Crystal structure of human T cell leukemia virus protease, a novel target for anticancer drug design

Mi Li*†, Gary S. Laco‡§, Mariusz Jaskolski¶, Jan Rozycki‡, Jerry Alexandratos*, Alexander Wlodawer*, and Alla Gustchina*‖

*Protein Structure Section, Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick, MD 21702; †Basic Research Program, Science Applications International Corporation, Frederick, MD 21702; ‡Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892; and ¶Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University and Center for Biocrystallographic Research, Institute of Bioorganic Chemistry, Polish Academy of Sciences, 61 704 Poznan, Poland

Communicated by David R. Davies, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, October 26, 2005 (received for review August 25, 2005)

The successful development of a number of HIV-1 protease (PR) inhibitors for the treatment of AIDS has validated the utilization of retroviral PRs as drug targets and necessitated their detailed structural study. Here we report the structure of a complex of human T cell leukemia virus type 1 (HTLV-1) PR with a substrate-based inhibitor bound in subsites P5 through P5′. Although HTLV-1 PR exhibits an overall fold similar to other retroviral PRs, significant structural differences are present in several loop areas, which include the functionally important flaps, previously considered to be structurally highly conserved. Potential key residues responsible for the resistance of HTLV-1 PR to anti-HIV drugs are identified. We expect that the knowledge accumulated during the development of anti-HIV drugs, particularly in overcoming drug resistance, will help in designing a novel class of antileukemia drugs targeting HTLV-1 PR and in predicting their drug-resistance profile. The structure presented here can be used as a starting point for the development of such anticancer therapies.

inhibitor | leukemia | retroviral protease

Human T cell leukemia virus type 1 (HTLV-1) is a retrovirus that is epidemiologically associated with mature $CD3^+CD4^+$ T cell-type leukemia/lymphoma (ATL), as well as with tropical spastic paraparesis/myelopathy (1, 2). It is estimated that up to 30 million people worldwide are infected with HTLV, with ATL being particularly prevalent in Japan (3). Only an estimated 3–5% of people infected with the virus develop ATL in their lifetime, but for those that do, the prognosis is poor (4). Although a number of treatments for ATL, such as combination chemotherapy, monoclonal antibodies directed against the α chain of the interleukin 2 receptor, and antiviral therapy involving IFN-α and zidovudine, are used clinically, they show only very limited efficacy (3). Novel approaches under investigation use proteasome inhibitors (5) and Tax-targeted immunotherapy (4), but they have not yet been tested in practice. It is clear that new anti-ATL targets need to be found.

In common with other retroviruses, HTLV-1 encodes a protease (PR) necessary for its maturation. Because inhibition of the enzyme has been shown to prevent viral proliferation, development of inhibitors targeting HTLV-1 PR is an attractive new path for chemotherapy (6). HTLV-1 PR is a homodimer, with each chain containing 125 residues. The enzymatic properties of HTLV-1 PR, including its substrate specificity, have already been studied in considerable detail (7, 8). Although the design and synthesis of inhibitors specific for HTLV-1 PR have been carried out, most of the compounds are active only in micromolar concentration (9, 10). The best statine-containing inhibitor has a K_i of 50 nM under high-salt conditions (7) but of only 2.3 μM in a low-salt buffer (11). In comparison, a number of subpicomolar inhibitors of HIV-1 PR have been developed by using the principles of rational drug design (12).

Structural investigations of HTLV-1 PR have not been successful in the past, due primarily to difficulties in expressing soluble protein with high and stable activity and in growing crystals. Thus, until now, only model structures could guide the development of specific inhibitors (6, 8). However, the limitations of the modeling approaches were clear, and the need for an experimental structure became obvious. We have now succeeded in crystallizing an *Escherichia coli*-expressed variant of HTLV-1 PR, with a nine-residue truncation at the C terminus. The protein is enzymatically active and can be inhibited by a compound that is a modification of the best-known HTLV-1 PR inhibitor. The structure explains the failure of HIV-1 PR inhibitors to inhibit HTLV-1 PR and defines a molecular target for the design of specific inhibitors for efficient therapies in HTLV-associated diseases.

Materials and Methods

Protein Expression and Purification. A plasmid containing the HTLV-1 PR gene (13) was modified via PCR in the following ways: (*i*) an NdeI restriction site was added to the 5′ end, resulting in an initiation Met codon being added 5′ to the PR Pro-1 codon; (*ii*) the Leu-40 codon was mutated to Ile to block autolysis (7); and (*iii*) a stop codon and a BamHI restriction site were introduced 3′ of the Pro-116 codon. The HTLV-1 PR gene was then cloned into pET-21 (Novagen) by using the NdeI and BamHI restriction sites to give pHTLVΔ9PR, and the construct was sequenced to confirm the mutations. pHTLVΔ9PR was transformed into *E. coli* BL21(DE3) pLysS cells (Novagen), and protein induction and inclusion body purification were performed as previously described, except that the inclusion bodies were washed with 0.5 M instead of 1 M, urea, and Nonidet P-40 was omitted (14). The HTLVΔ9PR inclusion bodies were solubilized in 8 M urea/10 mM Tris, pH 7.5/5 mM EDTA/5 mM 2-mercaptoethanol and were passed through a HiTrap Q column (Amersham Pharmacia) equilibrated with 6 M urea/20 mM Tris, pH 7.5/5 mM EDTA/5 mM 2-mercaptoethanol. The eluate was adjusted to pH 3.0 and loaded onto a HiTrap SP column equilibrated with buffer A (20 mM sodium acetate, pH 3.0/6 M urea/5 mM EDTA/5 mM 2-mercaptoethanol). The bound HTLVΔ9PR protein was eluted with a 0–1 M NaCl gradient in buffer A; dialyzed against 15 mM sodium acetate, pH 3.0/5% polyethylene glycol 300/5 mM DTT; and either stored at 5°C or made 50% glycerol and stored at −20°C. The HTLVΔ9PR

Conflict of interest statement: No conflicts declared.

Abbreviations: HTLV-1, human T cell leukemia virus type 1; PR, protease; ATL, T cell-type leukemia/lymphoma; SIV, simian immunodeficiency virus; EIAV, equine infectious anemia virus; FIV, feline immunodeficiency virus; RSV, Rous sarcoma virus.

Data deposition: The structural coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2B7F).

§Present address: Lake Erie College of Osteopathic Medicine, Bradenton, FL 34211.

‖To whom correspondence should be addressed. E-mail: alla@ncifcrf.gov.

protein was ≈95% pure, as judged by Coomassie blue-stained SDS/PAGE gels.

Synthesis and Purification of the Inhibitor Ac-Ala-Pro-Gln-Val-Sta-Val-Met-His-Pro. The inhibitor was synthesized on an ABI 431 Peptide Synthesizer (Applied Biosystems) (0.25 mM scale) starting with H-Pro-2-chlorotrityl resin. Standard FastMoc protocol was used for all synthetic cycles except for the Fmoc-Statine coupling reaction, which was carried out manually for ≈14 h with only 2-fold molar excess of Fmoc-Statine. The completeness of the coupling was confirmed by the ninhydrin test. After cleavage of the peptide from the resin, the crude product was purified by semipreparative RP-HPLC. Peptide purity was verified by analytical RP-HPLC and MALDI-TOF MS.

PR Assays. The HTLVΔ9PR was assayed for activity by using a fluorogenic substrate (acetyl-KDKTK-AbzVL/F-NO_2VQPKK-NH_2), where/indicates the scissile bond, and Abz and NO_2 are the donor and acceptor chromophores, respectively. Cleavage of the substrate was monitored at 37°C with an excitation wavelength of 325 nm and an emission wavelength of 410 nm. PR assay buffer contained 0.5 M NaCl; 50 mM NaAcetate, pH 5.5; and 5 mM DTT.

Preparation and Crystallization of HTLV-1 PR–Inhibitor Complex. The complex of HTLV-1 PR with the inhibitor was prepared by mixing the protein solution and the inhibitor (dissolved in 100% DMSO) at a molar ratio of 1:10 (protein monomer/inhibitor). The sample was concentrated in an Amicon (Millipore) stirred cell concentrator under nitrogen gas, at 5°C, by using a BioMax (Fairmouth, MA) polyethersulfone membrane with 100-kDa cutoff. This was necessitated by the aggregation of the protein, because a membrane with a lower cutoff was becoming clogged during the procedure. The eluate did not contain detectable amounts of protein. The sample was subsequently centrifuged for 4 min at 5°C in a table-top Eppendorf centrifuge. The final protein concentration was determined using a Bradford assay (Bio-Rad) with BSA as the standard and was typically 6–7 mg/ml. Because inhibitor that was not bound to the protein was lost during the concentration/dialysis step, the sample solution was supplemented with additional inhibitor, resulting in a 1:4 protein/inhibitor molar ratio (protein monomer/inhibitor) immediately before crystallization. Crystals of HTLV-1 PR were grown by the vapor diffusion method in hanging drops mixed from 4 μl of protein solution and 4 μl of well solution consisting of 17% polyethylene glycol (PEG) 8000, 16% PEG 300, and 10 mM DTT in 0.1 M acetate buffer, pH 5.2.

X-Ray Data Collection and Analysis. X-ray diffraction data extending to 2.6-Å resolution were collected at the Southeast Regional Collaborative Access Team beamline 22-ID (Advanced Photon Source, Argonne National Laboratory, Argonne, IL) on a MAR225 charge-coupled device detector (MAR-Research, Hamburg) at the wavelength of 1.0 Å. Data were processed and scaled with HKL2000 (HKL Research, Charlottesville, VA) (15) (Table 1).

Structure Solution and Refinement. The structure was solved by molecular replacement with the program PHASER (16). The structure was fitted and rebuilt with O (17) and refined with REFMAC5 (18) and CNS (19) (Table 1). More details of structure solution and refinement will be provided elsewhere (M.J., M.L., G.S.L., A.G., and A.W., unpublished results).

Results and Discussion

The Structure of HTLV-1 PR and Comparison with Other Retroviral PRs. Because numerous attempts to crystallize full-length HTLV-1 PR failed, several mutated forms of the PR have been constructed for crystallization purposes. Unlike most other retroviral PRs, HTLV-1 PR carries a C-terminal extension that is not essential for enzymatic activity *in vitro* (20). Among a series of C-truncated constructs (data not shown), a variant of HTLV-1 PR containing residues 1–116 yielded crystals of sufficient quality for structure determination. In our hands, this variant exhibited 60% of the activity of the wild-type enzyme. The enzyme was cocrystallized with the inhibitor Ac-Ala-Pro-Gln-Val-Sta-Val-Met-His-Pro, a modification of the substrate-based inhibitor with a reported K_i of 50 nM (7). The structure was solved by molecular replacement by using an atomic resolution model of HIV-1 PR (21) as a probe and was refined to 2.6-Å resolution. The crystallographic asymmetric unit contains three homodimeric molecules (AB, CD, and EF), and the main chain could be traced end to end in all of them. The inhibitor is fully ordered in two dimers (AB and EF), enabling its complete tracing. In dimer CD, the inhibitor shows 2-fold disorder that follows the pseudosymmetry of the enzyme. The PR dimer AB with bound inhibitor I is shown in Fig. 1*A*.

The three PR dimers are nearly identical and, when superimposed with the program ALIGN (22), show rms deviations of 0.31 Å, for the 232 Cα pairs between dimers AB and CD, and 0.25 Å for the 230 atom pairs between dimers AB and EF. The smaller deviation for the latter pair reflects the similarity of their interactions with the inhibitor, which is bound in the same orientation in both molecules, in contrast to its dual orientations in dimer CD. The slight asymmetry of dimers AB and EF, attributed to the polarity of the inhibitor, is noticeable in a comparison of the rms deviations for the superimposed monomers within these dimers (0.41 Å for both A/B and E/F) to that for monomers A and E, which interact with the same end of the

Table 1. Data collection and refinement statistics

Data collection	
Space group	*C*2
Cell dimensions	
a, *b*, *c*, Å; β, °	134.32, 77.79, 80.38; 99.3
Resolution, Å	50.0–2.6 (2.63–2.60)*
R_{merge}†	8.9 (30.4)
No. reflections (measured/unique)	161514/24645
$\langle I/\sigma I \rangle$	21.7 (2.8)
Completeness, %	98.1 (85.7)
Redundancy	6.55 (3.53)
Refinement	
Resolution, Å	10–2.6
No. reflections (refinement/R_{free})	23,030/1,143
R/R_{free}‡	0.198/0.278
No. atoms	
Protein	5,298
Ligand/ion	294
Water	172
B factors, Å²	
Protein	35.8
Ligand/ion	45.3
Water	38.3
rms deviations from ideal	
Bond lengths, Å	0.022
Bond angles, °	2.18
Ramachandran torsion pairs	
Allowed/additional/generous, %	89.0/9.9/1.1

*Highest-resolution shell is shown in parentheses.

†$R_{merge} = \Sigma_h \Sigma_i |I_i - \langle I \rangle| / \Sigma_h \Sigma_i I_i$, where I_i is the observed intensity of the *i*th measurement of reflection h, and $\langle I \rangle$ is the average intensity of that reflection obtained from multiple observations.

‡$R = \Sigma ||F_o| - |F_c|| / \Sigma |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively, calculated for all data. R_{free} is defined in ref. 33.

BIOCHEMISTRY

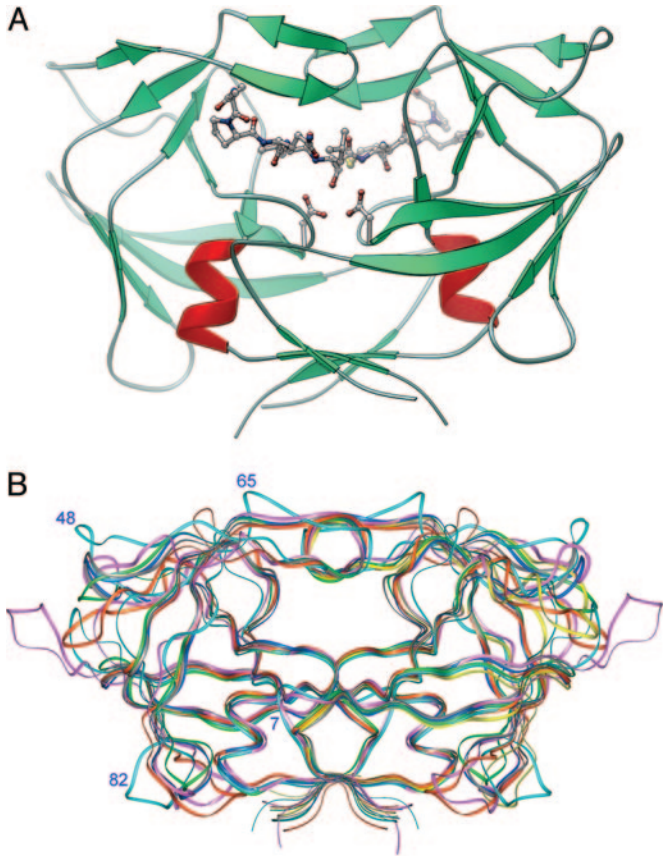

Fig. 1. The structure of HTLV-1 PR and a comparison with other retroviral PRs. (*A*) Overall view of a dimer of HTLV-PR. Helices are shown in red and β strands in pale green. The inhibitor and the catalytic aspartates are shown in stick representation. (*B*) Superposition of seven retroviral PRs shown in ribbon representation. HTLV-1 PR is colored blue; HIV-1 PR, green; HIV-2 PR, dark blue; SIV PR, gray; RSV PR, magenta; EIAV PR, yellow; and FIV PR, red. The numbers indicate residues within regions in HTLV-1 PR, with the most pronounced structural differences as compared with other retroviral enzymes.

inhibitor (0.24 Å). Side chains that are either directly or indirectly involved in crystal packing also have different orientations in the individual molecules. However, such differences between the monomers are not large, and thus any of them could be used for comparisons with other retroviral PRs. In the following, molecule A of HTLV-1 PR is discussed.

The overall fold of HTLV-1 PR is similar to that found in other retroviral PRs (Fig. 1*B*). Superposition of the dimer AB of HTLV-1 PR with other retroviral PRs shows rms deviations of 1.53 Å for HIV-1 PR (182 Cα pairs) (21), 1.64 Å for HIV-2 PR (190 Cα pairs) (23), 1.70 Å for SIV PR (191 Cα pairs) (24), 1.72 Å for equine infectious anemia virus (EIAV) PR (190 Cα pairs) (25), 1.77 Å for feline immunodeficiency virus (FIV) PR (187 Cα pairs) (26), and 1.93 Å for 220 Cα pairs of a nine-site mutant of Rous sarcoma virus (RSV) PR (27). A structure-based sequence alignment was created on the basis of those superpositions to evaluate the level of sequence similarity between HTLV-1 PR and the other enzymes (Fig. 2). When all seven enzymes are compared, only 15 residues are identical, whereas 19 are of a similar type. HTLV-1 PR has the highest level of identity with HIV-1 and EIAV PRs (31 residues) and of similarity with RSV PR (28 residues). The highest combined level of identity plus similarity is with RSV PR (56 residues vs. 50–53 for the other PRs).

The secondary structure of each polypeptide chain of HTLV-1 PR is closely related to that of the other enzymes, whereas the length of the strands and helices varies (Fig. 2). Residues 43–45 are found in a 3_{10} helix resembling that of FIV PR and an α helix of EIAV PR. A second helix, containing residues 103–110, is the longest observed so far in any retroviral PR. The conformations of the loops connecting the strands and helices are significantly different in HTLV-1 PR and likely determine the unique properties of the enzyme. The most dramatic changes were found in the flap area and in loop 91–100, equivalent to the so-called "polyproline" loop 76–84 of HIV-1 PR. Both regions carry functionally important residues that are involved in extensive interactions with the inhibitor and also participate in dimer stabilization.

The presence of two pseudosymmetric flaps in retroviral PRs that cover the substrates/inhibitors and form a number of intimate contacts with them is one of the most characteristic features of these homodimeric enzymes. With the exception of a few structures of unbound enzymes, in which the flaps were either open (28) or not seen due to disorder caused by their flexibility (29), the flaps assume a very similar conformation in all enzymes that have been studied so far. The tips of the two HIV-1 PR flaps approach each other in a parallel fashion, the distances between their Cα atoms being 4–5 Å (Fig. 3). High-resolution structures show that the symmetry of the interacting flaps is usually broken via a flip of the peptide bond between residues 50–51 (HIV-1 PR numbering), resulting in a hydrogen bond between flap residues 50 and 50′. In many structures, both peptide orientations are present, creating 2-fold disorder in the tips of the flaps.

Although the general features of the flaps, such as hydrogen bonds between the backbone atoms of the two strands within each hairpin, exemplified by the two hydrogen bonds between residues 58 and 61 (equivalent to 49 and 52 in HIV-1 PR) are preserved in HTLV-1 PR, the interactions between the tips of the flaps are very different. As in other complexes of retroviral PRs with peptidomimetic inhibitors, the flaps in HTLV-1 PR are locked in a closed conformation over the ligand. However, only the leading strands facing the inhibitor (residues 56–59) are in direct contact in a dimer, whereas the trailing strands of the flaps (residues 60–63) are far apart and do not interact with each other. The distances between the equivalent Cα atoms on the flaps of the two monomers vary from $\approx$5 Å on the leading strands to $\approx$8 Å on the trailing strands, making them significantly less parallel than in other retroviral enzymes (Fig. 3). Therefore, the chain leading to the tip of the flap follows a rather similar path in HTLV-1 PR and in other retroviral PRs (Cα–Cα distances between equivalent atoms are in the range of 0.5–2 Å), whereas the residues on the trailing end diverge more, with the Cα–Cα distances between Gly-61 and its equivalents in other PRs exceeding 3 Å. The hydrogen bond between the tips of the flaps, observed in the majority of the structures of other retroviral PRs, is not formed in HTLV-1 PR dimer (Fig. 3*A*).

Another unique feature of the flap region in HTLV-1 PR is an insertion of two residues into the stretch 64–68, which induces a nearly helical conformation of this region, thus disrupting the hydrogen-bonding pattern within the hairpin structure of the flap (Fig. 3*B*). The zigzag conformation of the backbone is stabilized by a weak hydrogen bond between the amide of residue 66 and the side-chain hydroxyl of Thr-63.

The loop 95–98 in HTLV-1 PR has only a one-residue insertion compared with the "polyproline" loop 79–81 in HIV-1 PR, but the conformation of that region partially resembles FIV PR, which has an insertion of three residues in the corresponding loop (Fig. 2). The unique conformation of loop 95–98 in HTLV-1 PR influences the architecture of the binding sites S1/S1′ and S3/S3′, which use residues from that structural element, such as Asn-97 and Trp-98. Residues from the segments with novel conformations in both loops, such as His-66, Phe-67, Lys-95, and

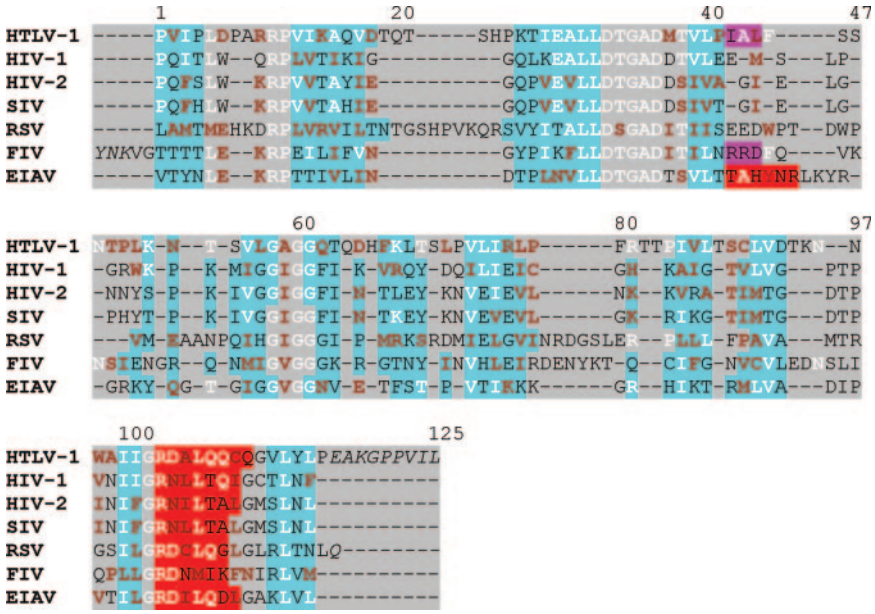

Fig. 2. Structure-based sequence alignment of the retroviral PRs with known structures. The color of the background indicates the secondary structure elements, with cyan denoting β strands; red, α helices; magenta, 3_{10} helices; and gray, loops and irregular structure. Conserved residues are bold and white, and residues similar to those in HTLV-1 PR are bold and brown, whereas the remaining residues are black. Residues that were either not present in the constructs used to solve the structures (HTLV-1 PR) or not visible in the electron density maps (FIV PR) are italicized.

Trp-98 (Fig. 6, which is published as supporting information on the PNAS web site), also contribute to dimer stabilization.

In common with other retroviral PRs, the dimer interface of HTLV-1 PR is a four-stranded antiparallel β-sheet composed of the N and C termini of both monomers. In all retroviral PRs, both termini are involved in electrostatic interactions either directly (HIV-1, HIV-2, SIV, and EIAV PRs) or indirectly (FIV and RSV PRs), and these interactions depend on the lengths of the termini. Thus, in HIV-1 PR, the matched lengths of the termini enable the formation of an ion pair between them. In contrast, the C terminus of the present HTLV-1 PR construct is one residue longer, precluding direct electrostatic interactions with the N terminus. However, on at least one side of each dimeric interface, the N terminus of one monomer interacts with the C terminus of the second monomer through a bridging phosphate anion. Thus, the N termini of molecules B, D, and F interact with phosphates that are also located within hydrogen bond distance to the C termini of molecules A, C, and E, respectively. One of these phosphates is located on a local two-fold axis, simultaneously mediating similar interactions between the termini of dimers AB and E′F′ and thus facilitating the formation of the crystal lattice (M.J., M.L., G.S.L., A.G., and A.W., unpublished results). That phosphate anion has almost perfect tetrahedral coordination, with the two remaining oxygen

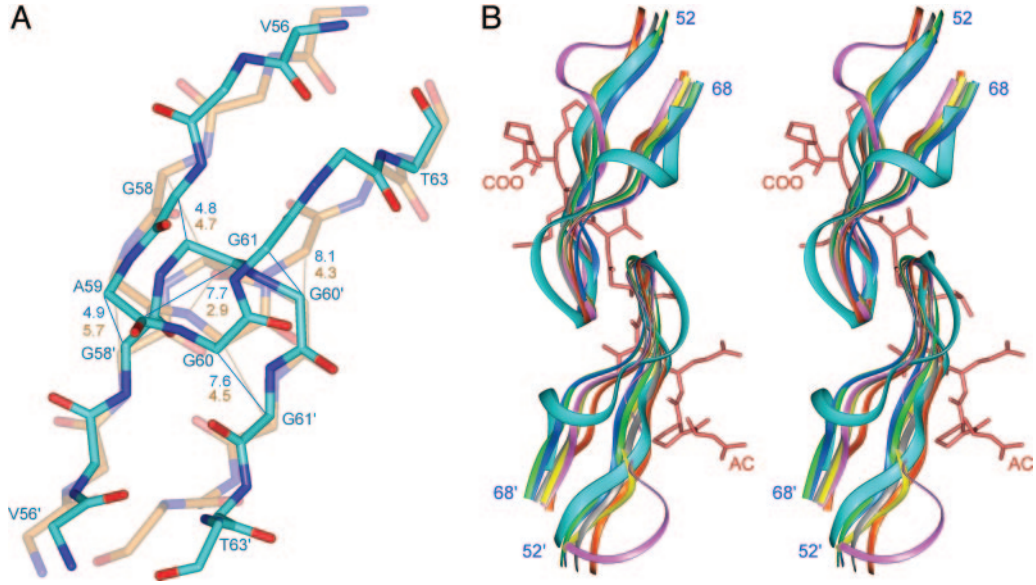

Fig. 3. The flaps of retroviral PRs. (*A*) Superposition of the flaps in HTLV-1 (blue) and HIV-1 (tan) PRs. The distances between the corresponding pairs of atoms on the leading and trailing edges of the flaps are indicated. (*B*) Superposition of the flaps of the seven retroviral PRs (colored as in Fig. 1*B*) shown in ribbon representation, in stereo. The inhibitor bound to HTLV-1 PR is shown for reference in stick representation.

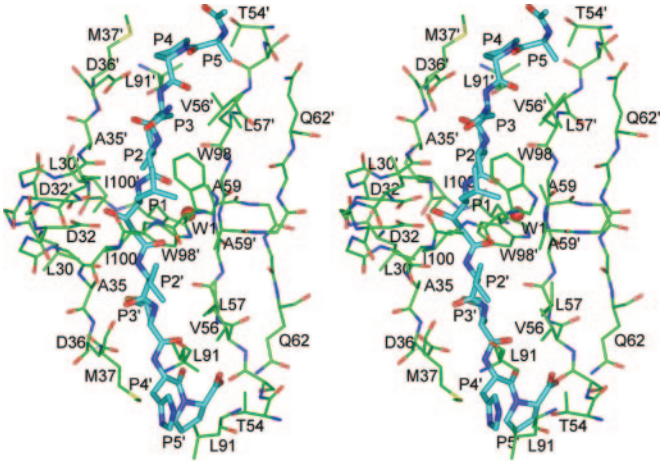

Fig. 4. Stereoview showing the binding of the inhibitor I (blue, with O atoms red and N, dark blue) to the AB dimer of HTLV-1 PR (green, with O atoms red; N, blue; and S, yellow). The conserved water located between the flaps and the inhibitor is shown as a red sphere.

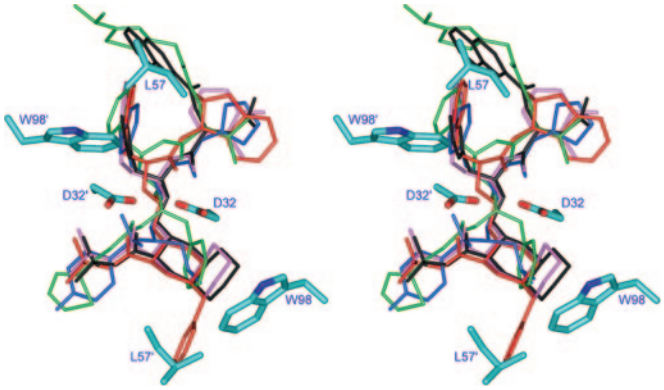

Fig. 5. Stereoview of the overlay of the inhibitor bound to HTLV-1 PR with the clinical inhibitors of HIV-1 PR. The alignment is based on the superposition of the Cα atoms of the proteins. Amprenavir is shown in blue, ritonavir in green, nelfinavir in pink, saquinavir in black, and indinavir in red. Selected residues of HTLV-1 PR that interfere with drug binding are shown in thick lines.

atoms hydrogen bonded to the hydroxyls of Tyr-114 in molecules A and E′. An equivalent type of interactions is found in dimer CD, with the exception that in the crystal this dimer interacts with its symmetry mate because the phosphate ion is located on the crystallographic two-fold axis. The N termini of molecules A, C, and E, as well as the C termini of B, D, and F, do not make contacts with any other molecules related by either local or crystallographic symmetry. Interestingly, in all monomers, the carbonyl oxygen of Leu-115 of one monomer is within hydrogen-bonding distance of the N terminus of the second monomer, thus indicating that a variant of HTLV-1 PR with truncation of 10 residues at the C terminus might have ionic interactions between its termini similar to those in HIV-1 PR. Additional electrostatic interactions between the C terminus of each monomer of HTLV-1 PR with the guanidinium group of Arg-81 from the other monomer contribute to dimer stabilization.

The Binding Mode of the Inhibitor. Because the inhibitor molecules bound to two of three HTLV-1 PR dimers are well ordered from end to end with a unique orientation, it is possible to describe in detail the inhibitor/substrate binding subsites P5-P5′. The inhibitor is bound in an extended conformation (Fig. 4), which is typical for peptidic and peptidomimetic ligands bound to retroviral PRs, and each carbonyl oxygen and amide group of its backbone participates in direct hydrogen bonds with the enzyme in an identical manner to those previously reported for the other retroviral PRs (30). Four hydrogen bonds between the enzyme and inhibitor, mediated by a water molecule, represent the canonical interactions between the flaps and backbone carbonyls of the P2 and P1′ inhibitor residues. The interactions between the inhibitor backbone and the enzyme are shown in Table 2 and Fig. 7, which are published as supporting information on the PNAS web site, whereas a detailed description of the binding subsites in HTLV-1 PR and their comparisons with those in other retroviral PRs are presented in Table 3, which is published as supporting information on the PNAS web site.

It should be noted that, in general, hydrogen bonding between the inhibitor and the interior of the molecule is tighter than with the flaps. The binding subsites from S2 through S2′ in HTLV-1 PR are very similar to the corresponding ones in other retroviral PRs. These subsites are predominantly hydrophobic, with the exception of the catalytic aspartates that interact with the statine hydroxyl. Only one residue, Trp-98 (equivalent to Val-82 in HIV-1 PR), is unique in the S1/S1′-binding subsites of HTLV-1 PR. Four of six residues that form the S1/S1′ pockets are identical in all retroviral enzymes, and Ala-59 at the tip of the flap is equivalent to either Ile or Val in the other PRs (Table 3). However, the presence of the bulky Trp-98 has a dramatic effect on the architecture of these binding sites, modifying their specificity.

Only a single unique residue, Met-37, is found in the S2/S2′-binding subsites of HTLV-1 PR. The remaining residues are either identical or similar to their structural equivalents in the other enzymes. The more flexible nature of the Met-37 side chain, as compared with Asn-30 in HIV-1 PR, combined with a smaller size of the Ala-59 side chain (Ile-50 in HIV-1 PR), allows the accommodation of larger residues at the P2/P2′ position of the ligand. These results confirm the prediction of Bagossi *et al.* (31).

The subsites S3/S3′ in HTLV-1 PR have three unique residues and three residues identical to other retroviral enzymes (Table 3). The presence of the unique Trp-98 and Leu-57 significantly changes the nature of these pockets. For example, Asn-97 is shielded by the large side chain of Trp-98 and does not interact with the inhibitor. Its structural analogs in other retroviral PRs (i.e., Pro-81 in HIV-1 PR) are important residues in the S3/S3′ subsites, always directly involved in inhibitor binding.

The HTLV-1 PR S4/S4′ subsites are large and mostly hydrophobic, in contrast to the corresponding small and hydrophilic subsites in HIV-1 and FIV PRs. However, the respective subsites in RSV and EIAV PRs are also large and open. The presence of the short Thr-54 in HTLV-1 and EIAV PRs S4/S4′ subsites, which substitutes for the long side chain of Lys-45 in HIV-1 PR, allows the inclusion of Leu-69 (equivalent to Gln-58 in HIV-1 PR) into these pockets, enhancing their hydrophobicity.

Subsites S5/S5′ are located on the surface of the molecule, and the residues that interact with the inhibitor are much more hydrophilic than their equivalents in HIV-1, HIV-2, and SIV PRs. The nature of these subsites appears to be more similar to those found in RSV and EIAV PRs. To summarize, distant subsites in HTLV-1 PR and other retroviral enzymes demonstrate higher variability than the first three subsites (S3-S3′) that are located closer to the cleavage site.

Modeling of the Fit of Anti-HIV Drugs. Superposition of the structure of HTLV-1 PR with five structures of HIV-1 PR, each representing a complex with a different anti-HIV drug in current clinical use, namely amprenavir [Protein Data Bank (PDB) ID code 1HPV], saquinavir (PDB ID code 1FB7), indinavir (PDB ID code 1HSH), ritonavir (PDB ID code 1HXW), and nelfinavir (PDB ID code 1OHR), reveals potential steric problems in the

accommodation of these molecules in the active site of the HTLV-1 enzyme. Residues Trp-98 and Leu-57 of HTLV-1 PR collide with the groups of the inhibitors that use the S1/S1′ and S3/S3′ pockets in HIV-1 PR (Fig. 5). That may explain the failure of these compounds to inhibit HTLV-1 PR (11). It is clear that the future inhibitors of HTLV-1 PR may need to be considerably different as compared with either the currently available drugs targeting HIV-1 PR or even novel HIV-1 PR inhibitors that are being introduced to overcome multidrug resistance.

Conclusion

The presented crystal structure of HTLV-1 PR complexed with a statine-based inhibitor reveals the similarity in the overall protein fold and in the inhibitor-binding mode to other retroviral PRs. However, distinctly unique features are identified in the areas of the flaps and the putative substrate-binding sites that can be correlated with the enzymatic properties of this molecule, such as substrate specificity and the resistance to anti-HIV drugs (8, 32). It will be necessary for rapid progress in future studies to overcome the propensity of the enzyme to aggregate, and the present structure will serve as a guide to surface mutations to alleviate that problem.

We have now provided a structural basis for the rational design of novel compounds that could serve as anticancer drugs in the treatment of HTLV-1-induced ATL. The extensive experience gained from utilization of PR inhibitors as anti-HIV drugs (12), coupled with the observations that antiviral compounds appear to provide therapeutic benefits for the treatment of pathological conditions caused by HTLV-1 (3), bodes well for the practical utilization of this novel drug target.

We thank Drs. Laura Moen (National Institutes of Health, Bethesda) and Tod Holler (Pfizer, Ann Arbor, MI) for the plasmid encoding HTLV-1 PR and Dr. Yves Pommier (National Cancer Institute) for his support. We acknowledge the use of beamline 22-ID of the Southeast Regional Collaborative Access Team, located at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract W-31-109-Eng-38. This project was supported in part by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research, and in part with federal funds from the National Institutes of Health, National Cancer Institute, under Contract NO1-CO-12400. The research of M.J. was supported by a Faculty Scholar fellowship from the National Cancer Institute and by a subsidy from the Foundation for Polish Science. The research of G.S.L. was supported by a National Cancer Institute postdoctoral fellowship.

1. Franchini, G., Fukumoto, R. & Fullen, J. R. (2003) *Int. J. Hematol.* **78,** 280–296.
2. Uchiyama, T. (1997) *Annu. Rev. Immunol.* **15,** 15–37.
3. Ishikawa, T. (2003) *Int. J. Hematol.* **78,** 304–311.
4. Kannagi, M., Harashima, N., Kurihara, K., Utsunomiya, A., Tanosaki, R. & Masuda, M. (2004) *Exp. Rev. Anticancer Ther.* **4,** 369–376.
5. Nasr, R., El Sabban, M. E., Karam, J. A., Dbaibo, G., Kfoury, Y., Arnulf, B., Lepelletier, Y., Bex, F., de The, H., Hermine, O., *et al*. (2005) *Oncogene* **24,** 419–430.
6. Shuker, S. B., Mariani, V. L., Herger, B. E. & Dennison, K. J. (2003) *Chem. Biol.* **10,** 373–380.
7. Louis, J. M., Oroszlan, S. & Tozser, J. (1999) *J. Biol. Chem.* **274,** 6660–6666.
8. Tozser, J., Zahuczky, G., Bagossi, P., Louis, J. M., Copeland, T. D., Oroszlan, S., Harrison, R. W. & Weber, I. T. (2000) *Eur. J. Biochem.* **267,** 6287–6295.
9. Abdel-Rahman, H. M., Kimura, T., Hidaka, K., Kiso, A., Nezami, A., Freire, E., Hayashi, Y. & Kiso, Y. (2004) *Biol. Chem.* **385,** 1035–1039.
10. Maegawa, H., Kimura, T., Arii, Y., Matsui, Y., Kasai, S., Hayashi, Y. & Kiso, Y. (2004) *Bioorg. Med. Chem. Lett.* **14,** 5925–5929.
11. Bagossi, P., Kadas, J., Miklossy, G., Boross, P., Weber, I. T. & Tozser, J. (2004) *J. Virol. Methods* **119,** 87–93.
12. Wlodawer, A. & Vondrasek, J. (1998) *Annu. Rev. Biophys. Biomol. Struct.* **27,** 249–284.
13. Agbuya, P. G., Sherman, N. E. & Moen, L. K. (2001) *Arch. Biochem. Biophys.* **392,** 93–102.
14. Laco, G. S., Fitzgerald, M. C., Morris, G. M., Olson, A. J., Kent, S. B. & Elder, J. H. (1997) *J. Virol.* **71,** 5505–5511.
15. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
16. Storoni, L. C., McCoy, A. J. & Read, R. J. (2004) *Acta Crystallogr. D* **60,** 432–438.
17. Jones, T. A. & Kjeldgaard, M. (1997) *Methods Enzymol.* **277,** 173–208.
18. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) *Acta Crystallogr. D* **53,** 240–255.
19. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al*. (1998) *Acta Crystallogr. D* **54,** 905–921.
20. Herger, B. E., Mariani, V. L., Dennison, K. & Shuker, S. B. (2004) *Biochem. Biophys. Res. Commun.* **320,** 1306–1308.
21. Brynda, J., Rezacova, P., Fabry, M., Horejsi, M., Stouracova, R., Soucek, M., Hradilek, M., Konvalinka, J. & Sedlacek, J. (2004) *Acta Crystallogr. D* **60,** 1943–1948.
22. Cohen, G. E. (1997) *J. Appl. Crystallogr.* **30,** 1160–1161.
23. Tong, L., Pav, S., Mui, S., Lamarre, D., Yoakim, C., Beaulieu, P. & Anderson, P. C. (1995) *Structure (Cambridge, U.K.)* **3,** 33–40.
24. Hoog, S. S., Towler, E. M., Zhao, B., Doyle, M. L., Debouck, C. & Abdel-Meguid, S. S. (1996) *Biochemistry* **35,** 10279–10286.
25. Gustchina, A., Kervinen, J., Powell, D. J., Zdanov, A., Kay, J. & Wlodawer, A. (1996) *Protein Sci.* **5,** 1453–1465.
26. Wlodawer, A., Gustchina, A., Reshetnikova, L., Lubkowski, J., Zdanov, A., Hui, K. Y., Angleton, E. L., Farmerie, W. G., Goodenow, M. M., Bhatt, D., *et al*. (1995) *Nat. Struct. Biol.* **2,** 480–488.
27. Wu, J., Adomat, J. M., Ridky, T. W., Louis, J. M., Leis, J., Harrison, R. W. & Weber, I. T. (1998) *Biochemistry* **37,** 4518–4526.
28. Wlodawer, A., Miller, M., Jaskólski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J. & Kent, S. B. H. (1989) *Science* **245,** 616–621.
29. Jaskólski, M., Miller, M., Rao, J. K. M., Leis, J. & Wlodawer, A. (1990) *Biochemistry* **29,** 5889–5898.
30. Gustchina, A., Sansom, C., Prevost, M., Richelle, J., Wodak, S. Y., Wlodawer, A. & Weber, I. T. (1994) *Protein Eng.* **7,** 309–317.
31. Bagossi, P., Sperka, T., Feher, A., Kadas, J., Zahuczky, G., Miklossy, G., Boross, P. & Tozser, J. (2005) *J. Virol.* **79,** 4213–4218.
32. Kadas, J., Weber, I. T., Bagossi, P., Miklossy, G., Boross, P., Oroszlan, S. & Tozser, J. (2004) *J. Biol. Chem.* **279,** 27148–27157.
33. Brünger, A. T. (1992) *Nature* **355,** 472–474.