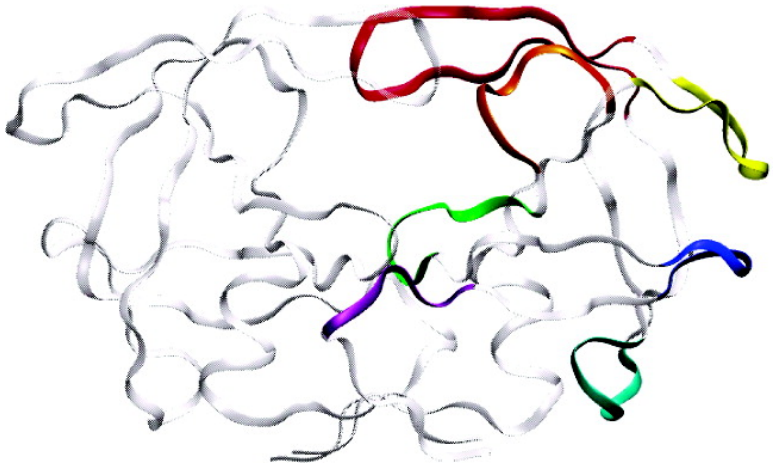
Communication

Fast Peptidyl cis–trans Isomerization within the Flexible Gly-Rich Flaps of HIV-1 Protease

Donald Hamelberg, and J. Andrew McCammon

J. Am. Chem. Soc., **2005**, 127 (40), 13778-13779• DOI: 10.1021/ja054338a • Publication Date (Web): 16 September 2005



J|A|C|S
COMMUNICATIONS

Published on Web 09/16/2005

Fast Peptidyl cis−trans Isomerization within the Flexible Gly-Rich Flaps of HIV-1 Protease

Donald Hamelberg* and J. Andrew McCammon

Howard Hughes Medical Institute, Center for Theoretical Biological Physics, Department of Chemistry and Biochemistry, Department of Pharmacology, University of California at San Diego, La Jolla, California 92093

Received June 30, 2005; E-mail: dhamelbe@mccammon.ucsd.edu

Human immunodeficiency virus (HIV), the retrovirus that causes AIDS, infects millions of people around the world. Several drugs have been developed to inhibit the actions of the most common HIV-1 virus by directly targeting its proteolytic machinery.[1] The HIV-1 protease is a homodimer with two flaps controlling access to the active site. The highly conserved flap tips sequence Ile47-Gly-Gly-Ile-Gly-Gly-Phe53 forms a β-hairpin loop[2,3] and is known to be flexible.[4] However, knowledge of the atomistic mechanism of the flexibility is lacking. The occasionally observed G48V mutation leads to some loss in catalytic activity.[3] We show that the Gly−Gly ω-bond in the flap tips undergoes fast cis−trans isomerization on the microsecond to millisecond time scale rather than on the usual seconds. Further study reveals that the fast isomerization is a direct consequence of the β-hairpin loop structure of the flap tips, which appears to be counterintuitive. After loop formation of a linear peptide containing the Gly−Gly motif, the rate of isomerization is shown to increase by many orders of magnitude.

In order for HIV-1 protease to carry out its function, the flaps have to be apart to allow the protein substrate into the active site and wrap around its substrate during proteolysis. The flexibility and mobility of the flaps are therefore vital to the survival of this virus and could therefore be taken advantage of in drug design. However, the established experimental techniques used to study protein structures fall a little short of providing a detailed atomistic dynamical mechanism of the flaps.

Crystal structures of the HIV-1 protease solved to date comprise closed and semi-open flap conformations.[2] These observed conformations, however, prevent substrates from entering the active site. Naturally, the flaps of the conformations poised to bind a protein substrate should be more flexible and farther apart in solution. Nuclear magnetic resonance experiments indicated that the flaps of free HIV-1 protease are mobile, and that the β-sheet structure at the tips is very irregular.[4] Flaps' conformational changes on the millisecond to subnanosecond time scale were observed. Conversely, the flap tip residues in the free protease that are flexible on the subnanosecond time scale are not flexible on this time scale in the bound protease.

The motions of the flaps have also been previously explored with molecular dynamics (MD) simulations at atomistic detail.[5] However, the motions that could be observed are on the order of nanoseconds for conventional MD. Consequently, none of the unrestrained MD simulations starting with the closed conformation was able to observe the opening of the flaps.

It would therefore be helpful to study the long time scale motions involved in the flaps of HIV-1 protease using simulations with atomistic resolution to fully understand the mechanism of its flexibility. To overcome the time scale limitation of conventional MD, we developed an accelerated molecular dynamics approach,[6] which we have used to carry out simulations on the HIV-1 protease in explicit water. This method, which is based on Voter's hyperdynamics scheme,[7] allows the system to evolve on a modified energy landscape and accelerates the escape of conformations trapped in potential energy wells. Previously, we have shown that this method samples the conformational space efficiently, and after re-weighting each conformation, the correct canonical distribution, as well as the kinetic information,[8] can be recovered.

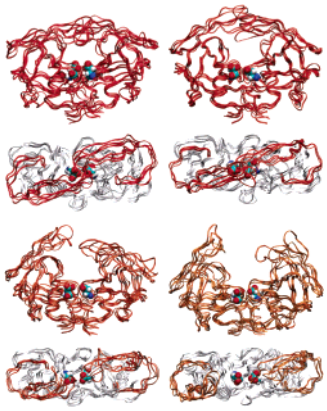


Figure 1. Snapshots showing the opening of the flaps. Within each of the four panels, a side view is on top and a view from above is on the bottom. The aspartic acids in the active site are shown using van der Waals spheres.

We started the free wild-type simulation using a crystal structure of an inhibitor-bound HIV-1 protease complex[9] with the inhibitor removed, thus having it in the closed conformation. This conformation was used so as to observe the flap opening mechanism, as shown in Figure 1.

The tips of the flaps curl in and out, and the whole flaps move from the closed to the semi-open and then to an open conformation. The flaps then asymmetrically curl in and out and now and again maintain an opening that is large enough to allow a substrate into the active site. We observe that the binding site and the catalytic aspartic dyad are intact during the whole simulation, despite the large motions of the flaps (Supporting Figure S1). Furthermore, we carried out simulations of the substrate-bound HIV-1 protease; the flaps stay directly over the peptide substrate in the closed conformation. Experimental studies have shown that substrate-like ligands bind to HIV-1 protease in a two-step process consisting of a relatively fast binding followed by slow rate-limiting conformational rearrangements.[10] Therefore, it could be inferred from these results that, upon binding of a substrate to a fully open conformation, the flaps would then undergo conformational rearrangements that would result in the closed conformation.

Seven flexible regions of the free HIV-1 protease were observed from the positional fluctuation of the backbone atoms, with the flap tips having the most mobility (Supporting Figure S2). Four regions are around the active site, while the others are loops on the periphery of the protein. The conserved glycines in the flaps

10.1021/ja054338a CCC: $30.25

are very flexible and sample almost all of the allowed $\varphi-\psi$ space (Supporting Figure S3) while providing inward and outward curling of the flap tips. Interestingly, all but one of the observed flexible domains contain one or more Gly residues.

In addition to the extensive $\varphi-\psi$ space sampled by the flexible Gly-rich flaps, we also observe frequent trans−cis isomerization of the Gly−Gly ω-bonds in the flap tips. This occurs both before and after the opening event. We mainly observed frequent trans−cis isomerization for the flexible regions that contain one or more Gly residues (Supporting Figure S4). The regions with the Gly−Gly motif, for example, the flaps and residues 15−18, undergo more frequent isomerization than the other flexible regions. The majority of the ω-bonds of the entire HIV-1 protease do not undergo a single isomerization during the simulation. Therefore, it appears that the time scale of trans−cis isomerization for these Gly-rich ω-bonds is shorter than that of regular ω-bonds, thus contributing to the flexibility of the flaps. We also carried out a simulation of the free G48V mutant, and the flaps took about 10 times longer to open compared to the wild-type HIV-1 protease. The results show that the $\varphi-\psi$ space sampled by V48 is much less than that sampled by G48 in the wild-type. Also, the ω-bond undergoes less frequent isomerization compared to that of the wild-type. This presumably renders the G48V mutant less catalytically active than the wild-type. Also, the other Gly residues in the flaps have been shown to be nearly intolerant of any substitution.[11] The flexibility of the Gly-rich flap tips contributes to the opening and motions of the flaps. Therefore, reduction of the flexibility of the flap tips by an inhibitor that interacts more favorable than the natural substrate would stabilize the closed conformation and potentially act as a drug.

The time scale of the trans−cis isomerization of the Gly−Gly ω-bond is typically very slow and on the order of seconds,[11] which is in the vicinity of that of other residues. As described below, however, additional simulations suggest that constraining the Gly−Gly motif in a loop increases the rate of cis−trans isomerization by many orders of magnitude. Therefore, the constrained β-hairpin loop at the flap tips results in the rapid isomerization and increased flexibility of Gly−Gly.

To study the detailed cis−trans isomerization mechanism of the Gly−Gly motif, we carried out additional simulations on a linear peptide, Ace-Ile-Gly-Gly-Ile-Nme, and three cyclic peptides, Ala-Ile-Gly-Gly-Ile, Ile-Gly-Gly-Ile, and Ala-Gly-Gly, of length 5, 4, and 3 amino acids, respectively. These peptide sequences represent residues within the flap tips of the HIV-1 protease. We observed cis−trans isomerization of the Gly−Gly ω-bond (Supporting Figure S5) and, using Gaussian kernel density estimation,[12] calculated the free energy profile along this bond (Figure 2). The activation free energy, $\Delta G^{\ddagger}$, of the Gly−Gly ω-bond of the linear peptide is calculated to be around 17.9 kcal/mol, similar to that in experiments.[11] We observed that going from the linear peptide to the cyclic peptides, $\Delta G^{\ddagger}$ decreases by as much as 18 kT. $\Delta G^{\ddagger}$ decreases by ~7 kT going from the linear to the cyclic pentapeptide, then by an additional ~4 kT going to the cyclic tetrapeptide, and a further ~7 kT for the cyclic tripeptide.

Using Kramers' kinetic rate theory, we estimate the rate of trans−cis isomerization of the Gly−Gly ω-bond of the cyclic tetrapeptide to be about 10^5 times faster than that of the linear form. The rate is even faster for the cyclic tripeptide. Therefore, since the cis−trans isomerization of the linear Gly−Gly occurs on the order of seconds, the time would decrease to around the microsecond to millisecond time scale due to loop formation. A similar effect on isomerization of X-Pro ω-bonds in cyclic peptides has previously been observed.[13]

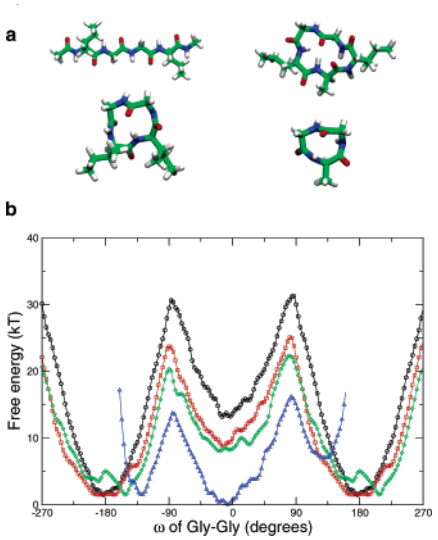


Figure 2. (a) Stick models of the linear and cyclic peptides simulated are also shown. (b) Free energy profile of the Gly−Gly ω-bond of the linear (black) and cyclic peptides of sizes 5 (red), 4 (green), and 3 (blue) amino acids.

Acknowledgment. This work was supported in part by grants from NSF, NIH, CTBP, NBCR, SDSC, and Accelrys, Inc. We thank Drs. N. Johnson, T. Shen, A. Perryman, and C. Chang for helpful discussions.

Supporting Information Available: Computational methods, snapshots of entire simulation, and variation of $\varphi-\psi$ and ω angles. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Wlodawer, A.; Erickson, J. W. *Annu. Rev. Biochem.* **1993**, *62*, 543−585.
(2) (a) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider, J.; Kent, S. B. H. *Science* **1989**, *245*, 616−621. (b) Lapatto, R.; Blundell, T.; Hemmings, A.; Overington, J.; Wilderspin, A.; Wood, S.; Merson, J. R.; Whittle, P. J.; Danley, D. E.; Geoghegan, K. F.; Hawrylik, S. J.; Lee, S. E.; Scheld, K. G.; Hobart, P. M. *Nature* **1989**, *342*, 299−302.
(3) Mahalingam, B.; Louis, J. M.; Reed, C. C.; Adomat, J. M.; Krouse, J.; Wang, Y. F.; Harrison, R. W.; Weber, I. T. *Eur. J. Biochem.* **1999**, *263*, 238−245.
(4) (a) Ishima, R.; Freedberg, D. I.; Wang, Y. X.; Louis, J. M.; Torchia, D. A. *Struct. Fold Des.* **1999**, *7*, 1047−1055. (b) Freedberg, D. I.; Ishima, R.; Jacob, J.; Wang, Y. X.; Kustanovich, I.; Louis, J. M.; Torchia, D. A. *Protein Sci.* **2002**, *11*, 221−232.
(5) (a) Scott, W. R.; Schiffer, C. A. *Struct. Fold Des.* **2000**, *8*, 1259−1265. (b) Perryman, A. L.; Lin, J. H.; McCammon, J. A. *Protein Sci.* **2004**, *13*, 1108−1123. (c) Collins, J. R.; Stanley, K. B.; Erickson, J. W. *Nat. Struct. Biol.* **1995**, *2*, 334−338.
(6) (a) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919. (b) Hamelberg, D.; Shen, T.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 1969−1974.
(7) (a) Voter, A. F. *Phys. Rev. Lett.* **1997**, *78*, 3908−3911. (b) Voter, A. F. *J. Chem. Phys.* **1997**, *106*, 4665−4677.
(8) Hamelberg, D.; Shen, T.; McCammon, J. A. *J. Chem. Phys.* **2005**, *122*, 241103.
(9) Reiling, K. K.; Endres, N. F.; Dauber, D. S.; Craik, C. S.; Stroud, R. M. *Biochemistry* **2002**, *41*, 4582−4594.
(10) Furfine, E. S.; Dsouza, E.; Ingold, K. J.; Leban, J. J.; Spector, T.; Porter, D. J. T. *Biochemistry* **1992**, *31*, 7886−7891.
(11) (a) Shao, W.; Everitt, L.; Manchester, M.; Loeb, D. D.; Hutchison, C. A.; Swanstrom, R. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2243−2248. (b) Schiene-Fischer, C.; Fischer, G. *J. Am. Chem. Soc.* **2001**, *123*, 6227−6231.
(12) Silverman, B. W. *Density Estimation for Statistics and Data Analysis*; Chapman and Hall: London, New York, 1986; p 175.
(13) Shi, T.; Spain, S. M.; Rabenstein, D. L. *J. Am. Chem. Soc.* **2004**, *126*, 790−796.

JA054338A