THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 267, No. 14, Issue of May 15, pp. 9491-9498, 1992
Printed in U.S.A.

Analysis of Substrate Interactions of the Rous Sarcoma Virus Wild Type and Mutant Proteases and Human Immunodeficiency Virus-1 Protease Using a Set of Systematically Altered Peptide Substrates*

(Received for publication, November 4, 1991)

Bjorn Grinde‡, Craig E. Cameron§, and Jonathan Leis¶

From the School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106

Irene T. Weber‖ and Alexander Wlodawer

From the National Cancer Institute-Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, Maryland 21702

Haim Burstein and Anna Marie Skalka

From the Fox Chase Cancer Center, Institute for Cancer Research, Philadelphia, Pennsylvania 19111

In the preceding study, mutant Rous sarcoma virus (RSV) proteases are described in which three amino acids found in the human immunodeficiency virus-1 (HIV-1) protease (PR) were substituted into structurally comparable positions (Grinde, B., Cameron, C. E., Leis, J., Weber, I., Wlodawer, A., Burstein, H., Bizub, D., and Skalka, A. M. (1992) *J. Biol. Chem.* 267, 9481-9490). In this report, the activity of the wild type and these mutant PRs are compared using a set of RSV NC-PR peptide substrates with single amino acid substitutions in each of the P4 to P3′ positions. With most substrates, the relative activities of the two active mutants followed that of the RSV PR. Substitutions in the P1 and P1′ positions were an exception; in this case, the mutants behaved more like the HIV-1 PR. These results confirm predictions from structural analyses which indicate that residues 105 and 106 of the RSV PR are important in forming the S1 and S1′ binding subsites. These results, further analyzed with the aid of computer modeling of the RSV PR with different substrates, provide an explanation for why only partial HIV-1 PR-like behavior was introduced into the above RSV PR mutants.

Crystal structures of two retroviral proteases (PRs)[1] are currently available. These include the native structure of the RSV PR, refined at 2-Å resolution (1), and three separate structures of the HIV-1 PR, refined at 2.7-3-Å resolution. In addition, several structures of HIV-1 PR with various substrate-based inhibitors bound in the active site are available (2-9). In the present study, crystallographic information has been extended by molecular modeling to predict the structure of the RSV PR with bound substrate in order to investigate the molecular basis of specificity.

Peptides that contain sequences corresponding to known target sites on retroviral precursor polyproteins can be cleaved precisely by these retroviral PRs *in vitro*. The main determinants of specificity reside in 7 amino acids of the substrates (10, 11) extending from P4-P3′ (notation according to Schechter and Berger (12)). Despite wide differences in preference for substrate, a detailed comparison of the structural data and sequence alignments has indicated that most HIV-1 PR amino acid residues in contact with substrate-based peptide inhibitors are remarkably conserved in the RSV enzyme. Only 3 residues, at RSV PR positions 65, 105, and 106, are significantly different (13). In order to test the significance of these differences, mutant RSV PRs were constructed where these residues in the RSV PR binding site were changed to those of the HIV-1 PR. Since the PR is a homodimer, mutations introduced into the monomeric subunit are *de facto* double mutations in the native PR. Analysis of two of these mutants, R105*P*,G106*V* and H65*G*,R105*P*,G106*V*, indicated that both enzymes showed a preference for a peptide substrate based on the reverse transcriptase-integrase cleavage site in the HIV-1 precursor protein (13). In this report, we present a more detailed analysis of the differences in substrate interactions caused by these mutations. A peptide substrate based on the NC-PR cleavage site in the RSV polyprotein was used for this purpose. This substrate was cleaved with similar efficiencies by both the AMV and HIV-1 wild type PRs as well as the mutant RSV PRs (13). The activity of the two mutant and two wild type PRs were examined using 40 different peptide substrates with single amino acid changes placed in the P4 through the P3′ positions. Results from this analysis show that amino acid substitutions in the PR at positions 105 and 106, which are primarily part of the S1 and S1′ enzyme subsites, cause changes which affect almost exclusively the recognition of the P1 and P1′ substrate positions.

* This work was supported in part by United States Public Health Service Grants CA38046 (to J. L.), CA06927, CA47486, and RR05539, an appropriation from the Commonwealth of Pennsylvania (to A. M. S.), and National Cancer Institute Contract N01 C01 74101 with ABL (to 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.

‡ Recipient of a fellowship from the American Cancer Society, Ohio Division, and support from the International Union Against Cancer and the Norwegian National Research Council. Present address: Dept. of Virology, National Institute of Public Health, Oslo, Norway.

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

¶ To whom correspondence and reprint requests should be addressed. Tel.: 216-368-3360.

‖ Present address: Jefferson Cancer Institute, Thomas Jefferson University, Philadelphia, PA 19107.

[1] The abbreviations used are: PR, protease; RSV, Rous sarcoma virus; AMV, avian myeloblastosis virus; HIV, human immunodeficiency virus; NC, nucleocapsid. Amino acids and numbers written in *italics* refer to the HIV-1 PR.

EXPERIMENTAL PROCEDURES

Bacterial Cells and DNA Constructs—Bacterial growth conditions, the expression vector used for mutagenesis, pPR, and preparation of purified RSV PRs containing the H65*G* and R105*P*,G106*V* and R105*P*,G106*V*,H65*G* substitutions from an inclusion body fraction derived from *Escherichia coli*, were as described previously (13–15). Preparation of HIV-1 PR, also from *E. coli*, and avian myeloblastosis virus (AMV) PR, from virions, has been described (16, 17). The RSV and AMV PRs, which differ by 2 amino acids in primary structure and are enzymatically indistinguishable, have been used interchangeably in these studies.

Peptides—Peptides were synthesized and purified to at least 95% purity by Advanced ChemTech (Louisville, KY). Peptides were characterized as described by Grinde *et al.* (13). All of these peptides were readily soluble in water.

Proteolytic Assay and Kinetic Analysis—The PR assay and kinetic analysis were as previously described (13–15).

Molecular Modeling—Atomic coordinates of RSV PR (1, 18) and HIV-1 PR with inhibitor (6) were examined on an Evans and Sutherland PS390 computer graphics system using the program FRODO (19).

RESULTS

Molecular Modeling of the Flaps of the RSV PR—Several residues, including N61 to G69, from one subunit and A59′ to G69′ from the other subunit in each flap of the RSV PR dimer are not visible in the crystal structure and are considered to be disordered. (Note that the prime notation in this context is used to denote the second subunit of the PR dimer.) Since the flap region is presumed to be very important for substrate binding and catalysis, and crystal structures of the RSV PR complexed to ligand are not available, we have modeled the structure of the RSV PR flaps based on the co-crystal structure of HIV-1 PR with the inhibitor MVT101 (6). The C_α atoms of the RSV PR dimer were superimposed on those of HIV-1 PR with inhibitor as described by Weber (20). Subsequently, the aligned structures and sequences (13) were examined in the flap region, and the missing residues were modeled (Fig. 1). There are 3 additional residues to accommodate in the flap of RSV PR compared to HIV-1 PR. In the structures of HIV-1 PR with ligands, the ends of the two flaps overlap and are linked by hydrogen bonds. In addition, the ends of both flaps interact with the inhibitor. It thus appeared unlikely that the extreme ends of the flaps could be extended by the 3 extra residues in the RSV PR. Also, a sequence alignment of the two enzymes showed many identical residues for the overlapping ends of the flaps (from I64 to I71, on either side of the centrally placed I50). Therefore, RSV PR sequences IHGIGGGI were modeled in the same positions as HIV-1 PR *IGGIGGFI* (Fig. 2). This included several conserved glycine and isoleucine residues, the straightforward removal of the *phenylalanine* side chain of residue *53* to form G70, and the change of *G48* to H65. The side chain of H65 was modeled as lying on the surface of the flap and extended across to the other strand of the antiparallel β pair where it may form a hydrogen bond with the carbonyl oxygen of glycine 69. H65 may stabilize the flap by interacting with G69 or it may interact with a substrate.

The additional 3 residues, N61 to Q63, were placed in a small surface loop connecting A60 and I64 in a continuous chain in which bond lengths were idealized. This leads into the established structural parts of the RSV PR flaps around residues 59 and 60. Comparison of the subunit structures of HIV and RSV PRs showed that small extensions of surface turns are common in RSV PR, while HIV PR has a more compact structure. One example of such a small insertion is provided by residues 7 and 8 of RSV PR, which extend as a surface loop. It is not clear why this enzyme has such extended surface features, although removal of several of the larger surface loops has been shown to drastically reduce the catalytic activity of the PR (13). The region leading to the flap, from residue 46, differs in conformation between RSV and HIV-1 PRs. Again, the RSV PR has two additional small surface loops compared to the same region in HIV-1 PR.

The flap was first modeled in one subunit and then a copy was rotated into the corresponding position in the other subunit of the dimer so that it superimposed on the HIV-1 PR second flap. The RSV PR dimer with modeled flaps is shown in Fig. 1. The modeled flap compared with the flap of the crystal structure of the HIV-1 PR with inhibitor is shown in Fig. 2. Once the flaps were built, it was possible to place a peptide substrate in a position similar to that of the HIV PR substrate-based inhibitors. Only the amino acid side chains were altered to model most of the substrate. It was not obvious how to position the scissile peptide bond between P1 and P1′ since the inhibitors have a reduced bond or a hydroxyl (representing a transition state analog) instead of the carbonyl oxygen lying near the catalytic aspartate residues. It appeared that a peptide bond between P1 and P1′ would be under strain due to the conformation of the substrate from P3 to P1 and P1′ to P3′. Therefore, no attempt was made to position this carbonyl oxygen accurately. The conformation of the main chain of the substrate was not changed, but the side chain conformations were adjusted when necessary to fit into the appropriate subsite. The model of the substrate is shown in Fig. 3.

RSV PR Subsites Deduced by Modeling the Flaps and Substrate—In HIV-1 PR, the inhibitors are bound by a series of hydrogen bond interactions. There is a set of conserved interactions between PR and main chain C=O and NH groups of the peptide substrate found in all available crystal structures (4–8). Also, each substrate side chain lies in a separate subsite or pocket in the PR dimer. In some cases, substrate side chains may participate in hydrogen bond or ionic interactions with amino acid residues lining the PR subsites. Equivalent RSV PR residues form the subsites and have the potential for hydrogen bond interactions with the main chain of the substrate or inhibitor. The interactions with the main chain peptide groups are independent of the substrate sequence, and energy calculations suggest that these contribute about 58% of the total substrate binding energy.[2]

A complete set of main chain hydrogen bond interactions is presented in Table I. These interactions are equivalent to those predicted for HIV-1 PR with bound substrate (21). In addition, each substrate side chain provides sequence-dependent interactions within individual subsites. These are mainly hydrophobic interactions, although in some cases hydrogen bonds or ionic interactions may be possible, as seen in the different co-crystal structures of HIV-1 PR with inhibitors. A list of amino acids predicted to form each RSV PR subsite is presented in Table II. Several residues contribute to more than one subsite and there is also some symmetry seen in the subsites on either side of the scissile bond. For example, subsites involving P2 and P2′ are similar. The residues interacting with a substrate side chain will depend upon the size of the side chain, since a longer side chain can contact subsite residues further away.

The overall features of each of the enzyme subsites predicted from the model are as follows. Subsite S4 is on the protein surface such that a variety of residues might be accommodated in the P4 position of substrate. Subsite S3 is polar, formed partially by the charged residues, R10′ and R105′. P2 is found in a smaller hydrophobic subsite while subsite S1 can accommodate large hydrophobic amino acid

[2] C. Sansom and I. Weber, unpublished data.

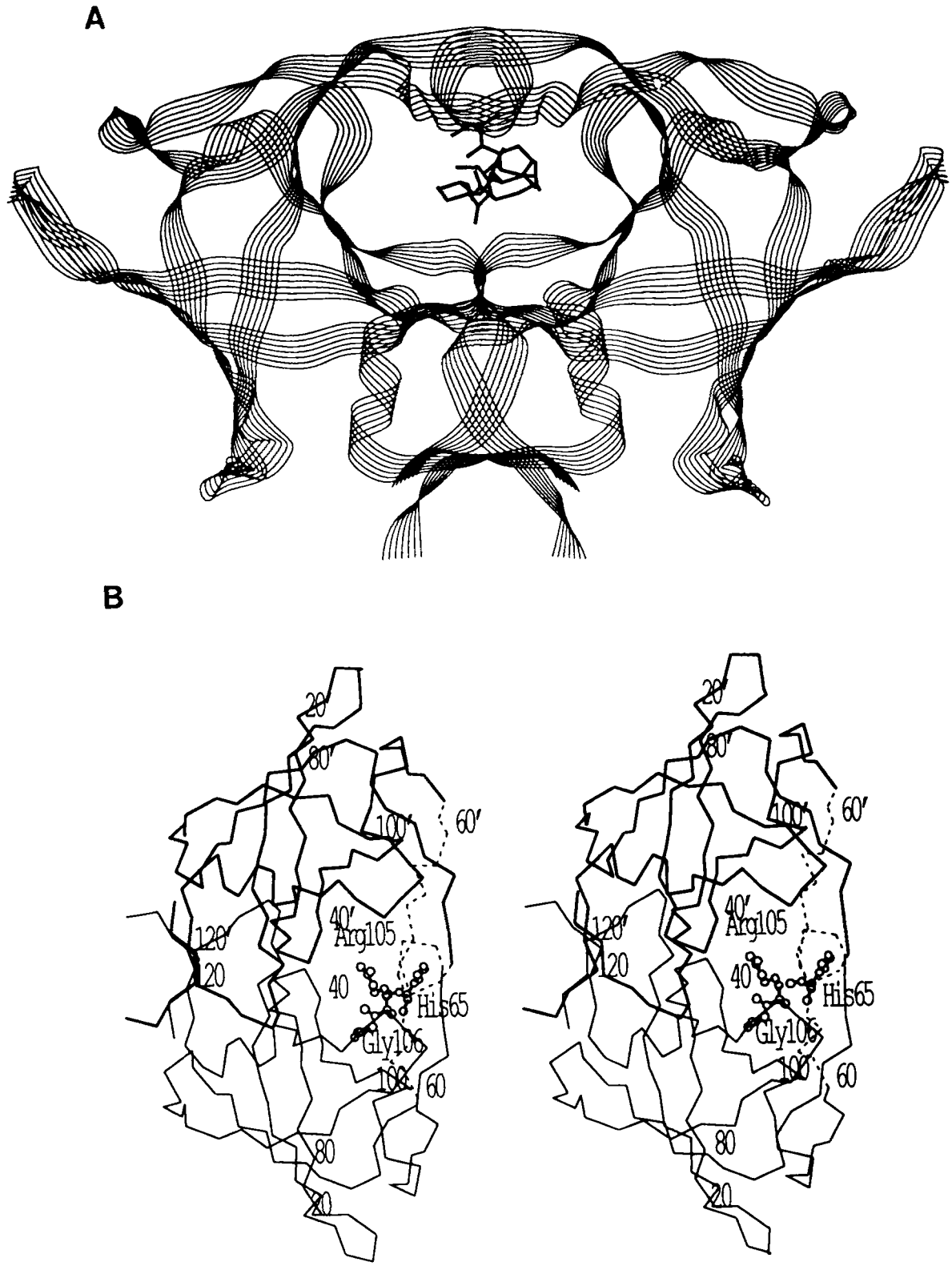

FIG. 1. **Structure of the RSV PR with substrate.** *A*, ribbon drawing of the polypeptide backbone of the model for RSV PR with substrate in the *center* in a line drawing. *B*, the position of the mutations in relation to the modeled flaps. Stereo view of the C_α backbone of the crystal structure of the RSV PR dimer with the two subunits in *thick* and *thin lines*. The C_α atoms of the modeled flaps are shown in *dashed lines*, residues 61 to 69 in one subunit and 59 to 69 in the other. The PR is labeled every 20 residues, and a *prime* indicates residues in the second subunit. The position of the mutations is indicated by the residue present in the wild type: H65, R105, and G106 in a ball and stick representation in one subunit only for clarity.

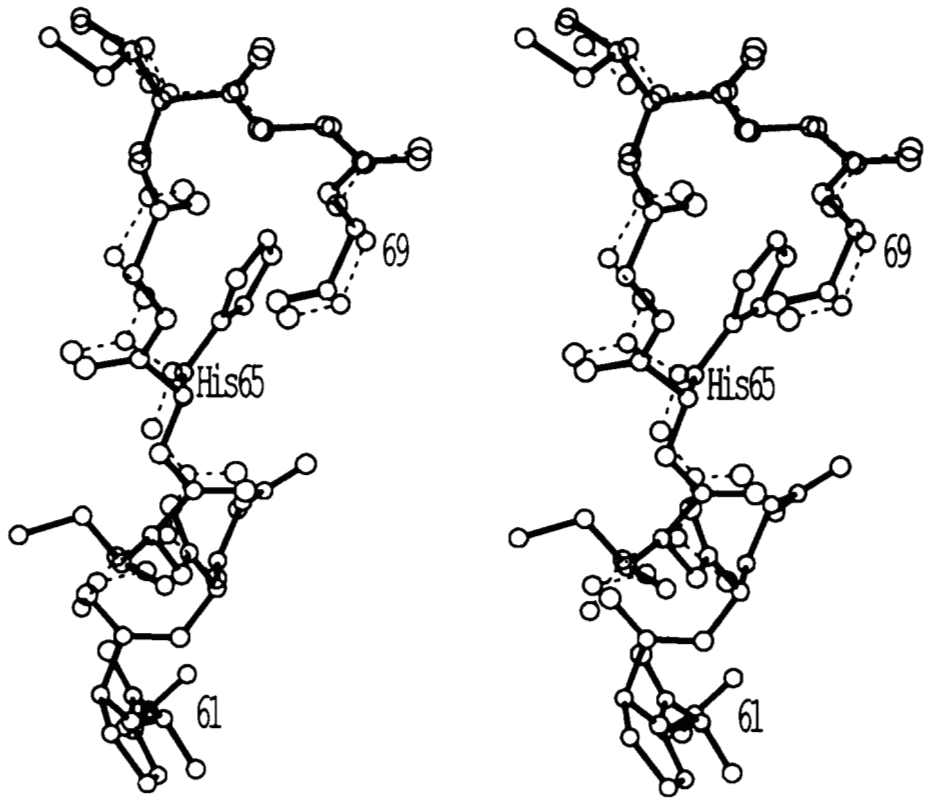

FIG. 2. **Modeled flaps of the RSV PR.** Stereo view of the atoms in the modeled portion, residues 61 to 69, of one flap of RSV PR (*thick lines*) compared to the structurally equivalent part of the HIV-1 PR flap from the crystal structure with inhibitor (*thin* or *dashed lines*). H65 is indicated and the side chain was modeled as interacting with residue 69 on the other strand of the flap.

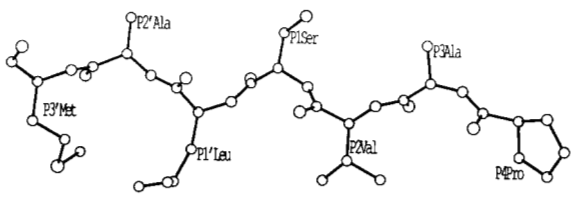

FIG. 3. **Modeled substrate.** Substrate residues P4 to P3′ from the model based on the co-crystal structures of HIV-1 PR with inhibitors.

TABLE I

Potential hydrogen bond interactions between substrate and protein

Substrate position	Hydrogen bond	PR residue[a]
P4	C=O ··· NH	**H65**
P3	NH ··· O	D41
	C=O ··· NH	D41
P2	NH ··· O=C	**H65**
	C=O ·· H_2O ·· NH	**I67**
P1	NH ··· O=C	G39
	C=O ··· O	D37
P1′	NH ··· O	D37′
	C=O ·· H_2O ·· NH	**I67′**
P2′	NH ··· O=C	G39′
	C=O ··· NH	D41′
P3′	NH ··· O=C	**H65′**
	C=O ··· NH	**H65′**

[a] Amino acids in **bold** are in the flaps. Prime indicates amino acids found in the second subunit.

residues. Subsite S1′ has a relatively medium-sized, hydrophobic pocket. In fact, the natural substrates do not show a residue larger than leucine or histidine at P1′. Subsite S2′, like S2, is smaller and hydrophobic. Subsite S3′ lies near the dimer surface, and the charged residues D41, H65, R10, and R105 contribute to it.

TABLE II

Amino acids in the substrate binding pocket of the RSV PR in proximity to the substrate

Subsite	Amino acid[a]
S4	**R10′, D41, I42, P62, Q63, R105′**, H7′, I64, M73
S3	**R10′, D41, H65, R105′**, L35′, G66, I67, V104′, G106, I108′
S2	**A40, I44, I64, H65, G66, I67′**, I42, M73, I108, G110
S1	**L35′, D37′, G39, G66**, H65, I67, V104′, R105′, G106′, S107′, I108′
S1′	**D37, G39′, G66′, I108**, R10, L35, D37′, H65′, I67, I67′, V104, R105, G106, S107
S2′	**G39′, A40′, I44′, I64′, H65′, I67**, D41′, I42′, I71′, M73′, V104′, I108′
S3′	**D41′, I64′, H65′, R105**, R10, V104, R111′

[a] Residues with atoms within 5 Å of the C_α or C_β of the substrate are in **bold type**. Residues with atoms between 5 and 10 Å from the C_α of substrate are in normal type. These residues were chosen since they are on the side of the side chain rather than nearer the main chain of the substrate. They are also predicted to form the distal portions of the various subsites. Underlined residues indicate that they are in the flaps. The prime denotes that the amino acid is derived from the second subunit.

Using the above model, we have noted that RSV PR residues 105 and 106 and the identical residues 105′ and 106′ in the second subunit of the homodimer form the roof of subsites S1 and S1′, but are also near subsites S3 and S3′. His-65 and -65′, which are in the flaps, are predicted to be associated with several subsites (Table II). These residues are also involved in hydrogen bonding to the P4, P3′, and P2 backbone carbonyl or amide groups in the substrate (Table I).

Strategy—In a previous investigation (13), a series of mutant RSV PRs was constructed that substituted the amino acid found in the HIV-1 PR into the analogous location in the RSV PR (positions 65, 105, and 106). Analysis of these mutant PRs (R105*P*,G106*V* and H65*G*,R105*P*,G106*V*) indi-

cated that they had a 10- to 20-fold increase in activity toward a substrate based on an HIV-1 cleavage site compared to the RSV wild type PR. Therefore, an effort was undertaken to investigate the change in substrate recognition caused by these mutations in more detail. To do so, a set of 40 peptide substrates was synthesized, each identical with the NC-PR sequence except for single amino acid substitutions in one of the positions from P4 to P3′ which interact with the 7 enzyme subsites (S4–S3′) in the substrate binding pocket. The choice of amino acid substituted into the NC-PR peptide was based on the above model and was designed to challenge predicted subsites with amino acids that were different from those found in the wild type substrate.

Activity of Wild Type PRs on Modified NC-PR Substrates—The specific activities of the AMV and HIV-1 PRs with these 40 peptides are summarized in Tables III and IV. An examination of these activity data indicates that the substrate binding pockets of the two PRs can be distinguished by this set of peptides. This is observed with amino acid substitutions placed in the P4, P3, P2, and P1 positions (Table III). Most substitutions placed in P4 and P2 positions were inactivating for the AMV PR while they had little effect on the HIV-1 PR. A similar observation was made with substitutions at the P3 position, except that some of these amino acid substitutions resulted in a 2.5- to 3.5-fold stimulation in activity with the AMV enzyme while having little effect on the HIV-1 PR (Table III). The increase in activity observed with P3 substitutions for the AMV PR was the result of increases in both affinity for binding substrate and catalytic rate (Table V).

Most amino acid replacements at the P1 position increased activity for the NC-PR substrate for both the AMV and HIV-1 PRs. In particular, when leucine was substituted at P1, there was an 11- and 6-fold increase in activity for the HIV-1 PR and AMV PR, respectively (Table III). When tryptophan was substituted at the same position, there was a 3- and 7-fold increase in activity for the HIV-1 PR and AMV PR, respectively. Thus, HIV-1 PR prefers leucine at this position, while AMV PR prefers tryptophan. The increase in activity with tryptophan at P1 is due to an increase in affinity for substrate and an increase in catalytic rate for both enzymes (Table V). Increases in activity and binding affinity are also observed for a peptide with leucine substituted at P1 for the AMV PR. There did not appear to be a significant change in K_m for the leucine-substituted substrate for HIV-1 PR. Instead, there appears to be a substantial increase in catalytic rate (Table V). Substitutions at P1′ to P3′ generally produced substrates with lower activity for both enzymes (Table IV). However, here again differences in activity between the two PRs can be noted with substitution of Phe at P1′ or tyrosine at P3′.

Activity of Mutant PRs on Modified NC-PR Substrates—

TABLE III

PR activities with NC-PR peptide substrates substituted at the P4 to P1 positions

Substitution[a]	Protease activity[b]			
	AMV	RSV (R105*P*,G106*V*)	RSV (H65*G*,R105*P*, G106*V*)	HIV-1
None	13.4	13.1	3.2	37.0
P4 (Pro)				
His	10.1	12.5	2.3	42.3
Asn	7.8	8.9	1.0	47.1
Gly	1.3	1.2	0.3	41.8
Leu	0.8	0.7	0.0	3.9
Phe	0.5	0.6	0.0	3.7
P3 (Ala)				
Arg	46.6	34.1	28.0	37.7
Phe	37.7	23.2	25.0	32.1
Asn	34.5	32.3	10.1	34.3
His	30.7	27.6	8.0	32.4
Asp	15.2	13.1	2.3	33.8
Gly	9.8	9.6	2.6	29.7
Pro	6.6	5.8	1.6	33.8
P2 (Val)				
Leu	1.3	0.6	0.3	30.6
Gly	1.2	0.8	0.1	31.4
His	1.2	0.8	0.3	3.6
Ser	1.1	0.7	0.1	29.3
Trp	0.9	1.1	0.5	12.5
P1 (Ser)				
Trp	96.5	30.6	68.0	104.0
Leu	73.9	75.7	157.2	421.2
Arg	48.4	49.6	19.3	41.3
Glu	41.3	46.2	14.8	40.2
His	32.1	69.0	44.5	49.7
Ala	23.5	20.6	5.6	51.0
Gly	2.1	4.1	1.3	25.1

[a] The NC-PR substrate (PPAVSLAMTMRR) was modified with single-amino acid substitutions in the P4 to P1 positions as indicated. The amino acid in the wild type substrate for each position is indicated in the parentheses.

[b] Activity with the purified AMV PR, HIV-1 PR, and RSV mutant PR containing the R105*P*,G106*V* and the H65*G*,R105*P*,G106*V* substitutions was assayed as described under "Experimental Procedures." The activity data are expressed as the initial rate of turnover in the presence of 100 μM substrate and measured as molecules of product formed per active enzyme dimer per min.

TABLE IV

PR activities with NC-PR peptide substrates substituted at the P1′ to P3′ positions

Substitution[a]	Protease activity[b]			
	AMV	RSV (R105*P*,G106*V*)	RSV (H65*G*,R105*P*, G106*V*)	HIV-1
None	13.4	13.1	3.2	37.0
P1′ (Leu)				
Arg	8.8	0.1	0.0	5.5
Phe	5.6	4.8	1.8	24.0
Gln	3.1	0.0	0.0	3.9
Glu	2.7	0.2	0.1	1.1
Gly	0.1	0.3	0.2	0.3
P2′ (Ala)				
Leu	1.0	0.7	0.6	11.4
Ser	0.9	1.0	0.4	9.1
Gly	0.6	1.0	0.5	8.6
His	0.4	0.4	0.5	2.3
Trp	0.0	0.0	0.0	0.0
P3′ (Met)				
Tyr	27.1	22.0	5.4	6.6
Leu	4.5	4.9	1.2	14.5
Asn	1.5	0.5	0.0	0.9
Gly	0.9	0.1	0.0	0.0
Asp	0.8	0.0	0.0	0.0

[a] The NC-PR substrate (PPAVSLAMTMRR) was modified with single-amino acid substitutions in the P1′ to P3′ positions as indicated. The amino acids in the wild type substrate are indicated in the parentheses.

[b] Activity with the purified AMV PR, HIV-1 PR, and RSV mutant PR containing the R105*P*,G106*V* and the H65*G*, *R*105*P*,G106*V* substitutions were assayed as described under "Experimental Procedures." The activity data are expressed as the initial rate of turnover in the presence of 100 μM substrate and measured as molecules of product formed per active enzyme dimer per min.

TABLE V
Kinetic data for NC-PR peptide substrates substituted at the P3 and P1 positions

Substitution[a]	Protease[b]			
	AMV	RSV (R105*P*,G106*V*)	RSV (H65*G*,R105*P*, G106*V*)	HIV-1
None				
K_m	1.0	1.0	1.0	1.0
k_{cat}	1.0	1.0	1.0	1.0
k_{cat}/K_m	1.0	1.0	1.0	1.0
P3 Arg				
K_m	0.52	0.64	0.76	0.63
k_{cat}	2.6	2.2	6.9	1.0
k_{cat}/K_m	5.0	3.3	9.1	1.6
P3 Phe				
K_m	0.46	0.32	0.45	0.66
k_{cat}	2.2	1.3	5.7	0.8
k_{cat}/K_m	4.7	4.0	12.8	1.2
P1 Trp				
K_m	0.23	0.19	0.27	0.33
k_{cat}	5.1	1.8	13.9	2.7
k_{cat}/K_m	21.8	9.4	51.5	8.1
P1 Leu				
K_m	0.42	0.22	0.3	1.1
k_{cat}	4.5	4.7	37.4	12.8
k_{cat}/K_m	10.4	21.6	124.0	11.5
P1 His				
K_m	0.64	0.63	0.71	0.84
k_{cat}	2.0	4.3	10.8	1.4
k_{cat}/K_m	3.0	6.8	15.1	1.6
P1 Gly				
K_m	0.67	1.0	1.1	1.18
k_{cat}	0.12	0.24	0.32	0.68
k_{cat}/K_m	0.2	0.25	0.25	0.57

[a] The NC-PR substrate as described in the legend to Table II was modified at the P1 and P3 positions as indicated.

[b] The activity of the wild type AMV and HIV-1 PRs and the RSV mutants was assayed, and K_m, k_{cat}, and k_{cat}/K_m values were calculated as described under "Experimental Procedures." The K_m values for the RSV, RSV R105*P*,G106*V*, RSV H65*G*,R105*P*,G106*V*, and HIV-1 PRs were 48.5, 45.4, 76.8, and 15.8 μM, respectively, for the NC-PR peptide. The k_{cat} for the RSV, RSV R105*P*,G106*V*, RSV H65*G*,R105*P*,G106*V*, and HIV-1 PRs were 21.6, 19.9, 6.1, and 43.8 min^{-1}, respectively, for the NC-PR peptide. The k_{cat}/K_m ratio for the RSV, RSV R105*P*,G106*V*, RSV H65*G*,R105*P*,G106*V*, and HIV-1 PRs were 0.45, 0.44, 0.08, and 2.77 μM^{-1} min^{-1}, respectively, for the NC-PR peptide. The kinetic data for substituted peptides are presented as the relative amount compared to that of the unmodified substrate.

The activity of the R105*P*,G106*V* and the H65*G*,R105*P*, G106*V* RSV PR mutants with the same set of modified NC-PR peptides is summarized in Tables III and IV. In P4 to P2 and P2′ to P3′ substitutions, where the activity of the AMV enzyme was different from that of the HIV-1 PR, the activity of the two mutants followed that of the wild-type AMV PR (Tables III and IV). For example, substitution of histidine, asparagine, and glycine, respectively, in the P4 position results in a progressive loss in activity for this substrate for the wild type AMV PR and for both mutants (Table III). In contrast, these same substitutions increased the activity for this substrate by the HIV-1 PR. Similar conclusions, that the mutants PRs behave like the wild type AMV enzyme, can be drawn with arginine, phenylalanine, asparagine, histidine, and proline substitutions at P3; leucine, glycine, and tryptophan substitutions at P2; leucine, serine, and glycine substitutions at P2′; and for tyrosine substitution at P3′. The other amino acid replacements described in Tables III and IV resulted in little or no discrimination between the two PRs.

Both mutant PRs behaved more like the HIV-1 PR when specific substitutions were placed at either the P1 or P1′ positions. This can be seen with P1′ substitutions of arginine, phenylalanine, and glutamine (Table IV) which resulted in a PR substrate that was less active than wild type. However, arginine substitution at P1′ produced a substrate that had 66% and 15% the activity for the AMV and HIV-1 PRs, respectively. The mutant enzymes resembled the HIV-1 PR in that they had very little detectable activity on this substrate (Table IV). Substitution of tryptophan and leucine at P1 produced the two most active substrates for both wild type PRs. Like the HIV-1 PR, both mutants preferred leucine rather than tryptophan at this position (Table III).

Substitution of histidine, glutamate, or arginine at P1 produced substrates that were more active than wild type but less active than with either the tryptophan or leucine substitutions. Stimulation of activity with the AMV PR was observed to be 2- to 3-fold greater than with the HIV-1 PR. The mutant enzymes showed substantial activity with these substrates, making them appear in these cases more like the AMV PR. However, a kinetic analysis of the histidine substitution at P1 in the substrate indicated that the increase in activity observed with the two mutants was due primarily to an increase in k_{cat}, with very little effect on K_m (Table IV). Also, the increase in k_{cat} was 2- to 5-fold greater than observed with the AMV PR. With all of the modified peptides analyzed so far, changes in activity with the AMV PR have always resulted in alterations in both binding affinity and catalytic rate. The finding that there is little change in the binding affinity for the histidine-substituted substrate concomitant with a large change in catalytic rate suggests that the mutant enzymes more closely resemble the HIV-1 PR rather than the AMV PR in their recognition of P1. As described above, leucine substitution in P1 resulted in a very large increase in k_{cat} without much effect on the binding of substrate by the HIV-1 PR. It should also be pointed out that the active sites in both mutants are hybrids derived from the parent enzymes and this may result in unique behavior with the various substrates.

Comparison of Activities of the Double and Triple Mutant—Introducing the *glycine* at 65 with the *proline* at 105 and *valine* at 106 has a synergistic effect on activity with some but not all P1-substituted peptides. This is seen directly with the tryptophan- and leucine-substituted peptides where the activity of the triple mutant is twice that of the sum of the activity of the double mutant and the H65*G* mutant alone (Table III). These data are even more striking when one takes into account that the H65*G*,R105*P*,G106*V* mutant is actually 4 times less active than the R105*P*,G106*V* mutant on the unmodified NC-PR substrate. Relative to the activity observed with the wild type NC-PR substrate, the triple mutant also has more activity than expected for other amino acid substitutions in P1. This is also true for at least 2 amino acids substituted in P3 positions (Table III).

A difference between the double- and triple-substituted mutants can also be seen by comparison of the kinetic data in Table V. There are very large increases in k_{cat} for cleavage of the P1-substituted peptides compared to wild type substrates for the H65*G*,R105*P*,G106*V* mutant that are not apparent for the R105*P*,G106*V* mutant. As argued above, the large effects on k_{cat} are associated with the S1 and S1′ subsites of the HIV-1 PR.

DISCUSSION

A complete molecular model of the RSV PR dimer with substrate bound in its active site has been developed. This was accomplished by combining the crystal structure of the RSV PR with modeled flaps and a substrate based on equivalent structures of the HIV-1 PR complexed with inhibitor. The positions of the parts of the flaps observed in the crystal structure of the RSV PR are more similar to the equivalent parts of the HIV PR with inhibitor than they are to the structure of the unliganded HIV PR. This suggested that the RSV PR crystal structure is closer to the conformation with substrate bound than the corresponding structure of the HIV-1 PR. This model was analyzed to predict both the amino acid residues forming the enzyme subsites that bind the 7 substrate residues required for cleavage and the potential hydrogen bond interactions between the peptide substrate and enzyme. Similar analyses of substrate or inhibitor binding using molecular models of HIV-2 PR have been described (21, 22).

In order to test the predictions of RSV PR subsites, a series of NC-PR peptide substrates were synthesized with single amino acid substitutions in the P4 to the P3′ positions. The substituted amino acids were chosen to challenge individual subsites with amino acids that vary in size, charge, and hydrophobicity. Glycine, for instance, which has no side chain, was substituted at each position to test the importance of the side chain interactions in general. It has previously been noted that there is a structural symmetry between subsites S3 and S3′, S2 and S2′, and S1 and S1′. Using the substituted NC-PR peptide substrates, it has now been shown that these paired enzyme subsites are functionally asymmetric. This is demonstrated by the substitution of asparagine in P3 and P3′ in the NC-PR substrate, which results in a 2.5-fold increase and a 9-fold decrease in activity, respectively, for the wild type AMV PR. Other examples of asymmetry can be found in Tables III and IV as well. In general, substitutions of amino acids in the P1′, P2′, and P3′ positions resulted in loss of activity for both enzymes with these substrates. The only exception is provided by tyrosine in P3′, which stimulates activity by the AMV PR.

Distinct differences can be seen not only with the structurally paired subsites, but between comparable subsites from the RSV and HIV-1 PRs. The set of 40 modified RSV NC-PR substrates distinguishes the substrate binding pockets of the two PRs. This is most clearly seen in the data presented in Table III. Substitutions of amino acids in P4 and P3 have little effect on HIV-1 PR activity. In contrast, substitutions of amino acids in P4 causes significant losses in activity for the AMV PR while substitutions in P3 result in a stimulation in activity for the same enzyme.

A comparison of the substrate binding pockets of RSV and HIV-1 PRs indicates that 58% of the amino acids are conserved. This is much higher than the 30% identity of residues found when comparing the total structures of the two PRs. This suggests a hypothesis that the large difference in catalytic rate and substrate selectivity of the two enzymes may be determined by a small number of different amino acid residues in the substrate binding pockets. We have identified three such amino acids where nonconservative substitutions exist. These are H65, R105, and G106 in RSV PR compared to the structurally equivalent *G48*, *P81*, and *V82* in HIV-1 PR.

The residues R105, R105′, G106, and G106′ are in a region that is seen clearly in the crystal structure of each subunit of the RSV PR. However, their relationship to the substrate must be deduced by modeling and comparison to inhibitor binding in HIV-1 PR crystal structures. Residues 105 and 106 form part of a surface turn, involving residues 104 to 108 from each subunit, that acts like a "roof" to the distal portions of the S1 and S1′ subsites. Thus, changes in residues 105 and 106 are predicted to affect mainly binding of amino acids in the P1 and P1′ substrate positions. The side chains of R105 and R105′ also lie near the amino acid side chains of residues in P3 and P3′. However, since R105 is directed toward the protein surface, these interactions are predicted to have less of an effect on the substrate binding. Also, R105 and G106 provide a large pocket for substrate side chains at P1 and P1′. They do not form hydrogen bond interactions with the main chain of the peptide substrate. By comparison, the *proline* and *valine* in the HIV-1 PR and in the R105*P*,G105*V* mutant provide a more hydrophobic environment for P1 and P1′. The *proline* and *valine* side chains protrude into the subsites and are predicted to restrict the binding of very large side chains

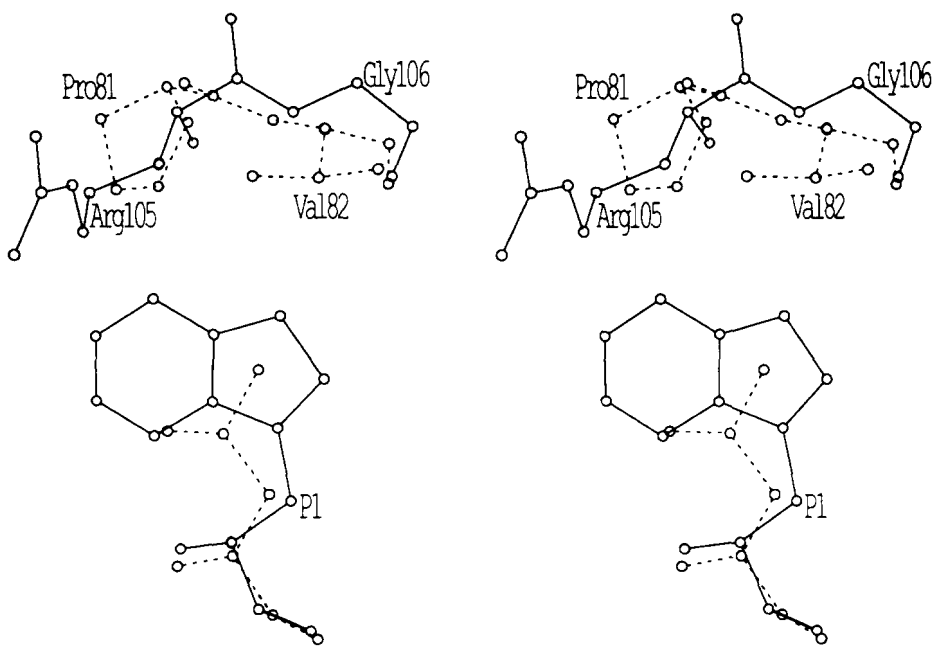

FIG. 4. **The effect of amino acid substitutions in the S1 enzyme subsite.** Stereo view of substrate residue P1 in relation to the position of RSV PR residues R105′ and G106′ (*continuous lines*) or HIV-1 PR residues P81′ and V82′ (*dashed lines*). The larger P1 tryptophan (*continuous lines*) fits better in the RSV PR subsite S1, and the smaller P1 leucine fits better in HIV-1 PR subsite S1. This is because the side chains of P81′ and V82′ protrude further into the subsite than R105′ and G106′.

at these substrate positions. As shown in Fig. 4 for P1, leucine binds well in the smaller subsite formed by the HIV-1 PR, while tryptophan binds well in the larger subsite formed by the RSV PR.

H65 is found in the flap region and was positioned by modeling and alignment with *G48* of HIV-1 PR (Fig. 2). The side chain of H65 was adjusted to form a hydrogen bond interaction with the carbonyl oxygen of G69 on the other strand of the two antiparallel β strands forming the flap. Therefore, H65 is predicted to help maintain the conformation of the flap. In addition, the amide and carbonyl oxygen of H65 are expected to form hydrogen bond interactions with the substrate (Table I) by exact analogy to *G48* of HIV-1 PR. H65 and H65′ are predicted to be part of the S3 and S3′ subsites, respectively, and they lie near the S1′ and S2 and S2′ enzyme subsites. Since H65 forms part of the same subsites as 105 and 106, additive effects may occur when all three amino acids are present in the substrate binding pocket.

It does not seem surprising that altering H65 would interfere with catalysis since it is needed for both maintaining the conformation of the flaps and for forming four hydrogen bond interactions with substrate. However, in the case of the triple mutant, when this mutation is combined with the changes in 105 and 106 in each subunit, the flaps apparently assume a near correct conformation that permits *G*65 to form hydrogen bonds with the substrate, resulting in an active enzyme. Nevertheless, this interaction is not optimal, since the triple mutant has about one-fourth the activity of the wild type RSV PR on the NC-PR peptide substrate. This interaction can be improved by modifying the substrate. For example, when leucine is substituted for serine in the P1 position of the RSV NC-PR substrate, the triple mutant, but not the double mutant, exhibits twice the activity compared to wild type RSV PR. It is thus possible that the side chain interactions introduced by substitution of leucine for serine bring the substrate sufficiently close to the flap to further enhance the interaction between *G*65 and substrate.

A further analysis of the modified NC-PR substrate data indicates that both RSV mutants exhibit behavior very similar to the wild type AMV PR for substitutions in the P4, P3, P2, P2′, and P3′ substrate positions. In contrast, the RSV mutants behave more like the HIV-1 PR in preference for amino acids substituted in the P1 and P1′ positions (Tables III and IV). This result is understandable since the 105- and 106-positions of the RSV PR are associated primarily with the S1 and S1′ enzyme subsites.

These results confirm the hypothesis that discrimination of substrate by PR is dependent upon side chain interactions between amino acids in the enzyme subsites and those in the substrate. As long as the main chain hydrogen bonding to substrate remains intact, then the subsites can act independently of one another in discrimination of substrates. This offers a logical explanation for why the exchange of amino acids at positions 65, 105, and 106 resulted in only a partial expansion of substrate preference for the reverse transcriptase-integrase HIV-1 substrate (13). Thus, to more fully change the substrate preference of the RSV PR for HIV-1 substrates and probably to increase the catalytic rate will require simultaneous amino acid exchanges in all seven of the enzyme subsites.

Acknowledgments—We thank Dr. Joe Giam, Case Western Reserve University, for inclusion body fractions containing the HIV-1 PR, Dr. Ed Houts, Molecular Genetic Resources, for AMV PR, and Dr. Terry Rosenberry for amino acid composition analysis.

REFERENCES

1. Jaskolski, M., Miller, M., Rao, J., Leis, J., and Wlodawer, J. (1990) *Biochemistry* **29,** 5889–5898
2. Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J., Whittle, P., Danley, D., Gerghegan, K., Havrylik, S., Lee, S., Scheld, K., and Hobart, P. (1989) *Nature* **342,** 299–302
3. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B., Baldwin, E., Weber, I., Selk, L., Clawson, L., Schneider, J., and Kent, S. (1989) *Science* **245,** 616–621
4. Fitzgerald, P., McKeever, B., van Middlesworth, J., Springer, J., Heimback, J., Leu, C.-T., Herber, W., Dixon, R., and Darke, P. (1990) *J. Biol. Chem.* **265,** 14209–14219
5. Erickson, J., Neidhart, D., VanDrie, J., Kempf, D., Wang, X., Norbeck, D., Plattner, J., Rittenhouse, J., Turon, M., Wideburg, N., Kohlbrenner, W., Simmer, R., Helfrich, R., Paul, D., and Knigge, M. (1990) *Science* **249,** 527–533
6. Miller, M., Sathyanarayana, B., Toth, M., Marshall, G., Clawson, L., Selk, L., Schneider, J., Kent, S., and Wlodawer, A. (1989) *Science* **246,** 1149–1152
7. Swain, A., Miller, M., Green, J., Rich, D., Schneider, J., Kent, S., and Wlodawer, A. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 8805–8809
8. Jaskolski, M., Tomasselli, A., Sawyer, T., Staples, D., Heinrikson, R., Schneider, J., Kent, S., and Wlodawer, A. (1991) *Biochemistry* **30,** 1600–1609
9. Wlodawer, A., Miller, M., Swain, A., and Jaskolski, M. (1991) in *Methods in Protein Sequence Analysis* (Jornvall, H., Hoog, J., and Gustavsson, A., eds) pp. 215–221, Verlag Wepf und Co., Basel
10. Kotler, M., Danho, W., Katz, R., Leis, J., and Skalka, A. (1989) *J. Biol. Chem.* **264,** 3428–3435
11. Darke, P., Nutt, R., Brady, S., Garsky, V., Ciccarone, T., Leu, C.-T., Lumma, P., Freidinger, R., Veber, D., and Signal, I. (1988) *Biochem. Biophys. Res. Commun.* **156,** 297–303
12. Schecter, I., and Berger A. (1967) *Biochem. Biophys. Res. Commun.* **27,** 157–162
13. Grinde, B., Cameron, C. E., Leis, J., Weber, I., Wlodawer, A., Burstein, H., Bizub, D., and Skalka, A. M. (1992) *J. Biol. Chem.* **267,** 9481–9490
14. Leis, J., Bizub, D., Weber, I., Cameron, C., Katz, R., Wlodawer, A., and Skalka, A. (1989) in *Current Communications in Molecular Biology: Viral Proteases as Targets for Chemotherapy* (Krausslich, H., Oroszlan, S., and Wimmer, E., eds) pp. 235–243, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
15. Bizub, D., Weber, I., Cameron, C., Leis, J., and Skalka, A. (1991) *J. Biol. Chem.* **266,** 4951–4958
16. Giam, C.-Z., and Boros, I. (1988) *J. Biol. Chem.* **263,** 14617–14629
17. Alexander, F., Leis, J., Soltis, D., Crowl, R., Danho, W., Poonian, M., Pan, Y.-C., and Skalka, A. (1987) *J. Virol.* **61,** 534–542
18. Miller, M., Jaskolski, M., Rao, M., Leis, J., and Wlodawer, A. (1989) *Nature* **337,** 576–579
19. Jones, A. T. (1978) *J. Appl. Crystallog.* **11,** 268–272
20. Weber, I. (1990) *J. Biol. Chem.* **265,** 10492–10496
21. Tozser, J., Gustchina, A., Weber, I., Blaha, E., Wondrak, E., and Oroszlan, S. (1991) *FEBS Lett.* **279,** 356–360
22. Gustchina, A. and Weber, I. (1991) *Proteins Struct. Funct. Genet.* **10,** 325–339