drug design was originally pointed out to Hayward by Dr Terry Stouch from the Pharmaceutical giant Bristol-Myers Squibb.

Research Team: Dr. Steven Hayward, G.Qi. Funding: BBSRC

Molecular Dynamics Simulation

Molecular Dynamics (MD) simulations are expensive in terms of cpu-time, but more importantly human time. Although high-throughput MD may one day be possible, for the foreseeable future, simulations will be performed by groups with interests in a particular target. However, the outlook for MD is looking particularly good as recent advances in computer hardware and MD software have made longer times scales accessible. Recent groundbreaking work by Grubmueller and coworkers [137] has shown that it can accurately reproduce experimental results, and more importantly reveal mechanism on the atomic scale. A recent study of the enzyme, citrate synthase has revealed for the first time that MD can reproduce a triggered conformational change in an enzyme [378], Figure 7.5. What is more, it was shown that there exists an energy barrier between the open state of that molecule and the closed state, bound to its substrate. This result was in contrast to earlier speculation that open and closed states were equally accessible for the unliganded enzyme [201].

Research Team: Dr. Steven Hayward, G.Qi. Funding: BBSRC

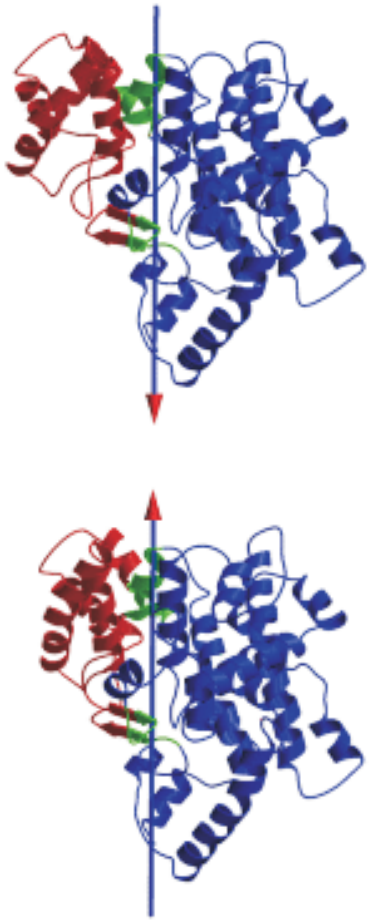

Figure 7.5: Citrate synthase in an open (top) and closed configuration .