Multiscale Modelling of the Interplay Between Global and Local Structural Changes in Viral Drug Target Proteins.

David W. Wright[1], Benjamin A. Hall[1], Milo Bem[1] & Peter V. Coveney[1§]

[1] Centre for Computational Science, Chemistry Department, University College London, London, UK

[§] Corresponding Author: email: p.v.coveney@ucl.ac.uk

Abstract: Information on protein structure and dynamics is vital to the understanding of their function and in the design of drugs which aim to perturb it. Here we aim to bring together atomistic insights from molecular dynamics simulations and more coarse grained models to help understand the architectural features of the proto foamy virus integrase and HIV-1 reverse transcriptase enzymes, both of which are large multi-domain proteins whose interactions with DNA are vital for their function and the activity of drugs.

Introduction

Molecular dynamics simulation is a widely used technique that allows protein motions to be derived from basic physical laws and the properties of their constituent atoms. Here we use MD to investigate the relationship between overall system organisation and local structural properties. To further assess the influence of global structure we construct anisotropic network models (ANM) of proteins which estimate the motions around an energy minimised structure [1]. This minimised structure is a coarse grained representation of the protein based on C_α positions alone.

Proto-Foamy Virus

One of the key processes in the viral lifecycle is the integration of a DNA copy of the viral genome into the host DNA. This process is performed by an enzyme called integrase (IN). It has long been a goal within structural biology to produce a structure of the complete HIV-1 IN. So far no such structure has been produced; however, one of the homologous Proto-Foamy Virus (PFV) has [2]. The core catalytic domain (CCD) of both proteins is very similar but sequence analogy beyond this region is more limited and significant structural differences are expected based on partial HIV-1 structures. Our goal is to use the PFV structure to model the impact of mutations upon the binding of antiviral drugs targeted at HIV-1. In order to do this we need to investigate whether we can create model systems which capture the key interactions between enzyme, drug and DNA but which are not strongly impacted by the areas of low homology between PFV and HIV-1 IN. The minimal biologically active unit of the PFV IN is thought to be a dimer, though higher order structures have been proposed to exist. The tetrameric crystal structure shows the core dimer structure, with DNA and drug molecules bound, and the core catalytic domain (CCD) of partially resolved monomeric units outside of the core dimer without any bound drug molecules or DNA resolved. Four distinct PFV construct systems were simulated with DNA and the drug raltegravir bound for 10 ns, in 11 replicas as summarised in Table 1. These constructs represent different, increasingly large subdivisions of the protein which could be expected to reproduce the binding site; the isolated core catalytic domain (CCD), the CCD and the C-terminal domain (CTD), the complete monomer and the complete dimer.

Construct	**Domains**	**Residue Range**
CCD	CCD	117-302
CCC	CCD + CTD	100-374
Complete monomer	CCD, CTD, NTD and NED	10-374
Complete dimer (2 x monomer)	2 x (CCD, CTD, NTD and NED)	10-374 x 2

Table 1: Description of the domains included in the various PFV constructs simulated

We find that whilst individual domains retain their fold over 10 ns periods, large inter-domain motions are observed in systems smaller than the complete dimer. The RMSD distribution relative to the original crystal structure for each construct is shown in Figure 1 (alongside a picture illustrating the motions observed in the

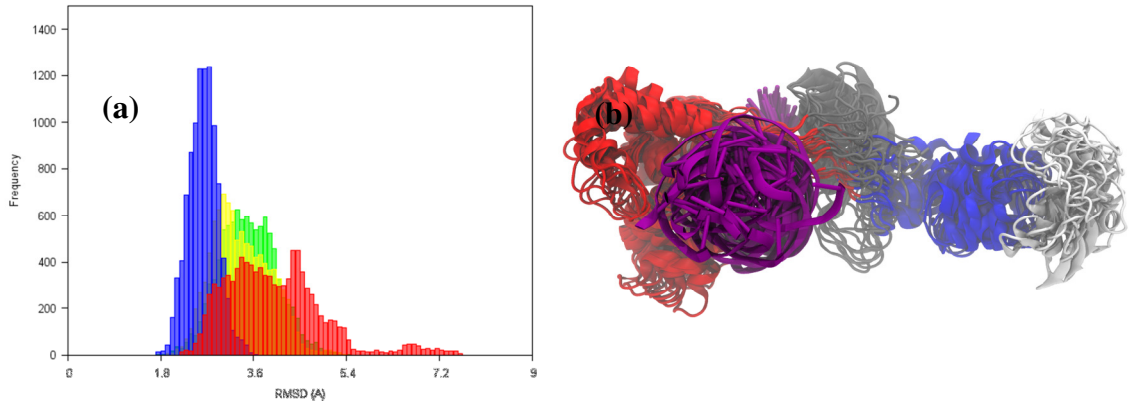

Figure 1: (a) RMSD distribution of the various constructs consider with respect to the same domains in the crystal structure. The values for the CCD are shown in green, CCC in yellow, the monomer in red and the dimer in blue. The narrow peak exhibited by the dimer illustrates its extra stability. (b) Superimposed structures from the ensemble simulations of the monomer. The domains are coloured; CCD in red, CTD in grey, NTD in blue and NED is white. DNA is shown in purple.

monomer simulations). The extra flexibility causes the breakup of the crystallographically defined binding site, leading to changes in specific contacts made by the drug molecule and the loss of magnesium ions from the binding site. In contrast, the structure of the integrase dimer retains the drug binding site, including magnesium ion positions, and reproduces tightly bound water molecules observed in high resolution crystal structures.

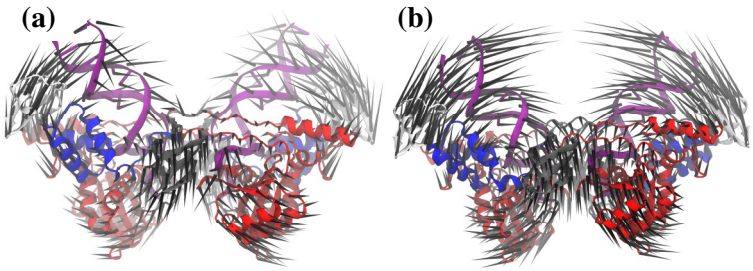

Figure 2: Motions of the PFV virus captured by (a) the first PC derived from ensemble molecular dynamics and (b) the slowest mode derived from ANM. Concerted domain motions are similar in both.

The concatenated dimer trajectory was analysed using principal component analysis in order to identify the most significant changes observed throughout the ensemble. The crystal dimer structure was also used to create and ANM. A comparison of the first PC and the slowest mode generated from the ANMs is shown in Figure 2. This indicates that simplistic models based mostly on the global architecture of proteins describe large scale equilibrium motions in this system.

HIV-1 Reverse Transcriptase

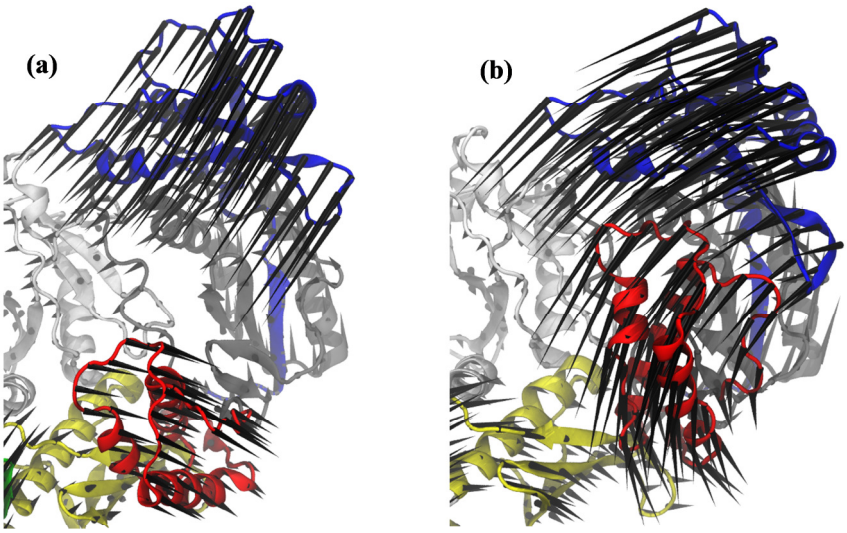

Figure 3: Fluctuations of the atoms in (a) the open and (b) the closed conformation of NNRTI bound RT indicated by the lowest frequency mode derived from anisotropic network models of the structure.

Our simulations of PFV IN indicate that overall structural architecture can be used to predict the dominant motions at equilibrium. Another key antiviral drug target enzyme is the HIV-1 reverse transcriptase (RT) which is responsible for the conversion of the viral RNA genome into DNA. The Non-Nucleoside RT Inhibitors (NNRTIs) are a class of drugs which act allosterically, binding away

from the active site but nonetheless compromising viral activity. One theory regarding their action is that they constrain the motion of a domain known as the 'thumb' [3]. Recent ensemble MD simulations within our group[4] have shown that the closing motion believed to be suppressed is in fact possible in NNRTI bound RT as suggested by previous network models[5] (although we have only captured it occurring in a single simulation indicating that it is a rare event). Figure 3 shows the results of ANM analysis of the initial and final snapshots from the simulation capturing the closing motion. In the open conformation the thumb and fingers move in independent directions whereas in closed conformations the thumb and fingers move in a coordinated manner. This behaviour resembles the results obtained for the closed apo structure.

In the case of the NNRTI nevirapine a water molecule is seen to bind alongside the drug in most crystal structure. However, this is not present in the 3HVT[6] structure. We conducted ensemble simulations of the 3HVT to see if water would enter the NNRTI binding pocket if not initially placed there. Not only did this occur but in each occasion it occurred through the same route and was accompanied by changes in the conformation of the fingers domain approximately 25 Å away from the binding site (see Figure 3). These results confirm the importance of atomistic level events and show that they can be rapidly communicated into larger scale motions in distant regions of proteins.

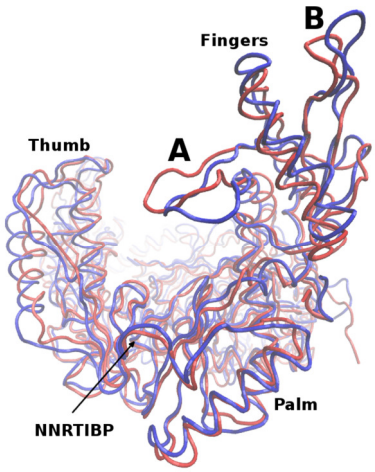

Figure 4: Conformational change exhibited by HIV-1 RT structure based on the 3HVT PDB upon water entry to the NNRTI binding pocket. Initial structure in red, final structure in blue.

Acknowledgements

This work has been supported by the EU FP6 ViroLab and the EU FP7 CHAIN, VPH-SHARE and P-Medicine projects. Access to US TeraGrid was made available by the National Science Foundation, and access to to the EU DEISA network by the Virtual Physiological Human Virtual Community managed by the VPH Network of Excellence (*www.vph-noe.eu*). We are grateful to EPSRC and CoMPLEX for funding the Ph.D. studentship of D.W.W.

References

1. Bakan, A. and I. Bahar, *The intrinsic dynamics of enzymes plays a dominant role in determining the structural changes induced upon inhibitor binding.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**: p. 14349-54.
2. Valkov, E., et al., *Functional and structural characterization of the integrase from the prototype foamy virus.* Nucleic acids research, 2009. **37**: p. 243-55.
3. Sluis-Cremer, N., N.A. Temiz, and I. Bahar, *Conformational changes in HIV-1 reverse transcriptase induced by nonnucleoside reverse transcriptase inhibitor binding.* Current HIV research, 2004. **2**: p. 323-32.
4. Wright, D.W., et al., *Thumbs down for HIV: Domain level rearrangements do occur in the NNRTI bound HIV-1 Reverse Transcriptase.* Submitted, 2012.
5. Temiz, N.A. and I. Bahar, *Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase.* Proteins, 2002. **49**: p. 61-70.
6. Wang, J., et al., *Structural basis of asymmetry in the human immunodeficiency virus type 1 reverse transcriptase heterodimer.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**: p. 7242-6.