Human Immunodeficiency Virus, Type 1 Protease Substrate Specificity Is Limited by Interactions between Substrate Amino Acids Bound in Adjacent Enzyme Subsites*

(Received for publication, September 18, 1995, and in revised form, November 14, 1995)

Todd W. Ridkytý, Craig E. Cameron®, J ohn Cameron, and J onathan Leis†
From the Case Western Reserve University School of Medicine, Cleveland, Ohio 44106-4935

Terry Copeland and Alexander Wlodawer
From the NCI-Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, Maryland 21702

Irene T. Weber and Robert W. Harrison
From the Jefferson Cancer Institute, Thomas Jefferson University, Philadelphia, Pennsylvania 19107

The specificity of the retroviral protease is determined by the ability of substrate amino acid side chains to bind into eight individual subsites within the enzyme. Although the subsites are able to act somewhat independently in selection of amino acid side chains that fit into each pocket, significant interactions exist between individual subsites that substantially limit the number of cleavable amino acid sequences. The substrate peptide binds within the enzyme in an extended anti-parallel β sheet conformation with substrate amino acid side chains adjacent in the linear sequence extending in opposite directions in the enzyme-substrate complex. From this geometry, we have defined both cis and trans steric interactions, which have been characterized by a steady state kinetic analysis of human immunodeficiency virus, type-1, protease using a series of peptide substrates that are derivatives of the avian leukosis/sarcoma virus nucleocapsid-protease cleavage site. These peptides contain both single and double amino acid substitutions in seven positions of the minimum length substrate required by the retroviral protease for specific and efficient cleavage. Steady state kinetic data from the single amino acid substituted peptides were used to predict effects on protease-catalyzed cleavage of corresponding double substituted peptide substrates. The calculated Gibbs' free energy changes were compared with actual experimental values in order to determine how the fit of a substrate amino acid in one subsite influences the fit of amino acids in adjacent subsites. Analysis of these data shows that substrate specificity is limited by steric interactions between pairs of enzyme subsites. Moreover, certain enzyme subsites are relatively tolerant of substitutions in the substrate and exert little effect on adjacent subsites, whereas others are more restrictive and have marked influence on adjacent cis and trans subsites.

The retrovirus protease (PR) is responsible for the post-translational processing of viral gag and gag-pol polyprotein precursors (1). This proteolytic processing is a necessary step in the replication of infectious virus and is a late event occurring as particles bud from infected cells. Cleavage of the viral polyproteins requires human immunodeficiency virus, type 1 (HIV-1), or avian myeloblastosis/Rous sarcoma virus (AMV/RSV) PR to act on nine unique sequences, each 8 amino acids in length. Consistent with this is the finding that the minimum length of a peptide substrate required for specific cleavage by either PR is 6–8 amino acids, depending upon the source of the enzyme (2–4). Substrates bind to HIV-1 PR in an extended anti-parallel β strand conformation with substrate amino acid side chains adjacent in the linear sequence extending in opposite directions in the enzyme-substrate complex (see Fig. 1). Interaction between substrate amino acid side chains and the corresponding binding pockets in the enzyme determines enzyme specificity. It has been shown previously that a variety of amino acid residues can be accommodated in each of the enzyme subsites, when single amino acid substitutions are placed in the context of an efficiently cleaved substrate (5). Additionally, it was found that individual enzyme subsites are capable of acting relatively independently in recognition of amino acids in the corresponding substrate position (6). If each of the eight subsites were able to accept n different amino acid side chains, where n = 4–7 amino acids as found in the naturally occurring gag and pol polyprotein cleavage sites and the subsites were acting completely independently in substrate amino acid selection, then PR would be able to cleave n8 different substrate sequences. However, in contrast to cellular proteases such as pepsin (7), the retroviral PR displays a remarkably limited substrate range, cleaving only a very select set of amino acid sequences.

In this report, we define substrate parameters that limit the possible combinations of amino acids that constitute a functional cleavage site. The activity of HIV-1 PR was analyzed with the use of a library of single and double substituted

---

* This work was supported in part by United States Public Health Service Grants CA52047 (to J. L.) and CA58166 (to L. T. W.) and National Cancer Institute, Department of Health and Human Services contract N01 C01 46000 with ABL (to T. C. and A. W.). 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.

† Medical Science Trainee supported by Public Health Service Grant GM07250 from the National Institutes of Health.

§ Recipient of a Predoctoral Fellowship GM13628 from the National Institutes of Health.

¶ To whom all correspondence should be addressed.

The abbreviations used are: PR, retroviral protease; RSV, Rous sarcoma virus; AMV, avian myeloblastosis virus; HIV, human immunodeficiency virus; NC, retroviral nucleocapsid protein.
synthetic peptide substrates, representing the cleavage junction between the naturally occurring RSV nucleocapsid (NC) and PR proteins in the gag precursor polyprotein. A steady state kinetic analysis was used to calculate ΔΔG values representing the difference in the Gibbs’ free energy changes for the proteolysis reactions resulting from amino acid substitutions in the wild type NC-PR-based substrate. A comparison was made between the experimentally observed ΔΔG values for the double substituted peptides and the predicted ΔΔG values calculated using the data derived from the single substituted peptides. This analysis indicates that there are steric interactions between amino acids in adjacent and alternate substrate positions that restrict the combinations of amino acids that comprise a functional cleavage site.

EXPERIMENTAL PROCEDURES

Purification of Retroviral Proteases—AMV PR was purified from virus obtained from Molecular Genetic Resources, Tampa, FL as described previously (11). HIV-1 PR was expressed in Escherichia coli and purified from the inclusion body fraction according to a procedure developed by Dr. C. Z. Giam, Case Western Reserve University.2 Briefly, E. coli JM105 harboring a plasmid with a lacZ2-HIV-1 PR fusion protein under control of the lac promoter was expressed with the addition of isopropyl-1-thio-D-galactopyranoside (2 mm) for 10 h. The inclusion body fraction was isolated and solubilized with 8 M urea and 150 mM 2-mercaptoethanol. The denatured protein was passed over a DEAE-Sepharose column equilibrated with 8 M urea, 150 mM 2-mercaptoethanol, 20 mM Tris-HCl, pH 8.8. The denatured fusion protein did not bind to the column. The protein was renatured by removal of the urea by dialysis against 20 mM HEPES, pH 7.0, 10 mM 2-mercaptoethanol. This activates the PR to cleave itself out of the fusion protein. The PR preparation was then dialyzed against 20 mM MES, pH 5.5, 10 mM mercaptoethanol and uncleaved fusion protein was separated from PR by-passing through a carboxymethyl-cellulose column equilibrated with 8 M urea, pH 5.5, 0.1 M KCl, and 10 mM mercaptoethanol. Protease was eluted from column by running a linear gradient from 0.1 M KCl to 20 mM MES, pH 5.5, and 10 mM 2-mercaptoethanol. Protease eluted at about 0.3 M KCl. The purified PR was concentrated with an amicon centrifuge concentrator to 0.2 mg/ml and frozen in aliquots at –80 °C. Aliquots of HIV-1 PR were used only once. The HIV-1 PR was greater than 95% pure as judged by SDS-polyacrylamide gel electrophoresis. The preparations specifically cleaved peptide substrates based on the natural polyprotein cleavage sites and had no detectable cleavage on nonrelated protein sequences or mature HIV-1 reverse transcriptase.

Peptides—The peptides used in this study were synthesized chemically and purified as described previously (5, 50). The peptides were analogs of the RSV NC-PR cleavage junction. The natural sequence is indicated in bold letters, and the substituted substrates contained either one or two amino acid changes in the underlined positions of this sequence: PPAVS-LAMTMR. Peptides were solubilized in 1 mM 2-mercaptoethanol or 1 mM dithioerythritol, and their concentrations were determined by amino acid composition analysis.

Assay of PR Activity—The reaction mixture contained 80 mM sodium phosphate, pH 5.9, 0.8 M sodium chloride, 10–400 μM peptide as indicated, and 0.5–5 μM HIV-1 PR. Reaction volumes were 25 μl. Incubations times varied from 2 to 10 min at 37 °C depending upon the substrate. Reactions were initiated by the addition of PR and stopped by the addition of 300 μl of 0.5 M sodium borate, pH 8.5. 20 μl of 0.05% (w/v) fluorescamine was then added. HIV-1 PR was never incubated more than 10 min due to its instability, presumably a result of auto-degradation. After reaction with fluorescamine, the relative fluorescence was determined on a Perkin-Elmer LS-50B spectrophluorometer using an excitation wavelength of 386 nm and an emission wavelength of 477 nm. Excitation and emission slit widths were 5 and 10 nm, respectively. Relative fluorescence intensity was converted to moles of product using a standard curve described by the following equation: nmoi of product = relative fluorescence intensity/400. The standard curve was obtained using a hexapeptide with a free amino terminus (5). The AMV PR was assayed as described previously (5). The peptides used in this study were designed with prolines at their amino termini so that the relative fluorescence intensity represents only the newly formed amino termini produced as a result of proteolytic cleavage. Arginine residues were added to the carboxyl terminus to improve solubility of the peptides, without changing their kinetic parameters.

Steady State Kinetic Analysis—Kinetic constants were determined using the assay described above. Concentrations of peptide ranged from 0.25 to 4 times the K_m value. No more than 20% of the substrate was allowed to be consumed during the course of any given experiment. Initial velocity data used to calculate kinetic constants were obtained from at least three experiments performed in duplicate. Kinetic constants were determined by a nonlinear fit of the data to the Michaelis-Menten equation using the NFIIT program (9, 10). Correlation coefficients of the fit were greater than 0.98, and the standard deviation of the constants reported was less than 20%.

ΔΔG Calculations—The ΔΔG values for single and double substituted peptides were determined from steady state kinetic data (k_cat/K_m) values by the relationships described below. The equations were derived from relationships described by Fersht (12). ΔΔG, or ΔΔG, represent the deviation in the Gibbs’ free energy change from the wild type NC-PR peptide for single substituted peptides. ΔΔG represents deviation in the Gibbs’ free energy for a double substituted peptide. The ΔΔG was determined experimentally and also predicted mathematically by adding the ΔΔG values from the corresponding single substituted peptides.

\[
\Delta \Delta G = \Delta \Delta G + \Delta \Delta G \quad \text{(predicted)}
\]

where

\[
\Delta \Delta G = - \frac{RT \ln(\frac{k_m}{k_m^0})}{x} + \ln(\frac{k_m^0}{k_m^0}) \quad \text{(experimental)}
\]

\[
\Delta \Delta G = - \frac{RT \ln(\frac{k_m}{k_m^0})}{x} \quad \text{(3)}
\]

\[
\Delta \Delta G = - \frac{RT \ln(\frac{k_m}{k_m^0})}{x} \quad \text{(4)}
\]

\(k_m^0\) represents the catalytic efficiency for the wild type NC-PR peptide; \(k_m^0\) and \(k_m^0\) represent the catalytic efficiency for different single substituted peptides, respectively; \(k_m^0\) represent the catalytic efficiency for double substituted peptides.

Molecular Modeling—The starting structure consisted of the protease dimer from the crystal structure of HIV-1 protease with the inhibitor, J365, (13), and a model for a peptide substrate. The structures were examined on a Silicon Graphics SGI 316 computer graphics system running the program CHAIN (14). The peptide substrate was built by altering amino acid side chains in the inhibitor and forming a peptide group (CONH) instead of the nonhydrolyzable bond of the inhibitor. The modeled substrate consists of variations of the central eight residues (P4-P4'), Pro-Ala-Val-Ser-Leu-Ala-Met-Thr, of the peptide representing the NC-PR cleavage site of RSV, where the peptide bond between Ser and Leu is hydrolyzed. All metallographic water molecules were included because several appeared to be structurally important. A transition state substrate model was built in which the scissile peptide CO-NH was replaced by the transition state (COH)NH2, and the water that interacts with either the substrate or inhibitor and with the flaps of the protease was removed.

The atomic coordinates for the substituted residues were produced with the program AMMP (15), and the energy of the HIV protease-substrate complex was minimized. A modified version of the UFF potential set (16) was used. Infrared spectral data were not included in the original UFF parameterization and have been used to improve the parameters for proteins and nucleic acids. These modifications do not significantly change the performance of the potential set on small molecules but result in consistently smaller root mean square deviations between minimized and observed protein and nucleic acid structures. One of the strengths of the UFF potential is that the new terms were easy to add in a manner that is consistent with the rest of the potential set. The atomic charges from the AMBER all atom set were used for the protein and water (18). Charges for the nonstandard groups in the transition state were produced as described in Harrison et al. (19).

No screening dielectric term or bulk solvent correction was included. No cut-off was applied for nonbonded and electrostatic terms, which were calculated with an algorithm that amortizes or spreads the cost of calculation over many simpler calculations, which results in lower

2 C. Z. Giam, personal communication.

3 I. T. Weber, and R. W. Harrison, submitted for publication.
RESULTS AND DISCUSSION

We have used a simple method to determine the extent to which amino acids in given substrate positions influence the fit of adjacent and alternate substrate amino acids into their corresponding enzyme subsites. This method involves a steady state kinetic analysis of HIV-1 PR on a series of peptide substrates based on the RSV NC-PR cleavage sequence that have two residues altered from wild type. The amino acid substitutions chosen were those that were analyzed as single substitu-
tions chosen were those that were analyzed as single substitu-
tions, and the amino acid positions were initially tethered to those in the crystal structure of HIV-1 protease in order to calculate and minimize the hydrogen atom positions. The side chain atoms were

average cost, as described in Harrison and Weber (20). The atomic positions for the protein and water molecules were initially tethered to those in the crystal structure of HIV-1 protease in order to calculate and minimize the hydrogen atom positions. The side chain atoms were removed down to the C α atom, or the Ca for substitution of Gly, for the substituted amino acids, and the new atomic positions were created by a variation on distance geometry (19). The new atoms were substituted with respect to bond, angle, torsion, and hybrid potentials. The protease structure with nonhydrogen atoms from the crystal structure and minimized hydrogen atoms was combined with each of the different peptides with single or double amino acid substitutions. Then, each of the side chain torsion angles for substituted residues in the peptide sub-
strate was rotated through 360° in steps of 15° to search for alternate conformations. This torsion search finds the angle(s) that have a minimum in the nonbonded energy. Finally, each model of HIV protease with a different substrate was optimized by a longer minimization using 100 steps of conjugate gradients followed by eight cycles of alternating conjugate gradients (30 steps) and short runs of molecular dyna-

mics (20 fs steps at 300 K).

TABLE I

Comparison of HIV-1 PR kinetic data for AMV/RSV NC-PR substrates with single substitutions in the P3 to P1′ positions

Varying concentrations of the NC-PR (PPAVS-LAMTMRR) or NC-PR substrates with amino acid substitutions in the P3 to P1′ positions as indicated were incubated with purified HIV-1 PR (50 ng), and the extent of cleavage was determined using the fluorescamine assay described under “Experimental Procedures.” The wild type amino acid in the NC-PR peptide is indicated in parentheses in the first column.

| Substrate position | Substituted amino acid | kcat\(^a\) | Km | Relative kcat/Km | ΔΔG\(^b\) observed\(^b\) |
|--------------------|------------------------|-----------|-----|----------------|--------------------------|
| P3 (Ala)           | Phe                    | 1.8       | 0.9 | 2.1            | -0.45                    |
|                    | Ser                    | 1.2       | 1.3 | 0.92           | 0.05                     |
|                    | Thr                    | 0.87      | 1.3 | 0.7            | 0.22                     |
| P2 (Val)           | Ala                    | 1.0       | 3.6 | 0.3            | 0.8                      |
|                    | His                    | 0.9       | 3.6 | 0.25           | 0.85                     |
| P1 (Ser)           | Ile                    | 1.2       | 1.2 | 1.6            | -0.29                    |
|                    | Val                    | 1.4       | 2.3 | 0.62           | 0.29                     |
| P1′ (Leu)          | Ala                    | 0.7       | 6.6 | 0.1            | 1.42                     |

\(^a\) The kinetic parameters for the P3 to P1′-substituted NC-PR substrates are given relative to the unmodified wild type substrate, which is defined as equal to 1 in each case. For HIV-1 PR, the Km varied from 13.4 to 16 μM, and kcat varied from 14.9 to 43.8 min\(^{-1}\), with calculated kcat/Km values from 1.1 to 2.7 min\(^{-1}\) μM\(^{-1}\) dependent upon the enzyme preparation used.

\(^b\) The ΔΔG values were calculated from the kinetic data as described under “Experimental Procedures.”

We propose that this lower activity is the result of steric interference that places one or both of the amino acids in the cis subsites in an altered conformation not favorable for binding and cleavage. This is suggested by analysis of the peptide substrates modeled in the HIV-1 PR structure in Fig. 4, which
shows the relative positions of the P3 to P1 amino acids of the NC-PR peptide substrate predicted by energy minimalization. Shown are two peptides, both of which have His substituted for Ala in P3 and one that also has Trp substituted for Ser in P1 (Fig. 4, thin lines and balls). The bulky Trp in P1 appears to directly effect the position of the His in P3, which is pushed back toward Phe53. The only exception to this steric argument observed so far involves Gly in P3 with Trp in P1. Glycine, which does not have a side chain to contribute to substrate binding energy, is a special case and will be discussed later in the context of several Gly substituted peptides.

Asimilar steric interaction can be seen with cis substitutions in P2 and P1’ (Fig. 2B). When Gly or Ala is placed in P2 with Phe in P1’, the predicted and experimental values agree. When the larger Leu is placed in P2, there appears to be insufficient flexibility in the binding pockets to accommodate the combination of the two large groups extending into the same side of the enzyme. Peptides that fixed Leu in P2 and vary the amino acid in P1’ also have been analyzed. However, these peptides, which have Ala or Gly in P1’, were predicted to be cleaved poorly because the single substituted substrates are cleaved with low efficiency. Indeed, this was observed, and experimental ΔG values could not be calculated accurately because product did not accumulate to any measurable extent under the assay conditions used. In the above argument, the size of an amino acid relative to its given subsite determines the magnitude of the steric effect. For instance, Leu would be considered a large residue in the small S2 subsite, although it would be a medium sized residue in the larger S1 or S3 subsites (see Fig. 1) (2).

In the data shown in Fig. 2B, it was surprising to find that the presence of His in P2 and Phe in P1’ are predictive. This is in contrast with what was observed when His was in P3 and Trp in P1 (Fig. 2A). The disparity between these two results may reflect differences between substrates that interact with substrate amino acids that span across the scissile bond and those that do not. Each half of the inhibitor forms a short β sheet with the two anti-parallel strands of the flap and residues 27–29 of one subunit (22, 23, 24) (Fig. 5). There is a set of hydrogen bond interactions between the carbonyl oxygens and amides of the inhibitor and the main chain carbonyl oxygens and amides of PR. The β sheets are interrupted near the nonhydrolyzable group of the inhibitors, where there is a kink in the extended conformation of the inhibitor. The interruption in the β sheet near the scissile bond means that the P2-P1’ side chain interactions are not the same as those of P1-P3 or P1’-P3’. In fact, the side chains of P1 and P1’ tend to be directed away from each other so that P1 interacts more closely with P3 than P1’ with P2. Therefore, steric interactions involving P1 and P3 may be more pronounced than those involving P1’ and P2.

**trans Interactions**—Although steric relationships between adjacent cis subsites were expected, it was not clear whether such relationships would also be present for adjacent trans subsites. To examine potential trans interactions, we extended the double substituted peptide study to include substitutions of amino acids adjacent in the linear substrate sequence. The results of substitutions made in P3 and P2 are summarized in Fig. 3A. In contrast to what is observed with the cis interaction data (Fig. 2, A and B), there does not appear to be a significant steric influence with any of these tested substituted substrates (Fig. 3A). The predicted and experimental ΔG values were in close agreement over the entire range of substitutions including the presence of two large residues, Arg in P3 with Leu in P2. Leu in P2 is a strong determinant limiting the amino acids that can be accommodated in P1 or P1’. 
although there appear to be limited interactions between amino acids in P3 and P2, there are significant interactions observed when similar substitutions are placed in the P2 and P1 trans positions. With Trp fixed in P1, Ala substituted for the natural Val in P2 was predictive, whereas the substitution of the larger His or Leu was not (Fig. 3B). The mechanism by which amino acids in the trans configuration interact is not clear. However, it is likely that the presence of a large substrate amino acid in an enzyme subsite will distort the position of the substrate peptide backbone in a way that can be catalytically compensated for by placing a smaller rather than a larger residue in the trans subsite. This can be seen in the structural model shown in Fig. 6. Shown are two peptides in which Leu has been substituted in P2. One of the two peptides also has Trp substituted for Ser in P1. The presence of the Trp in P1 results in substantial movement of the substrate peptide backbone with Leu in P2 being pushed deeper into the S2 subsite resulting in loss of activity. A similar α-carbon backbone distortion could also contribute to the differences in the ΔG values observed with the cis substituted peptides.

Trp in P1 and His or Leu in P2 represent about the largest structural model shown in Fig. 6. Shown are two peptides in which Leu has been substituted in P2. One of the two peptides also has Trp substituted for Ser in P1. The presence of the Trp in P1 results in substantial movement of the substrate peptide backbone with Leu in P2 being pushed deeper into the S2 subsite resulting in loss of activity. A similar α-carbon backbone distortion could also contribute to the differences in the ΔG values observed with the cis substituted peptides.
active as predicted, whereas the placement of a larger Phe in P1 resulted in an efficiency of cleavage that was considerably less than predicted. Similar results are obtained if the bulky group is fixed in P1 and the size of the residue in P1 is varied (Fig. 3C).

As mentioned above, we have detected only two of the peptides so far that were more active than predicted by the single substituted peptides. These are the P1 Trp-P1' Ala and the P1 Leu-P1' Ala peptides. These peptides were predicted to be poor substrates primarily because single P1' Ala substituted peptides are cleaved poorly. The presence of a larger Leu or Trp substituted for Ser in P1, however, seems to restore cleavage of the P1' Ala substituted substrates to levels similar to those observed with the unmodified reference peptide (Fig. 3C). The P1' position has a preference for a large hydrophobic amino acid side chain. Therefore, an Ala residue in this position is too small to fit well into the HIV-1 PR S1' subsite to form the requisite stabilizing van der Waals' interactions with subsite amino acids. The presence of a bulky group in P1 may position the P1' Ala deeper into the S1' subsite allowing it form van der Waals interactions and thereby producing a substrate with more activity than predicted by the single substituted peptides. This interpretation is consistent with the observation that double substituted peptides containing Gly in P1' and Trp or Leu in P1 are also poor substrates. However, in this instance, an adjacent Trp or Leu in P1 does not "rescue" the cleavage defect. A Gly residue cannot provide a side chain for van der Waals interactions, even if the backbone of the peptide is shifted due to the larger residue in P1. Restoration of cleavage of a peptide containing Ala in P1' is predicted to be stronger when a bulky group is in a trans rather than the cis configuration. This is what is observed as shown in Fig. 2B. When Leu is placed into the adjacent cis position, as in the P2 Leu-P1' Ala peptide, there is very little cleavage of this peptide detected. Of the
FIG. 3. Comparison of experimental and predicted ΔΔG values for cleavage by HIV-1 PR of the RSV NC-PR peptide substrates containing double amino acid substitutions in the trans orientation. The calculated and experimentally determined ΔΔG values listed in Table II for cleavage of a peptide substrate representing the RSV NC-PR but containing substitutions in P3 and P2 (A), P2 and P1 (B), and P1 and P1′ (C) were plotted as a function of the amino acids substituted (top of panels) as described in the legend to Fig. 2. ●, predicted ΔΔG values; △, observed ΔΔG values.
three trans interactions examined, P3-P2, P2-P1, and P1-P1', the two involving exclusively the S1, S1', or S2 subsite showed interaction effects, whereas the pair involving S3 and S2 did not. This may reflect the fact that S2, S1, and S1' are internal subsites near the scissile bond, whereas S3 is found near the enzyme surface and therefore able to accommodate a variety of larger amino acids with little steric interaction with the adjacent subsite.

Glycine Substitutions in Peptides—In this study, we have examined 10 double substituted peptides with a Gly substituted in one of the substrate positions. Of these, seven were reasonably predictive and three were not. These data include a P4 substituted Gly with Trp in P1 and Gly substituted in P3, P2, or P1' positions (Table II). The effect of substitution of a Gly residue into a peptide substrate is complicated by the potential loss of a side chain that could form van der Waals interactions with key amino acids in each of the subsites. The data suggest that enzyme subsites unoccupied by substrate side chains may be more sensitive to substitutions in adjacent subsites. Normally, an occupied subsite filled with a side chain could be “buffered” from perturbations in structure by stabilizing van der Waals' interactions between the side chain and the enzyme subsite. This would not be the case with an empty subsite. Thus the single Gly substituted peptides may not be as predictive of the activity on the corresponding double substituted peptides as are single substituted peptides containing side chains that contribute to the binding.

Other Considerations—Substrate steric relationships also apply to the AMV PR. In this case, we have analyzed a small number of double substituted peptides. These data are presented in Table III for P3 and P1 substituted peptides. There are large discrepancies between the predicted and experimentally determined ΔΔG values when large residues are substituted in P3 and P1. Moreover, the larger the size of the substituted residues, the larger the discrepancy in the ΔΔG values. For instance, there is a 0.33 and 1.28 kcal/mol discrepancy for the peptide with Arg in P3 and Gly or Trp, respectively, in P1. When Phe is placed in P3 with the Trp in P1, the discrepancy...
increases to 2.9 kcal/mol.

Of the double substituted peptide set analyzed in this study, about half had experimentally determined ΔG values that were predicted well from the single substituted data. These peptides involve substitutions of at least one small sized residue relative to a given subsite in one of the two substituted substrate positions. In contrast, about 45% of the double substituted peptides were cleaved with catalytic efficiencies that were significantly worse than predicted. These discrepancies can be explained by steric interference. Only two of the peptides tested so far had a higher efficiency than predicted. Thus, substitutions of various amino acids into the different substrate positions limit the number of amino acid combinations that constitute a cleavable site. Although the data set analyzed in this report has focused primarily on nonprime substrate positions, we predict that similar relationships probably exist for the prime side (see Fig. 1). Also, there may be effects of substitutions of amino acid residues at substrate positions 3 or 4 amino acids apart in the linear sequence, on both sides of the scissile bond, that would restrict further the choice of amino acids that could constitute a cleavage site. These latter interactions, if they are important, may be of less magnitude than those involved in adjacent and alternate subsites reported in this study. Taken together, these results indicate that although many different side chains can bind effectively into each enzyme subsite, enzyme specificity is limited by interactions between substrate amino acids bound in both cis and trans positions. Interactions between at least one pair of subsites examined appears to be minimal. Clearly, an understanding of these relationships will be very important to the rational design of HIV-1 PR inhibitors as potential therapeutic agents for AIDS. For a potential compound to bind effectively in the enzyme subsites, it must not violate any of the adjacent occupancy rules defined in this study.

Acknowledgments—We thank Dr. E. Houts, Molecular Genetic Resources, for AMV PR and the Molecular Biology Core facility at Case Western Reserve University for amino acid composition analysis. We also thank Dr. Chou Giam for the HIV-1 PR expression clone and the purification procedure. We also thank Jin Wu for preparation of stereo figures and Eva Majerova and Suzanne Specht for assistance with peptide synthesis.

REFERENCES

1. Skalka, A. M. (1988) Cell 56, 911–913
2. Kotler, M., Katz, R., Danho, W., Leis, J., and Skalka, A. M. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 4185–4189
3. Darke, P., Nutt, R., Brady, S., Garsky, V., Ciccone, T., Lumma, C., Pryor, R., Veber, D., and Sigal, I. (1988) Biochem. Biophys. Res. Commun. 156, 297–303
4. Tozer, J., Frieden, D., Weber, I., Blaha, I., and Orozco-Soto, S. (1993) Biochemistry 32, 3347–3353
5. Cameron, C. E., Grinde, B., Jentoft, J., Leis, J., Wlodawer, A., and Weber, I. (1993) J. Biol. Chem. 268, 11711–11717
6. Cameron, C. E., Ridky, T., Shulein, S., Leis, J., Weber, I., Copeland, T., Wlodawer, A., Buxton, B., Bibicki, B., and Skalka, A. M. (1994) J. Biol. Chem. 269, 11170–11177
7. Huang, W., and Tang J. (1996) J. Biol. Chem. 271, 5085–5091
8. Cameron, C. E., Grinde, B., Jentoft, J., Leis, J., Weber, I., Copeland, T., and Wlodawer, A. (1992) J. Biol. Chem. 267, 27375–27371
9. Grinde, B., Cameron, C. E., Leis, J., Weber, I., Wlodawer, A., Buxton, B., Bibicki, B., and Skalka, A. M. (1992) J. Biol. Chem. 267, 9481–9490
10. Grinde, B., Cameron, C. E., Leis, J., Weber, I., Wlodawer, A., Buxton, B., and Skalka, A. M. (1992) J. Biol. Chem. 267, 9491–9498
11. Alexander, F., Leis, J., Saitis, D., Crow, R., Danho, W., Poonian, M., Pan, Y.-C., and Skalka, A. M. (1987) J. Virol. 61, 534–542
12. Fersht, A. R. (1985) Enzyme Structure and Mechanism, 2nd Ed., pp. 350, W. Freeman and Co., New York
13. Swain, A. L., Miller, M., Green, J., Rich, D. H., Schneider, J., Kent, S., and Wlodawer, A. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 8805–8809
14. Sack, S. I. (1988) J. Med. Gen. 6, 224–229
15. Harrison, R. W. (1993) Comp. Chem. 14, 1112–1122
16. Rapp, A. K., Case, R. J., Colwell, K. S., Goddard, W. A., III, and Skiff, W. M. (1992) J. Am. Chem. Soc. 114, 10092–10095
17. Jaskelioff, M., Tomasselli, A., Sawyer, T., Staples, D., Heinrickson, R., Schneider, J., Kent, S., and Wlodawer, A. (1992) Biochemistry 30, 1600–1609
18. Weiner, S. J., Kohli, P. A., Nguyen, D. T., and Case, D. A. (1995) J. Comp. Chem. 7, 230–238
19. Harrison, R. W., Chatterjee, D., and Weber, I. T. (1995) Proteins Struct. Funct. Genet. 23, 463–471
20. Harrison, R. W., and Weber, I. T. (1994) Protein Eng. 7, 1353–1363
21. Wlodawer, A., and Erickson, J. W. (1993) Annu. Rev. Biochem. 62, 543–585
22. Gustchina, A., and Weber, I. (1990) FEBS Lett. 283, 180–194
23. Gustchina, A., Sansom, C., Prevost, M., Richelle, J., Wodak, S. Y., Wlodawer, A., and Weber, I. T. (1994) Protein Eng. 7, 309–317
24. Wlodawer, A., Machin, J., Gilliland, G. L., Gallagher, W., and Woodward, C. (1987) J. Mol. Biol. 196, 469–480
25. Zegers, I., Maes, D., Dao-Thi, M.-H., Poortmans, F., Palmer, R., and Wyns, L. (1994) Protein Sci. 3, 2322–2339
Human Immunodeficiency Virus, Type 1 Protease Substrate Specificity Is Limited by Interactions between Substrate Amino Acids Bound in Adjacent Enzyme Subsites
Todd W. Ridky, Craig E. Cameron, John Cameron, Jonathan Leis, Terry Copeland, Alexander Wlodawer, Irene T. Weber and Robert W. Harrison

J. Biol. Chem. 1996, 271:4709-4717.
doi: 10.1074/jbc.271.9.4709

Access the most updated version of this article at http://www.jbc.org/content/271/9/4709

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 24 references, 9 of which can be accessed free at http://www.jbc.org/content/271/9/4709.full.html#ref-list-1