Studies on the Symmetry and Sequence Context Dependence of the HIV-1 Proteinase Specificity*

József Tózser‡‡, Péter Bagossi‡, Irene T. Weber§, John M. Louis‡, Terry D. Copeland**, and Stephen Oroszlan‡‡

From the ‡Department of Biochemistry, University Medical School of Debrecen, H-4012 Debrecen, Hungary, the ¶Department of Pharmacology, Jefferson Cancer Institute, Thomas Jefferson University, Philadelphia, Pennsylvania 19120, the §Laboratory of Cellular and Developmental Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892, the **Special Program in Protein Chemistry and Carcinogenesis Laboratory, Advanced BioScience Laboratories—Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702-1201

Two major types of cleavage sites with different sequence preferences have been proposed for the human immunodeficiency virus type 1 (HIV-1) proteinase. To understand the nature of these sequence preferences better, single and multiple amino acid substitutions were introduced into a type 1 cleavage site peptide, thus changing it to a naturally occurring type 2 cleavage site sequence. Our results indicated that the previous classification of the retroviral cleavage sites may not be generally valid and that the preference for a residue at a particular position in the substrate depends strongly on the neighboring residues, including both those at the same side and at the opposite side of the peptide backbone of the substrate. Based on these results, pseudo-symmetric (palindromic) substrates were designed. The retroviral proteinases are symmetrical dimers of two identical subunits; however, the residues of naturally occurring cleavage sites do not show symmetrical arrangements, and no obvious symmetrical substrate preference has been observed for the specificity of HIV proteinase. To examine the role of the asymmetry created by the peptide bonds on the specificity of the respective primed and nonprimed halves of the binding site, amino acid substitutions were introduced into a palindromic sequence. In general, the results suggested that the asymmetry does not result in substantial differences in specificity of the $S_2$ and $S_2'$ subsites, whereas its effect is more pronounced for the $S_3$ and $S_3'$ subsites. Although it was possible to design several good palindromic substrates, asymmetrical arrangements may be preferred by the HIV proteinase.

The specificity of retroviral proteinases has been studied intensively using both polyproteins and oligopeptides as substrates (for review, see Refs. 1–3). These studies have provided a basis for the rational design of potent, selective inhibitors. Various proteinase inhibitors are now in clinical trials or approved for therapy (for review, see Refs. 4–6). Comparison of cleavage site sequences of human immunodeficiency virus type 1 (HIV-1) and type 2 (HIV-2) suggested that the enzyme had a broad specificity and lacked consensus substrate sequence (7). Initially three types of cleavage sites were proposed for HIV-1, HIV-2, and simian immunodeficiency virus (8). Subsequently, two major types of cleavage sites were proposed for retroviral proteinases, type 1 having -Tyr(Phe)*Pro- and type 2 having hydrophobic residues (excluding Pro) at the site of cleavage (9–11). These two types of cleavage sites were proposed to have different preferences for the $P_2$ and $P_2'$ positions, where the peptide bond between $P_1$ and $P_1'$ is cleaved (notation is according to Ref. 12). Our studies with type 1 substrates indicated a preference for small residues like Cys or Asn at the $P_2$ position and a preference for $\beta$-branched Val or Ile at the $P_2'$ position (10). The lower catalytic constants with $P_2$, $\beta$-branched residues were predicted to be due to steric collision with $P_1$, Pro (10). On the other hand, using a series of peptides based on a type 2 cleavage site, $\beta$-branched residues, especially Val, were found to be favorable at $P_2$, whereas Glu was preferred at $P_2'$ (11). Interestingly, Griffith et al. (11) found Glu as the preferred $P_2'$ residue in a peptide series, when the $P_2$-$P_2'$ sequence of a type 1 cleavage site (Asn-Tyr*Pro-) was substituted into a type 2 substrate. Although some of the differences of the subsite specificity in type 1 and type 2 cleavage sites were explained by molecular modeling, most of the dependence of the specificity of proteinases on the sequence context is unexplored.

The HIV-1 proteinase is a dimer of two identical subunits. It exhibits an exact crystallographic, 2-fold rotational ($C_2$) symmetry in the structure without inhibitor (for review, see Ref. 4). Based on this symmetry, the potential advantages of $C_2$ symmetric HIV-1 proteinase inhibitors including high selectivity, potency, and stability were proposed, and structurally symmetric HIV-1 proteinase inhibitors were designed containing two amino-terminal halves of a putative substrate (13). Crystal structures of HIV-1 proteinase with inhibitors can have either symmetric or asymmetric (4). The symmetry or asymmetry was initially thought to arise from the symmetry or asymmetry of the inhibitor, but even crystal structures of HIV proteinase with symmetric inhibitors can have asymmetric proteinase substrates (14).

Considering the symmetry of the HIV proteinase, a symmetrical preference for substrate residues would be expected for

* This research was sponsored in part by the Hungarian Science and Research Fund (OTKA F 5104, T 16396), the United States-Hungarian Science and Technology Fund JP320, NIH Grant CA58166, a grant from the Intramural AIDS Targeted Program of the Office of the Director of the National Institutes of Health, and by the NCI, Department of Health and Human Services, under contract with Advanced BioScience Laboratories. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel./Fax: 36-52-416-432; E. mail: tozer@indi.biochem.dote.hu.

§ The abbreviations used are: HIV-1 and HIV-2, human immunodeficiency virus type 1 and type 2; MA, matrix protein; CA, capsid protein; NC, nucleocapsid protein. The nomenclature of viral proteins is according to Leis et al. (36).
naturally occurring cleavage sites, since the high mutation rate could readily evolve such sequences. However, there is no obvious preference for symmetrical sequences in HIV protease cleavage sites (listed in Ref. 15) or in other retroviral proteinase cleavage sites (see Ref. 1). Using a series of oligopeptides containing single amino acid substitutions in a naturally occurring type 1 cleavage site peptide, the respective P and P' positions (for example, P$_2$ and P$_2'$) appeared to be similar, but substantial differences were also found. For example, the peptide containing P$_2$ Asn was a very good substrate; however, Asn at P$_2'$ resulted in a poor substrate (10). It was not clear whether these differences were because of the asymmetrical interactions of the peptide amides and carbonyl oxygens in the substrate or intramolecular interactions of the substrate side chains (10).

To explore further the dependence of HIV protease specificity on the sequence context of its substrates we introduced single or multiple substitutions into a type 1 cleavage site peptide and changed it to a naturally occurring type 2 cleavage site sequence. Based on the results obtained, we designed a pseudosymmetric (palindromic) substrate and introduced amino acid substitutions into this sequence to explore the effect of asymmetry created by the peptide backbone on the different specificities of the respective primed and nonprimed protease sites. To study whether the enzyme prefers pseudosymmetric (palindromic) or asymmetric arrangements of the substrate residues, we have also studied the doubly substituted (also palindromic) versions of the starting pseudosymmetric substrate.

**MATERIALS AND METHODS**

**Oligopeptide Synthesis and Characterization—**Oligopeptides were synthesized by standard tert-butyloxycarbonyl or $\beta$-fluorenylmethoxy- carbonyl chemistry on a model 430A automated peptide synthesizer (Applied Biosystems, Inc.) or a semiautomatic Vega peptide synthesizer (Vega-Fox Biochemicals). All peptides were synthesized with an amide end. Amino acid composition of the peptides was determined with either a Durrum D-500 or a Waters Pico-Tag amino acid analyzer. Stock solutions and dilutions were made in distilled water (or in 10 mM sodium phosphate buffer, pH 7.4, containing 150 mM NaCl). All peptides were synthesized with an amide end, and their amino acid composition was determined by amino acid analysis.

**Enzyme Assay—**Purified HIV-1 protease was prepared as described previously (16). Active site titration for the HIV-1 protease was performed with compound 3 (17). The proteinase assays were performed in 0.25 M potassium phosphate buffer, pH 5.6, containing 8–140 nM enzyme. The reaction mixture was incubated at 37 °C for 1 h, and the reaction was stopped by the addition of guanidine HCl (6 M final concentration). The solution was acidified by the addition of trifluoroacetic acid, and an aliquot was injected onto a Durrum D-500 or a Waters Pico-Tag amino acid analyzer. Stock solutions and dilutions were made in distilled water (or in 10 mM sodium phosphate buffer, pH 7.4, containing 150 mM NaCl). The substrate or intramolecular interactions of the substrate side chains resulted in a poor substrate (10). It was not clear whether these differences were because of the asymmetrical interactions of the peptide amides and carbonyl oxygens in the substrate or intramolecular interactions of the substrate side chains (10).

Energy minimization and molecular dynamics of the modified substrates were run using the program AMMP (21), as described previously (19). Finally, the model structures were examined in the computer graphics system.

**RESULTS AND DISCUSSION**

**Substitution of Amino Acids of a Type 1 Cleavage Site Peptide with Residues of a Type 2 Substrate**

Previously, we performed extensive comparisons of the specificities of HIV-1 and HIV-2 proteases using oligopeptides representing naturally occurring cleavage sites in their Gag and Gag-Pol polyproteins (15). These cleavage sites have been classified as type 1, which contains an aromatic amino acid and Pro at P$_1$ and P'$_1$', respectively, and type 2, which has mainly hydrophobic residues but not Pro at the site of cleavage. We showed that an oligopeptide (peptide 1 in Table I) representing the cleavage site in p66 of HIV-1 for generating the p51 subunit of the heterodimeric reverse transcriptase of HIV-1 and another peptide (peptide 2 in Table I) representing the homologous sequence in p68 of HIV-2, and therefore proposed to be the cleavage site (22), were substrates of the HIV proteases (15). These peptides, which match type 2 cleavage site sequences, were the starting points for our design of palindromic substrates since they are partly symmetric with aromatic amino acids at the P$_1$, P'$_1$', and P$_2$, and they contain negatively charged residues at the P$_3$ and P'$_3$ positions. In addition, peptide 2 also contains Thr at both P$_2$ and P'$_2$ positions. Furthermore, peptides 1 and 2 with the exception of the P'$_4$ residues share the same sequence in the P$_4$-P'$_4$ region, which is the major determinant for specificity (23). However, peptide 2 was found to be a much poorer substrate of the HIV proteases than peptide 1 (see Table I and Ref. 15).

Subsequently Fan et al. (24) demonstrated that in fact the sequence of peptide 2 does not represent the actual cleavage site required to be cleaved to produce the smaller subunit of the HIV-2 reverse transcriptase. They found that the real cleavage site has the sequence of AFAM*ALTD and is downstream from the one proposed by Le Grice et al. (22). Nevertheless, we in this study and Fan et al. (24) have confirmed our initial finding that the HIV-2-derived peptide 2 or its shorter octapeptide homolog is a substrate of the HIV-1 protease.

We have also compared the specificity of the HIV-1 and HIV-2 proteases using a series of oligopeptide substrates containing single amino acid substitutions in the sequence of SP-211 (see peptide 3 in Table I), a peptide that corresponds to the type 1 MA/CA cleavage site in HIV-1 (10, 23). In these studies it was found that substitution of Pro at the P$_1$ position or in other amino acid tested, including Tyr, formed nonhydrolyzable or very poor substrates of HIV proteases. These P$_1$-substituted peptides inhibited the hydrolysis of SP-211 by HIV protease, which suggested that they were able to bind to the enzyme (10). The best inhibition was obtained with the P$_1$ Tyr-substituted peptide, suggesting its high affinity for the HIV-1 protease. In good agreement with these preliminary findings, both the $K_m$ and $k_{cat}$ values determined in the present study were substantially lower for the P$_1$ Tyr-substituted peptide compared with the unmodified one (compare peptides 3 and 4 in Table I). Substitution of P$_4$ Pro of SP-211 with Tyr converts a type 1 substrate to a type 2 substrate. However, the P$_1$ Tyr-substituted peptide is a poor substrate of HIV-1 protease (peptide 4 in Table I). To understand better the specificity of HIV-1 protease and the differences of the subsite preference in type 1 and type 2 cleavage sites, further single and multiple substitutions were carried out in the type 1 MA/CA site substrate (peptide 3), introducing residues characteristic of the type 2 cleavage site peptides 1 and 2.

---

2 J. Töszér and S. Oroszlan, unpublished results.
Pseudosymmetric Substrates of HIV Proteinase

(16809)

TABLE I

Assay of substrates having the sequence of naturally occurring cleavage sites and those containing single or multiple substitutions in the cleavage site sequence between the matrix and capsid proteins of HIV-1 by HIV-1 proteinase

| Number | Sequencea | Kcat | kcat/Km | kcat/Km |
|--------|-----------|------|---------|---------|
| 1      | AETF*YVDDGAAb | 0.046 ± 0.006 | 0.42 ± 0.02 | 9.1 |
| 2      | GAEF*YDDGSb | 0.33 ± 0.06 | 0.21 ± 0.03 | 0.64 |
| 3      | VSQNY*P1YVQa | 0.15 ± 0.03 | 6.8 ± 0.07 | 45.3 |
| 4      | VSQNY*YVQQ | 0.020 ± 0.003 | 0.010 ± 0.001 | 0.50 |
| 5      | VSQTY*P1YVQ | 0.41 ± 0.05 | 0.70 ± 0.02 | 1.7 |
| 6      | VSQNY*YVQa | 0.85 ± 0.20 | 2.07 ± 0.37 | 2.43 |
| 7      | VSQNY*YVQa | 1.77 ± 0.32 | 2.62 ± 0.15 | 1.48 |
| 8      | VSQNY*YVQa | 6.84 ± 0.98 | 3.09 ± 0.31 | 0.45 |
| 9      | VSQTY*YVQA | 0.035 ± 0.005 | 0.020 ± 0.001 | 0.57 |
| 10     | VSQNY*YVQA | 0.026 ± 0.007 | 0.18 ± 0.01 | 6.92 |
| 11     | VSQTY*YVQA | 0.021 ± 0.003 | 0.050 ± 0.001 | 2.38 |
| 12     | VSQNY*YDDO | <0.02 | NDf | 0.09 |
| 13     | VSQTY*YDDO | 0.006 ± 0.001 | 0.10 ± 0.01 | 16.67 |
| 14     | VSQNY*YDDO | 0.20 ± 0.03 | 0.20 ± 0.02 | 1.00 |
| 15     | VSQTY*YDDO | 0.17 ± 0.03 | 0.10 ± 0.01 | 0.59 |
| 16     | VSQTY*YDDO | 0.19 ± 0.03 | 0.030 ± 0.003 | 0.16 |

a Amino acids substituted in the sequence of peptide 3 are underlined.

b Peptide representing the determined cleavage site in reverse transcriptase of HIV-1. Kinetic parameters for this peptide were also reported in Ref. 15.

c Peptide representing a proposed cleavage site in reverse transcriptase of HIV-2 (22).

d Peptide representing the MA/CA cleavage site of HIV-1 (peptide 3, previously designated as SP-211). Kinetic parameters for this peptide were also reported in Ref. 15.

e ND, not determined.

f Determined as competitive substrate with peptide 45 (Table V).

In summary, results on the preference of Asn over Thr at P2 by Tyr analog of SP-211 (peptide 9 in Table I) did not yield substantial changes in the kinetic parameters, whereas substitution of P2 Thr for Ile in the same sequence context (peptide 10) yielded an approximately 10-fold increase in kcat/Km. The same substitution was very unfavorable in the P1 Thr-containing peptides (compare peptides 3 and 6 in Table I), suggesting a strong influence of the P1 residue on the preference for the P2 residue.

Substitution of Thr at both the P2 and P2 residues of peptide 4 was less effective than the single P2 substitution (compare peptides 10 and 11 of Table I). Interestingly, whereas the P2 Asp substitution of peptide 4 yielded a substrate that was even less susceptible to hydrolysis (peptide 12 in Table I), a further substitution for P2 Thr (peptide 13) yielded a substrate that was even more resistant to hydrolysis (peptide 14 in Table I). Conversely, with the larger Asn at P2 the smaller Thr is preferred over Asn at P2 (peptides 14/15, see Table II). In the -P2-Tyr*P2-Asp context there is a preference for one larger and one smaller residue at P2 and P2 suggesting complementarity of these sites.

In the case of Ile at P2, the smaller Thr is preferred over Asn at P2 (peptides 12/13 in Table II), whereas in case of Thr at P2 the larger Asn is preferred over Thr at P2 (peptides 14/15 in Table II). Conversely, with the larger Asn at P2 the smaller Thr is preferred at P2 (peptides 12/14), whereas with the smaller Thr at P2 the larger Ile is preferred at P2 (peptides 13/15). In the case of the -P2-Tyr*P2-Val context the preference is less distinct. In accordance with the predicted P2-P2 complementarity, with P2 Thr there is a preference for the larger Asn at P2 (peptides 10/11 in Table II), whereas with Asn at P2 there is a preference for the smaller Thr at P2 (peptides 4/10 in Table II). However, when P2 is Thr, P2 Thr was preferred over Ile (peptides 9/11 in Table II). Apparently the hydrophobic, b-branched Val at P2 may restrict the type of amino acid which can be accommodated optimally in the S2 subsite.

In summary, results on the preference of Asn over Thr at P2 and Ile over Thr at P2 positions (Table II) suggest a very strong dependence on sequence context. For example, P2 Thr preference depends not only on the P1 residue, but also on the P2 and P2 residues. Shrop et al. (29) suggested a context dependence in the P2-P2 residues in the specificity of avian myeloblastosis virus proteinase.
ase, which was also demonstrated by Ridky et al. (30). Our results indicate that P3 and likely other outer distal residues like P4 and P3 may also substantially influence the preference for a subsite. The change of preference at the P1 position as a function of P1 and P3 residues (Table II) also suggests that the context dependence of the specificity of HIV-1 proteinase is not restricted only to residues located at the same side of the peptide backbone. Molecular modeling and inspection of the crystal structure of HIV proteinase-inhibitor complexes suggest that the peptide backbone of the substrate does not occupy a rigid position, and depending on the substitutions, it could move toward the S or S' sites. SP-211 contains the relatively small Pro at P1, which may allow a relatively larger S2' subsite, and this peptide showed Ile over Thr preference (peptide 3/6 in Table II). A large hydrophobic residue at P1', such as Tyr, could force the peptide backbone toward the S2' site, making it smaller. This could be the reason for the Thr over Ile preference at P2' in the case of the P1' Tyr-substituted peptides (see peptides 4/10 in Table II).

Substitutions in a Palindromic Substrate Sequence
S3 and S3' Substitutions—Peptide 16 (Table I) contains a palindromic sequence at the P3-P3' positions of substrates.

Elimination of P5 Val and substitution of P4 Gln for Ser resulted in a completely palindromic substrate (peptide 17 in Table III). It is important to note that this shorter peptide showed 10-fold higher kcat and kcat/Km values than peptide 16. The effect of residues P5 and P4 on substrate hydrolysis has been reported previously (11, 23). To compare the specificity of the S3 and S3' substrate binding pockets in an identical sequence context, the Asp residues of the palindromic peptide 17 were substituted to Gly, Gln, Phe and Leu residues (Table III). Although the Gly and Gln residues were also doubly substituted, kinetic measurements were not possible for the doubly substituted Phe and Leu peptides because of low solubility (peptides 26 and 29 in Table III).

Single substitutions to Gln also resulted in moderate changes in the kinetic parameters, whereas the double substitution yielded a substantial reduction of the kcat value. Whereas substitution to P3 Phe resulted in a 4-fold decrease in Km and kcat (peptide 24), the P3' Phe substitution did not cause such a big effect (peptide 25). Substitution of P2 or P3' Asp to Leu resulted in a substantial decrease in catalytic constants.

### Table II

Preferences for Asn over Thr at P2 and Ile over Thr at P2'

| Peptides | Sequencea | kcat/Km ratio |
|----------|-----------|---------------|
| 3/5      | Asn/Thr   | P2 P1 P1' P2' P3' | 26.6 |
| 4/9      | Asn/Thr   | Tyr Pro Ile Val  | 0.9 |
| 10/11    | Asn/Thr   | Tyr Tyr Thr Val  | 2.9 |
| 12/13    | Asn/Thr   | Tyr Tyr Ile Asp  | 0.005 |
| 14/15    | Asn/Thr   | Tyr Tyr Thr Asp  | 1.7 |
| 3/6      | Asn       | Tyr Pro Ile/Thr Val  | 18.6 |
| 4/10     | Asn       | Tyr Tyr Ile/Thr Val  | 0.07 |
| 9/11     | Thr       | Thr Tyr Ile/Thr Val  | 0.24 |
| 12/14    | Asn       | Tyr Tyr Ile/Thr Asp | 0.09 |
| 13/15    | Thr       | Thr Tyr Ile/Thr Asp | 28.25 |

a Only P2-P3 segment of the substrates are shown; the outer residues were identical. The preferred residue is underlined.

FIG. 1. Panel A, schematic representation of the P1' Tyr-substituted HIV-1 MA/CA substrate (peptide 4 in Table I) in the S3-S3' subsites of proteinase. The relative size of each subsite is indicated approximately by the area enclosed by the curved line around each substrate side chain. Panel B, stereoview of residues P3-P3' of peptides 1 (thin lines) and 4 (thick lines).
Modeling showed that the same residues of the enzyme interact with the appropriate subsite at both sides of the scissile bond (Table IV). S3 and S3' subsites are near the surface, and they are formed by hydrophobic residues and two charged residues (Table IV). In the starting palindromic substrate (peptide 17) the P3 and P3' Asp residues can interact with Arg-8' and Arg-8, respectively (Fig. 2) and compensate for the lower van der Waals energy compared with Phe in P3 and P3' positions. Surprisingly, the catalytic constants of the Gly-substituted peptides are comparable to those obtained with Phe and Asp substitutions. Apparently, P3 and P3' in this sequence context do not have a substantial effect on the efficiency of hydrolysis. The largest differences between the P3 and P3' mutated peptides were obtained with Leu substitutions. Also, these peptides had the lowest k_cat/K_m values. Previously, we had proposed that the interaction between the hydrophobic P3' residue of SP-211 (see peptide 3 in Table I) and Phe-53 of the enzyme may play a crucial role in proper closing of the flexible flap (10). In a molecular dynamics simulation of the flap movement, Phe-53 and its symmetry counterpart were among the few residues suggested to be involved in the “triggering” event (31). In the palindromic sequence context, although peptides with hydrophobic residues at P3' have lower K_m values than those with polar residues or Gly (Table III), the differences are not as dramatic as in the case of SP-211, where at least 10-fold differences were found (peptides 3 and 8 in Table I; Ref. 10). In our palindromic sequence context, the effect of a hydrophobic residue at P3' resembles that of P3 in SP-211 in which only a moderate increase of K_m was obtained by introducing nonhydrophobic residues at P3 (10). A P3' Tyr residue in the pseudosymmetric context may partially occupy the overlapping S3 subsite as P1 Tyr can partially occupy the S3 pocket as P1 Tyr can partially occupy the S3 pocket (Fig. 3), resulting in less dependence on P3' hydrophobicity for the proper closing of the flap.

Only moderate differences were found in the kinetic parameters of substrates having identical substitutions at the P3 and P3' sites, except for the Leu substitutions. These results suggest that the asymmetry of the peptide backbone does not play a major role in the large differences of the specificities of the respective S3 and S3' subsites we had previously established (10). Instead, the specificity differences are a consequence of the different sequence context of the substrates.

**S2 and S2' Substitutions**—The P3 or P3' Thr residues of peptide 17 (Tables III and V) were substituted with Gly, Glu, Cys, Ala, Ile, and Leu residues (Table V). Because of the low solubility, the doubly substituted Ile and Leu substrates were not assayed (peptides 44 and 47, respectively). Single substitutions of the Thr residues with Gly resulted in substantial increases in K_m values and decreases of k_cat/k_m values compared with peptide 17, whereas the doubly substituted peptide was not hydrolyzed (peptides 31–33 in Table V). Previous studies showed that Gly in these positions does not form good substrates, independently of the sequence context (10, 32). Substitution of P3 Thr to Glu resulted in a large increase in K_m value; however, this peptide showed the highest k_cat value among all of the substituted peptides (peptide 33 in Table V). P3' Glu substitution (peptide 34) resulted in a less dramatic increase in K_m and unchanged k_cat values. P3' Glu was found to be optimal in both type 1 and type 2 sequence context by Griffith et al. (11). This preference was predicted to be because of hydrogen bonding of the side chain carboxyl to the backbone NH of residues 29' and 30' of the enzyme (11). Also, P3' Glu is frequently found in viral and cellular protein cleavage sites of HIV-1 protease (33). At the P3-P3' region the type 1 substrate used by Griffith et al. (11) contained residues (Asn-Tyr'Pro-Ile') identical to those in SP-211; however, P3' Glu substitution...

---

**Fig. 2. Stereoview of residues P3-P3' of the palindromic peptide 17 (Table V).** Hydrogen bond interactions between P3 Asp and Arg-8' as well as P3' Asp and Arg-8 in the enzyme-substrate complex are shown by dashed lines.
was an unfavorable substitution in the SP-211, suggesting the importance of residues outside the P2-P2′ determining P2′ preference, as we have also found with the multisubstituted peptides (Table II). Therefore, preference for Glu at P2 was an unfavorable substitution in the SP-21, suggesting the asymmetry of the peptide bonds, the pseudo-symmetry center of the substrates is shifted approximately by 0.5 Å, positioning the carbonyl oxygen rather than the center of the scissile bond into the line of the C2 symmetry axis of the proteinase dimer. This deviation puts the P2′ residue farther from the center of the enzyme compared with P2 (Fig. 4), toward the open end of the binding site, which may be a reason why larger residues are better at P2′ than at P2. The effect of this deviation is stronger close to the center, which may be the reason for larger differences obtained for the P2 and P2′ substitutions than for the P3 and P3′ substitutions. Furthermore, whereas S3, S3′ subsites are exposed at the surface, the S2, S2′ subsites are more restricted because they are smaller and inside the proteinase.

**Hydrolysis of Substrates Having Palindromic Sequences**

The hydrolysis of peptide 17 and its doubly substituted derivatives in Tables III and V showed that peptides with completely palindromic sequences could be substrates of HIV-1 proteinase. Based on mirroring the NH2-terminal side of the sequence of the naturally occurring MA/CA (VSQNY*PIVQ) and CA/NC cleavage site (TATIM*MQRGN), we synthesized other palindromic peptides, and these were also found to be substrates of the enzyme (peptides 48 and 49 of Table VI). Simultaneous substitution of P2 and P2′ residues of peptide 49 to Ala (peptide 50) produced little change in kcat/Km. Since Tyr is better in P1 than Met in both type 1 (10) and type 2 cleavage site sequences (2), Tyr was substituted for Met at both P1 and P1′, which resulted in the most efficient palindromic substrate in our series (peptide 51). These results suggest that various palindromic substrates could be designed for HIV-1 proteinase,
and specificity studies using these substrates could eliminate the differences caused by the sequence context on the preference of the respective S and S’ sites.

It is also of interest to note that exchange of P₁ and P₁’ residues of SP-211 as well as mirroring the P’ residues resulted in nonhydrolyzable peptides (peptides 52 and 53 of Table VI). These results also indicate the nonequivalent nature of the respective S and S’ subsites.

### Conclusions

The dependence of the specificity of HIV-1 proteinase on the sequence context of its substrate peptides was studied by making single and multiple substitutions of amino acids in naturally occurring cleavage site peptides, and palindromic substrates were designed. Although it was possible to design palindromic substrates, our results suggest that a symmetric arrangement of residues may not be favorable. The S₁-S₃’ region of the HIV proteinase is generally hydrophobic. However, a completely hydrophobic sequence may not provide soluble oligopeptides. Furthermore, the naturally occurring cleavage sites in retroviruses are expected to be in regions connecting folded protein domains of polyproteins, where a high degree of hydrophobicity may be unfavorable for the correct conformation. Besides the hydrophobicity, the size of residues also appears to be important. Our results suggest that in the case of tight packing of S₁ and S₁’ with Tyr residues, a relatively larger residue at either P₂ or P₂’ is preferable, but not at both positions. These factors may explain why symmetric subsite arrangements have not evolved at the naturally occurring cleavage sites, despite the symmetric nature of the retroviral proteinases.

Previous studies did not explain whether the rather large differences in specificity of respective S and S’ subsites are attributed to the asymmetry introduced by the binding of the substrate and/or the different amino acid sequence context. Design of a palindromic substrate and substitutions in its respective positions allowed us to study these effects. In general, our results suggest that the previously established differences in specificity of P₂ and P₂’ positions could be mainly due to the different sequence context: in these positions the asymmetry introduced by the peptide backbone of the substrate does not seem to play an important role. For P₂ and P₂’ positions the effect of asymmetry is not negligible; however, the different sequence context could be still the predominant cause of the substantial differences found previously (10).

Our results also show that the earlier classification of retroviral cleavage sites into two types is an oversimplification. Preference of Asn in P₂ and Ile in P₂’ in type 1 cleavage sites as well as Ile (Val) in P₂ and Glu at P₂’ in type 2 cleavage site substrates seems to be a function of the residues outside of the P₂-P₂’ region of the substrate. The enzyme-substrate interaction of the HIV-1 proteinase as well as other retroviral proteinases cannot be described residue by residue, but the neighboring residues of the substrate should also be taken into account. These neighboring residues include not only the residues at the same side of the peptide backbone but also residues at the opposite side as suggested by Strop et al. (29) from analysis of avian myeloblastosis virus proteinase specificity. Similar results were obtained by Ridky et al. (30) from kinetic analysis of substrates containing double substitutions from P₂ to P₂’ in the NC/proteinase cleavage site peptide of Rous sarcoma virus. However, the specificity pattern is even more complex than was originally thought, not only the internal P₂-P₂’ residues, but outer residues like P₂’ may also substantially influence the preference for a subsite, even for positions separated by four residues in the peptide sequence.

The most important features of the sequence context dependence are the size and β-branch in a residue. P₁’ Tyr, unlike Pro, may fill completely the large S₁’ subsite and could restrict the space for substrate residues at other subsites. On the other hand, substitution of Pro at P₁’ could eliminate unfavorable interactions with β-branched residues at P₂ (10, 30). Even small changes in the substrate sequence might cause rearrangement of the substrate residues, as seen in some enzyme-inhibitor complexes as well as predicted by molecular modeling. This strong sequence context dependence also raises difficulties for predicting cleavage sites. For example, an early prediction method developed by Poorman et al. (7) failed to predict the cleavage site in Nef protein (34). New prediction methods that take into account the context-dependent nature...
of the specificity (35) might give better results. Also, molecular
modeling and improved energy calculation methods may help
to develop better methods for predictions.

Acknowledgments—We thank Tunde Caire for technical assistance;
Cathy Hixson and Suzanne Specht for help with the amino acid anal-
ysis; Ivo Bláha, Eva Majerova, and Patrick Wedlock for peptide syn-
thesis; David Cavanough for help in molecular modeling; Jin Wu for
preparing the figures; and Carol Shawver for preparing the manuscript.

REFERENCES
1. Oroszlan, S., and Luftig, R. B. (1990) "Curr. Top. Microbiol. Immunol." 157,
153–185
2. Dunn, B. M., Gustchina, A., Wlodawer, A., and Kay, J. (1994) "Methods
Enzymol." 241, 254–278
3. Tomasselli, A. G., and Heinrickson, R. L. (1994) "Methods Enzymol." 241,
279–301
4. Wlodawer, A., and Ericksson, J. W. (1995) "Annu. Rev. Biochem." 62, 543–585
5. Winslow, D. L., and Otto, M. J. (1995) "AIDS" 9, Suppl. A, S183–S192
6. Mellors, J. W. (1996) "Nat. Med." 2, 274–275
7. Poorman, R. A., Tomasselli, A., Heinrikson, R. L., and Kezdy, F. J. (1991)
"J. Biol. Chem." 266, 14554–14561
8. Henderson, L. E., Benveniste, R. E., Sawder, R., Copeland, T. D., Schultz, A.
M., and Oroszlan, S. (1988) "J. Virol." 62, 2587–2595
9. Pettit, S. C., Simsac, J., Loeb, D. D., Evricter, L., Hutchison, C. A., III, and
Swanson, R. (1991) "J. Biol. Chem." 266, 14539–14547
10. Tozsér, J., Weber, I. T., Gustchina, A., Bláha, I., Copeland, T. D., Louis, J. M.,
and Oroszlan, S. (1992) "Biochemistry" 31, 4793–4800
11. Griffith, J. T., Phylip, L. H., Konvalinka, J., Strop, P., Gustchina, A.,
Wlodawer, A., Davenport, R. J., Briggs, R., Dunn, B. M., and Kay, J. (1992)
"Biochemistry" 31, 5193–5200
12. Schechter, I., and Berger, A. (1967) "Biochem. Biophys. Res. Commun." 27,
157–162
13. Erickson, J., Neidhart, D. J., VanDrie, J., Kempf, D. J., Wand, X. C., Norbeck,
D. W., Plattner, J. J., Rittenhouse, J. W., Turon, M., Widelburg, N.,
Kohlbrenner, W. E., Simmer, E., Helfrich, R., Paul, D. A., and Knigge, M.
(1990) "Science" 249, 527–533
14. Dreyer, G. B., Boehm, J. C., Chenera, B., Desjarlais, R. L., Hassell, A. M.,
Meek, T. D., and Tomasszek, T. A., Jr. (1993) "Biochemistry" 32, 957–967
15. Tozsér, J., Bláha, I., Copeland, T. D., Wondrak, E. M., and Oroszlan, S. (1991)
"FEBS Lett." 281, 77–80
16. Louis, J., McDonald, R., Nashed, N., Wondrak, E. M., Jerina, D., Oroszlan, S.,
and Mora, P. (1991) "Eur. J. Biochem." 199, 361–369
17. Grobelny, D., Wondrak, E. M., Galardy, R. E., and Oroszlan, S. (1990)
"Biochem. Biophys. Res. Commun." 169, 1111–1116
18. Sack, J. S. (1988) "J. Mol. Graph." 6, 224–225
19. Bagossi, P., Cheng, E. Y. S., Oroszlan, S., and Tozsér, J. (1996) "Protein Eng." 9,
997–1003
20. Weber, I. T., and Harrison, R. W. (1996) "Protein Eng." 9, 679–690
21. Harrison, R. W. (1993) "J. Comp. Chem." 14, 1112–1122
22. Le Grice, S. F. J., Este, R., Mills, J., and Moas, J. (1989) "J. Biol. Chem." 264,
14902–14908
23. Tozsér, J., Gustchina, A., Weber, I. T., Bláha, I., Wondrak, E. M., and
Oroszlan, S. (1991) "FEBS Lett." 279, 356–360
24. Fan, N., Rank, K. B., Leone, J. W., Heinrikson, R. L., Bannow, C. A., Smith, C.
W., Evans, D. B., Poppe, S. M., Tarpley, W. G., Rothrock, D. J., Tomasselli,
A. G., and Sharma, S. K. (1995) "J. Biol. Chem." 270, 13573–13579
25. Perseh, A. (1985) "Enzyme Structure and Mechanism," 2nd Ed., pp. 311–317,
W. H. Freeman and Co., New York
26. Pozsgay, M., Cs-Szabó, G., Bajusz, S., Simonsson, R., Gáspár, R., and Elodi, P.
(1981) "Eur. J. Biochem." 115, 497–502
27. Tozsér, J., Cs-Szabó, G., Pozsgay, M., Aurell, L., and Elodi, P. (1986) "Acta
Biochim. Biophys. Hung." 21, 335–348
28. Baldwin, E. T., Bhat, T. N., Liu, B., Pattabiraman, N., and Ericksson, J. W.
(1995) "Nat. Struct. Biol." 2, 244–249
29. Strop, P., Konvalinka, J., Stys, D., Pavlickova, L., Bláha, I., Velek, J.,
Travnick, M., Kostka, V., and Sedlacek, J. (1991) "Biochemistry" 30,
3437–3443
30. Ridky, T. W., Cameron, C. E., Cameron, J., Leis, J., Copeland, T., Wlodawer,
A., Weber, I. T., and Harrison, R. W. (1996) "J. Biol. Chem." 271,
4709–4917
31. Collins, J. R., Burt, S. K., and Ericksson, J. W. (1995) "Nat. Struct. Biol." 2,
334–338
32. Cameron, C. E., Grinde, B., Jacques, P., Jentoft, J., Leis, J., Wlodawer, A.,
and Weber, I. T. (1993) "J. Biol. Chem." 268, 11711–11720
33. Tomasselli, A. G., Sarach, J. L., Barrett, L. J., Reardon, I. M., Howe, W. J.,
Evans, D. B., Sharma, S. K., and Heinrikson, R. L. (1993) "Protein Sci."
2167–2176
34. Gaedigk-Nitschko, K., Schoen, A., Wachinger, G., Erfe, V., and Kohleisen, B.
(1995) "FEBS Lett." 357, 275–278
35. Chou, K.-C., Tomasselli, A. G., Reardon, I. M., and Heinrikson, R. L. (1996)
"Proteins" 24, 51–72
36. Leis, J., Baltimore, D., Bishop, J. M., Coffin, J., Fleissner, E., Goff, S. P.,
Oroszlan, S., Robinson, H., Skalka, A. M., Temin, H. M., and Vogt, V. (1988)
"J. Virol." 62, 1808–1809