Allosteric dynamics of SAMHD1 studied by molecular dynamics simulations

K K Patra1, A Bhattacharya2 and S Bhattacharya1

1Department of Physics, Indian Institute of Technology Guwahati, Guwahati, Assam, India 781039
2Department of Biochemistry, MSC 7760, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX, USA 78229-3900

swaticb@iitg.ernet.in

Abstract. SAMHD1 is a human cellular enzyme that blocks HIV-1 infection in myeloid cells and non-cycling CD4+ T cells. The enzyme is an allosterically regulated triphosphohydrolase that modulates the level of cellular dNTP. The virus restriction is attributed to the lowering of the pool of dNTP in the cell to a point where reverse-transcription is impaired. Mutations in SAMHD1 are also implicated in Aicardi-Goutières syndrome. A mechanistic understanding of the allosteric activation of the enzyme is still elusive. We have performed molecular dynamics simulations to examine the allosteric site dynamics of the protein and to examine the connection between the stability of the tetrameric complex and the Allosite occupancy.

1. Introduction

The SAMHD1 is a cellular enzyme in humans with phosphohydrolase activity, i.e. it converts deoxynucleotide triphosphates (dNTPs) into nucleosides and inorganic triphosphate and in the process depletes the cellular pool of dNTPs thereby preventing viral replication. The role of SAMHD1 in blocking HIV-1 infection in dendritic cells has generated considerable interest in its activity and regulation[1–9].

Previous studies have suggested that the catalytically active protein is a complex allosteric tetramer[9,10]. Tetramerization can occur only in the presence of GTP and dNTPs. The enzyme has two putative allosteric sites and a catalytic site (see Figure 1 b). Allosite 1 selects specifically for GTP or dGTP. Five hydrogen bonds exist between the base edge of the guanine and residues D137, Q142 and R145. GTP can provide one extra hydrogen bond with V117 carbonyl group that is not possible for dGTP. Other nucleotides do not bind to Allosite 1. Allosite 2 selects specifically for dNTPs but not NTPs. The catalytic site can accommodate any dNTP with water mediated non-specific contacts between the dNTP base and the enzyme. X-ray crystallographic structure based analysis revealed that the binding affinity of the various dNTPs to Allosite 2 (dATP>dGTP>dTTP>dCTP) was in the reverse order as that to the catalytic site (dCTP>dGTP/dTTP>dATP). Under physiological conditions, Allosite 1 is occupied by GTP while any of the four dNTPs can bind to Allosite 2 although, under equal concentrations, dATP is most likely to occupy Allosite 2 due to highest affinity. Enzyme assays also indicate a complex matrix of activation cross-talk. The catalysis rate in single dNTP experiments (dA>dT>dC>dG) is exactly the opposite to that in mixed dNTP experiments dG>dC>dT>dA, i.e the rate of phosphohydrolysis at the catsite depends on which dNTP occupies Allosite 2. However, a
mechanistic understanding of the allosteric activation has been elusive. As a first step, we examine the question of how the presence or absence of the bound ligand influences the stability of the tetramer.

Figure 1.(a) A snapshot showing the SAMHD1 immersed in a box of water. The four subunits of the tetrameric complex are represented by violet, red, green and ochre ribbons. (b) Snapshot showing the allosteric pocket with the GTP and dATP molecules in stick representation. The protein is shown in ribbon representation with the different colors corresponding to different subunits (chain C in violet, B in red and D in green). The residues directly interacting with the GTP and the dATP molecules are shown in stick representation. (c) The RMSD (root mean square deviation) of the protein backbone in the five systems plotted as a function of time.

2. Results.
To examine the dynamics at the allosteric pocket of the SAMHD1 protein, we set up five systems based on the crystal structure (protein databank entry 4TNR ) with various occupancies of the Allosites and the catalytic sites summarized in Table 1. The five wt systems generated with various occupancies of the Allosites and the catsite were simulated for more than 200 ns each, with the aggregate MD simulation time exceeding 1µs.

| System | GTP at Allosite 1 | dATP at Allosite 2 | dATP at catalytic site | NVT simulation length (ns) |
|--------|------------------|--------------------|------------------------|---------------------------|
| 1      | Absent           | Absent             | Absent                 | 210                       |
| 2      | Absent           | Present            | Absent                 | 210                       |
| 3      | Present          | Absent             | Absent                 | 210                       |
| 4      | Present          | Present            | Absent                 | 210                       |
| 5      | Present          | Present            | Present (3 out of 4 Cat sites) | 210                      |

Table 1. List of the five systems simulated summarizing the occupancy/vacancy at the Allosites and the catalytic sites as well as the length of the MD trajectories generated.
An analysis of the root mean square deviation (RMSD) of the protein backbone (see Figure 1 (c)) shows that the two systems with both allosteric sites occupied (systems 4 and 5) having a consistently smaller RMSD compared to the remaining three. We therefore conclude that the binding of the nucleotides at both allosteric sites imparts stability to the tetrameric complex. The removal of either of the nucleotides at the Allosites results in an increased deviation of the protein backbone from the initial crystal structure.

In order to examine the local effects of the Allosite occupation, we calculated the RMSD of short segments of the protein containing the residues in direct interaction with the Allosite bound GTP/dATP molecule (see Figure 2). Four segments were studied: (i) Met115 to Pro130, (ii) Leu131 to Lys148, (iii) Gly151 to Pro158 and (iv) Asp330 to Val340. The first segment contains Val117 and Asn119, which are the residues in direct contact with Allosite 2 dATP. The second segment contains Asp137, Gln142 and Arg145 which interact with the GTP at Allosite 1. The third segment contains the residue Val156 that interact with dATP at Allosite 2. The fourth segment includes residue Arg333 that is has a stacking interaction with the dATP according to the crystal structure. The segments (iii) and (iv) belong to a different subunit compared to the residues of segments (i) and (ii) since the Allosites are at the interface of three different subunits. Hence the plots in Figure 2 are grouped according to the proximity to a given allosteric pocket. As expected, the RMSD of the segments in systems 4 and 5 is small indicating little deviation from the crystal structure. The most prominent deviations are observed in systems 3 and 1. In both these systems, dATP at Allosite 2 is missing. It is interesting to note that the absence of Allosite 2 dATP also causes a prominent increase of the RMSD in segment (ii) that interacts directly with Allosite 1. The results show that the vacancy at Allosite 2 causes local perturbations to the structure. A vacancy at Allosite 1 also causes increased perturbations (see panels (i) and (l) in Figure 2). However, the overall effect is not as large as that caused by Allosite 2 vacancy.
3. Discussion

We have studied the allosteric site dynamics of the SAMHD1 enzyme which is an HIV-1 restriction factor. Previous experimental studies have indicated a complex cross-talk between the allosteric sites in the protein, which is not understood as yet. A series of simulations designed to explore the effect of the presence/absence of the preferred Allosite 1 and 2 ligands show that the stability of the tetramer is adversely affected when either or both allosteric sites are vacant. The destabilizing effect is more pronounced when the vacancy is in Allosite 2, i.e., in the dATP-lost form of the protein. The dATP bound to Allosite 2 therefore appears to play a role in pinning the subunits together. A complex network of hydrogen bonds between the dATP/GTP and nearby protein residues are crucial to the stability of the complex. However the exact role of the two Allosite nucleotides in the catalysis remains to be explored.

4. Methods

4.1 Simulation System Setup The crystal structure of SAMHD1 complexed with GTP and dATP obtained from the Protein Data Bank entry 4TNR[11] was used to generate the starting conformation for the all atom MD simulations in an explicit water environment. Three of the four catalytic sites are occupied by dATP while the fourth (in subunit A) is vacant. The pdb structure was used to generate initial structures for five systems as shown in Table 1. In all five systems, the crystallographic waters were retained. The Mg+2 ions coordinated by allosteric site molecules were deleted in System 1. However, systems 4 and 5 contained both allosteric site molecules along with the Mg ions. The missing residues in the loop 278-283 were inserted in the protein structures whereas the missing N terminal and C terminal residues were ignored. The four R206 and N207 residues were mutated back to histidine and aspartate in accord with the sequence of the wt SAMHD1 (Uniprot Q9Y3Z3-1). Disulfide bonds were introduced between residues 341 and 350. Each system was solvated in a cubic waterbox with in ~59000 pre-equilibrated TIP3P water molecules. An appropriate number of sodium ions were added to produce a neutral system. Each system consisted of ~210,000 atoms measured 13×12×14 nm.

4.2 General MD Methods. Energy minimizations and molecular dynamics simulations were performed using NAMD 2.9[12,13]. All simulations employed periodic boundary conditions and multiple time-stepping wherein local interactions were calculated every 2 fs and full electrostatic evaluations were performed every two timesteps. The particle mesh Ewald[14] method was adopted for treating long range electrostatic calculations. CHARMM31[15,16] force fields were employed along with the TIP3P water model. During simulations, the SETTLE[17] and RATTLE[18] algorithms were applied to keep the covalent bonds involving hydrogen in water and other molecules rigid. A cut-off of 12 Å and a switching distance of 10 Å were used for treating van der Waals and short-ranged interactions. Conjugate gradient minimization was performed for 3000 steps for each system following which each system was equilibrated in the NPT ensemble using the Nosé-Hoover Langevin piston pressure control at 295 K for at least 5 ns. Finally NVT simulations were performed with the temperature maintained at 295 K using the Langevin thermostat. The coordinates was recorded at 10 ps intervals. Each of the five systems were simulated for about 210 ns bringing the total simulation time to 1.2 μs.
Acknowledgements  We are deeply thankful to the National PARAM Supercomputing facility (NPSF), CDAC, for providing us the computational facilities for this study.

References

[1] Behrendt R, Schumann T, Gerbaulet A, Nguyen L a., Schubert N, Alexopoulou D, Berka U, Lienenklaus S, Peschke K, Gibbert K, Wittmann S, Lindemann W, Weiss S, Dahl A, Naumann R, Dittmer U, Kim B, Mueller W, Gramberg T and Roers A 2013 Mouse SAMHD1 has antiretroviral activity and suppresses a spontaneous cell-intrinsic antiviral response Cell Rep. 4 689–96

[2] Hollenbaugh J a, Tao S, Lenzi G M, Ryu S, Kim D-H, Diaz-Griffero F, Schinazi R F and Kim B 2014 dNTP pool modulation dynamics by SAMHD1 protein in monocyte-derived macrophages. Retrovirology 11 63

[3] Hollenbaugh J a, Gee P, Baker J, Daly M B, Amie S M, Tate J, Kasai N, Kanemura Y, Kim D-H, Ward B M, Koyanagi Y and Kim B 2013 Host factor SAMHD1 restricts DNA viruses in non-dividing myeloid cells. PLoS Pathog. 9 e1003481

[4] Yan N and Lieberman J 2013 SAMHD1 does it again, in resting T cells Nat. Med. 18 1199–216

[5] Cribier A, Descours B, Valadão A, Laguette N and Benkirane M 2013 Phosphorylation of SAMHD1 by Cyclin A2/CDK1 Regulates Its Restriction Activity toward HIV-1 Cell Rep. 3 1036–43

[6] Baldauf H-M, Pan X, Erikson E, Schmidt S, Daddacha W, Burggraf M, Schenkova K, Ambiel I, Wabnitz G and Gramberg T 2012 SAMHD1 restricts HIV-1 infection in resting CD4+ T cells Nat. Med. 18 1682–9

[7] Beloglazova N, Flick R, Tchigvintsev A, Brown G, Popovic A, Nocek B and Yakunin A F 2013 Nuclease activity of the human SAMHD1 protein implicated in the aicardi-goutieres syndrome and HIV-1 restriction J. Biol. Chem. 288 8101–10

[8] Seamon K J, Sun Z, Shlyakhtenko L S, Lyubchenko Y L and Stivers J T 2015 SAMHD1 is a single-stranded nucleic acid binding protein with no active site-associated nuclease activity Nucleic Acids Res. 1–14

[9] Ji X, Wu Y, Yan J, Mehrens J, Yang H, DeLucia M, Hao C, Gronenborn A M, Skowronsksi J, Ahn J and Xiong Y 2013 Mechanism of allosteric activation of SAMHD1 by dGTP Nat. Struct. Mol. Biol. 20 1304–9

[10] Yan J, Kaur S, DeLucia M, Hao C, Mehrens J, Wang C, Golczak M, Palczewski K, Gronenborn A M, Ahn J and Skowronsksi J 2013 Tetramerization of SAMHD1 is required for biological activity and inhibition of HIV infection. J. Biol. Chem. 288 10406–17

[11] Ji X, Tang C, Zhao Q, Wang W and Xiong Y 2014 Structural basis of cellular dNTP regulation by SAMHD1 Proc. Natl. Acad. Sci. 111 E4305–14

[12] Kal L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, Phillips J, Shinozaki A, Varadarajan K and Schulten K 1999 NAMD2 : Greater Scalability for Parallel Molecular Dynamics J. Comput. Phys. 151 283–312

[13] Phillips J C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R D and Kalé L 2005 Scalable Molecular Dynamics with NAMD J. Comput. Chem. 26 1781–802

[14] Batcho P F, Case D a. and Schlick T 2001 Optimized particle-mesh Ewald/multiple-time step integration for molecular dynamics simulations J. Chem. Phys. 115 4003–18

[15] Mackerell A D, Bashford D, Bellott M, Dunbrack R L, Evansenck J D, Field M J, Fischer S, Gao J, Guo H, Ha S, Kuczner L, Kuczera K, Lau F T K, Mattos C, Michnick S, Ngo T, Nguyen D T, Prodhom B, Reider W E, Roux B, Schlenkrich M, Smith J C, Stote R, Straub J, Watanabe M, Wio J, Yim D and Karplus M 1998 All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins † J. Phys. Chem. B 5647 3586–616

[16] Klauda J B, Venable R M, Freites J A, O’Connor J W, Tobias D J, Mondragon-Ramirez C,
Vorobyov I, MacKerell A D and Pastor R W 2010 Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114** 7830–43

[17] Miyamoto S and Kollman P a. 1992 Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models *J. Comput. Chem.* **13** 952–62

[18] Andersen C 1983 Rattle: A “Velocity” Molecular Version of the Shake Dynamics Calculations for *J. Comput. Phys.* **52** 24–34