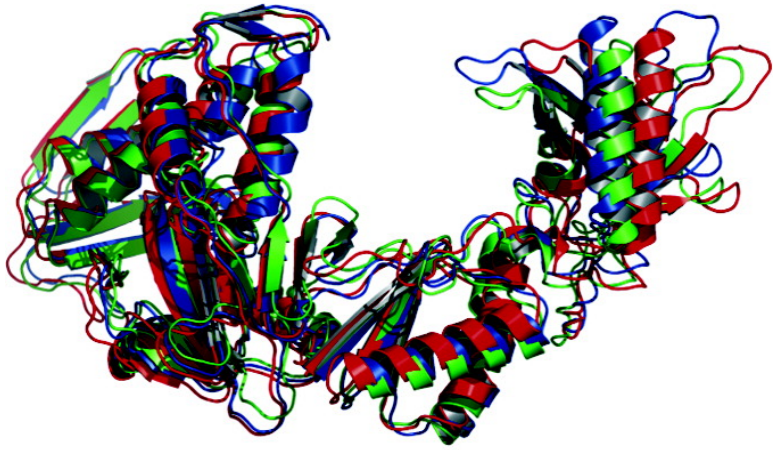
Article

Effect of a Bound Non-Nucleoside RT Inhibitor on the Dynamics of Wild-Type and Mutant HIV-1 Reverse Transcriptase

Zhigang Zhou, Marcela Madrid, Jeffrey D. Evanseck, and Jeffry D. Madura

J. Am. Chem. Soc., **2005**, 127 (49), 17253-17260• DOI: 10.1021/ja053973d • Publication Date (Web): 17 November 2005



J|A|C|S
ARTICLES
Published on Web 11/17/2005

Effect of a Bound Non-Nucleoside RT Inhibitor on the Dynamics of Wild-Type and Mutant HIV-1 Reverse Transcriptase

Zhigang Zhou,[†] Marcela Madrid,[‡] Jeffrey D. Evanseck,[†] and Jeffry D. Madura*[,†]

Contribution from the Department of Chemistry and Biochemistry and Center for Computational Sciences, Duquesne University, Pittsburgh, Pennsylvania 15282, and Pittsburgh Supercomputing Center, Pittsburgh, Pennsylvania 15213

Received June 15, 2005; E-mail: madura@duq.edu

Abstract: HIV-1 reverse transcriptase (RT) is an important target for drugs used in the treatment of AIDS. Drugs known as non-nucleoside RT inhibitors (NNRTI) appear to alter the structural and dynamical properties of RT which in turn inhibit RT's ability to transcribe. Molecular dynamics (MD), principal component analysis (PCA), and binding free energy simulations are employed to explore the dynamics of RT and its interaction with the bound NNRTI nevirapine, for both wild-type and mutant (V106A, Y181C, Y188C) RT. These three mutations commonly arise in the presence of nevirapine and result in resistance to the drug. We show that a bound NNRTI hinders the motion of almost all RT amino acids. The mutations, located in the non-nucleoside RT inhibitor binding pocket, partially restore RT flexibility. The binding affinities calculated by molecular mechanics/Poisson−Boltzmann surface accessibility (MM-PBSA) show that nevirapine interacts stronger with wild-type RT than with mutant RT. The mutations cause a loss of van der Waals interactions between the drug and the binding pocket. The results from this study suggest that a good inhibitor should efficiently enter and maximally occupy the binding pocket, thereby interacting effectively with the amino acids around the binding pocket.

Introduction

The Acquired Immunodeficiency Syndrome (AIDS) is one of the most serious threats to human health. Important advances in its treatment have been made with the discovery of new drugs and the administration of combinations of drugs. The Reverse Transcriptase of the Human Immunodeficiency virus type 1 (HIV-1 RT) is one of the main targets of drugs used in the treatment of AIDS. Several RT inhibitors have been developed and approved by the FDA and are in clinical use.[1,2] In particular, the non-nucleoside RT inhibitors (NNRTIs) are highly effective and produce few side effects. However, RT mutations rapidly emerge that confer resistance to all known NNRTIs, reducing the efficiency of the drugs.[3,4] Therefore, it is essential to understand the detailed interactions of the inhibitors with wild-type and mutant RT, to design drugs that are effective across mutations.

HIV-1 RT is part of the HIV capsid and has an essential role in the replication of the AIDS virus. It transcribes the single-stranded RNA genome of the AIDS retrovirus to a double-stranded DNA. RT is a protein dimer consisting of two subunits of 66 kDa (p66) and 51 kDa (p51).[4−9] The p66 and p51 subunits are composed of four subdomains called thumb, palm, fingers, and connection. The p66 subunit also contains the RNase H subdomain. The polymerase active site is located in the p66 palm and contains three aspartic acids (Asp110, Asp185, Asp186). The p66 fingers, palm, and thumb subdomains form a large cleft that binds the template-primer DNA.[10]

NNRTIs bind to a common hydrophobic site, the non-nucleoside inhibitor binding pocket (NNIBP), located in the p66 palm subdomain approximately 10 Å away from the polymerase active site.[3,4,11−14] Mutations that make RT resistant to NNRTIs are known to arise for most of the amino acids that comprise

[†] Duquesne University.
[‡] Pittsburgh Supercomputing Center.

(1) Koup, R. A.; Merluzzi, V. J.; Hargrave, K. D.; Adems, J.; Grozinger, K.; Eckner, R. J.; Saivan, J. L. *J. Infect. Dis.* **1991**, *163*, 966−970.
(2) Richman, D.; Rosenthal, A. S.; Shoog, M.; Eckner, R. J.; Chou, T. C.; Sabo, J. P.; Merluzzi, V. J. *Antimicrob. Agents Chemother.* **1991**, *35*, 305−308.
(3) De Clercq, E. *Clin. Microbiol. Rev.* **1995**, *8*, 200−239.
(4) Tantillo, C.; Ding, J.; Jacobo-Molina, A.; Nanni, R. G.; Boyer, P. L.; Hughes, S. H.; Pauwels, R.; Andries, K.; Janssen, P. A. J.; Arnold, E. *J. Mol. Biol.* **1994**, *243*, 369−387.
(5) Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Science* **1992**, *256*, 1783−1790.
(6) Smerdon, S. J.; Jager, J.; Wang, J.; Kohlstaedt, L. A.; Chirino, A. J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3911−3915.
(7) Ding, J.; et al. *Structure* **1995**, *3*, 365−379.
(8) Ding, J.; Das, K.; Moereels, H.; Koymans, L.; Andries, K.; Janssen, P. A.; Hughes, S. H.; Arnold, E. *Nat. Struct. Biol.* **1995**, *2*, 407−415.
(9) Jacobo-Molina, A.; Ding, J.; Nanni, R. G.; Jr., A. D. C.; Lu, X.; Tamtilo, C.; Williams, R. L.; Kamer, G.; Ferris, A. L.; Clark, P.; Hizi, A.; Hughes, S. H.; Arnold, E. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 6320−6324.
(10) Huang, H.; Chopra, R.; Verdine, G. L.; Harrison, S. C. *Science* **1998**, *282*, 1669−1675.
(11) Larder, B. A. In *Reverse transcriptase*; Skalka, A. M., G. S., Eds.; Cold spring Harbor Laboratory Press: Plainview, NY, 1993; pp 205−222.
(12) Larder, B. A. In *Reverse Transcriptase*; Skalka, A. M., G. S., Eds.; Cold spring Harbor Laboratory Press: Plainview, NY, 1993; pp 163−191.
(13) Ding, J.; Das, K.; Hsiou, Y.; Zhang, W.; Arnold, E.; Yadav, P. N. S.; Hughes, S. H. *Struct.-Based Drug Des.* **1997**, 41−82.
(14) Pedersen, O. S.; Pedersen, E. B. *Antiviral Chem. Chemother.* **1999**, *10*, 285−314.

the binding pocket. Crystal structures of RT complexed with different NNRTIs[8,15−18] and our own docking calculations[19] of different NNRTIs show that structurally homologous inhibitors have similar binding patterns and interaction modes to RT. Although the overall shape of the NNRTI binding site is similar among the different crystal structures of RT/NNRTI complexes, there are subtle differences among them, showing that the binding pocket is flexible and can adapt to the shape of different NNRTIs.[20−22] Amino acids in the NNIBP, such as Tyr 181, Tyr 188, Val 179, and Phe 227, interact with the bound NNRTI via van der Waals interactions. Hydrogen bonds can form between some inhibitors and amino acids Val 189 and Tyr 318. Water molecules also can form a hydrogen bond bridge network between the inhibitor and amino acids at the mouth of the NNIBP.[19]

Crystallographic studies have shown that the binding of an NNRTI induces several conformational changes.[23] While the p66 thumb subdomain of unliganded RT is in a "closed" configuration almost touching the p66 fingers subdomain,[20,24] upon ligation of an NNRTI the p66 thumb subdomain moves to an "open" or upright position.[5] Other structural changes induced by the bound NNRTI include displacements of the base of the p66 thumb, the p66 connection, and RNase H subdomains[20] and of the region known as the "primer grip".[4,7,8,17,20,24,25] The primer grip positions the primer terminus near the polymerase active site.[9] Smaller structural changes include the reorientation of side chains to form the non-nucleoside inhibitor binding pocket (the NNIBP does not exist in the crystal structures of RT without an NNRTI).[4,7,8,24] There are also differences at the catalytic site between the crystallographic structures of RT/DNA and RT/NNRTI.[7,8]

Based on structural and biochemical information, three models have been proposed for the mechanisms of inhibition of RT by NNRTIs (for a review, see Sarafianos et al.[23]). The model known as "molecular arthritis" postulates that the bound NNRTI impairs the mobility of the p66 thumb subdomain.[5] The "primer grip" and the "active site" models propose that the bound NNRTI distorts the exact configuration of the primer grip[4,7,8,20,24] or of the catalytic aspartic acids,[5,26] respectively. Biochemical studies have shown that NNRTIs block the chemical step of the polymerization reaction, the formation of the phosphodiester bond.[23,27−29] The "molecular arthritis" model does not appear adequate to explain the inhibition of this chemical step.[23] However, the bound NNRTI could inhibit RT by a combination of structural and dynamical factors.[18]

Comparison of crystallographic structures and site-directed spin labeling experiments[30] have shown that RT has regions of extreme flexibility, and it has been proposed that this flexibility is essential for the polymerization process.[7−9,20,31] Molecular dynamics (MD) simulations of unliganded RT and RT complexed to double-stranded DNA have shown that the flexibility of RT depends on its ligation state, increasing upon DNA binding.[32,33] The two ligation states studied (unliganded and complexed to DNA) presented different patterns of concerted motions. The motions of the amino acids that form the non-nucleoside binding pocket upon binding of the NNRTI are anticorrelated to the p66 fingers (in the RT/DNA complex) and correlated to the RNase H subdomain (in unliganded RT). Studies using the Gaussian network model have also shown that RT has the potential to undergo large concerted motions.[34,35] Recently, targeted MD simulations, with a water shell around the NNIBP, have been used to study the conformational changes during the association/dissociation of ligand from a K103N mutant. A series of 0.5 ns MD simulations showed that the hydrogen bond formed between N103 and Y188 plays a role on inhibitor entry.[36,37]

Motivated by the important role that flexibility appears to play on RT function and on the inhibition of RT by NNRTIs, in this study we present molecular dynamics simulations of RT complexed to an NNRTI, solvated with an explicit water box, with periodic symmetry. Simulations were performed for 2.5 ns on both wild-type (WT) and mutant RT (MT) complexed to the NNRTI nevirapine (Nev). Multi-nanosecond simulations with explicit water and periodic symmetry conditions are essential to explore the dynamics of protein affected by ligation. The mutations were simulated by changing three amino acids in the binding pocket (V106A, Y181C, and Y188C). These mutations were chosen because they commonly arise in the presence of nevirapine and result in resistance to the drug.[38−49]

(15) Wang, J.; Smerdon, S. J.; Jager, J.; Kohlstaedt, L. A.; Rice, P. A.; Friedman, J. M.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7242.
(16) Ren, J.; Esnouf, R.; Garman, E.; Somers, D.; Ross, C.; Kirby, I.; Keeling, J.; Darby, G.; Jones, Y.; Stuart, D.; Stammers, D.; et al. *Nat. Struct. Biol.* **1995**, *2*, 293.
(17) Das, K.; Ding, J.; Hsiou, Y.; Clark, A. J.; Moereels, H.; Koymans, L.; Andries, K.; Pauwels, R.; Janssen, P.; Boyer, P.; Clark, P.; Smith, R. J.; Kroeger, S. M.; Michejda, C.; Hughes, S.; Arnold, E. *J. Mol. Biol.* **1996**, *264*, 1085−1100.
(18) Ding, J.; Das, K.; Ydev, P. N. S.; Hsiou, Y.; Zhang, W.; Hughes, S. H.; Arnold, E. In *Structure Based Drug Design*; P., W., Ed.; Marcel Dekker: New York, 1997; pp 41−82.
(19) Zhou, Z.; Madrid, M.; Madura, J. D. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 529−542.
(20) Hsiou, Y.; Ding, J.; Das, K.; Jr., A. D. C.; Hughes, S. H.; Arnold, E. *Structure* **1996**, *4*, 853−860.
(21) Jager, J.; Smerdon, S. J.; Wang, J.; Boisvert, D. C.; Steitz, T. A. *Structure* **1994**, *2*, 869−876.
(22) Ding, J.; Hughes, S. H.; Arnold, E. *Biopolymers* **1997**, *44*, 125−138.
(23) Sarafianos, S. G.; Das, K.; Ding, J.; Hsiou, Y.; Hughes, S. H.; Arnold, E. Current perspectives on mechanisms of HIV-1 reverse transcriptase inhibition by nonnucleoside inhibitors. In *Anti-Infectives: Recent Advances in Chemistry and Structure-Activity Relationships*; Bentley, P. H., O'Hanlon, P. J., Eds.; 1997; Vol. 198, pp 328−334.
(24) Rodgers, D. W.; Gamblin, S. J.; Harris, B. A.; Ray, S.; Culp, J. S.; Hellmig, B.; Woolf, D. J.; Debouck, C.; Harrison, S. C.; Ray, S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1222−1226.
(25) Kroeger Smith, M. B.; Hose, B. M.; Hawkins, A.; Lipchock, J.; Farnsworth, D. W.; Rizzo, R. C.; Tirado-Rives, J.; Arnold, E.; Zhang, W.; Hughes, S. H.; Jorgensen, W. L.; Michejda, C. J.; Smith, R. H., Jr. *J. Med. Chem.* **2003**, *46*, 1940−1947.
(26) Ren, J.; Esnouf, R.; Hopkins, A.; Ross, C.; Jones, Y.; Stammers, D.; Stuart, D. *Structure* **1995**, *3*, 915−926.
(27) Rittinger, K.; Divita, G.; Goody, R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 8046−8049.
(28) Spence, R. A.; Kati, W. M.; Anderson, K. S.; Johnson, K. A. *Science* **1995**, *267*, 988−993.
(29) Spence, R. A.; Anderson, K. S.; Johnson, K. A. *Biochemistry* **1996**, *35*, 1054−1063.
(30) Kensch, O.; Restle, T.; Wohrl, B. M.; Goody, R. S.; Steinhoff, H.-J. *J. Mol. Biol.* **2000**, *301*, 1029−1039.
(31) Hermann, T.; Heumann, H. *Eur. J. Biochem.* **1996**, *242*, 98−103.
(32) Madrid, M.; Jacobo-Molina, A.; Ding, J.; Arnold, E. *Protein: Struct., Funct., Genet.* **1999**, *35*, 332−337.
(33) Madrid, M.; Lukin, J. A.; Madura, J. D.; Ding, J.; Arnold, E. *Proteins: Struct., Funct., Genet.* **2001**, *45*, 176−182.
(34) Bahar, I.; Erman, B.; Jernigan, R. L.; Atilgan, A. R.; Covell, D. G. *J. Mol. Biol.* **1999**, *285*, 1023−1037.
(35) Temiz, N. A.; Bahar, I. *Proteins: Struct., Funct., Genet.* **2002**, *49*, 61−70.
(36) Rodríguez-Barrios, F.; Balzarini, J.; Gago, F. *J. Am. Chem. Soc.* **2005**, *127*, 7570−7578.
(37) Rodríguez-Barrios, F.; Gago, F. *J. Am. Chem. Soc.* **2004**, *126*, 15386−15387.
(38) Richman, D. D.; Havlir, D.; Corbeil, J.; Looney, D.; Ignacio, C.; Spector, S. A.; Sullivan, J.; Cheeseman, S.; Barringer, K.; Pauletti, D. *J. Virol.* **1994**, *68*, 1660−1666.
(39) Richman, D. D. *Antimicrob. Agents Chemother.* **1993**, *37*, 1207−1213.
(40) Richman, D. D. *Annu. Rev. Pharmacol. Toxicol.* **1993**, *33*, 149−164.
(41) Richman, D. D. *AIDS Res. Rev.* **1992**, *2*, 237−248.

RT without nevirapine was also simulated for comparison purposes. The effect that a bound NNRTI has on the dynamics of wild-type (WT) and mutant RT was investigated by comparing the three molecular dynamics simulations using principal component analysis (PCA) methods. Free energy simulations of nevirapine with wild-type and mutant RT were done to identify the intermolecular interactions important upon binding.

Methods

Preparation of the Initial Structures. The starting coordinates for the RT/nevirapine complex were obtained from the Protein Data Bank (www.rcsb.org), PDB code 1VRT (2.2 Å resolution).[16] Hydrogen atoms were added to the nevirapine molecule with the Molecular Operating Environment (MOE) software.[50] Partial charges for nevirapine were calculated with AM1[51] in MOPAC7. The calculated charges were fitted into the AM1-bcc type of charge.[52] The force field parameters of nevirapine were prepared with the module Antechamber[53] of AMBER.[54,55]

Hydrogen atoms were added to the crystallographic protein with the AMBER module leap. Five chloride counterions were added to neutralize the system. The system was then solvated with a rectangular box of TIP3P[56] water molecules. This water model was used for comparison of previous simulations. The minimum distance from the surface of the protein to the faces of the box was 11 Å. The final box dimensions were $\sim 100 \times 112 \times 116$ Å^3.

RT without nevirapine (unliganded) was started from the RT/nevirapine complex, from which the nevirapine molecule was removed. This starting structure differs from the crystallographically determined unliganded structures (which have the p66 thumb subdomain in a "closed" configuration and no non-nucleoside inhibitor binding pocket) but was chosen for this study in order to compare directly the effect of a bound nevirapine and of an open, empty binding pocket on RT mobility.

RT with three point mutations (V106A, Y181C, Y188C) was built using MOE. The mutations were performed using the homology tool of MOE. The backbone of the system was assumed not to change during the mutation since in all cases the side chain of the mutated amino acid is shorter than that of the wild-type. After mutation, minimization was performed on the side chains. All three molecular systems, (unliganded RT, nevirapine complexed to wild-type RT, and nevirapine complexed to the mutant RT) were prepared in the same manner and each was simulated for 2.5 ns. Each solvated system contains approximately 142 000 atoms and 42 000 water molecules.

MD Protocol. MD simulations were carried out with AMBER7 on the Terascale Computing System at the Pittsburgh Supercomputing Center. The force fields used were Amber99 for the protein and GAFF for nevirapine.[57] The particle mesh Ewald (PME) method was used to treat long-range electrostatic interactions, with a cubic B-spline interpolation. The cutoff for nonbonded interactions was 9 Å. SHAKE was used to constrain all bonds involving hydrogen atoms with a geometric tolerance of 0.000 01 Å. Periodic boundary conditions were used. The time steps were 1 fs during the equilibration and 2 fs during the production dynamics. The translational and rotational motions of the center of mass were removed every 1000 ps. The temperature was rescaled with the Berendsen weak-coupling algorithm,[58] with a time constant of 0.5 ps for the heat bath. After the system was heated to 300 K from the initial temperature of 100 K, molecular dynamics were performed at a constant temperature of 300 K.

First, the water atoms and counterions were minimized for 500 steps of steepest descent and 1500 steps of conjugate gradient while restraining all other atoms with a force constant of 5000 kcal/mol Å^2. This was followed by three series of minimizations, during which the hydrogen atoms, side chains, and finally the whole protein were successively allowed to move, while restraining the rest with a force constant of 500 kcal/mol Å^2. In each step, 100 steps of steepest descent and 400 steps of conjugate gradient minimizations were performed. Using the module leap of AMBER7 to solvate a protein, we usually obtain a low-density solution. The dimensions of the box were adjusted so as to increase the density to 0.93−0.94 g/cm^3 prior to performing minimizations. The density of the system quickly reached $\sim$1 g/cm^3 during the heating stage of the equilibration simulation.

The production stage consisted of 300 ps at constant pressure (with a pressure relaxation time of 0.2 ps) followed by 2.1 ns at constant volume, for a total of 2.4 ns. All molecular structure visualization and analysis were performed with MOE.

Free Energy Calculations. The free energy of ligand binding in the non-nucleoside inhibitor binding pocket was analyzed by the Molecular Mechanics-Poisson−Boltzmann Surface Area (MM-PBSA) method,[53,59,60] integrated in AMBER7. The binding free energy was calculated as the average of the binding free energy of the last 150 snapshots, taken at 2 ps intervals from the trajectories of each simulation. Similar results where obtained when a different range of snapshots was considered for the wild-type RT/nevirapine complex. The free energy was calculated as

$$\Delta G_{\mathrm{tot}} = \Delta E_{\mathrm{MM}} + \Delta G_{\mathrm{sol}} - T\Delta S \qquad (1)$$

where ΔE_{MM} is the molecular mechanics energy, comprised of a van der Waals and an electrostatic contribution; ΔG_{sol} is the solvation energy, which consists of electrostatic and nonpolar interactions. The electrostatic solvation energy is determined using the finite difference Poisson−Boltzmann method, and the nonpolar contribution is estimated by the solvent-accessibility surface method. $T\Delta S$ is the entropy contribution to the free energy, which is estimated using normal-mode analysis. The binding free energy for ligand in a protein can be estimated using

$$\Delta G_{\mathrm{b}} = \Delta G_{\mathrm{c}} - (\Delta G_{\mathrm{p}} + \Delta G_{\mathrm{l}}) \qquad (2)$$

where c, p, and l denote the protein/ligand complex, the protein, and the ligand, respectively.

The same force field used in the MD simulations was used in the free energy calculations. The entropy contribution to the binding free

(42) Richman, D. D. *AIDS Res. Hum. Retroviruses* **1992**, *8*, 1065−1071.
(43) Mayers, D. L.; McCutchan, F. E.; Sanders-Buell, E. E.; Merritt, L. I.; Dilworth, S.; Fowler, A. K.; Marks, C. A.; Ruiz, N. M.; Richman, D. D.; Roberts, C. R. *J. Acquired Immune Defic. Syndr.* **1992**, *5*, 749−759.
(44) Richman, D. D. *Curr. Top. Microbiol. Immunol.* **1992**, *176*, 131−142.
(45) Richman, D. D. *J. Enzyme Inhib.* **1992**, *6*, 55−64.
(46) Richman, D. D.; Guatelli, J. C.; Grimes, J.; Tsiatis, A.; Gingeras, T. *J. Infect. Dis.* **1991**, *164*, 1075−1081.
(47) Richman, D.; Shih, C. K.; Lowy, I.; Rose, J.; Prodanovich, P.; Goff, S.; Griffin, J. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 11241−11245.
(48) Mellors, J. W.; Dutschman, G. E.; Im, G. J.; Tramontano, E.; Winkler, S. R.; Cheng, Y. C. *Mol. Pharmacol.* **1992**, *41*, 446−451.
(49) Mellors, J. W.; Im, G. J.; Tramontano, E.; Winkler, S. R.; Medina, D. J.; Dutschman, G. E.; Bazmi, H. Z.; Piras, G.; Gonzalez, C. J.; Cheng, Y. C. *Mol. Pharmacol.* **1993**, *43*, 11−16.
(50) Shen, L.; Shen, J.; Luo, X.; Cheng, F.; Xu, Y.; Chen, K.; Arnold, E.; Ding, J.; Jiang, H. *Biophys. J.* **2003**, *84*, 3547−3563.
(51) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(52) Akalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132−146.
(53) Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. *J. Am. Chem. Soc.* **2001**, *123*, 5221−5230.
(54) Case, D. A.; et al. *AMBER7*; University of Califonia: San Francisco, 2002.
(55) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; III, T. E. C.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. A. *Comput. Phys. Commun.* **1995**, *91*, 1−41.
(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(57) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(58) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(59) Wang, W.; Kollman, P. A. *J. Mol. Biol.* **2000**, *303*, 567−582.
(60) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E., III. *Acc. Chem. Res.* **2000**, *33*, 889−897.

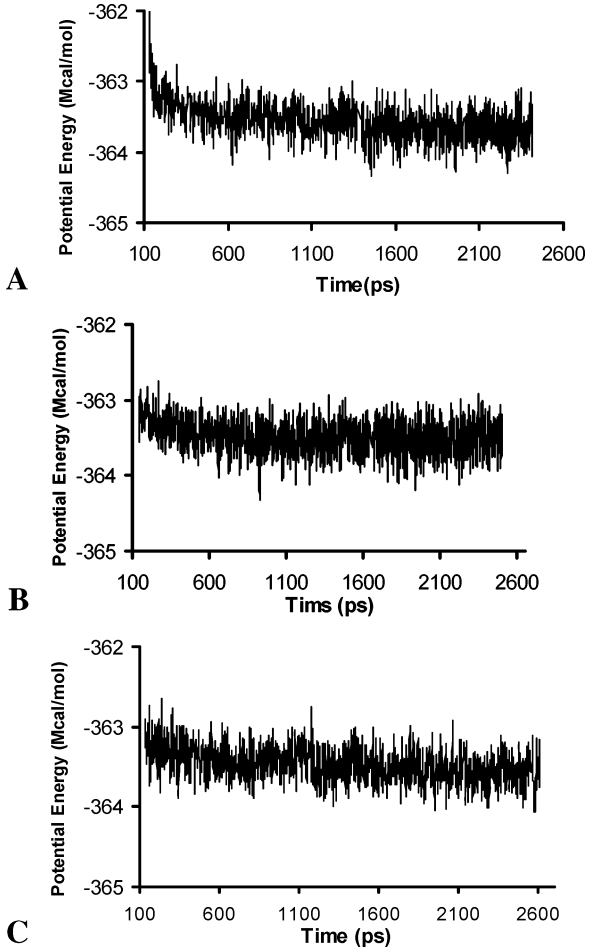


Figure 1. Potential energies during the molecular dynamics simulation for the three systems studied: (A) wild-type RT/nevirapine; (B) mutant RT/nevirapine; (C) unliganded RT.

energy was neglected because it has been shown that it mostly cancels when it is estimated from the same set of coordinates for the protein, ligand, and the protein/ligand complex.[61−70]

Principal Component Analysis (PCA). PCA can be used to simplify a large set of data by decreasing the number of variables using multivariate analysis. PCA has been used to analyze the structures from MD simulations and has been previously described in detail.[61−71]

A symmetric matrix **A** with elements equal to the pairwise RMS coordinate difference between superimposed structures along the trajectory is employed in this work:

$$A_{ij} = \mathrm{RMSD}_{ij} \tag{3}$$

where RMSD is the structural RMS deviation of two snapshots. The RMS coordinate difference is a measure of structure similarity, and

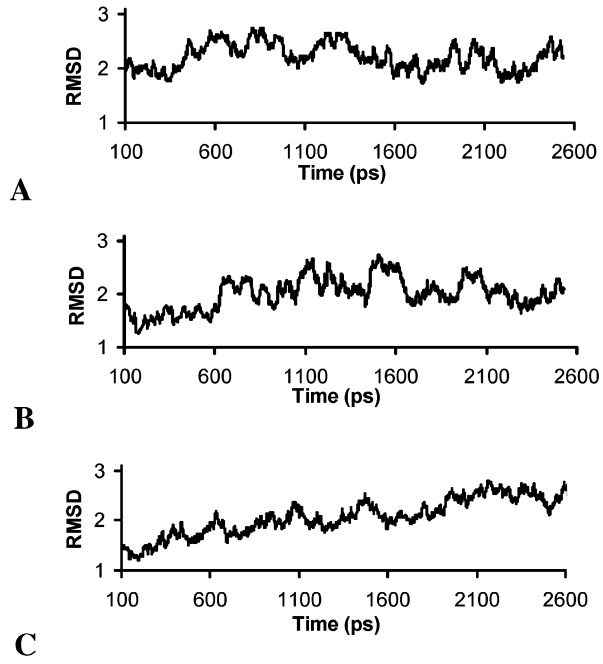


Figure 2. RMDs of the Cα atoms as a function of time, for each of the systems studied: (A) wild-type RT/nevirapine; (B) mutant RT/nevirapine; (C) unliganded RT.

the matrix **A** allows a representation of the conformational sampling of the protein during simulation. In this work, the RMSD matrix was generated using 100 snapshots from each set of MD trajectories, sampled in equal intervals. The total 300 sets of structures were then used to build an RMSD matrix. All structures were aligned based on the C_α atoms onto the first structure of WT, followed by a 2-D RMSD calculation. The source code of Ptraj module of AMBER7 was modified to output a 2-D RMSD matrix.

A "similarity" matrix **S** is constructed with elements S_{ij}.[71]

$$S_{ij} = -0.5 \times A_{ij}^{\ 2} \tag{4}$$

The eigenvectors and eigenvalues are obtained by diagonalizing the **S** matrix.[72,73] After diagonalization the eigenvalues are ranked from largest to smallest. Since it was found that the first three eigenvalues account for ∼86% of the variation, these three eigenvalues were projected on a 2D plot. All PCAs were carried out using MatLab 6.5.[74]

Results

Analysis of the MD Simulations. The potential energies of the three systems studied (wild-type RT/nevirapine, mutant RT/nevirapine, and RT without nevirapine) are plotted as a function of time in Figure 1A−C. The root-mean-square deviations (RMSDs) between the Cα atoms of the structures obtained during the trajectories and the initial structures are shown in Figure 2 for the three systems. Figures 1 and 2 show that the potential energies and the structures remain stable after approximately the first 800 ps of the simulation.

The B-factors of the wild-type RT/nevirapine complex (calculated by multiplying the atomic positional fluctuations from the MD simulation by $(^8/_3)\pi^2$ are compared to the crystallographic B-factors (Figure 3). Overall agreement is observed between the calculated and crystallographic B-factors.

(61) Schulze, B. G.; Evanseck, J. D. *J. Am. Chem. Soc.* **1999**, *121*, 6444−6454.
(62) Balsera, M. A.; Wriggers, W.; Oono, Y.; Schulten, K. *J. Phys. Chem.* **1996**, *100*, 2567−2572.
(63) Saarela, J. T. A.; Tuppurainen, K.; Perakyla, M.; Santa, H.; Laatikainen, R. *Biophys. Chem.* **2002**, *95*, 49−57.
(64) Kitao, A.; Go, N. *Curr. Opin. Struct. Biol.* **1999**, *9*, 164−169.
(65) Nolde, S. B.; Arseniev, A. S.; Orekhov, V. Y.; Billeter, M. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 250−258.
(66) Tournier, A. L.; Smith, J. C. *Phys. Rev. Lett.* **2003**, *14*, 208106/208101−208106/208104.
(67) Caves, L. S. D.; Evanseck, J. D.; Karplus, M. *Protein Sci.* **1998**, *7*, 649−666.
(68) Prompers, J. J.; Bruschweiler, R. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 177−189.
(69) Gower, J. C. *Biometrics* **1967**, *23*, 623−637.
(70) Gower, J. C. *BiometriK* **1966**, *53*, 325−338.
(71) Manly, B. F. J. *Multivariate Statistical Methods: A Primer*, 3rd ed.; Chapman & Hall/CRC: Boca Raton, FL 33431 USA, 2004.
(72) Troyer, J. M.; Cohen, F. E. *Proteins: Struct., Funct., Genet.* **1995**, *23*, 97−110.
(73) Johnson, R. A.; Wichern, D. W. *Applied Multivariate Statistical Analysis*, 5th ed.; Prentice Hall: Upper Saddle River, New Jersey, 2002.
(74) *Matlab*, version 6.5 ed.: Natick, Massachusetts, 2003.

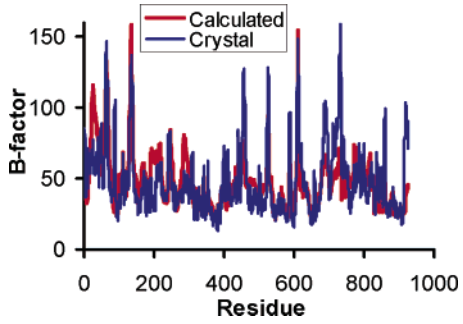


Figure 3. B-factors calculated from the molecular dynamics simulation (red) and reported for the crystallographic structure of the RT/nevirapine complex (pdb code 1vrt) (blue).

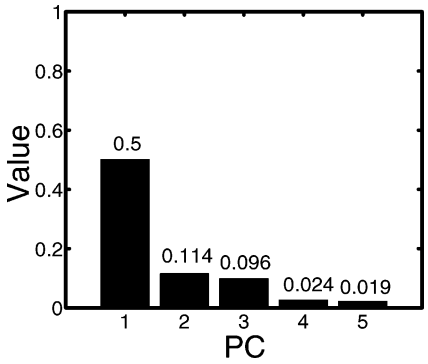


Figure 4. Values of the first five eigenvalues calculated.

In particular, the simulations reproduce the sharp peaks observed in the crystallographic structure around amino acids 50, 70, and 140 (p66 fingers subdomain), 445−450 and 527−532 (RNase H subdomain), 620−630 (p55 fingers), and 720−730 (p55 palm). In most regions, the MD simulation predicts B-factors similar to those observed experimentally.

Since hydrogen bonds can play an important role in ligand binding, water proton and protein hydrogen bonds around the NNIBP were checked. Water molecules were observed to form hydrogen bonds around the putative entrance of NNIBP of the wild-type and mutant RT/nevirapine complexes during the 2.5 ns molecular dynamics trajectories. The hydrogen bond bridges formed by these water molecules between nevirapine and hydrophilic amino acids around the putative entrance help to stabilize the binding of nevirapine in the NNIBP. This water network is consistent with the water bridge network observed in the crystallographic structure[16] and in our docking results.[19] During molecular dynamic simulation, these water molecules are dynamically moving around (see Supporting Information Figure S1 for the comparison of the water molecules from a simulation snapshot with the crystal water molecules and the hydrogen bond network).

Principal Component Analysis. To cluster the conformational spaces and compare the feature structures from each different structure space, PCA is employed to identify the conformational clusters for the simulations. The structure clustering analysis was done with snapshots taken from three MD trajectories at equal intervals and should represent all major conformational spacing.

The first three principal components (PCs) of the Cα atoms account for approximately 86% of the total sample variance for the three sets of structures (Figure 4). Therefore, these three principal components (PC1, PC2, and PC3) are used to analyze the structures of the three simulations, and their projections are shown in Figure 5. It can be seen that unliganded RT has two

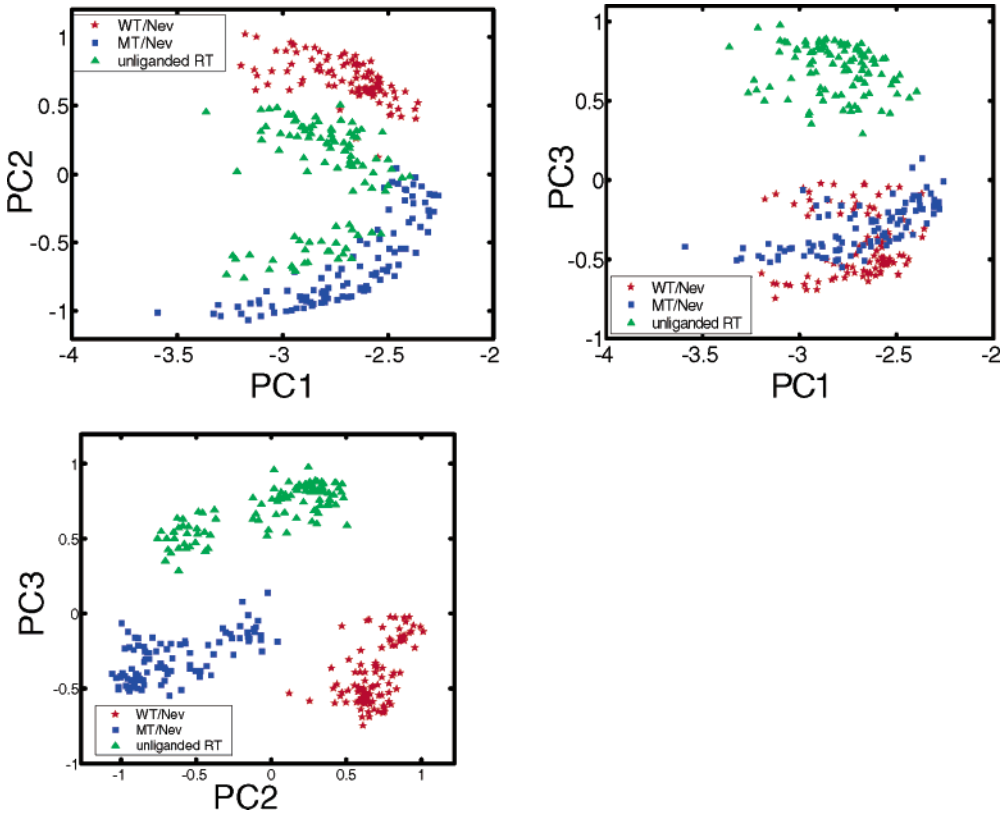


Figure 5. Charts of the first three principal components (PC1, PC2, and PC3) calculated from 300 structures of the three MD trajectories. Wild-type RT/nevirapine complex (red stars), mutant RT/nevirapine complex (blue squares), and unliganded RT (green triangles).

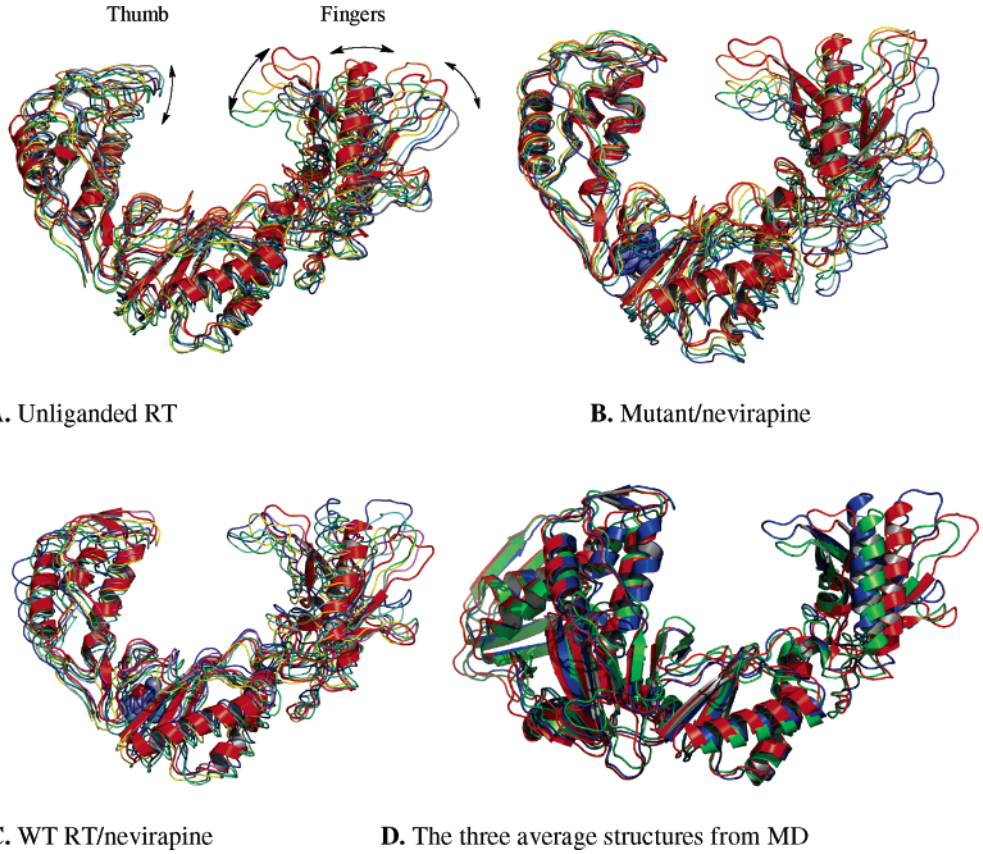


Figure 6. Superposition of snapshots picked from different areas of the PCA charts (Figure 5) for the three MD trajectories: (A) unliganded RT; (B) mutant RT/nevirapine complex; (C) wild-type RT/nevirapine complex; (D) superposition of the average structure from each MD trajectory: wild-type RT/nevirapine (red), mutant RT/nevirapine (blue), and unliganded RT (green). The ligand, nevirapine, is shown in the space-filled ball model in the RT/nevirapine complexes.

well-defined clusters, while the two liganded systems do not present well separated clusters. These results indicate that the unliganded system samples two distinct minima during the molecular dynamics trajectory. In addition, the size of each cluster in Figure 5 appears to indicate that unliganded RT undergoes larger conformational changes than the liganded systems.

Motions of the p66 Thumb, Fingers, and Palm Subdomains. To illustrate the movements undergone by the p66 fingers, palm, and thumb subdomains, six structures were taken from each MD trajectory of unliganded RT (Figure 6A), mutant RT complexed to nevirapine (Figure 6B), and wild-type RT complexed to nevirapine (Figure 6C), from different areas of the PCA plots shown in Figure 5. It is seen that the p66 fingers of unliganded RT undergo the largest amplitude movements during the trajectories, while these motions are smallest in the wild-type RT/nevirapine complex. The average structures from each of the simulations are superimposed in Figure 6D, showing that the inside tip of the p66 fingers subdomain of unliganded RT (green) closes down toward the p66 palm subdomain the most during the MD trajectories, while that of wild-type RT/nevirapine (red) moves down the least.

Table 1 compares the RMSD between the average structures from each of the three simulations and the unliganded crystal structure (PDB code 1HMV).[24] The table shows the RMSDs between the Cα atoms of the p66 subunit and, in parentheses, of the whole protein. It is seen that the three simulated structures (WT/Nev, MT/Nev, and unliganded RT) remain similar to each other. The RMSD between the p66 subunits of the simulated unliganded RT and the crystal structure is 5.1 Å, indicating that 2.5 ns of molecular dynamics in explicit solvent, at room temperature, are not sufficient to overcome the potential barriers needed to reach the experimentally observed crystallographic structure starting from the WT RT/nevirapine structure, from which the nevirapine molecule was removed.

Table 1. RMSD between the Cα Atoms of the Average Structures from the Three Simulations and the Unliganded Crystal Structure (1HMV),[24] for the P66 Subunit (and the Whole Protein)

RMSD (Å)	MT/Nev	unliganded RT	unliganded crystal RT
WT/Nev	2.6 (2.8)	2.7 (3.0)	6.0 (8.4)
MT/Nev		2.2 (2.4)	5.6 (7.8)
unliganded RT			5.1 (7.0)

Table 2. RMSD of Cα (and All Heavy Atoms) between the Polymerase Active Sites of the Average Structures from the Three Simulations and the Crystal Structure of WT RT/Nevirapine Complex (1VRT)[16]

RMSD (Å)	MT/Nev	unliganded RT	WT/Nev crystal RT
WT/Nev	1.2 (1.6)	1.3 (1.7)	1.5 (2.1)
MT/Nev		0.5 (1.0)	1.5 (1.7)
unliganded RT			1.7 (1.9)

Table 2 compares the RMSD of amino acids Asp110−Tyr115 and Met184−Asp186, which are part of the polymerase active site, between the average structures of the simulations and the crystal structure of RT/Nev. The RMSDs are listed for the Cα atoms and, in parentheses, for all heavy atoms, based on the superposition of the p66 subunits. It is seen that these amino

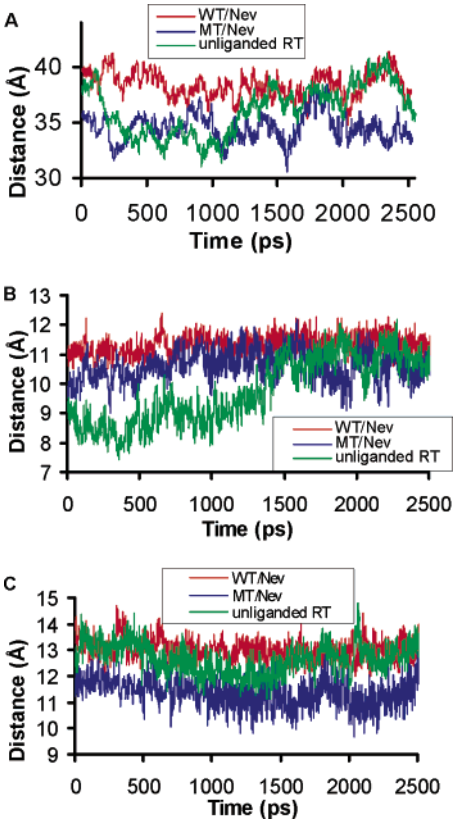


Figure 7. Distances as a function of time for wild-type RT/nevirapine (red), mutant RT/nevirapine (blue), and unliganded RT (green). (A) Distance between the geometric centers of the Cα atoms of amino acids 254-266 (p66 thumb) and amino acids 75-80 (p66 fingers). (B) Distance between the Cα atoms of Leu 100 and Tyr (NNIBP). (C) Distance between the Cα atoms of Ile 94 and Tyr 188 (NNIBP).

acids have similar conformations in the three structures simulated and in the crystal structure.

The distances between the geometric centers of the Cα atoms of amino acids 254−266 (located in the p66 thumb subdomain) and amino acids 75−80 (located in the p66 fingers) are monitored during the trajectory for unliganded RT, RT/nevirapine, and mutant RT/nevirapine (Figure 7A). The large fluctuations in these distances correspond to motions of the p66 thumb and fingers relative to each other. The faster motions, on the picosecond time scale, correspond to local vibrations. Of the three systems studied, unliganded RT shows the larger variations in these distances, between ∼31 and 41 Å (in green in Figure 7A). Conversely, the RT/nevirapine complex shows the smallest variation in the distances between the p66 thumb and p66 fingers, with distances ranging between ∼35 and 41 Å (in red in Figure 7A). The distances corresponding to the mutant RT/nevirapine complex (in blue in Figure 7A) fluctuate between ∼31 and 39 Å.

The distances between amino acids that form part of the binding pocket were monitored in order to study the dynamics of the binding pocket, as follows. The distances between the Cα atoms of Leu 100 and Tyr 181 and between Ile 94 and Tyr 188 that are located on two subdomains forming the binding poeckt are shown in Figure 7B and 7C, respectively. It is seen that the largest fluctuations in these distances correspond to unliganded RT (in green in Figure 7B and 7C), while these distances remain approximately constant for wild-type RT/nevirapine (in red in Figure 7B and 7C).

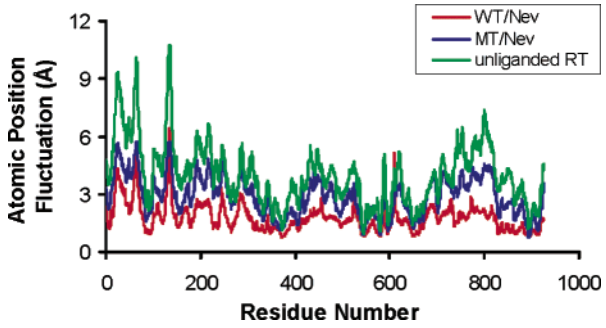


Figure 8. Atomic positional fluctuations of the Cα atoms of the three systems studied.

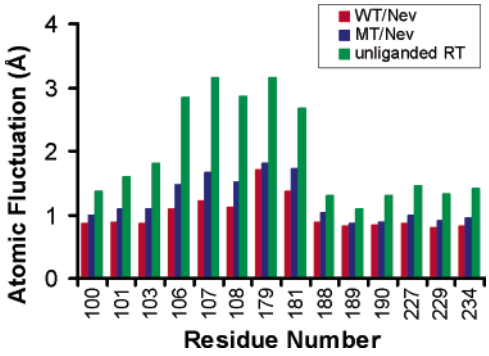


Figure 9. Atomic positional fluctuations of amino acids in the NNIBP, for the three systems studied.

In summary, the distances studied indicate that, for the atoms involved, the wild-type RT/nevirapine complex is the most rigid (in red in Figure 7A−C), while unliganded RT is the most flexible (in green in Figure 7A−C). After the first 1000 ps, the distance between Leu 100 and Tyr 181 increases for unliganded RT (Figure 7B green). During the same period of time, the distances between the p66 thumb and fingers for unliganded RT also increase (Figure 7A green), indicating that the motions of these portions of the p66 thumb, fingers, and palm subdomains in unliganded RT might be correlated.

The atomic positional fluctuations (APF) of the Cα atoms are shown in Figure 8. It is seen that, of the three systems studied, unliganded RT (with an open NNIBP) has the largest fluctuations for almost all Cα atoms. The most flexible regions correspond to amino acids 20−75 and 130−140 of the p66 fingers subdomain. The presence of nevirapine in the binding pocket of the RT/nevirapine complex results in the smallest atomic fluctuations for both the p66 and p51 subunits. The mutations produce a partial restoration of the fluctuations of all Cα atoms. In particular, most of the fluctuations of the p66 thumb subdomain (amino acids 244−322) are restored by the presence of the mutations.

Figure 9 shows the atomic positional fluctuations of the Cα atoms of amino acids that compose the NNIBP, for the last 1.5 ns of the trajectory. As expected, the amino acids surrounding the binding pocket are most flexible when the binding pocket is open but without a bound NNRTI, as in the case of RT without nevirapine. The presence of an NNRTI in the binding pocket diminishes this flexibility, while the mutations in the binding pocket slightly increase the flexibility of the amino acids surrounding the binding pocket.

Binding Free Energy. The predicted free energies of nevirapine bound to wild-type and mutant RT are −54.78 kcal/

mol and −48.84 kcal/mol, respectively, indicating that the interaction of nevirapine with wild-type RT is stronger than that with mutant RT. This is consistent with the fact that nevirapine is less effective in the presence of mutations. The van der Waals (vdW) interaction energies between nevirapine and wild-type and mutant RT are −43.90 kcal/mol and −40.24 kcal/mol, respectively. Thus, vdW interactions are the largest contribution to the free energy (∼80%), suggesting that nevirapine fills the NNIBP and that there is good contact between the inhibitor and the receptor. The calculated difference in free energy of binding between wild-type and mutant RT is −5.94 kcal/mol. The relative free energy difference can be used to measure the binding affinity difference of nevirapine in wild-type and mutant RT and can be correlated to nevirapine's activity. The binding free energy difference of nevirapine bound to wild-type RT and to a one-point mutant has been predicted, by means of Monte Carlo simulations,[75,76] to be between 3.33 and 3.88 kcal/mol. In addition, biochemical experiments suggest that nevirapine has a 500 times higher dissociation constant (K_d) value with Y181C mutant/DNA complex, which corresponds to a free energy of 3.7 kcal/mol for the binding/dissociation reaction.[29] The cavity change of the NNIBP in our study is larger than that in this previous study, because we have mutated three amino acids to ones with shorter side chains. Thus, it is reasonable that the interaction energy change between nevirapine and RT in our calculation is larger than previous computational and experimental results.

Conclusions

We have studied the flexibility and free energy of binding of wild type and mutant RT complexed to the non-nucleoside RT inhibitor nevirapine. In addition, unliganded RT was simulated from the crystallographic coordinates of the RT/nevirapine complex, from which the nevirapine molecule was removed, leaving an empty non-nucleoside binding pocket and the p66 thumb subdomain in an "open" configuration.

Comparisons of the atomic positional fluctuations of the Cα atoms during the molecular dynamics trajectories show that the above-mentioned unliganded RT undergoes the largest atomic positional fluctuations. Particularly large fluctuations are observed for amino acids 20−75 and 130−140, located in the p66 fingers subdomain. Large mobility of the p66 fingers subdomain has also been previously observed in the molecular dynamics simulations of RT complexed to dsDNA.[33] This flexibility is consistent with the ability of the p66 fingers to close down on the dNTP binding site, as evidenced by the RT/dsDNA/dNTP crystal structure.[10,77]

The bound nevirapine in the binding pocket of wild-type RT decreases the flexibility of almost *all* amino acids of HIV-1 RT, both in the p66 and p51 subunits, providing support for the theory of "molecular arthritis", which postulates that the bound NNRTI decreases the flexibility of the p66 thumb subdomain. The amino acid mutations V106A, Y181C, and Y188C in the binding pocket partially restore RT flexibility, which can be attributed to the fact that the mutated, smaller, amino acid side chains provide more space in the NNIBP for the amino acids to move without steric collisions. Evidence for this larger NNIBP is given by the predicted loss of van der Waals interactions between nevirapine and the mutated RT with respect to the wild-type RT of 3.66 kcal/mol. The increased flexibility of the binding pocket results in an overall increase of RT flexibility due to correlated motions.

In summary, our simulations indicate that RT flexibility depends on the volume of the binding pocket occupied by the NNRTI, with the flexibility increasing in the presence of the three mutations studied, and becoming most flexible when the nonnucleoside inhibitor binding pocket is empty. These results are consistent with recent molecular dynamics simulations, which show a decrease in the distances between the p66 fingers and thumb subdomains while the NNRTI αAPA is pulled out of the binding pocket and once it has left the binding pocket.[50]

Based on the molecular dynamics and principal component analyses, we hypothesize that unliganded RT samples a larger conformational space during the trajectories than the two liganded systems.

It is suggested that the larger the NNRTI and the stronger the interaction with amino acids around the NNIBP, the stronger it will inhibit RT, provided that it can enter efficiently and bind effectively inside of the NNIBP. This is partly supported by our docking study[19] results that highly active neotripterifordin inhibitors can bind in the NNIBP well and occupy more space in the direction from amino acids 92−94 to amino acids 185−188 than the less active and smaller nevirapine. Thus, the ligand can interact well with these amino acids and affect the dynamics of RT.

Acknowledgment. The authors would like to thank Dr. Pranav Dalal for his help. This research was partially supported by a grant at the Pittsburgh Supercomputing Center (PSC) through the NIH National Center for Research Resources (RR06009) and MCB010008P and the Commonwealth of Pennsylvania (22-125-0001) and an NIH pre-NPEBC Grant to the University of Pittsburgh.

Supporting Information Available: Complete refs 7, 54 and Figure S1. This material is available free of charge via the Internet at http://pubs.acs.org.

JA053973D

(75) Smith, M. B. K.; Lamb, M. L.; Tirado-Rives, J.; Jorgensen, W. L.; Michejda, C. J.; Ruby, S. K.; Smith, R. H., Jr. *Protein Eng.* **2000**, *13*, 413−421.
(76) Rizzo, R. C.; Wang, D.-P.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 12898−12900.
(77) Sarafianos, S. G.; Das, K.; Ding, J.; Boyer, P. L.; Hughes, S. H.; Arnold, E. *Chem. Biol.* **1999**, *6*, 137−146.