Role of invariant Thr80 in human immunodeficiency virus type 1 protease structure, function, and viral infectivity

Jennifer E. Foulkes-Murzycki
University of Massachusetts Medical School

Let us know how access to this document benefits you.
Follow this and additional works at: https://escholarship.umassmed.edu/oapubs

Part of the Life Sciences Commons, and the Medicine and Health Sciences Commons

Repository Citation

This material is brought to you by eScholarship@UMassChan. It has been accepted for inclusion in Open Access Publications by UMass Chan Authors by an authorized administrator of eScholarship@UMassChan. For more information, please contact Lisa.Palmer@umassmed.edu.
Role of Invariant Thr80 in Human Immunodeficiency Virus Type 1 Protease Structure, Function, and Viral Infectivity

Jennifer E. Foulkes, Moses Prabu-Jeyabalan, Deyna Cooper, Gavin J. Henderson, Janera Harris, Ronald Swanstrom, and Celia A. Schiffer

Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, and UNC Center for AIDS Research and Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7295

Received 8 September 2005/Accepted 1 May 2006

Sequence variability associated with human immunodeficiency virus type 1 (HIV-1) is useful for inferring structural and/or functional constraints at specific residues within the viral protease. Positions that are invariant even in the presence of drug selection define critically important residues for protease function. While the importance of conserved active-site residues is easily understood, the role of other invariant residues is not. This work focuses on invariant Thr80 at the apex of the P1 loop of HIV-1, HIV-2, and simian immunodeficiency virus protease. In a previous study, we postulated, on the basis of a molecular dynamics simulation of the unliganded protease, that Thr80 may play a role in the mobility of the flaps of the protease. In the present study, both experimental and computational methods were used to study the role of Thr80 in HIV protease. Three protease variants (T80V, T80N, and T80S) were examined for changes in structure, dynamics, enzymatic activity, affinity for protease inhibitors, and viral infectivity. While all three variants were structurally similar to the wild type, only T80S was functionally similar. Both T80V and T80N had decreased the affinity for saquinavir. T80V significantly decreased the ability of the enzyme to cleave a peptide substrate but maintained infectivity, while T80N abolished both activity and viral infectivity. Additionally, T80N decreased the conformational flexibility of the flap region, as observed by simulations of molecular dynamics. Taken together, these data indicate that HIV-1 protease functions best when residue 80 is a small polar residue and that mutations to other amino acids significantly impair enzyme function, possibly by affecting the flexibility of the flap domain.

Human immunodeficiency virus type 1 (HIV-1) protease is an enzyme crucial to the life cycle of the virus (28). Inhibition of the protease prevents viral maturation, and thus further infection, making it a prime target in the treatment of HIV-1-infected patients (8, 18, 53). Unfortunately, HIV-1 mutates frequently because of the high replication rate of the virus (9) and the infidelity of the reverse transcriptase (21, 50, 51). The viral protease tolerates mutations in many areas of the sequence, especially mutations that decrease the binding affinity of inhibitors without affecting the ability of the protease to recognize and cleave its nine processing sites within the Gag and Gag-Pro-Pol polyproteins. Such mutations cause drug resistance. Many studies have explored residues susceptible to mutation and their role in the development of drug resistance (2, 14, 23, 24, 39, 41, 55). However, some amino acids in the protease have not been observed to mutate, even in the presence of inhibitor therapy (58) (Fig. 1). While these conserved residues potentially play a crucial role in either the structure or the function of the enzyme, few studies have explored the purpose of the invariant amino acids outside the active site. Utilizing these residues could assist in the design of new classes of protease inhibitors that may be less susceptible to developing resistance.

HIV protease, a homodimeric aspartyl protease, envelops its substrate when the flap region of each monomer closes down on the substrate sites within the Gag and Gag-Pro-Pol polyproteins. In a 10-ns molecular dynamics simulation (MD) of the unliganded protease, the tips of the flaps (residues 47 to 53) were observed to curl toward the protease core like fingers curling toward the palm of a hand (57) (Fig. 1b). When the tip of the flap curls inward, Ile50 contacts a number of predominantly hydrophobic residues in its own P1 loop (residues 78 to 83). The only polar residue in this region is virtually invariant Thr80. Thr80 is one of only 27 (out of 99) residues that are invariant in 99.9% of the sequences in the Stanford University HIV Drug Resistance Database (58). The remaining 26 invariant residues include the catalytic Asp-Thr-Gly triad, five other glycines, three prolines, both tryptophans, and the tyrosine. Of the 536 HIV-1, HIV-2, and simian immunodeficiency virus protease sequences found in the Los Alamos National Laboratory Database, residue 80 is 1 of only 10 residues invariant in 99.8% of the population. Additionally, numerous crystal structures show that the side chain hydroxyl group forms highly conserved hydrogen bonds with the backbone of residue 82. Other studies have observed structural changes in the P1 loop when binding inhibitors and drug resistance mutations at Val82 and Ile84 (1, 3, 17, 25–27, 37, 48). By examining invariant residues within this loop, we may discover new ways to target protease inhibitor therapy. Interactions between Thr80 and Ile50 could be important for destabilizing the burial of Ile50 in a hydrophobic pocket and maintaining the mobility of the flap tips. Thr80 could be important for stabilizing the P1 loop through its highly conserved hydrogen bond with Val82. While the role of this residue has not been previously studied,
the observation that Thr80 is invariant indicates that it is likely important to protease structure and/or function.

In the present study, we used various methods to study the role of Thr80. Native Thr80 was mutated to asparagine, serine, and valine to generate protease variants referred to as T80N, T80S, and T80V. The impact of Thr80 on HIV-1 protease structure and function was investigated by a variety of biophysical methods, including circular dichroism (CD) spectroscopy, tryptophan fluorescence, protein crystallography, MD simulations, fluorescence resonance energy transfer, and isothermal titration calorimetry. Additionally, the impact of these mutations on the virus was examined by using a virus specific-infectivity (SpIn) assay. Taken together, these data indicate that Thr80 plays an important role in enzyme activity and help to clarify why Thr80 does not mutate in vivo.

MATERIALS AND METHODS

Protease gene construction. The synthetic gene for the wild-type (WT) HIV-1 protease sequence was made with codons optimized for protein expression in Escherichia coli and included a substitution of Gln7Lys to prevent autoproteolysis (54). Residue 80 was mutated from threonine to three different amino acids, asparagine, serine, and valine. Mutagenesis was performed with the Stratagene QuickChange site-directed mutagenesis kit and confirmed by sequencing.

Protease expression and purification. The expression and purification of the HIV-1 protease were previously described (26, 46). Briefly, the HIV protease gene was cloned into plasmid pUC-35 (American Type Culture Collection, Manassas, Va.) (7), which was transfected into E. coli TAP106. Transfected cells were grown in a 12-liter fermentor and, following protein expression, lysised to release the inclusion bodies containing the protease (20). Inclusion bodies were isolated by centrifugation, and the pellet was dissolved in 50% acetic acid to extract protease. High-molecular-weight proteins were separated from the desired protease by size exclusion chromatography on a 2.1-liter Sephadex G-75 superfine (Sigma Chemical) column equilibrated with 50% acetic acid. Refolding was accomplished by rapidly diluting the protease solution into a 100-fold excess of refolding buffer. Excess acetic acid was removed through dialysis. Protein used for crystallization was further purified with a Pharmacia Superdex 75 fast-performance liquid chromatography column equilibrated with refolding buffer.

Far-UV CD spectroscopy. To determine the effects of mutations on the secondary structure of the protease, far-UV CD spectra were measured for the WT and variant proteases with a JASCO J-810 spectrometer. The temperature was maintained at 20°C with a Peltier temperature control device. Measurements were made in a rectangular cell with a 1-mm path length. Each spectrum was obtained by averaging three spectra recorded from 250 to 200 nm at a 0.5-nm data pitch and a rate of 20 nm min⁻¹. The response time for each point was 8 s, and the bandwidth was 2.3 nm. Protease concentrations were between 5 and 11 μM in a buffer containing 10 mM sodium acetate at pH 5 and 2 mM Tris(2-carboxyethyl)phosphine. Buffer scans were subtracted from the sample scans. The mean residue ellipticity (MRE), in millidegrees per square centimeter per decimole, was calculated with the equation MRE = θ/(10(MV × C × l)), where θ is the measured signal in millidegrees, M is the number of residues, V is the molar concentration of the protein, and l is the path length of the cuvette.

Tryptophan fluorescence. To further explore structural changes caused by mutations at residue 80, tryptophan fluorescence was measured in the presence and absence of saquinavir (SQV). Fluorescence was measured with a Photon Technology International spectrofluorimeter at 25°C. The excitation wavelength was 295 nm, and the emission spectrum was scanned from 300 to 500 nm in 1-nm steps. Protease concentrations were between 5 and 5.5 μM in a buffer containing 10 mM sodium acetate and 2 mM Tris(2-carboxyethyl)phosphine at pH 5. Buffer scans were subtracted from the sample scans. Each variant was also examined for fluorescence changes upon inhibitor binding. Protease solutions were equilibrated with 30 μM SQV prior to measurement. Buffer scans containing SQV were subtracted from the sample scans.

Protein crystallization. To look for more localized changes in the structures of the proteases, T80N and T80S were crystallized with SQV, yielding the structures SQVT80N and SQVT80S. Protein solutions between 1.16 and 1.75 mg ml⁻¹ were equilibrated with a fivefold molar excess of the inhibitor SQV. Crystals were grown by the vapor diffusion hanging-drop method. The reservoir solution contained 126 mM sodium phosphate (pH 6.2), 63 mM sodium citrate, and 23 to 24% ammonium sulfate (25). The crystallization screens yielded rectangular crystals, with the longest dimension being 0.1 mm.

Crystallographic data collection. The crystals were mounted within 0.2-mm Hampton cryoloops and flash frozen over a nitrogen stream. Data were collected on an in-house Rigaku Kuai K-14 image plate. Two hundred fifty frames of 5-min exposures were collected per crystal with an angular separation of 1° and no overlap between frames. Frames containing unprocessed raw data were indexed with Denzo and scaled with ScalePack (35, 43). The SQVT80N and SQVT80S crystals diffracted to 1.5 Å and 2.0 Å, respectively. Complete data collection statistics are listed in Table 1.

Structure solution and crystallographic refinement. The crystal structures were solved and refined with the programs within the CCP4 interface (10). The structures of SQVT80N and SQVT80S were solved with the molecular replacement program AMoRe (40). The coordinates of the WT protease in complex with the substrate (Protein Data Bank [PDB] code 1F7A) (47) were used as the starting model. Upon obtaining the rotation and translation solutions, the molecular replacement phases were further improved by using ARP/wARP (36) to build solvent molecules into the unaccounted for regions of electron density. Subsequently, interactive model building was carried out with O (22). Initial 2Fo-Fc and Fo-Fc maps unambiguously yielded the positions of SQV in each complex. Conjugate gradient refinement with ReLmac (38) was performed by incorporating the Schomaker and Trueblood tensor formulation of translation, libration, and screw rotation parameters (30, 56, 61). The working R (Rwork) and its cross validation (Rcross) were monitored throughout refinement. The geometry of the structure was assessed with Procheck (31) at the end of each refinement round. The refinement statistics are also shown in Table 1.

Isothermal titration calorimetry. To assess changes in the interaction between each variant and SQV relative to that of the WT, the binding affinity and enthalpy of interaction between each protease variant and SQV were measured by competitive displacement isothermal titration calorimetry experiment (26, 42,
TABLE 1. Crystallographic statistics of variant HIV-1 proteases

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SOV&lt;sub&gt;T80S&lt;/sub&gt;</th>
<th>SOV&lt;sub&gt;T80N&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt; (%)</td>
<td>7.3</td>
<td>4.2</td>
</tr>
<tr>
<td>Total no. of reflections</td>
<td>88,158</td>
<td>123,033</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt; (%)</td>
<td>23.2</td>
<td>19.8</td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.008</td>
<td>0.005</td>
</tr>
<tr>
<td>Bond angle (°)</td>
<td>1.6</td>
<td>1.4</td>
</tr>
<tr>
<td>No. of waters</td>
<td>130</td>
<td>193</td>
</tr>
<tr>
<td>No. of phosphates</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>PDB code</td>
<td>1FGU</td>
<td>1FGV</td>
</tr>
</tbody>
</table>

* RMSD, root mean square deviation.

The role of residue 80 in the structure, function, and viral infectivity of the HIV-1 protease was investigated in three variants, T80N, T80V, and T80S.
The spectrum of T80V was almost identical to that of the WT protease, and within the accuracy of protein concentration measurements, those of T80S and T80N were similar to that of the WT. The overall resemblance among these spectra indicates that these proteases were all folded similarly.

To further examine any structural changes caused by mutations at residue 80, tryptophan fluorescence was measured for all variants. Each protease monomer has two tryptophans, Trp6 at the dimer interface and Trp42 at the end of the flap region (Fig. 1). The fluorescence of the native tryptophans was measured for the WT protease and each variant in the presence and absence of SQV. The emission spectra for T80V, T80S, and T80N were similar to that of the WT, indicating that the tryptophans in the Thr80 protease variants are in an environment similar to that of those in the WT protease (Fig. 2b).

For both the WT and variant proteases, the presence of SQV led to a 10% decrease in the observed tryptophan fluorescence, suggesting that SQV interacts with all of these proteases. However, these results did not provide information on how this interaction may have been altered at Thr80.

Thermodynamic parameters of inhibitor binding. To reveal any altered interactions between SQV and the Thr80 variant proteases, the binding affinity and thermodynamic parameters of each variant and WT with respect to SQV were determined by isothermal titration calorimetry (Table 2). SQV was chosen because it is a well-characterized protease inhibitor (12, 15, 19, 29, 33, 52, 64). SQV is an entropically driven inhibitor. As SQV is a fairly hydrophobic molecule, it likely causes surrounding water molecules to become ordered. The favorable entropy observed in the reaction between the HIV-1 protease and SQV is presumably due to the release of this ordered water when SQV is sequestered within the active site of the protease. The free energy of binding of T80N and SQV was 4.2 kcal mol\(^{-1}\) worse than that between the WT and SQV. This decreased binding affinity was associated with larger changes in the enthalpy and entropy of the reaction. The entropy of T80N binding to SQV was 5.0 kcal mol\(^{-1}\) less favorable than that of the WT, while the enthalpic contribution was 17.7 kcal mol\(^{-1}\) less favorable. Similarly, a small decrease in binding affinity between T80V and SQV was associated with larger changes in the enthalpy and entropy of the reaction. While the free energy of the reaction between T80V and SQV decreased by 0.5 kcal mol\(^{-1}\) relative to that of the WT, the entropy of SQV binding was 4.5 kcal mol\(^{-1}\) more favorable and the enthalpy of the reaction was 5.0 kcal mol\(^{-1}\) less favorable than its binding to the WT. In contrast to the other two variants, there was no change within error in the free energy of the reaction between T80S and SQV relative to that of the WT. While the entropy of T80S binding to SQV was 0.5 kcal mol\(^{-1}\) more favorable than that of WT binding to SQV, the free energy of the reac-

### General structural changes
Effects of the three mutations on the secondary structure of the HIV-1 protease were probed by far-UV CD spectroscopy. Since HIV-1 protease is a predominantly β-sheet protein, any mutation that dramatically disrupted the structure would decrease the β-sheet content of the protein and increase its MRE at 216 nm. As shown in Fig. 2a, the three variants had secondary structures similar to that of the WT, comparable to previously published results (65).

![General structural features of WT and variant proteases](image)

**FIG. 2.** General structural features of WT and variant proteases (WT, black; T80V, red; T80N, yellow; T80S, green). (a) Far-UV CD spectra of WT and variant proteases. (b) Tryptophan fluorescence spectra of WT and variant proteases, unbound and bound to SQV. The unliganded WT protease is shown as a solid line, and the unliganded variant proteases are shown as filled symbols. The WT protease bound to SQV is shown as a dotted line, and the variant proteases are shown as empty symbols. AU, arbitrary units.

### Table 2. Binding thermodynamics of SQV at 20°C

<table>
<thead>
<tr>
<th>Protease</th>
<th>(K_a (M^{-1}))</th>
<th>(K_d (nM))</th>
<th>(K_d) ratio</th>
<th>(\Delta H (kcal\ mol^{-1}))</th>
<th>(\Delta G (kcal\ mol^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>(2.0 ± 0.1) × 10^9</td>
<td>0.50 ± 0.03</td>
<td>0.36 ± 0.06</td>
<td>3.6 ± 0.1</td>
<td>-16.1</td>
</tr>
<tr>
<td>T80S</td>
<td>(1.5 ± 0.1) × 10^9</td>
<td>0.66 ± 0.05</td>
<td>1.3</td>
<td>3.6 ± 0.1</td>
<td>-16.1</td>
</tr>
<tr>
<td>T80V</td>
<td>(8.8 ± 0.1) × 10^8</td>
<td>1.1 ± 0.02</td>
<td>2.3</td>
<td>3.6 ± 0.1</td>
<td>-16.1</td>
</tr>
<tr>
<td>T80N</td>
<td>(1.3 ± 0.1) × 10^8</td>
<td>760 ± 43</td>
<td>1.5 × 10^3</td>
<td>21 ± 1.7</td>
<td>-29.5</td>
</tr>
</tbody>
</table>

The spectrum of T80V was almost identical to that of the WT protease, and within the accuracy of protein concentration measurements, those of T80S and T80N were similar to that of the WT. The overall resemblance among these spectra indicates that these proteases were all folded similarly.
tion was maintained at approximately the same value by 0.7 kcal mol\(^{-1}\) less favorable enthalpy. The thermodynamic parameters of all reactions were measured with SQV. Therefore, the changes in enthalpy and entropy were likely due exclusively to changes in the free and bound states of the enzyme between the variants. Both the Thr80Val and Thr80Asn mutations decreased the binding affinity of protease for SQV, and all three mutations affected the thermodynamics of binding between the protease and SQV, demonstrating the importance of Thr80 in the activity of the protease.

**Enzymatic activity.** An assay of enzymatic activity was used to further verify the importance of Thr80 to protease function. Each variant protease was examined for deleterious effects on activity relative to the WT through comparison of the ability of each variant to cleave a fluorogenic MA-CA substrate. Both the WT and T80S completely cleaved all of the fluorescent substrate within 10 min (Fig. 3a). T80V required more than 2 h to completely cleave all of the substrate. T80N was not able to cleave the substrate. The catalytic efficiencies, \(k_{\text{cat}}/K_m\) values, of the WT and T80S were \(7.4 \times 10^{-2} \mu\text{M}^{-1}\text{s}^{-1}\) and \(8.4 \times 10^{-2} \mu\text{M}^{-1}\text{s}^{-1}\), respectively (Table 3), indicating greater catalytic activity for the T80S variant. The \(k_{\text{cat}}/K_m\) of T80V was \(4.5 \times 10^{-3} \mu\text{M}^{-1}\text{s}^{-1}\), a 16-fold decrease relative to that of the WT. The impact of the Thr80Val and Thr80Asn mutations on protease activity further confirms the importance of Thr80 in protease function.

**Viral infectivity.** A SpIn assay of viral fitness was performed with the Thr80 mutants to provide information about the proteases’ abilities to cleave the substrate sites necessary for viral maturation. The measured infectiousness of both the T80S virus and the T80V virus was within twofold of that of the WT virus. The apparent discrepancy between the enzymatic activity and the SpIn was probably due to the fact that the MA-CA cleavage site is not the rate-determining step in viral maturation (32, 44, 45, 63). This indicates that, despite impaired activity versus the MA-CA site, T80V was still able to cleave other sites with sufficient efficiency to maintain nearly WT levels of infectivity. In contrast, the T80N virus had very low levels of infectivity. When we examined the extent of processing by Western analysis (Fig. 3b), we found that the T80N virus was composed of mostly unprocessed Gag, similar to a D25A protease active-site mutant virus. In contrast, the T80S virus and the T80V virus were composed of Gag that had been largely processed although to a lesser extent than the WT virus.

**Structures of protease-SQV complexes.** The Thr80Asn mutation consistently impaired protease activity, while the Thr80Ser mutation did not affect protease activity. To determine what structural changes were causing the observed effect on protease activity, T80N and T80S were crystallized with SQV, yielding the structures SQVT80N and SQVT80S, with 2.0-Å and 1.5-Å resolutions, respectively. These structures were submitted to the PDB and assigned codes 1FGU and 1FGV, respectively. The electron density of both SQV and the mutated side chains was unambiguous, and there were no crystal contacts between the mutated side chains and surrounding protease molecules. The crystallographic statistics for these structures are listed in Table 1. The PDB structure of the WT HIV-1 protease bound to SQV (PDB code 1HXB) (29) has SQV in two orientations. In order to obtain the most accurate comparison between the structures and because T80S behaved just as the WT did, the SQVT80N and SQVT80S structures were compared with each other. To look for changes between the SQVT80N and SQVT80S structures, difference distance plots with the \(C_\alpha\) backbone of each structure were generated comparing the two protein variants (Fig. 4). This plot shows that one monomer of the SQVT80N structure underwent significant

### TABLE 3. Catalytic efficiencies of WT and variant proteases

<table>
<thead>
<tr>
<th>Protease</th>
<th>(k_{\text{cat}}/K_m) ((\mu\text{M}^{-1}\text{s}^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>((7.4 \pm 0.5) \times 10^{-2})</td>
</tr>
<tr>
<td>T80S</td>
<td>((8.4 \pm 1.3) \times 10^{-2})</td>
</tr>
<tr>
<td>T80V</td>
<td>((4.5 \pm 2.0) \times 10^{-3})</td>
</tr>
<tr>
<td>T80N</td>
<td>No detectable activity</td>
</tr>
</tbody>
</table>
structural rearrangement to accommodate asparagine, while the other monomer was able to accommodate the larger side chain without altering the position of the main chain atoms of the protease.

While residue 80 is not in the active site of the protease, it is bound to and contacts residues directly involved in ligand binding. Therefore, a mutation at this residue could cause significant rearrangement of the protease with inhibitor bound. As shown in the double difference distance plot, the distance between residues Pro79 through Val82 and most regions of the protein increased by more than 0.5 Å in SQVT80N relative to SQVT80S. The most significant increases were with residues in both flap regions, Lys43-Gln58 and Lys43’-Gln58’ and the P1 loop in the opposite monomer, Pro79’-Val82’. In SQVT80N, the side chain of Asn80 folded back on itself (Fig. 5a). Relative to Ser80, this conformation of Asn80 increased its interaction with Val32 and Leu33 and formed a hydrogen bond between Asn80 OD and Pro79 N. The side chain amide nitrogen of Asn80 and the side chain carboxyl oxygen of Ser80 maintained the conserved hydrogen bond with the backbone oxygen of Val82. However, the conformation of Asn80 altered the P1 loop relative to its position in SQVT80S, pulling Pro81 back from the 2-methyl-decahydroisoquinoline-3-carboxylic acid-P1’ (DiqP1’) group of SQV and Val82 toward DiqP1’. In SQVT80N, SQV lost a contact with Pro81, which was above it, when the distance between Pro81 CG and DiqP1’ C5 increased by 0.7 Å and gained a contact with Val82, which was below it, when the distance between Val82 CG1 and DiqP1’ C6 decreased by 0.9 Å in SQVT80N relative to SQVT80S. Figure 6b shows contacts that decreased by more than 0.2 Å in SQVT80N relative to SQVT80S and the opposite information is displayed in Fig. 6a. These results emphasize that structural rearrangement of the P1 loop can significantly alter the interactions between the protease and SQV.

Although one monomer of the SQVT80N structure had rearranged the P1 loop relative to SQVT80S, the other monomer did not undergo any repositioning of its main chain atoms relative to the SQVT80S structure. As shown in the double difference plot, the distance between 80 Cα and other residues was not significantly different between the SQVT80N and SQVT80S structures, indicating that when the protease variants were bound to SQV, the protease was able to accommodate asparagine at residue 80. Asn80’ adopted a conformation similar to that of Ser80’ (Fig. 5b), and both residues formed the same hydrogen bonds with the main chain atoms of Pro81’ and Val82’. A significant difference between the two structures was the presence of a water molecule in the SQVT80N structure between Ile50 and residues Pro79’ and Pro81’ that was not observed in the SQVT80S structure. This water molecule forms
one hydrogen bond with the side chain amine group of Asn80\textsuperscript{H11032}, one contact with SQV, and numerous contacts with both main chain and side chain atoms of Ile50, Pro79\textsuperscript{H11032}, Asn80\textsuperscript{H11032}, and Pro81\textsuperscript{H11032}. Despite the increased size of asparagine and the presence of the extra water molecule, there was remarkable similarity of this monomer between the structures. However, there was some structural rearrangement of SQV and a few variations in the van der Waals contacts made between the protease and SQV. There was one contact between Asn80\textsuperscript{H11032} and the phenylalaninol-P1 group of SQV not observed in the SQV T80S structure. Conversely, the distance between Pro81\textsuperscript{H11032} and the phenylalaninol-P1 and 2-carbonylquinoline-P3 groups was increased by more than 0.2 Å in SQV T80N relative to SQV T80S. Two van der Waals contacts were lost between SQV and the SQVT80N variant protease because of this increased distance (Fig. 6). This indicates that, in contrast to the structural rearrangements observed by residues Pro79 and Val82, the altered interactions between SQV and protease in this monomer are due to rearrangements of SQV to accommodate Asn80\textsuperscript{H11032}.

Interactions between SQV and the protease flap region differed between SQVT80N and SQVT80S. These changes were caused by the structural modifications that were made by either the P1 loops or SQV. For example, in the SQVT80S structure, the DiqP1\textsuperscript{H11032} group has the most van der Waals contacts with the flap region while in the SQVT80N structure, the asparagine-P2 group of SQV had the most contact with the flap region. The 2-carbonylquinoline-P3 group of SQV was closer to Gly48, and asparagine-P2 moved closer to Ile50\textsuperscript{H11032} in SQVT80N relative to their distance in SQVT80S. Thus, substitution of asparagine at residue 80 was able to indirectly affect the flap region in the bound protease.

**Changes in flap dynamics.** In order to further examine the role of residue 80 in the structure and conformational flexibility of the flap region of the unbound protease, a separate MD simulation was performed for the WT and each variant protease, T80V, T80N, and T80S. Each simulation was started from the 9-ns structure of the previous MD simulation (57) since the flap region was already folded in toward the P1 loop.
in this structure. Because these were relatively short simulations and nothing was restraining the protease, the two monomers in each simulation exhibit two conformational samplings of protease monomers. Figure 7 shows four time snapshots from the four simulations at 0 ns, 1.5 ns, 3 ns, and 4.5 ns. The WT, T80S, and T80V simulations sampled similar conformational spaces. Although each monomer moves separately, in all three simulations the protease flap region unfolds into the solvent. To quantify the range of motion of the protease flaps in each simulation, the distance between residue 50 and residue 80 in both monomers in all of the simulations was monitored over time. In the WT, T80V, and T80S simulations, the distance between these two residues fluctuated between 5 and 22 Å in both monomers. In one monomer of the T80S simulation, the space sampled by the flap region was significantly decreased, relative to the WT and its other monomer, because of a hydrogen bond between the backbone oxygen of Pro79 and the backbone nitrogen of Ile50. After approximately 3.5 ns, that hydrogen bond was broken and the flap began to unfold into the solvent. In contrast to the other simulations, the flaps within the T80N simulation did not unfold into the solvent. As shown in Fig. 7, there was decreased flexibility in both flap regions in the T80N simulation relative to the other protease variants. The distance between Asn80 and Ile50 fluctuated between 7 and 17 Å in one monomer and between 6 and 8 Å in the other monomer. In this simulation, the flap region with decreased movement was bound to the P1 loop by two hydrogen bonds, one between the side chain oxygen of Asn80 and the backbone nitrogen of Gly49 and the other between the backbone oxygen of Pro79 and the same backbone nitrogen of Gly49. In the other monomer, there was not a specific hydrogen bond between the flap region and the P1 loop, yet this flap remained curled toward the P1 loop for most of the simulation. This was probably due to water bridging hydrogen bonds between the flap and Asn80. These results demonstrate that substituting a large polar residue for threonine at residue 80 can significantly and detrimentally affect the conformational flexibility of the flap region. However, substitution of a non-polar residue did not impact the ability of the protease flap to unfold into the solvent.

As mentioned previously, there is a conserved hydrogen bond between the side chain of residue 80 and the backbone of residue 82. In these simulations, the existence of this hydrogen bond over time was monitored, except in the T80V simulation, where no hydrogen bond can exist. In the WT simulation, this hydrogen bond existed less than 10% of time, and in the T80S simulation it existed less than 1% of the time. It did not exist in the T80N simulation. As shown in Fig. 7, this hydrogen bond was not required to maintain the stability of the P1 loop. These simulations provide interesting insights into the impact of variation at residue 80 on protease dynamics and may help to explain the results seen in the various activity assays described above.

**DISCUSSION**

Many residues within the HIV-1 protease are susceptible to mutation, especially under the selective pressure of therapy;
therefore, those residues that are invariant despite therapy are likely important to enzyme structure and/or function. Thr80, which is invariant in the HIV-1, HIV-2, and simian immunodeficiency virus proteases, regardless of protease inhibitor treatment, is one such residue. While not directly involved in ligand binding, it is located on the edge of the active-site cavity in the middle of the P1 loop. To examine the role of this residue in protease, we constructed three protease variants, T80S, T80N, and T80V, and used a variety of biochemical techniques to characterize its role in the structure, dynamics, and function of protease. Additionally, a viral infectivity assay was used to examine the impact of changes at residue 80 on the virus as a whole. The results indicate that mutation at this site can detrimentally impact the enzymatic activity, decrease viral infectivity, alter the enthalpy and entropy contributions to inhibitor binding, and likely decrease the conformational flexibility of the HIV-1 protease.

Of the three variants studied, T80S had the least impact on protease structure and function. This variant had a slightly increased \( k_{cat}/K_m \) relative to the WT but no substantial change in its affinity for SQV and only an approximately twofold effect on viral infectivity. Thus, this mutation does not appear to give the virus any selective advantage or cause any resistance to SQV. Further studies are needed to clarify why Thr80Ser does not occur as a natural variant. Regardless of the amino acid at residue 80, all of the protease variants had structures similar to that of the WT (Fig. 2). Two variants, T80S and T80N, were crystallized with the inhibitor SQV. The structures were remarkably similar, except that the Thr80Asn mutation led to a structural rearrangement of the P1 loop in one monomer relative to the SQV-T80N structure (Fig. 5) that altered the van der Waals contacts of T80N relative to those of T80S (Fig. 6). Despite the rearrangement of this loop, the asparagine at this site could be involved in hydrogen bonds with Val82. In fact, maintenance of this hydrogen bond maybe what led to the rearrangement of the P1 loop in this monomer. The hydrogen bond between the side chain of residue 80 and the backbone of Val82 was also observed in the other monomer of SQV-T80S and in both monomers of SQV-T80N. Thus, an important role of Thr80 in the bound protease appears to be the formation of this hydrogen bond.

While this hydrogen bond was conserved in the crystal structures, it was not conserved in the MD simulations of unbound protease. MD simulations were performed on all variant proteases and on the WT. In the WT, T80S, and T80N simulations, residue 80 was capable of forming a hydrogen bond between its side chain and the backbone of Val82. However, this hydrogen bond existed only intermittently in the WT and T80S simulations and did not exist in the T80N simulation. Without the hydrogen bond between residue 80 and Val82, the P1 loop itself was still structurally stable. Actually, residue 80 appeared to have the most significant impact on the flexibility of the flap region. In the WT, T80S, and T80V simulations, the flap region uncurled into the solvent (Fig. 7). This was somewhat surprising in the T80V simulation because our original hypothesis was that polar Thr80 prevented Ile50 from packing into the hydrophobic pocket formed by other residues surrounding the P1 loop. Unexpectedly, the polar asparagine at residue 80 prevented the flap from uncurling by forming electrostatic interactions with the flap region, limiting its flexibility.
SVO, and the solvent were constant between the reactions. For T80V, the loss of polar threonine to hydrophobic valine could explain the observed changes in enthalpy and entropy. This mechanism may therefore account for the im-
more water would be released from the active site of T80V when binding SVO. Additionally, a valine would have more conformational freedom as it is not restricted by a hydrogen bond to the backbone. Both aspara-
gine and threonine are polar residues, but asparagine is larger and capable of forming polar interactions with residues in the flap region. In the case of T80N, the –13.4 kcal mol⁻¹ more favorable entropy could be due to the change in the flap dynamics that was observed in the MD simulation of the unli-
ganded form of the enzyme. In contrast to the other protease variants, where the flaps are very flexible and a large loss of conformational entropy within the protease occurs upon form-
ing a complex, the flaps of T80N are much less flexible. There-
to, there is less conformational entropy to lose with T80N and the binding would be entropically more favorable than with the other variants. Thus, combining the results of these biophysical measurements, the molecular role of Thr80 in the HIV-1 protease is that of a small hydrophobic residue that facilitates flap flexibility in the unliganded state while allowing tight packing of the active site in the liganded state of the enzyme. This mechanism may therefore account for the im-
portance and observed invariance of Thr80 within the HIV-1 protease.

ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of Osman Bilsel, Jim Caldwell, Hong Cao, Anthony Arruthers, Amanda Fitzgerald, Nancy King, William Kobertz, Mary Munson, Ellen Nailvalka, Snohob, O'Brien, John Osterhout, and Ying Wu in the acquisition and analysis of data and Claire Baldwin in the preparation of the manuscript.

This research was supported by NIH grant P01-GM66524. This work was supported in part by NIH grant R01-AI50485 to R.S. G.J.H. was supported by training grant T32-AI07419 and by fellowship award F30-DA019379. We also acknowledge the use of reagents obtained from Drs. K. Ariyoshi, M. Matsuda, H. Miura, S. Tateishi, K. Yamada, and W. Sugiura.

REFERENCES


tallogr. 50:760–763.

anism, and data analysis. Wiley-VCH, Weinheim, Germany.

deficiency virus type 1 protease cleavage site mutations associated with protease inhibitor cross-resistance selected by indinavir, ritonavir, and/or saquinavir. J. Virol. 70:589–594.


vention and therapy. Liss, New York, N.Y.


tallographic and modeling methods in molecular design. Springer-Verlag, Berlin, Germany.


35. Minor, W. 1993. XDISPLAYF program. Purdue University, West Lafayette, Ind.


