Specificity and Inhibition of Proteases from Human Immunodeficiency Viruses 1 and 2*

Alfredo G. Tomasselli, John O. Hui, Tomi K. Sawyer, Douglas J. Staples, Carol Bannow, Ilene M. Reardon, W. Jeffrey Howe, Dianne L. DeCamp§, Charles S. Craik‡, and Robert L. Heinrikson

From the Biopolymer Chemistry and Computational Chemistry Unit, Discovery Research Division, The Upjohn Company, Kalamazoo, Michigan 49001 and the ‡Department of Pharmaceutical Chemistry, The University of California, San Francisco, California 94143-0446

Highly purified, recombinant preparations of the virally encoded proteases from human immunodeficiency viruses (HIV) 1 and 2 have been compared relative to 1) their specificities toward non-viral protein and synthetic peptide substrates, and 2) their inhibition by several P1-P1′ pseudodipeptidyl-modified substrate analogs. Hydrolysis of the Leu-Leu and Leu-Ala bonds in the *Pseudomonas* exotoxin derivative, LysPE40, is qualitatively the same for HIV-2 protease as published earlier for the HIV-1 enzyme (Tomasselli, A. G., Hui, J. O., Sawyer, T. K., Staples, D. J., Fitzgerald, D. J., Chaudhuri, V. K., Pastan, I., and Heinrikson, R. L. (1990) J. Biol. Chem. 265, 408–413). However, the rates of cleavage at these two sites are reversed for the HIV-2 protease which prefers the Leu-Ala bond. The kinetics of hydrolysis of this protein substrate by both enzymes are mirrored by those obtained from cleavage of model peptides. Hydrolysis by the two proteases of other synthetic peptides modeled after processing sites in HIV-1 and HIV-2 *gag* polyproteins and selected analogs thereof demonstrated differences, as well as similarities, in selectivity. For example, while the two proteases were nearly identical in their rates of cleavage of the Tyr-Pro bond in the HIV-1 *gag* fragment, Val-Ser-Gln-Asn-Tyr-Pro-Ile-Val, the HIV-1 protease showed a 64-fold enhancement over the HIV-2 enzyme in hydrolysis of a Tyr-Val bond in the same template. Accordingly, the HIV-2 protease appears to have a different specificity than the HIV-1 enzyme; it is better able to hydrolyze substrates with small amino acids in P1 and P1′, but is variable in its rate of hydrolysis of peptides with bulky substituents in these positions. In addition to these comparisons of the two proteases with respect to substrate specificity, we present inhibitor structure–activity data for the HIV-2 protease. Relative to P1-P1′ statine or PheΨ(CH2N)Pro-modified pseudopeptidyl inhibitors, compounds having XaaΨ(CH(OH)CH2)yaa inserts were found to show significantly higher affinities to both enzymes, generally binding from 10 to 100 times stronger to HIV-1 protease than to the HIV-2 enzyme. Molecular modeling comparisons based upon the sequence homology of the two enzymes and x-ray crystal structures of HIV-1 protease suggest that most of the nonconservative amino acid replacements occur in regions well outside the catalytic cleft, while only subtle structural differences exist within the active site. In addition, energy-based modeling of a peptide fragment patterned after one of the observed substrates indicated that class 3 substrates may be interacting with charged protease side chains clustered at the ends of the binding cleft.

The aspartyl protease encoded in the *pol* gene of retroviruses is essential for viral maturation (1), and the protease from human immunodeficiency virus (HIV) has been targeted as a possible therapeutic intervention point in the treatment of acquired immunodeficiency syndrome (AIDS). Numerous reports have appeared in the recent literature describing the purification and characterization of HIV protease produced by recombinant or synthetic means (2–11), the tertiary structural analysis of this (12–14), and a related retroviral enzyme (15), and studies of enzyme chemistry (6, 16), specificity, and inhibition (6, 7, 17–23). In addition, an x-ray crystallographic model of the HIV-1 protease bound to a synthetic pseudopeptide inhibitor has been reported recently by Wlodawer and co-workers (24). This level of structural detail in our understanding of enzyme-inhibitor interactions has been complemented on the functional side by recent descriptions of HIV protease inhibitors that block viral maturation and polyprotein processing, respectively, in HIV-1-infected cell culture assays (25, 26) and in a cell model system which mimics HIV-1 infectivity (25). Therefore, both from a structural and functional point of view, the protease has been characterized at a high level of sophistication, and this serves as a basis for drug design and implementation in AIDS therapy.

Most of the findings published thus far with respect to the protease have been obtained from studies of the enzyme from HIV-1, that form of the virus that is associated with AIDS in most of the Western world. However, it is now clear that HIV-2 is a distinct etiologic factor that is also of importance in human AIDS. Since the first description of its genomic sequence (27), HIV-2 was recognized as a separate virus that closely resembles simian immunodeficiency virus (SIV). Evidence for the possible origins of HIV-1 and HIV-2 from a reservoir of African lentiviruses, collectively termed SIV, was presented in a recent paper by Hirsch et al. (28), and the implications of these findings were discussed further by Doollittle (29). Whatever the evolutionary history of these viruses,
therapeutic approaches to AIDS must take into account infections derived from HIV-2, and if the protease is the target, it is crucial to establish similarities and differences between enzymes from HIV-1 and HIV-2. Moreover, at the present time, the concept of a protease inhibitor as a drug against AIDS has yet to be validated in an animal model. One reasonable choice here is the monkey (30), and the close similarity between the sequences both of the HIV-2 and SIV proteases (28) and of their polyprotein substrates (31) would imply that a study of the HIV-2 enzyme should provide important information for design of inhibitors for testing in non-human primate models.

The present paper describes a comparison of the proteases from HIV-1 and HIV-2 relative to their substrate specificities and with respect to their inhibition by a variety of compounds. Furthermore, the recent availability of x-ray crystal structures of HIV-1 protease has enabled modeling studies to be undertaken in conjunction with the experimental work. The combined results suggest structural factors that may contribute to the observed substrate preferences.

**EXPERIMENTAL PROCEDURES**

**Materials**

Recombinant HIV-1 protease was prepared from Escherichia coli inclusion bodies as described by Tommasselli et al. (6). The recombinant enzyme from HIV-2, expressed in a soluble form in yeast, was purified according to similar methods; the lyophilized, homogeneous protein was eluted from the reverse-phase HPLC column and refolded by the same methods employed for HIV-1 protease (6). Concentrations of the enzymes were established both by amino acid analysis and by titration with the substrate-based inhibitor Val-Ser-Gln-Asn-Leu-Phe-Pro-Ile-Val (6). Pseudomonas exotoxin, PE66, and derivatives thereof in which domain I is either missing (Lys-PE40), or replaced by the first two domains of soluble CD4 (CD4(178)-PE40), were the same as described in detail in an earlier publication (18), and were the generous gift of Dr. David J. FitzGerald, National Cancer Institute, Frederick Cancer Research Facility), and 2) uncomplexed HIV-1 protease monomer (Brookhaven Protein Data Bank entry 3HVP (12)). A model of the protease dimer was constructed from the coordinates of the monomer by applying the specific (Z)-transformation.

One of the cleavage sites observed for both HIV-1 and HIV-2 proteases had a Glu-Glu pair in the substrate P1\'-P2\' positions. To examine the extent to which these residues could interact with charged side chains that are clustered at the end of the HIV-1 protease active site cleft (24), the following approach was used. A Leu-Ala-Glu-Glu fragment was constructed, using Mosaic, with all backbone amino acid residues 180 degrees.

**Methods**

**HIV Protease Assay**—both HIV-1 and HIV-2 proteases were assayed against the octapeptide, Lys-Ser-Gln-Asn-Tyr-Pro-Ile-Val, corresponding to the sequence of the natural HIV-1 gag polyprotein sequence from residue 128 to 135 (33). Conditions for the assays and for monitoring the course of hydrolysis of the Tyr-Pro bond by HPLC were detailed in earlier publications (6, 19).

Protein substrates, PE66, Lys-PE40, and CD4(178)-PE40 were cleaved by the two proteases under essentially the same conditions, but hydrolysis was monitored in these cases by quantitative SDS-polyacrylamide gel electrophoresis (19). CD4(178)-PE40 (100 μg) was incubated for 7 h at 30 °C with 0.40 μg of HIV-1 protease or 0.65 μg of the HIV-2 enzyme in 200 μl of 0.2 M sodium acetate, pH 5.5, containing 10% glycerol and 5% ethylene glycol. Samples were withdrawn at various times and subjected to reducing SDS-polyacrylamide gel electrophoresis in order to follow the progress of reaction. Sequence analysis was performed on samples from the last time points of each hydrolysis mixture to determine if cleavage could be documented. Lys-PE40 was cleaved under similar conditions, except that reactions were carried out for 4 h in a volume of 100 μl with 0.7 μg of protease.

With all of the studies involving protein substrates, control reactions were included which contained all of the ingredients of the assays except for the protease. Another set of controls included the tight binding HIV protease inhibitor Val-Ser-Gln-Asn-Leu-Phe-Pro-Ile-Val (U-85548E) in order to prove that any observed proteolysis was due, exclusively, to the retroviral protease activity (6).

**Amino Acid Analysis**—Compositional analysis of proteins and peptides was performed by conventional automated ion-exchange chromatography on a Beckman model 6300 analyzer. Samples were hydrolyzed in 6 N HCl for 24 h in vacuo at 110 °C. Hydrazides were dried in a Speed Vac Concentrator (Savant), and the residues were dissolved in buffer at pH 2.2 (Na-S, Beckman) for application to the analyzer.

**Sequence Analysis**—Protein and peptide sequencing was performed by automated Edman degradation in an Applied Biosystems Inc. model 470 sequencer fitted with an on-line HPLC analyzer (model 120-A) for phenylthiohydantoin amino acids. Peaks from the latter were integrated by a Nelson Analytical 3000 Series chromatography data system connected in parallel with the recorder to the output of the HPLC system.

**Quantitative SDS-Polyacrylamide Gel Electrophoresis**—Laemmli (34) gels (12%) were run under reducing conditions in either mercaptoethanol or diethiothreitol and stained with Coomassie Blue for detection of proteins. Stained gels were analyzed in a Visage 110 scanner from BioImage, Ann Arbor, MI, which integrates the optical density of the gel bands and thus provides a means of quantitation for assaying the time course of cleavage of protein substrates by the HIV proteases.

**Molecular Modeling**—Molecular modeling studies made use of various components of the Mosaic software package for model building, graphical visualization, and energy refinements. Two x-ray crystal structures were used in the modeling work: 1) complexed HIV-1 protease/MVT101 reduced-bond inhibitor (Ac-Thr-Ile-Nle-Glu[CH2-NH]-Nle-Gln-Arg-NH2 (24); coordinates supplied by Dr. A. Wlodawer, National Cancer Institute, Frederick Cancer Research Facility), and 2) uncomplexed HIV-1 protease monomer (Brookhaven Protein Data Bank entry 3HVP (12)). A model of the protease dimer was constructed from the coordinates of the monomer by applying the specific (Z)-transformation.

One of the cleavage sites observed for both HIV-1 and HIV-2 proteases had a Glu-Glu pair in the substrate P1\'-P2\' positions. To examine the extent to which these residues could interact with charged side chains that are clustered at the end of the HIV-1 protease active site cleft (24), the following approach was used. A Leu-Ala-Glu-Glu fragment was constructed, using Mosaic, with all backbone amino acid residues 180 degrees.

**RESULTS**

**Hydrolysis of Pseudomonas Exotoxin Derivatives**—In an earlier paper, we demonstrated that two Pseudomonas exotoxin derivatives, Lys-PE40 and CD4(178)-PE40, were hydrolyzed by the HIV-1 protease at bonds within interdomain regions (19). Interestingly, cleavage occurred at two sites in Lys-PE40, a Leu-Leu and a Leu-Ala bond representative of class 3 substrates (31). The expected Tyr-Pro bond in the interdomain region was not hydrolyzed (19). Since we had

---

8 S. Pichuanets, L. Babé, P. J. Barr, and C. S. Craik, manuscript submitted.
9 L. Babé, S. Pichuanets, and C. S. Craik, manuscript submitted.
already characterized these novel, non-viral proteins as substrates for the HIV-1 protease, we were interested in determining the course of their cleavage by the HIV-2 enzyme. Qualitatively speaking, the two proteases are identical in their cleavage of the Leu-Ala and Leu-Leu bonds in LysPE40 and at only the latter site in the chimeric CD4(178)PE40 (Fig. 1 (19, 36)). In both protein substrates, this Leu-Leu bond is hydrolyzed twice as fast by the HIV-2 protease as compared to the HIV-1 enzyme. These kinetic data were obtained by scanning densitometry of SDS-polyacrylamide gel electrophoresis gels of reaction products after various times (data not shown). Although the rate of cleavage of the Leu-Ala bond by the HIV-1 protease was shown by sequence analysis of the products to be about one-third that of the Leu-Leu bond (19), it was actually the preferred site of hydrolysis by the HIV-2 enzyme, being cleaved 1.5 times faster than Leu-Leu. Thus, the HIV-2 protease cleaves the Leu-Ala bond about 10 times faster than the HIV-1 enzyme. We showed earlier that octa- and nonapeptides corresponding to these susceptible regions were cleaved by the HIV-1 protease with kinetics mirroring those with the protein substrates (19). As shown in Table I (compounds 1 and 2), the general course of hydrolysis of the LysPE40 derivatives by both proteases is reflected in the kinetic analysis of peptide hydrolysis. The nonapeptide containing the Leu-Leu bond, and the octapeptide with the Leu-Ala bond are cleaved 2.5, and 17 times faster, respectively, by the HIV-2 protease as compared with the HIV-1 enzyme. Interestingly, a peptide identical to compound 1 in Table I, but having His-Met in place of Asn-Leu, is cleaved only very slowly by the HIV-1 protease. This peptide corresponds to the linkage region in the chimeric protein CD4(178)PE40 (Fig. 1), and thus the failure to cleave the Met-Ala bond in the protein is, once again, reflected by results with the peptide substrate. Other parallels with the earlier study were documented; neither enzyme was able to hydrolyze PE66, nor was cleavage seen at the Tyr-Pro bond (Fig. 1) originally expected to be the most likely site of hydrolysis ((19, Table 1). Phospholipase A of the HIV proteases in the presence of 25 μM U-88548E, an active site titrant, completely abolished cleavage of any of the protein or synthetic peptide substrates.

**Hydrolysis of Peptides Modeled after Viral Polypeptides**

Comparison of HIV-1 and HIV-2 proteases relative to their kinetics of hydrolysis of the HIV-1 gag polypeptide-based model peptide Val-Ser-Gln-Asn-Tyr-Pro-Ile-Val, reveals that they have identical $K_m$ values, but the $V_{max}$ for cleavage of the Tyr-Pro bond by the HIV-1 enzyme is twice that of the HIV-2 protease (Table I, compound 8). Thus, with respect to this substrate, and compound 9 in which the Tyr is replaced by Phe, the enzymes are quite similar. This also holds true for compounds 3 and 4 which are modeled from class 3 cleavage sites (31). However, replacement of the Tyr (or Phe) at P$_i$ in the template of Val-Ser-Gln-Asn-Xaa-Yaa-Ile-Val by cyclohexylalanine (Cha) or Leu (compounds 10 and 11, respectively) or the P$_i$' Pro by Val (compound 12) yields peptide substrates much more efficiently cleaved by the HIV-1 protease. In fact, the latter peptide is bound equally well by both enzymes ($K_m$ is lowered from 2.0 to 0.6 mM for both relative to the parent peptide, compound 8) but is hydrolyzed 64 times faster by HIV-1 protease. Compounds 5, 6, and 13 (Table I) are not cleaved by either enzyme despite obvious similarities to other compounds listed that are good substrates. Compound 5 represents a sequence processed in the pol gene by avian myeloblastosis viral protease. It is difficult to say what subtle differences account for its lack of hydrolysis. Compound 6 is modeled after a putative site of cleavage (27, 31) at the Ph$_{hel}$ Tyr-Pro$_{hel}$ bond in the HIV-2 gag region, but was not hydrolyzed by either enzyme. Since it is generally true that bonds cleaved in a protein are also hydrolyzed in model peptides (7, 19), this result was a surprise. Compound 14, containing an Ala-Ala scissile bond and modeled after an HIV-2 polypeptide processing site, represents one case of a peptide that is a substrate for the HIV-2 protease, but not for the HIV-1 enzyme. However, replacement of the Tyr-Pro sequence in compound 8 by Ala-Ala (compound 13) destroys the ability of the peptide to serve as substrate for either enzyme.

In summary, the ability of the HIV-2 protease to cleave Ala-Ala bonds in particular substrates underscores a basic difference from the HIV-1 enzyme in that the former is able to cleave subunits with small segments of Lys$_{hel}$ and P$_i$'. Substrates having β-branched amino acids at P$_i$' (compound 12), however, exhibited significantly lower rates of hydrolysis by the HIV-2 protease as compared with the HIV-1 enzyme. These findings would suggest that the HIV-2 protease binding pockets for P$_i$ and P$_i$' side chains may be more occluded, and perhaps less efficient in the binding of bulky groups. In general, the HIV-2 enzyme appears to display a broader substrate specificity than the HIV-1 protease. This conclusion finds further support from the course of hydrolysis of calmodulin by the two proteases described by Tomasselli et al.5

---

**Fig. 1.** Diagrammatic representation of two derivatives of the *Pseudomonas* exotoxin PE66 (19, 45, 44) in which domain I is missing (LysPE40, lower dashed line), or replaced by the 178 amino acids of the two NH$_2$-terminal domains of CD4 (upper dashed line). That portion of the OmpA gene product remaining as part of the recombinant LysPE40 is indicated, together with Roman numerals referring to parts of, or whole domains defined crystallographically (44). Arrows above bonds indicate sites of cleavage of these protein substrates by proteases from HIV-1 and HIV-2. The Tyr-Pro bond enclosed in a box is not cleaved by either enzyme. Residues in parentheses are present only in LysPE40 and not in the chimeric CD4(178)PE40 (36).

---

5 A. G. Tomasselli, W. J. Howe, J. O. Hui, T. K. Sawyer, I. M. Reardon, D. DeCamp, C. S. Craik, and R. L. Heinrikson, manuscript submitted.
TABLE I
Comparison of substrate specificity of HIV-1 and HIV-2 proteases relative to hydrolysis of synthetic peptides based upon viral polyproteins from HIV and avian myeloblastosis virus (AMV) and from the non-viral protein substrate Lys PE40

<table>
<thead>
<tr>
<th>Compounds</th>
<th>HIV-1 protease</th>
<th>HIV-2 protease</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_m$ (mM)</td>
<td>$V_{max}$ (μmol × min⁻¹ × mg⁻¹)</td>
</tr>
<tr>
<td>1 H-Ala-Asn-Leu-Ala-Glu-Glu-Ala-Phe-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>1.3</td>
<td>0.3</td>
</tr>
<tr>
<td>2 H-Ser-Gly-Asp-Ala-Leu-Leu-Glu-Arg-Asn-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>1.6</td>
<td>1.2</td>
</tr>
<tr>
<td>3 H-Thr-Ala-Thr-Ile-Met-Met-Gln-Arg-Gly-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>1.3</td>
<td>3.6</td>
</tr>
<tr>
<td>4 H-Thr-Ala-Thr-Ile-Nle-Nle-Gln-Arg-Gly-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>1.7</td>
<td>3.0</td>
</tr>
<tr>
<td>5 H-Thr-Phe-Glu-Ala-Tyr-Pro-Leu-Leu-Glu-Ala-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6 H-Lys-Pro-Arg-Asn-Phe-Pro-Val-Ala-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>1.9</td>
<td>3.9</td>
</tr>
<tr>
<td>7 H-Tyr-Val-Ser-Glu-Phe-Pro-Ile-Val-Gln-Gln-Arg-Asn-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>2.0</td>
<td>4.9</td>
</tr>
<tr>
<td>8 H-Val-Ser-Gln-Glu-Ile-Ala-Val-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>6.7</td>
<td>3.9</td>
</tr>
<tr>
<td>9 H-Val-Ser-Gln-Glu-Ile-Ala-Val-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>2.0</td>
<td>3.9</td>
</tr>
<tr>
<td>10 H-Val-Ser-Gln-Leu-Pro-Ile-Val-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>10</td>
<td>0.9</td>
</tr>
<tr>
<td>11 H-Val-Ser-Gln-Leu-Pro-Ile-Val-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>0.6</td>
<td>2.7</td>
</tr>
<tr>
<td>12 H-Val-Ser-Ala-Asn-Met-Ala-Val-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13 H-Val-Ser-Ala-Asn-Met-Ala-Val-OH&lt;sup&gt;Ⅵ&lt;/sup&gt;</td>
<td>4.6</td>
<td>4.6</td>
</tr>
</tbody>
</table>

<sup>Ⅵ</sup>LysPE40, <sup>Ⅵ</sup>HIV-1 gag fragment (or analog), <sup>Ⅵ</sup>AMV pol fragment, <sup>Ⅵ</sup>HIV-2 gag fragment (or analog), ND, not determined.

TABLE II
Comparative inhibition of HIV-1 and HIV-2 proteases

<table>
<thead>
<tr>
<th>Entry</th>
<th>Compound</th>
<th>HIV-1</th>
<th>HIV-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepstatin</td>
<td>Iva-Val-Val-Sta-Ala-Sta-OH</td>
<td>362</td>
<td>720</td>
</tr>
<tr>
<td>U-85549E</td>
<td>H-Val-Ser-Gln-Asn-Sta-Ile-Val-OH</td>
<td>3,690</td>
<td>9,000</td>
</tr>
<tr>
<td>U-84645E</td>
<td>H-Val-Ser-Gln-Phe-Phe[CH₃,N]Pro-Ile-Val-OH</td>
<td>3,520</td>
<td>26,100</td>
</tr>
<tr>
<td>U-85548E</td>
<td>H-Val-Ser-Gln-Leu-Phe[CH₃,N]CH₃Val-Ile-Val-OH</td>
<td>&lt;1</td>
<td>9</td>
</tr>
<tr>
<td>U-71038</td>
<td>Boc-Pro-Phe-NMe-His-Leu[CH₃,N]CH₃Val-Ile-Val-OH</td>
<td>10</td>
<td>&gt;1,000</td>
</tr>
<tr>
<td>U-81749</td>
<td>Tba-Cha[Phe[CH₃,N]CH₃]Val-Ile-Val-OH</td>
<td>80</td>
<td>1,000</td>
</tr>
</tbody>
</table>

Inhibitor Structure-Activity Relationships—Thus far, we have compared the HIV-1 and HIV-2 proteases according to their cleavage site preferences in native proteins and modeled peptides. In Table II are shown $K_i$ values determined for a number of inhibitors relative to the two enzymes. The natural product, pepstatin, is only moderately inhibitory toward both HIV-1 and HIV-2 proteases. Relative to pepstatin, we evaluated a series of HIV-gag/pol precursor-based inhibitors with the generic structure, Val-Ser-Gln-Asn-Xaa-Yaa-Ile-Val, having Pr-P₃ substitutions by Sta, Phe[CH₃,N]Pro, and Leu[CH₃,N]CH₃Val and found them to be of greater potency and selectivity. Noteworthy was U-85548E, a high affinity inhibitor of both enzymes ($K_i < 1$ nM for HIV-1 protease, and 9 nM for the HIV-2 enzyme). Previous studies by Richards et al. (21, 38) showed that the Leu[CH₃,N]CH₃Val-substituted renin inhibitor, H-261, is a strong inhibitor of both HIV-1 and HIV-2 proteases ($K_i = 15$ and 90 nM, respectively). To extend this study, we evaluated another high affinity renin inhibitor, U-71038, against both proteases (Table II). Overall, these findings show that U-85548E is the most potent inhibitor of both proteases yet described and that U-71038 is the most selective of those tested thus far.

The small inhibitor U-81749 (Tba-Cha[Phe[CH₃,N]CH₃]Val-Ile-Val-OH (25); $M_r = 572$) is of special interest in this group. A Dixon plot for inhibition of HIV-1 protease by this compound is shown in Fig. 2; values of 83 and 70 nM were determined for HIV-1 protease at pH 5.5 and 7.1, respectively. A replot of the slopes of the Dixon plots (inset, Fig. 2) gives a straight line through the origin, indicating that U-81749 is a competitive inhibitor of the protease. This inhibitory activity was demonstrated not only against the pure HIV-1 protease but against viral maturation in a cell culture system consisting.
of HIV-infected human peripheral blood lymphocytes (25). In this latter study, an IC_{50} value was determined to be between 0.1 and 1 μM concentrations of inhibitor (25). Processing of HIV-1 gag and gag/pol polyproteins to p24 in cells infected with a recombinant vaccinia virus expressing the HIV-1 pre-cursors (39) was also blocked by 10 μM concentrations of U-81749. However, U-85548E, an inhibitor bound about 10 times more strongly to the enzyme (Table II) showed little activity in these cell culture assays. These findings underscore the importance of considerations relative to cell or particle delivery as well as K_{i} when designing protease-targeted drugs against HIV.

**Structural Comparisons of HIV-1 and HIV-2 Proteases**

Comparison of the HIV-1 and HIV-2 protease amino acid sequences (Fig. 3) reveals numerous similarities, including an 8-residue stretch of exact identity in the 23-30 region surrounding the catalytic Asp-25. Overall, the identity between the two sequences is 50%, and when conservative substitutions are taken into account the similarity approaches 75%. Therefore, although the three-dimensional structure of HIV-2 protease is not yet known, the structure of HIV-1 protease can serve as a useful starting point from which to examine the observed substrate preferences of the two enzymes.

In Fig. 4 is depicted a smoothed backbone representation of the complexed form of HIV-1 protease from which the MVT101 inhibitor (Ac-Thr-Ile-Nle\(\text{ψ}\)CH_{3}NH)Lle-Gln-Arg NH_{2}) has been removed. The orientation of the structure is such that the 2-fold symmetry of the molecule is readily apparent. The flap regions are at the top center of the diagram, and the view is down the length of the substrate binding cleft which is located just below the flaps. Color mapping of the

---

**Fig. 2.** A Dixon plot for inhibition of HIV-1 protease by U-81749 shown on the figure. The substrate was Val-Ser-Glu-Asn-Tyr-Pro-Ile-Val at concentrations of 2.8 mM (■), 1.0 mM (○), and 0.5 mM (△). The inset shows a replot of the slopes from the Dixon plot versus 1/S; since this line passes through the origin, U-81749 is a competitive inhibitor of the protease.

**Fig. 3.** Comparison of the sequences of proteases from HIV-1 (33, 37, 46), HIV-2 (27), and SIV (45).
HIV-2 protease sequence has been applied to the structure to enable a comparison of the two enzymes. Regions of exact sequence match are shown in blue, conservative replacements are shown in yellow, and nonconservative replacements are in red. It should be stressed that this construction is not intended to be a model of HIV-2 protease; it is simply the HIV-1 protease structure with sequence information for the HIV-2 enzyme mapped onto it. Nonetheless, it is a useful homology representation technique with which one can begin to infer some of the structural factors that may be influencing the relative specificities of the enzymes.

Of particular interest in this representation is the fact that most of the nonconservative replacements occur at the periphery of the structure, far removed from the binding site, yet the residues which define the binding site surface are those which would be expected to have the greatest influence on substrate recognition. With regard to those residues actually in contact with the inhibitor in the x-ray structure (24), the two enzymes differ in only three symmetrically located pairs. The first is Val-82 in the A and B chains of the HIV-1 protease dimer, which is replaced by an Ile in the HIV-2 enzyme. In the crystal structure, this residue contributes surfaces which are in contact with inhibitor side chains at the P$_1$ and P$_1'$ positions. The second replacement is Val-32 to Ile-32 in HIV-2 protease; this residue is in contact with the P$_2$ and P$_2'$ side chains of the inhibitor. The third is Ile-47 to Val in the HIV-2 enzyme, a residue also in contact with the P$_3$ and P$_3'$ side chains. All three are conservative replacements which involve nothing more than addition or deletion of a single methyl group at each of the 3 pairs of residues. Assuming, then, that the binding site of HIV-2 enzyme is defined by the same residue positions as are seen in the HIV-1 protease structure, it can be said that there are differences in the shape of the two binding sites, but they are likely to be subtle ones. Since this is a map and not a model, however, the possibility cannot be ignored that the more extreme sequence substitu-
Compared Specificities of HIV-1 and HIV-2 Proteases

Leu-Ala-Glu-Glu-Ala-Phe (compound 1, Table I) may be due to enhanced binding of the Glu-Glu pair in the HIV-2 protease.

DISCUSSION

The present study has shown that the HIV-1 and HIV-2 proteases may be distinguished according to their specificities toward particular substrates and inhibitors, despite the fact that their natural polyproteins are similar, and the enzymes display similar selectivities. A recent report based upon maturation of chimeric viral polyproteins came to essentially the same conclusions (40), although specific sites hydrolyzed were not identified and no kinetic analyses was undertaken. In defining sites of cleavage shared in common, it is important to consider kinetic parameters for hydrolysis as well. With LysPE40, both proteases cleave the same two bonds, but with dramatically different preferences (Table I, compounds 1 and 2). This difference cannot be attributed solely to the relative sizes of the P1' residues (Ala versus Leu) since the compounds differ at other positions as well. However, the modeling comparison of the HIV-1 and HIV-2 protease-binding sites does suggest a greater occlusion at the S1/S2' subsites in HIV-2 protease, which could favor smaller P1' side chains in corresponding substrates. This would be in keeping with the processing function of this enzyme in cleaving Ala-Ala bonds in HIV-2 polyproteins (31). Such processing sites do not exist in the HIV-1 polyproteins. Accordingly, we have demonstrated in the present paper that a peptide substrate based upon the HIV-2 gag polyprotein-processing site that contains an Ala-Ala scissile bond is cleaved by HIV-2 protease, but not by the HIV-1 enzyme (Table I, compound 14). However, it is of interest that one cannot replace the Tyr-Pro bond in a class 1 substrate (31) by an Ala-Ala sequence characteristic of class 3 substrates; the resulting peptide (Table I, compound 13) is cleaved by neither enzyme.

Both proteases are able to hydrolyze peculiar particular peptide bonds having bulky, hydrophobic amino acids in P1, and P1' (Table I). Thus, the subtle differences at S1/S2' suggested by the binding site models certainly do not rule out HIV-2 protease binding of bulky side chains at those positions, but they may have an effect on the efficiency of hydrolysis. For example, changing the Phe at P1' (compound 9) to a Cha (compound 10) causes a substantial reduction in hydrolysis by HIV-2 protease. A similar effect is seen when a P1' Pro (compound 8) is changed to the y-branched Val (compound 12). Interpretation of this latter observation, however, is complicated by the possibility that the P1' Pro may be inducing a backbone conformation in the P1' to P2' positions different from that which exists in compound 12. This possibility deserves further study, especially in light of the fact that substrate activity of the class 1 compound 8 (Tyr-Pro insert) is abolished in compound 13 (Ala-Ala insert) for both enzymes, even though the two peptides are identical in all other respects. Additionally, our modeling results suggest that class 3 substrates, which often contain charged residues at P1', P2', P3, or P4', bind to both proteases in a manner that involves salt bridging with one or more of the charged protease residues at the ends of the clef. Such interactions may not be possible when a Pro is at the P1' position and could explain the general lack of charged residues downstream from the Pro in class 1 substrates.

As an approach to rationalizing the differences we have observed between the proteases in specificity and inhibitor binding, we have made extensive use of structural models of the HIV-1 enzyme. We were particularly interested in the observation (24) that the MVT101 peptide binds in an ex-
tended conformation. This fact is easily reconciled with results of studies in solution. Our observations from protease hydrolysis of PE40 derivatives are in accord with the view that cleavage takes place at flexible, extended structures in the protein substrates that behave, essentially, the same as small modeled peptides (19). Indeed, restriction of processing by the protease to interdomain or interpeptide regions of the natural viral polypeptides provides further support for the idea that the enzyme prefers substrates with an extended conformation. A similar conclusion was reached in modeling studies of HIV-1 and HIV-2 protease cleavage site in calmodulin, described by Tommasselli et al.

From the foregoing discussion, it would appear that if one were to design inhibitors of the HIV-2 enzyme with selectivity over the HIV-1 protease, scissile dipeptide substitutions with small amino acids such as Ala might contribute in an important way to their differential activity. This is certainly an interesting topic to pursue from the point of view of delineating differences in binding of specific compounds to the two proteases and in elucidating facets of their enzymology. In a more practical sense, the increasing impact of HIV-2 on the global AIDS epidemic makes it clear that therapeutic approaches to the disease must take into account whatever similarities or differences may exist between targets in the two major viral forms. With our focus on the protease as an intervention point, it appears that we already have inhibitors that may prove effective against HIV-1 in a clinical setting, and it could be that a single inhibitor will serve to block viral maturation in both HIV-1 and HIV-2. However, in the event that such a universal drug lacks the activity required to block maturation in HIV-2, the present work provides the basis for development of inhibitors with improved binding characteristics relative to the HIV-2 protease. In any case, it is now clear that tightly bound inhibitors of the protease in vitro are not necessarily effective antiviral agents in cell culture. If inhibitors are unable to penetrate cells or barded viral particles, they will not find their way to the target. Such considerations underscore the importance of the delivery characteristics of candidate compounds as well as $K_i$ in drug design.

Finally, it should be stressed that therapeutic approaches to AIDS treatment that are based upon recombinant CD4 strategies or “suicide substrates” such as AZT have already been validated in monkeys or humans (41, 42). Such is not the case for protease inhibitors as drugs, and it would be desirable to test the concept in an animal model such as SIV. Because of the close similarity between the genomic structures of SIV and HIV-2, it is logical that the latter disease should provide more definitive clues about compounds for testing in SIV. In considering the virally encoded proteases, it is interesting to note that although the HIV-1 and HIV-2 enzymes are about 50% identical in sequence, the latter is about 90% identical to the SIV protease (Fig. 3). Moreover, the processing sites in SIV and HIV-2 gag polypeptides are remarkably similar (51). One would infer, based upon sequence homology, that the tertiary structures of proteases from SIV and HIV-2 will be essentially identical. Indeed, we find that there are no differences in sequence between the HIV-2 and SIV proteases that would translate to a structural difference anywhere within 9 Å of the active site cleft, except for a very conservative Lys-to-Arg replacement at residue 7. One might well expect that this active site similarity between the SIV and HIV-2 proteases will extend, as well, to their substrate specificity. In any case, what we have learned about the substrate preferences of the HIV-2 protease suggests strategies for design of inhibitors that could be used to validate the concept of such enzymes as AIDS therapeutics in the monkey model.

Acknowledgments—We thank Drs. A. Wlodawer of the Frederick Cancer Research Facility, and Drs. D. B. McKay of Stanford University for providing the x-ray crystallographic coordinates for HIV-1 protease and Pseudomonas exotoxin, respectively. We wish to acknowledge D. J. Fitzgerald and I. Pastan of the National Cancer Institute for their generous gifts of exotoxin and recombinant derivatives, and for helpful discussions of the work.

REFERENCES

24. Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V.,
Compared Specificities of HIV-1 and HIV-2 Proteases