Introduction. Biomolecular simulation

Adrian J. Mulholland*

School of Chemistry, Centre for Computational Chemistry, University of Bristol, Bristol BS8 1TS, UK

‘Everything that living things do can be understood in terms of the jigglings and wigglings of atoms’ as Richard Feynman provocatively stated nearly 50 years ago. But how can we ‘see’ this wiggling and jiggling and understand how it drives biology? Increasingly, computer simulations of biological macromolecules are helping to meet this challenge.

Keywords: molecular dynamics; protein; enzyme; DNA; multi-level modelling; quantum mechanics/molecular mechanics

Experiments can provide detailed structures of biological macromolecules such as proteins, but it is hard to study directly how the structures of individual molecules change on short time scales as they function. Similarly, it is not yet possible to study directly by experiment alone the molecular mechanisms of fast processes such as chemical reactions in enzymes or ion transport through membranes. Simulations based on fundamental physics offer the potential of filling in these crucial ‘gaps’, modelling how proteins and other biomolecules move, fluctuate, interact, react and function. Simulated structures can be analysed in atomic detail, to dissect what is important in determining biological activity. Ever more powerful computers, and sophisticated techniques, mean that biomolecular simulations can now do far more than provide pretty pictures: they can show biomolecular machines at work, and analyse the mechanisms by which they function. Physics-based simulations complement experiments in building a molecular-level understanding of biology: they can test hypotheses and interpret and analyse experimental data in terms of interactions at the atomic level. A wide variety of simulation techniques have been developed, applicable to a range of different problems in biomolecular science. Simulations have already shown their worth in helping to analyse how enzymes catalyse biochemical reactions (Garcia-Viloca et al. 2004; Mulholland 2005; Warshel et al. 2006), and how proteins adopt their functional structures (Daggett 2006). They can help in the design of drugs (Jorgensen 2004; Michel et al. 2007) and catalysts (Jiang et al. 2008; Röthlisberger et al. 2008), and in understanding the molecular basis of disease. And simulations have played a key role in developing the conceptual framework now at the heart of biomolecular science, that is, the understanding that the way that biological molecules move and flex—their dynamics—is central to their function (Karplus et al. 2005), demonstrating the truth of Feynman’s (1963) assertion.

Biomolecular simulation is a vibrant and growing area, making increasingly significant contributions to biology. The papers assembled in this issue of Interface Focus provide good examples of a number of different types of applications across the field. van der Kamp et al. (2008) provide an overview of biomolecular simulation from its origins over 30 years ago to current state-of-the-art studies and exciting prospects for future developments. Biological macromolecules, such as proteins and nucleic acids, are complex and challenging to model. They are large, often function in heterogeneous environments (e.g. cell membranes) and show complex conformational behaviour over a variety of time scales. Simulations of biomolecular motions—molecular dynamics simulations—are the most widespread type of biomolecular simulation. Such simulations calculate how the atoms of a biomolecule move according to classical (e.g. Newtonian) mechanics. Atomistic simulations can now reach the microsecond time scale for small proteins or oligonucleotides in solution (Freddolino et al. 2008; Mura & McCammon 2008); for larger systems, simulations are limited to tens to hundreds of nanoseconds. Using techniques of this sort on supercomputers, it is now possible to simulate, in atomic detail, the dynamics of multi-million-atom biological macromolecular assemblies such as ion channels (Gunbart et al. 2005; Biggin & Bond 2008; Lindahl & Sansom 2008) and the protein synthesis machinery of the ribosome (Trobro & Åqvist 2007; Trabuco et al. 2008), and to identify functionally relevant large-scale conformational changes in proteins (Karplus et al. 2005; Pentikäinen et al. 2008).

This Interface Focus contains a number of examples of molecular dynamics simulations of biomolecules. Jiranusornkul & Laughton (2008) study the effects of oxidation of guanine in DNA in molecular dynamics simulations of double-stranded DNA by molecular dynamics simulations. They find a significant change in flexibility at the oxidation site, with implications for

---

*adrian.mulholland@bristol.ac.uk

One contribution of 9 to a Theme Supplement ‘Biomolecular simulation’.
how repair proteins locate such regions of damaged DNA. Yan et al. (2008) investigate a drug target, human heat shock protein 90, using molecular dynamics simulations to study the binding of water molecules and a small molecule ligand to the protein. They find that four tightly bound water molecules should be included when modelling the interactions of ligands (potential drugs) to the protein. The water molecules appear to be an essential part of the binding site.

Simulations of the dynamics of large biomolecular systems on long time scales are feasible because they typically use fairly simple, though atomically detailed, ‘ball and spring’ type models. Empirical ‘molecular mechanics’ (MM) potential energy functions of this type represent atoms as spheres with point charges, with chemical bonds treated as springs. Years of careful parameter development have produced MM models that can give good descriptions of protein structure and dynamics. The simplified physical description, though, means that these methods cannot easily be applied to some types of problem: for example modelling chemical reactions (e.g. in enzymes) where treating the bonds as ideal springs would not allow them to break. Chemical bonding is a quantum mechanical phenomenon, so to model the making and breaking of bonds in a chemical reaction, ideally methods taking into account the essential quantum mechanics (QM) are required. QM methods, i.e. which calculate molecular electronic structure, can give very accurate results for small molecules, but are too computationally intensive for large biological systems. A solution is provided by hybrid QM/MM methods, which combine a QM (electronic structure) treatment of a small region (e.g. an enzyme active site) with a MM description of the surroundings (e.g. protein and solvent). Approximate (e.g. semi-empirical) QM/MM methods are now fast enough to be used in biomolecular dynamics simulations. At the other end of the spectrum, QM/MM calculations can now be carried out with highly accurate QM methods (Claeyssens 2006), a development that promises to revolutionize enzymology (Mulholland 2007).

This Interface Focus contains several studies that use QM/MM methods. Senthilkumar et al. (2008) test QM/MM methods for modelling biomolecular hydrogen bonds (with density functional theory level QM methods), in particular the treatment of polarization, and find that these methods give a good description of these interactions, which are central to biomolecular recognition. The widest application of QM/MM methods in biomolecular modelling to date has been in the area of enzyme catalysis (Friesner & Guallar 2005; Mulholland 2005, 2008; Senn & Thiel 2007), and this issue contains two examples of such applications. Ferrer et al. (2008) study the important model enzyme lactate dehydrogenase. They examine the effect of an active site mutation, showing how biomolecular modelling can help to interpret the effects of site-directed mutagenesis on enzyme activity. Johannissen et al. (2008) examine conformational effects on the proton transfer reaction in aromatic amine dehydrogenase, which is a subject of high interest today (Lodola et al. 2007). This enzyme reaction involves quantum tunneling, an area that is attracting considerable debate (Allemann et al. 2006; Klinman 2006; Sutcliffe et al. 2006; Wang et al. 2006), a debate in which biomolecular simulation has a crucial role to play (Hammes-Schiffer & Watney 2006; Masgrau et al. 2006; Olsson et al. 2006). In this Interface Focus, Guallar & Wallrapp (2008) demonstrate a different type of application of QM/MM methods, namely a QM/MM approach to finding electron transfer pathways in proteins. They describe investigations of protein–substrate interactions in cytochrome P450cam, ascorbate peroxidase and cytochrome c peroxidase, and propose a new approach for studying long range electron transfer between proteins, applied here to the cytochrome c peroxidase–cytochrome c complex.

To model larger systems and/or longer time scales (e.g. over a microsecond), even simpler models than atomistic molecular dynamics are needed. Coarse-grained models use reduced representations, in which a single particle represents a group of atoms (e.g. an amino acid side chain). Here, Khalid et al. (2008) use coarse-grained models in molecular dynamics simulations of DNA–lipid complexes, which are potentially useful as delivery vectors in gene therapy. These workers study the self-assembly of lipid bilayers in the presence of a DNA dodecamer, observing spontaneous bilayer formation of lipid bilayers from initially random systems. They also study transfer of the DNA through the bilayer, finding a high energy barrier to DNA insertion into the bilayer hydrophobic core of the bilayer.

No single biomolecular simulation technique is capable of addressing the whole range of time scales and length scales, and range of types of processes, which arise in biomolecular science. Different techniques have to be used for different types of problem, or for different spatial or temporal scales. Combining different levels of calculation—multi-scale or multi-level modelling—is a hot and increasingly important topic in biomolecular simulation (Sherwood et al. in press; Woods & Mulholland in press). QM/MM methods provide an example of a hybrid, multi-level approach (Woods et al. 2008). Multi-level biomolecular simulation promises in future to provide exciting new insight into the workings of ‘molecular machines’ such as biological motors and pumps.

This themed supplement arose from the recent CCPB conference ‘Biomolecular Simulation 2008: Frontiers of Biomolecular Simulation’ held in Bristol, UK, 7–9 January 2008. CCPB is a collaborative project, supported by BBSRC, which aims to strengthen and broaden the UK biomolecular simulation community, and also to build links internationally (www.ccpb.ac.uk). Membership of CCPB is free and open to all. Among events organized by CCPB are specialist training workshops to train (particularly younger) researchers in this growing field. CCPB also aims to foster links between computational and experimental scientists: close interaction with experiment will be vital in the ongoing development and exploitation of modelling methods and the results they produce. Given the rapid pace of development (Woods et al. 2005), and the sophisticated techniques now available, it seems clear that biomolecular simulation will make
an increasingly important contribution to twenty-first century science.

A.J.M. would like to thank all the contributors to this themed supplement, and Dr Tim Holt (Publishing Editor, J. R. Soc. Interface) for handling all the administrative work, and his assistance. A.J.M. thanks EPSRC and BBSRC for support.

REFERENCES


