JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Thumbs Down for HIV: Domain Level Rearrangements Do Occur in the NNRTI-Bound HIV-1 Reverse Transcriptase

David W. Wright,[†,§] S. Kashif Sadiq,[‡,§] Gianni De Fabritiis,[‡] and Peter V. Coveney*[,†]

[†]Centre for Computational Science, Department of Chemistry, University College London, London WC1H 0AJ, U.K.

[‡]Computational Biochemistry and Biophysics Lab, GRIB, IMIM-UPF, Barcelona Biomedical Research Park, C/Doctor Aiguader, 88, P/C 08003 Barcelona, Spain

Supporting Information

ABSTRACT: One of the principal targets in human immunodeficiency virus type 1 (HIV-1) therapy is the reverse transcriptase (RT) enzyme. Non-nucleoside RT inhibitors (NNRTIs) are a class of highly specific drugs which bind to a pocket approximately 10 Å from the polymerase active site, inhibiting the enzyme allosterically. It is widely believed that NNRTIs function as "molecular wedges", disrupting the region between thumb and palm subdomains of the p66 subunit and locking the thumb in a wide-open conformation. Crystal structure data suggest that the binding of NNRTIs forces RT into a wide-open conformation in which the separation between the thumb and fingers subdomains is much higher than in the apo structure. Using ensemble molecular dynamics simulations (aggregate sampling ~600 ns), we have captured RT bound to the NNRTI efavirenz in a closed conformation similar to that of the apo enzyme, suggesting the constraint of thumb motion is not as complete as previously believed. Rather, our investigation confirms that a conformational distribution across open and closed states must exist in the drug-bound enzyme and that allosteric modulation is effected via the alteration of the kinetic landscape of conformational transitions upon drug-binding. A more detailed understanding of the mechanism of NNRTI inhibition and the effect of binding upon domain motion could aid the design of more effective inhibitors and help identify novel allosteric sites.

Allosteric modulation is conventionally described as the structural alteration of multi-conformational proteins into either active or inactive states upon binding of modulators to sites distal to protein function. It is becoming increasingly evident, however, that many proteins exist in a conformational equilibrium of states with inter-transitions on a variety of time scales.[1,2] This requires a more comprehensive view of the process of modulation that acknowledges the existence of a conformational distribution even among the drug-bound states of the protein. Allostery is then seen as a change in the energetic landscape of accessible conformations.[3−5]

Molecular dynamics simulations provide a detailed atomic resolution view of conformational dynamics in proteins,[6−8] but kinetic sampling of conformational transitions is limited by the time scale of the transition. Thus for fast time scale processes, including rapid secondary structure transitions or domain motions[9] and even fast-folding processes,[10] accurate determination of the kinetic transitions between various states is achievable; longer time scale transitions, by contrast, are occasionally observed but no kinetic information is yielded.[9] Proteins also exhibit conformational changes on time scales well beyond the current capabilities of simulation.[11] Computational demand scales nonlinearly with system size. Therefore, even accessing rare conformations for larger proteins poses significant challenges. This complicates the explanation of the underlying causes of allosteric modulation.

A good example is HIV-1 RT, a multifunctional enzyme that contains DNA polymerase and RNaseH active sites and is formed of two subunits, p66 and p51. The p66 subunit contains polymerase and RNaseH domains joined by a connection domain. Subdomains in the polymerase domain are known as the fingers, palm, and thumb due to their resemblance to a right hand (see Figure 1). The p51 subunit contains only the first four of these subdomains and adopts a different folded conformation. The polymerase catalytic triad consisting of Asp110, Asp185, and Asp186 is located in the p66 palm.[12] Numerous crystal structures of HIV-1 RT bound to NNRTIs

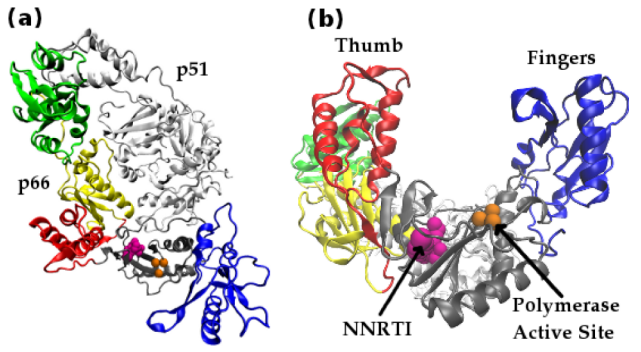

Figure 1. Structure of the NNRTI-bound HIV-RT taken from the 1IKW crystal structure. (a) HIV-1 RT is a heterodimer consisting of a p66 and a p51 subunit. The domains of p66 are individually colored: the fingers are shown in blue, the palm gray, the thumb red, the connection yellow, and the RNaseH green. The p51 subunit is shown in white. The location of the polymerase active site triad is depicted in orange. The location of the bound NNRTI efavirenz is shown in magenta. (b) A view along the HIV-1 RT from the polymerase active site illustrates the resemblance of the p66 subunit to an opened hand.

Received: February 16, 2012
Published: July 24, 2012

are available, all of which show the inhibitors bound to a pocket approximately 10 Å from the polymerase active site, indicating that they inhibit the enzyme allosterically. This pocket is located near the junction between thumb and palm and opens onto the p66/p51 interface.[13]

One widely believed hypothesis as to the mode of function of NNRTIs is the "arthritic thumb" model in which NNRTIs act as a "molecular wedge", disrupting the region between the p66 thumb and palm subdomains, locking the thumb in a wide-open conformation. All available structures of the NNRTI-bound HIV-1 RT enzyme adopt this conformation, while crystal structures of the apo enzyme show the enzyme in a closed conformation with contacts made between the p66 thumb and fingers domains.[14,15] However, despite the lack of corresponding crystal structure information, the apo enzyme must open to function, while the possibility that the drug-bound enzyme can also exhibit conformational change has remained unknown.

We investigated this theory by conducting an ensemble of all-atom molecular dynamics simulations of HIV-1 RT bound to the NNRTI efavirenz (EFV), with a single simulation of the apo enzyme for comparison. Initial structures were taken from the 1IKW[16] and 1DLO[17] crystal structures for the EFV-bound and apo simulations, respectively. Hydrogen atoms and solvent water molecules were added and topologies generated using the leap module of AMBER 9.[18] The force field parameters for the inhibitors were completely described by the general AMBER force field (GAFF).[19] The standard AMBER force field for bioorganic systems (ff03)[20] was used to describe the protein parameters. All simulations were conducted in the NVT ensemble with temperature maintained at 300 K and pressure at 1 bar. Initial equilibration was conducted for 6 ns (see Figure S1) using NAMD2[21] before production simulation was conducted using ACEMD.[22] A full description of the equilibration and simulation protocol is provided in the Supporting Information. Five simulations were performed of the drug-bound and one of the apo HIV-1 RT, each for a length of 100 ns. We shall henceforth refer to the simulations of the drug-bound HIV-1 RT as EFV1 to EFV5 and that of the apo system as APO1. In all trajectories, structural convergence of inflexible residues (Table S1) was preserved (Figure S2), consistent with the outcome of the equilibration phase and with previous measurements of structural root mean squared deviation (RMSD) by Ivetac and McCammon.[23] This permitted analysis of local structural changes within the enzyme.

In four of the five EFV-bound simulations (EFV2 to EFV5) RT maintains the open conformation seen in the crystal structure, with a separation of 36 Å or greater between p66 thumb and fingers at all times. In the remaining run (EFV1) the p66 thumb is seen to move across the DNA binding cleft toward the fingers (see Figure 2). This results in a conformation similar to that observed in the apo enzyme, as shown in Figure 2. The rearrangement occurs rapidly over approximately 2 ns, roughly 20 ns into the trajectory. Some further relaxation occurs after 50 ns. The closing process is shown in full in the movie provided as Supporting Information. The apo enzyme remains in the closed conformation for the duration of the simulation, in agreement with experimental evidence that this is the dominant conformation of the enzyme in the absence of DNA.[24] A comparison of the conformations adopted in all our simulations with those observed in crystal structures is provided in Figures S3–S5.

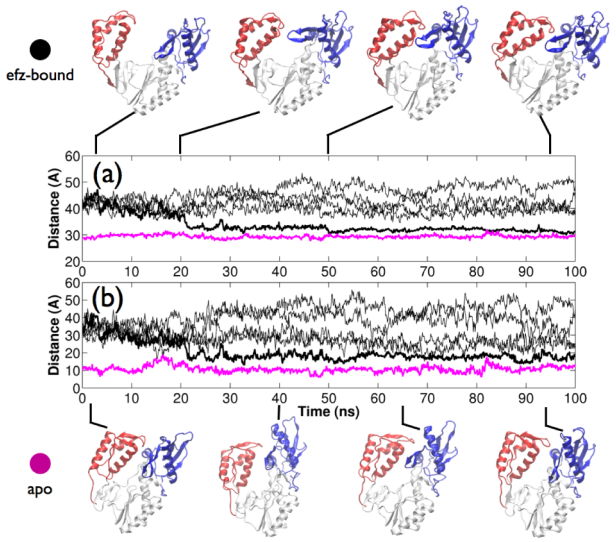

Figure 2. Thumb–finger distances of HIV-1 reverse transcriptase measured over our six simulations: five simulations of the efavirenz-bound enzyme (black) and one of the closed apo-HIV-1 RT (magenta) for comparison. The efavirenz-bound simulation that exhibited thumb–finger closure (EFV1) is highlighted by a thicker line. Metrics are (a) the distance between the backbone centers of mass of the thumb (residues 244–318) and fingers (residues 1–84 and 120–150) and (b) the distance between the non-heavy-atom centers of mass of thumb residue 287 and finger residue 24. Structures of apo and drug-bound systems at corresponding times are also shown.

Comparison of the closed structure observed during EFV1 with that of the apo enzyme indicates that the p66 thumb bends to a similar degree across the palm subdomain in both cases, resulting in a similar level of occlusion of the binding cleft. However, the thumb bends approximately 10° farther away from the RNaseH domain in the EFV-bound closed structure. This change alters significantly the interactions between the p66 thumb and fingers domains (see Figure 2). Figure 3 shows contacts made between the two subdomains for the APO1 and EFV1 trajectories. In the APO1 simulation contacts appear between residue 289 and residues 76 and 63. Once closed, the EFV1 structure exhibits contacts in a similar region, between 288 and 75. Coincident with the relaxation observed 50 ns into the EFV1 trajectory, a hydrogen bond between Thr290 and Asp76 is formed; this bond is broken and re-formed repeatedly during the remainder of the simulation (see Figure S6). The closed structure in EFV1 also gains a strong contact between the tip of the fingers, around residue 65 (which is believed to play a vital role in the polymerase activity of RT[25,26]), and residues 249–251. Transient hydrogen bonding between subdomains is also observed coincident with their rearrangements (see Figure S7). This represents a contact between the thumb and the loop running from β3 to β4 which is involved in dNTP binding. Direct contact between these domains may play an important role in altering the recognition of RNA or DNA templates and primers and consequently the ability of the enzyme to perform its polymerase function. Furthermore, a number of mutations in the fingers subdomain have been linked to NNRTI hypersusceptibility, and the mechanism of this interaction is poorly understood.[27−29] Direct contacts between these two subdomains provide a pathway for communication between seemingly distant loci within the RT structure. The single contact found in the apo crystal structure, between residues 24

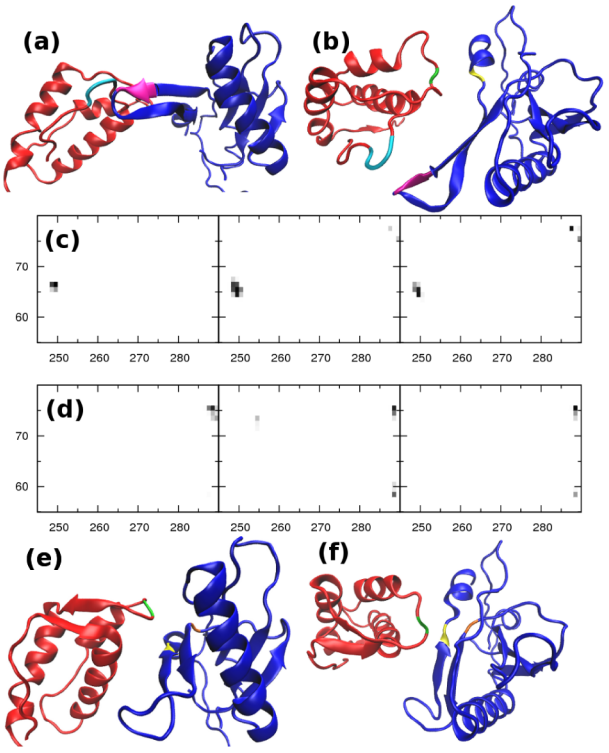

Figure 3. Contact between the p66 fingers and thumb subdomains for efavirenz-bound and apo HIV-1 RT. (a,b) Location of residues 65 (pink), 249−251 (cyan), 75 (yellow), and 288 (green) which are involved in contacts in the efavirenz-bound system. (c) Contact map for the drug-bound trajectory as averages over 1−20, 40−60, and 80−100 ns (the darker the pixel, the more frequently the contact is made). (d) Contact maps for the simulation of the apo enzyme. (e,f) Locations of residues 63 (yellow), 76 (orange), and 289 (green) involved in these contacts.

and 288, is not pictured but is sporadically present throughout the APO1 simulation.

Previous molecular dynamics simulations have not seen any large-scale motions of the p66 thumb in NNRTI-bound simulations, which has been taken to be support for the "arthritic thumb" model.[23] Analyses of the available crystal structures (reproduced using the software ProDy[30] in the Supporting Information) also identify only local rearrangement of this subdomain.[15] However, coarse-grained network models have suggested that NNRTI binding alters the direction of domain motions rather than suppressing them.[31] Such models are constrained to describe equilibrium motions around a structure assumed to represent a minimum in the conformational landscape and describe the direction but not magnitude of these fluctuations. The novel observation of a fully closed EFV-bound structure in our simulations demonstrates that the modes of motion identified by these simple models can result in global conformational changes. Network models are unable to resolve atomistic effects and incorporate drug molecules as single nodes. One of the consequences of this is that dynamically significant rearrangements of local interactions are not accounted for in these models. Our results can be reconciled with both previous MD simulations and network models if both apo and EFV-bound HIV-1 RT explore similar ensembles of conformers, the effect of NNRTI binding being to strongly bias the exploration of these conformers toward the most open configurations. Our simulations suggest that to understand the method of operation of NNRTIs will require a more complete understanding of the conformational landscape of apo and drug-bound systems and the kinetics of conformational transitions. Different NNRTIs may differentially alter the landscape, potentially altering both the available states and the rates of interconversion between them.

Accurate kinetic characterization of conformational transitions using computational simulation is an extremely challenging yet eventually necessary task for such a large enzyme, more so because the rates of conformational transitions are likely to be on the microsecond time scale or above.[24,32] However, our confirmation that both open and closed conformations are accessible to the drug-bound reverse transcriptase provides information that new NNRTIs would be more effectively designed not only on the basis of optimizing the binding free energy but, more importantly, on the basis of maximizing the thermodynamic difference between the drug-bound functional and dysfunctional conformations. The closed state of the enzyme is not functional, and it therefore should serve as a new target for anti-retroviral therapy, where the aim should be focused on designing complementary allosteric inhibitors that increase the conformational equilibrium in favor of this dysfunctional state.

■ ASSOCIATED CONTENT

Supporting Information

Details of the equilibration and simulation protocols used, along with a video showing the closing of the EFV-bound HIV-1 RT system and a PDB file of the closed EFV-bound enzyme. This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

p.v.coveney@ucl.ac.uk

Author Contributions

§D.W.W. and S.K.S. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the support of the U.S. TeraGrid/XSEDE (access to which was made available by the National Science Foundation under NRAC grant MCA04N014). Additional computation was performed on the UCL Legion cluster. D.W.W. and P.V.C. thank EPSRC, CoMPLEX, and the EU FP7 CHAIN project (HEALTH-2007-2.3.2-7) for their support. S.K.S. acknowledges support from a European Commission FP7 Marie Curie Intra-European Fellowship. G.D.F. acknowledges support from the Ramón y Cajal scheme, Acellera Ltd., and from the Spanish Ministry of Science and Innovation (Ref. BIO2011-27450).

■ REFERENCES

(1) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789−796.
(2) Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D. *Nature* **2007**, *450*, 913−916.
(3) Hilser, V. J. *Science* **2010**, *327*, 653−654.
(4) Wrabl, J. O.; Gu, J.; Liu, T.; Schrank, T. P.; Whitten, S. T.; Hilser, V. J. *Biophys. Chem.* **2011**, *159*, 129−141.
(5) Pan, H.; Lee, J. C.; Hilser, V. J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 12020−12025.

(6) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646−652.
(7) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, M. *Science* **2010**, *330*, 341−346.
(8) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. *Science* **2011**, *334*, 517−520.
(9) Sadiq, S. K.; Fabritiis, G. D. *Proteins* **2010**, *78*, 2873−2885.
(10) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. *J. Am. Chem. Soc.* **2010**, *132*, 1526−1528.
(11) Vandenberg, J. I.; Torres, A. M.; Campbell, T. J.; Kuchel, P. W. *Eur. Biophys. J.* **2004**, *33*, 89−97.
(12) Jacobo-Molina, A.; Arnold, E. *Biochemistry* **1991**, *30*, 6351−6356.
(13) Lawtrakul, L.; Beyer, A.; Hannongbua, S.; Wolschann, P. *Monatsh. Chem.* **2004**, *135*, 1033−1046.
(14) Sluis-Cremer, N.; Temiz, N. A.; Bahar, I. *Curr. HIV Res.* **2004**, *2*, 323−332.
(15) Bakan, A.; Bahar, I. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 14349−14354.
(16) Lindberg, J.; Sigurdsson, S.; Löwgren, S.; Andersson, H. O.; Sahlberg, C.; Noréen, R.; Fridborg, K.; Zhang, H.; Unge, T. *Eur. J. Biochem.* **2002**, *269*, 1670−1677.
(17) Hsiou, Y.; Ding, J.; Das, K.; Clark, A. D.; Hughes, S. H.; Arnold, E. *Structure* **1996**, *4*, 853−860.
(18) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(19) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(20) Duan, Y.; Wu., C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(21) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(22) Harvey, M. J.; Giupponi, G.; de Fabritiis, G. *J. Chem. Theory Comput.* **2009**, *5*, 1632−1639.
(23) Ivetac, A.; McCammon, J. A. *J. Mol. Biol.* **2009**, *388*, 644−658.
(24) Kensch, O.; Restle, T.; Wöhrl, B. M.; Goody, R. S.; Steinhoff, H. J. *J. Mol. Biol.* **2000**, *301*, 1029−1039.
(25) Spence, R. A.; Kati, W. M.; Anderson, K. S.; Johnson, K. A. *Science* **1995**, *267*, 988−993.
(26) Xia, Q.; Radzio, J.; Anderson, K. S.; Sluis-Cremer, N. *Protein Sci.* **2007**, *16*, 1728−1737.
(27) Clark, S. A.; Shulman, N. S.; Bosch, R. J.; Mellors, J. W. *AIDS* **2006**, *20*, 981−984.
(28) Haubrich, R. H.; Kemper, C. A.; Hellmann, N. S.; Keiser, P. H.; Witt, M. D.; Forthal, D. N.; Leedom, J.; Leibowitz, M.; Whitcomb, J. M.; Richman, D.; McCutchan, J. A.; Group, C. C. T. *AIDS* **2002**, *16*, F33−F40.
(29) Whitcomb, J. M.; Huang, W.; Limoli, K.; Paxinos, E.; Wrin, T.; Skowron, G.; Deeks, S. G.; Bates, M.; Hellmann, N. S.; Petropoulos, C. J. *AIDS* **2002**, *16*, F41−F47.
(30) Bakan, A.; Meireles, L. M.; Bahar, I. *Bioinformatics* **2011**, *27*, 1575−1577.
(31) Temiz, N. A.; Bahar, I. *Proteins* **2002**, *49*, 61−70.
(32) Seckler, J. M.; Howard, K. J.; Barkley, M. D.; Wintrode, P. L. *Biochemistry* **2009**, *48*, 7646−7655.