Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 2243–2248, March 1997
Biochemistry

Sequence requirements of the HIV-1 protease flap region determined by saturation mutagenesis and kinetic analysis of flap mutants

WEI SHAO*†, LORRAINE EVERITT*‡, MARIANNE MANCHESTER*§¶, DANIEL D. LOEB*‖**, CLYDE A. HUTCHISON III‖, AND RONALD SWANSTROM*††‡‡

*Lineberger Comprehensive Cancer Center, §Curriculum in Genetics and Molecular Biology, and Departments of ‖Microbiology and Immunology and ††Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

Contributed by Clyde A. Hutchison III, December 24, 1996

ABSTRACT **The retroviral proteases (PRs) have a structural feature called the flap, which consists of a short antiparallel β-sheet with a turn. The flap extends over the substrate binding cleft and must be flexible to allow entry and exit of the polypeptide substrates and products. We analyzed the sequence requirements of the amino acids within the flap region (positions 46–56) of the HIV-1 PR. The phenotypes of 131 substitution mutants were determined using a bacterial expression system. Four of the mutant PRs with mutations in different regions of the flap were selected for kinetic analysis. Our phenotypic analysis, considered in the context of published structures of the HIV-1 PR with a bound substrate analogs, shows that: (*i*) Met-46 and Phe-53 participate in hydrophobic interactions on the solvent-exposed face of the flap; (*ii*) Ile-47, Ile-54, and Val-56 participate in hydrophobic interactions on the inner face of the flap; (*iii*) Ile-50 has hydrophobic interactions at the distance of both the δ and γ carbons; (*iv*) the three glycine residues in the β-turn of the flap are virtually intolerant of substitutions. Among these mutant PRs, we have identified changes in both k_{cat} and K_m. These results establish the nature of the side chain requirements at each position in the flap and document a role for the flap in both substrate binding and catalysis.**

The retroviral protease (PR) is encoded in the viral *pro* gene for all retroviruses, including HIV-1 (reviewed in refs. 1 and 2). Proteolytic processing of the viral Gag and Gag–Pro–Pol precursor proteins by the viral PR is a required step in the virus life cycle. The active form of PR is a homodimer with an active site composed of the carboxylate side chains of two aspartic acid residues, one from each subunit of the dimer (3–7). The presence of aspartic acid residues at the active site makes the retroviral PRs members of the aspartic proteinase family of enzymes (8). The aspartic proteinases are a large family of enzymes with diverse functional roles that also share a number of structural features (reviewed in refs. 9 and 10). One such feature is called the flap, which lies above the active site cleft.

A short region of sequence similarity can be discerned within the flap of distantly related retroviral PR sequences (11). By sequence alignment, the conserved sequence domain of the flap begins at position 47 of the HIV-1 sequence and extends through the glycine at position 52. Sequences between these points align in length and in a pattern: aliphatic-Xaa-Gly-aliphatic-small-Gly. In the HIV-1 PR, these residues combine to give a short stretch of β-sheet followed by a turn that ends with the conserved glycine at position 52 (12).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

0027-8424/97/942243-6$2.00/0
PNAS is available online at **http://www.pnas.org**.

One powerful approach to the study of function is the use of large-scale genetic screens. Expression of a protein or an RNA, or use of a DNA control element, either in its homologous setting or in a convenient heterologous setting under conditions where function can be monitored, can represent a useful starting point for a detailed functional analysis using mutagenesis. Such approaches usually allow a breadth of analysis that is not possible using traditional assays. The strategy of a large-scale genetic screen is generally applicable and has been applied to a diverse set of biological systems (for examples, see refs. 13–22).

We have previously examined the sensitivity of the entire HIV-1 PR to randomly incorporated mutations and observed that the flap is one of three regions within the PR that display consecutively placed, mutationally sensitive amino acids (23). Using saturation mutagenesis, we have extended those studies to determine in detail the sequence requirements of each position of the HIV-1 flap. In this report, we describe the activity of 131 PR mutants with single amino acid substitutions in the HIV-1 flap region. In addition, we have examined the catalytic properties of several of the PR flap mutants. The properties of these mutants indicate that the flap participates in both substrate binding and catalysis.

METHODS

Mutagenesis and Expression of the pART2 Vector Expressing *pro* and *pol*. Mutations within the HIV-1 PR domain were introduced into the *pro/pol* expression vector pART2 (23, 24). This plasmid contains a 3.7-kb fragment of the HIV-1 genome (*Bgl*II to *Sal*I from the HXB2 infectious clone of HIV-1; ref. 25) inserted into the bacterial expression vector pIBI20. The oligonucleotide-directed mutagenesis procedures used have been described previously (23, 26, 27). The mutagenic oligoase long nucleotide in each reaction was a 27-base long oligonucleotide complementary to the HIV-1 PR coding domain. Three nucleotides which corresponded to a particular codon within the flap coding region were randomized during oligonucleotide synthesis; the three randomized positions were always placed 14 nt downstream of the 5′ end of the oligonucleotide.

Abbreviations: PR, protease; RT, reverse transcriptase.

†Present address: Structural Biochemistry Program, National Cancer Institute Frederick Research and Development Center, Frederick, MD 21702.
‡Present address: Department of Pathology, Duke University Medical Center, Durham, NC 27710.
¶Present address: Department of Neuropharmacology, Division of Virology, Scripps Research Institute, La Jolla, CA 92037.
**Present address: McArdle Laboratory for Cancer Research, University of Wisconsin Medical School, Madison, WI 53706.
‡‡To whom reprint requests should be addressed at: CB7295, Lineberger Building, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599.

The analysis of PR activity after induction of expression was as described (23, 24). Western analysis of bacterial lysates from bacteria expressing the Pro–Pol fusion protein was performed using a monoclonal antibody (NEA-9304; DuPont/NEN) to detect the extent of processing of reverse transcriptase within the Pol domain. We have included in our present analysis 40 mutants that were previously reported (23). In addition, mutants with changes at position 56 have been screened for temperature sensitivity (28). Some of the mutants were evaluated with an alternative bacterial growth protocol to improve the regulation of the *lac* promoter before induction (29), which resulted in several minor changes in phenotypic assignment (23).

Mutagenesis of pETPR Vector Expressing PR. To produce HIV-1 PR for the analysis of enzyme kinetics, the HIV-1 PR coding domain was inserted into *Xba*I and *Hin*dIII sites of pET-24a(+) (Novagen). The inserted PR-encoding fragment included an initiation codon and termination codon flanking the 99 codons of the *pro* gene. Site-directed mutagenesis to incorporate selected mutations was conducted as described above.

Enzyme Activity Assays. Expression of the HIV-1 PR was induced with isopropyl β-D-thiogalactoside (IPTG) in *Escherichia coli* transformed with pETPR expression vector. The PR was purified from inclusion bodies and refolded as described previously (30). Active site titration was done using the HIV-1 PR inhibitor SKF922 (K_i, ≈1 nM) with E_o and K_i determined by fitting the Morrison equation (31). Two substrates were used to assay enzyme activity. For the substrate KARVL/AEAMSF(NO_2) [Leu/Ala substrate; F(NO_2) is used to designate nitrophenylalanine, which functions as a chromophore], the reactions were performed at 37°C in 100 mM NaOAc, 1 M NaCl, 1 mM EDTA, and 0.1% polyethylene glycol (PEG) 8000 (32), pH 6.0, for 30 min, with <20% substrate cleavage. Reactions were terminated by adding 500 μl of 3% trifluoroacetic acid (TFA), and 200 μl of the reaction mixture was analyzed by HPLC. Substrate and product were separated using a Waters DELTA PAK 5-micron C18 100A column on a Waters HPLC, and integrated peak areas were calculated using the Waters Millennium Chromatography Manager. Kinetic parameters were calculated using ENZFIT.

For the fluorogenic substrate KAVY/F(NO_2)EANle(NH_2), the reactions were performed in 1.25 M $(NH_4)_2SO_4$, 6.25 mM DTT, 0.125 M *N*-(2-acetamido)-2-aminoethane-sulfonic acid (ACES), 0.1% PEG 8000, pH 6.2 (33, 34). The reactions were monitored with a Perkin–Elmer Luminescence Spectrometer LS50B at 277-nm excitation wavelength and 306-nm emission wavelength. The total reaction volume was 600 μl, and reactions were monitored for 2 min.

RESULTS

Activity of HIV-1 PR Mutants Assayed with a Bacterial Expression System. We previously showed that several positions in the flap region of the HIV-1 PR are sensitive to nonconservative amino acid substitutions, resulting in the loss of PR activity (23). We have extended these studies to determine the sequence requirement at each position of the flap, starting with the methionine at position 46 (Met-46) and extending through the valine at position 56 (Val-56). These 11 aa form an antiparallel β-sheet that extends over bound substrate/inhibitor, with positions 49–52 forming a turn at the end of the flap (Fig. 1). A total of 131 independent substitution mutations were identified within this 11-codon stretch; this represents ≈63% of all possible single amino acid substitutions within this region.

The mutations were introduced into the PR coding domain as part of the HIV-1 *pro–pol* expression vector pART2 (23, 24). When expression of this plasmid was induced in *E. coli*, a polypeptide containing the viral PR, reverse transcriptase (RT), and integrase was synthesized. When the viral PR is active, this large protein is processed to give mature Pro and Pol protein products of the same size as those found in virus particles (23, 24, 35, 36). The activity of the viral PR was easily monitored by the appearance of the two processed forms of the reverse transcriptase (p66 and p51; Fig. 2, lanes 1 and 2). In the absence of processing, the complete translation product encoded in the *pro* and *pol* genes (p120) was seen along with a series of aberrant products that do not have the expected sizes of the viral RT (Fig. 2, lane 3). For some mutants, intermediate processing of the Pro–Pol precursor protein was seen, with some of the precursor remaining unprocessed (see Fig. 2, lane 6). The range of phenotypes seen with substitutions at each position is shown in Fig. 2, and the phenotypes of all of the mutants are listed in Table 1.

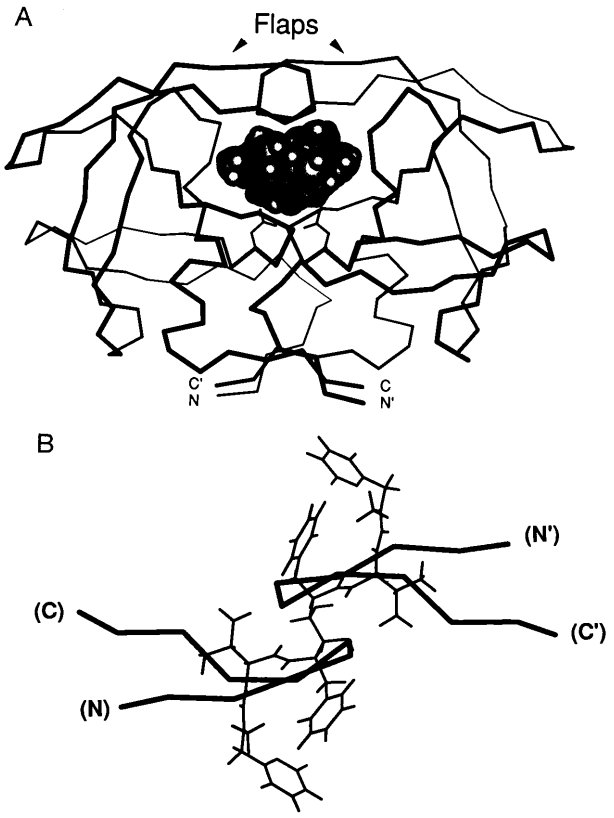

FIG. 1. Flap region of the HIV-1 PR. (*A*) The α carbon tracing of the HIV-1 PR is shown with the positions of the flaps indicated. The active site aspartic acid residues are shown in wireframe under the bound inhibitor (shown in space filling), and the N and C termini of the two chains of the dimer PR are shown, with the prime (′) used to indicate one of the two chains. This figure was generated using the coordinates of the structure of the HIV-1 PR bound to the inhibitor A-77003 (57). (*B*) The HIV-1 flap regions in thick line overlaying the inhibitor A-77003 (57). The chain orientation is shown in parentheses. Structures were evaluated using MACIMDAD.

In this discussion, the positions within the flap are divided into three groups. The first group includes those with side chains extending outward toward solvent: Met-46, Phe-53, and Lys-55. The second group includes those in the Gly-rich region: Gly-48, Gly-49, Gly-51, and Gly-52. The third group includes those with a hydrophobic chain extending inward: Ile-47, Ile-50, Ile-54, and Val-56. As shown in Table 1, those positions with side chains extending out were most tolerant to substitution, suggesting that these residues do not participate in critical interactions. For Met-46, the most active substitutions at this position were hydrophobic amino acids, while polar and charged amino acids gave only intermediate activity. Similarly for Phe-53, hydrophobic amino acids gave wild-type activity

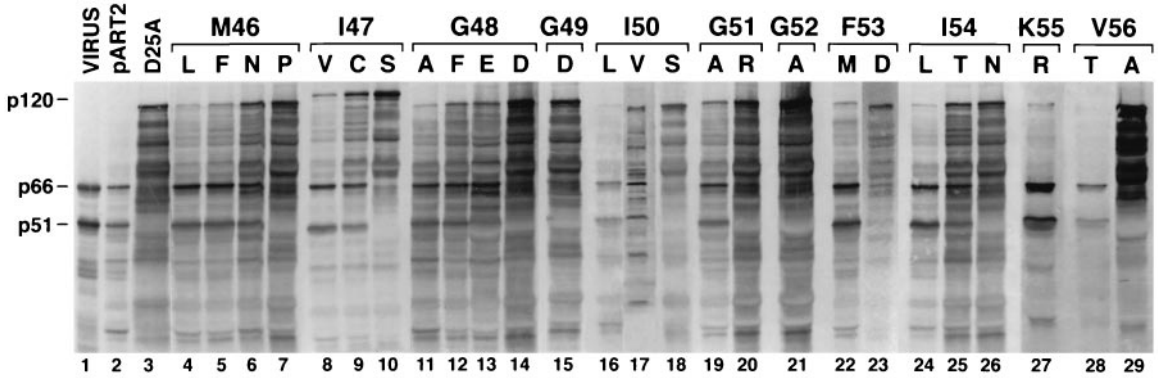

FIG. 2. Western analysis of HIV-1 PR mutants. Extracts from bacteria expressing either the wild type or a PR mutant of pART2 were electrophoresed in a polyacrylamide gel under denaturing conditions. The proteins in the gel were transferred to nitrocellulose and analyzed for the presence of processed reverse transcriptase. p120 indicates the position of the unprocessed Pro–Pol protein precursor, and p66 and p51 indicate the positions of the processed forms of HIV-1 reverse transcriptase. Lane 1 shows staining of reverse transcriptase isolated from virus particles. Lane 2 is the extract from a bacterial culture expressing the wild-type Pro–Pol from the pART2 expression vector. Lane 3 is a negative control for processing in which the codon of the catalytic aspartic acid was changed to encode alanine (D25A) in the pART2 expression vector. Lanes 4–29 are examples of PR mutants with changes at each of 11 positions in the flap [Met-46 (M46) through Val-56 (V56)]. The amino acid substituted in each mutant is shown at the top of the lane using the single letter abbreviation for amino acids.

while either charged, polar, or small amino acids gave only intermediate activity (Table 1). However, for Lys-55, charged, polar, and hydrophobic amino acids all gave full activity (Table 1).

Those positions belonging to the Gly-rich region, with the exception of Gly-48, are the most sensitive to substitution. None of the amino acids tested gave activity except for Gly-51 → Ala. For Gly-48, however, a variety of amino acid substitutions, including polar and basic amino acids, gave full activity. Hydrophobic amino acids tended to give lower activity. Acidic amino acids and proline were the least tolerated. An aspartic acid substitution destroyed activity, while substitution with glutamic acid altered the specificity of the PR. As noted previously, the Gly-48 → Glu substitution showed reduced processing at the internal cleavage site of the reverse transcriptase, giving reduced amounts of the p51 RT product (Fig. 2, lane 13; see ref. 23).

For those positions with a hydrophobic side chain extending inward, Ile-47, Ile-50, Ile-54, and Val-56, only a few conservative substitutions were tolerated. None of the 15 aa tested at Ile-47 gave wild-type activity, although cysteine and valine gave intermediate activity (Table 1). The isoleucine at position 50 could be substituted with leucine or valine and retain substantial activity. For Ile-54, two conservative substitutions, leucine and valine, gave full activity, while two nonconservative substitutions, cysteine and threonine, gave intermediate activity. For Val–56, only two amino acid substitutions, threonine and cysteine, gave wild type-like activity (Table 1). For this group of positions, the tolerated substitutions are aliphatic and hydrophobic, and the pattern of sensitivity is distinctive for each position suggesting nonequivalent hydrophobic environments.

Kinetic Analysis of Selected HIV-1 PR Mutants. PR mutants representing the three groups were selected for kinetic analysis. Ile-50 → Val and Ile-54 → Thr are mutants of the group with side chains extending inward, Gly-51 → Ala represents a mutant in the Gly-rich region, and Phe-53 → Asp represents a member of the group with side chains extending outward toward solvent. Two substrates were used in the analysis of enzyme kinetics. Data for the substrate Leu/Ala, which represents the processing site between CA and p2 in the Gag polyprotein precursor, are shown in Table 2. Wild-type PR has a K_m of 0.49 mM and k_{cat} of 8.5 s^{-1}. Ile-50 → Val has an ≈6-fold increase in K_m and no change in k_{cat}. Gly-51 → Ala has a <2-fold increase in K_m and an ≈6-fold decrease in k_{cat}. For Phe-53 → Asp, K_m and k_{cat} are slightly lower than those of wild type. The K_m of Ile-54 → Thr is too high to be measured accurately due to limited solubility of the substrate.

As an alternative way to assess the activity of the Ile-54 → Thr mutant on the Leu/Ala substrate, the initial velocity of the reaction catalyzed by the Ile-54 → Thr mutant was measured at 0.05 mM substrate concentration to derive k_{cat}/K_m from the equation $V = (k_{cat}/K_m) \times E \times S$. The Ile-54 → Thr mutant has a k_{cat}/K_m of 0.2, which is ≈30-fold less than the wild-type enzyme.

Kinetic analysis of the mutant PRs was also carried out using the fluorogenic substrate, KAVY/F(NO_2)EANle-NH2, and the results are also shown in Table 2. All of the mutants had similar differences compared with the wild-type enzyme using either substrate.

Activity of a PR Mutant with an Alternative Flap Sequence. A comparison of the HIV-1 PR sequence with the sequences of other retroviral PR flap regions (11) reveals that at multiple positions within the flap there are substitutions in other retroviruses that are incompatible with activity of the HIV-1 PR. To examine how the retroviral PR flap can accommodate a different sequence and still function, we made a flap mutant with the entire sequence of human leukemia virus type 2 flap (GGIGG to LGASG). No PR activity was detected with this mutant PR using the fluorogenic substrate (data not shown). This result indicates that this different flap sequence is not equivalent and that other changes must be present elsewhere within either the PR or the substrate to compensate for the alternative flap sequence present in the human leukemia virus type 2 flap. A similar conclusion of nonequivalence of flap sequences has been drawn from studies comparing the PRs of Rous sarcoma virus and HIV-1 (37–39).

DISCUSSION

The flap plays a central role in the activity of the retroviral PR. It must move away from the active site to allow the entry of substrate and the exit of products. In the down position, interacting with substrate, the two flaps of the dimer PR coordinate a water molecule that helps position the substrate for catalysis. Alteration of the backbone atoms of the HIV-1 PR flap that coordinate this water molecule profoundly affect k_{cat} (40). The crystal structure of the PR suggests other important backbone and side chain interactions involving the flap contribute to PR activity and substrate specificity (12, 41, 42). We have used large-scale mutagenesis in an effort to define the side chain requirement at each position of the

Table 1. Activity of HIV-1 PR mutants

Level of activity	Substitutions at indicated flap positions									
	M46	I47	G48	G49	I50	G51	G52	F53	I54	K55	V56
Wild-type	Val		Ser		Leu			Tyr	Leu	Thr	Thr
	Trp		Thr					Leu	Val	Asn	Cys
	Phe		His					Ile		Ile	
	Leu		Gln					Val		Arg	
			Arg					Trp		Gln	
			Ala					Met			
			Lys								
Intermediate											
Nearly wild-type	His	Val	Phe		Val			Ser			
	Arg		Val								
In between	Asn	Cys	Asn			Ala		Glu	Cys		
	Ser		Ile					Gly			
			Glu					Asp			
			Trp					Arg			
Nearly negative	Gln								Thr		Gly
	Gly										
Negative	Pro	Asp	Asp	Trp	Tyr	Leu	Asn		His		Leu
		Ala	Pro	Ala	Asn	Thr	His		Asp		Ile
		Leu		Asp	Thr	Arg	Lys		Asn		Trp
		Arg		Ser	Ala	Pro	Asp		Arg		Ser
		Met		Phe	Trp	Cys	Ala		Lys		Ala
		Asn		Gln	Ser	Lys	Phe		Gly		Glu
		Gln		Arg	Gln	Val	Val		Pro		Arg
		Gly		Val	Lys	Met	Ser		Ala		Lys
		Tyr		Glu	Met	Phe	Pro		Ser		
		Glu		Lys	Arg	Ile	Arg		Phe		
		Ser		Leu			Cys				
		Lys		His			Met				
		Thr					Gln				
							Leu				

The level of protease activity was assessed based on results from Western analysis. A mutant was considered to have a wild-type level of activity if the Pro–Pol precursor was completely processed. A negative phenotype was assigned when no correct processing to generate either the p66 or p51 form of RT was seen. An intermediate phentoype was assigned when some, but not complete, processing was observed. The intermediate phentoype was further subdivided based on whether the activity was nearly wild-type, nearly negative, or in between. One-letter abbrevations for amino acids are used, followed by the position number.

HIV-1 flap as a way of assessing the environment and contribution of each of these residues. Changes at 4 of these 11 positions within the flap have been associated with resistance to PR inhibitors. Several of the mutants presented here have been described previously by others (43–46), and there is good but not complete agreement with these previous results.

The Solvent-Exposed Face. Only one of the three residues that are oriented toward solvent on the face of the flap behaves as a surface residue (Fig. 3). There is little change in activity

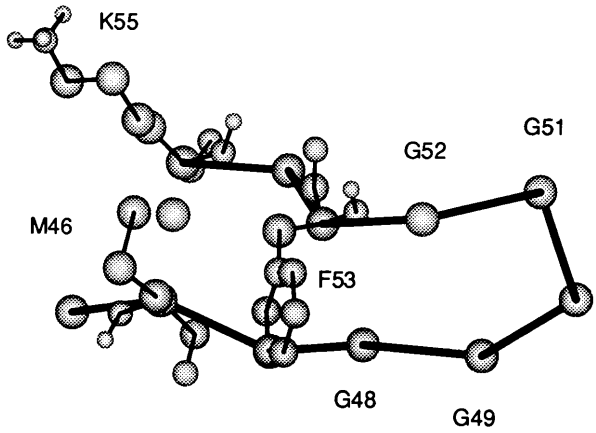

FIG. 3. The outer face of the flap. α carbon tracing of the flap region with the side chains for residues Met-46, Phe-53, and Lys-55 is shown. In addition, the locations of the glycines at positions 48, 49, 51, and 52 are shown. This figure was made using the structure of the PR bound to the inhibitor A-77003 (57).

when position 55 is substituted from lysine to a variety of other amino acids, as would be expected for a typical solvent-accessible surface residue. In contrast, Met-46 and Phe-53 both behave as if they are involved in important hydrophobic interactions (Table 1), although this effect was more apparent in the bacterial expression system compared with the activity of the purified PR mutant Phe-53 → Asp using two peptide substrates (Table 2).

There are several possibilities that could explain the hydrophobic nature of these residues. First, Met-46 and Phe-53 form a cleft between them that could interact with substrate. The substrate sequence beyond the P4 amino acid (which has a shallow pocket at the edge of the substrate binding cleft) has been shown to affect PR cleavage (47), possibly indicating the existence of interactions on the surface of the PR beyond the substrate binding cleft, although such distal side chain interactions have not been observed using structural analysis (41). Another possibility is that these hydrophobic residues affect movement of the flap and dehydration of the active site (48). Changes at position 46 to other hydrophobic amino acids are frequently found after selection for resistance to PR inhibitors that themselves do not extend beyond the substrate binding cleft (49, 50). Taken together, these results indicate that the residues at positions 46 and 53 contribute as hydrophobic residues to PR activity in an important but as yet poorly understood way.

A Hydrophobic Patch on the Inside of the Flap. Ile-47, Ile-54, and Val-56 are involved in hydrophobic interactions, with Ile-47 sitting just above Val-32 in the hydrophobic core (Fig. 4*A*). This is a highly organized structure as can be inferred by the distinctive patterns of activity seen after substitution at each position. Substitution with leucine had no apparent effect on position 54 but was incompatible with activity at position 47 (Table 1; see also ref. 45). In contrast, mutants with a

Table 2. Kinetics of flap mutants

Enzyme	Substrate: KARVL/AEAMSF (NO_2)			Substrate: KAVY/F (NO_2) EANle-NH_2		
	K_m, mM	k_{cat}, s^{-1}	k_{cat}/K_m	K_m, μM	k_{cat}, s^{-1}	k_{cat}/K_m
Wild-type	0.49	8.5	17.3	3.7	13.4	3.6
Ile-50 → Val	3.13	8.3	2.7	16.4	19.5	1.2
Gly-51 → Ala	0.73	1.5	2.1	2.2	2.6	1.2
Phe-53 → Asp	0.21	3.9	18.6	6.9	15.0	2.2
Ile-54 → Thr	>5	ND	ND	20.4	1.5	0.1

ND, not determined. One-letter abbreviations for amino acids are used.

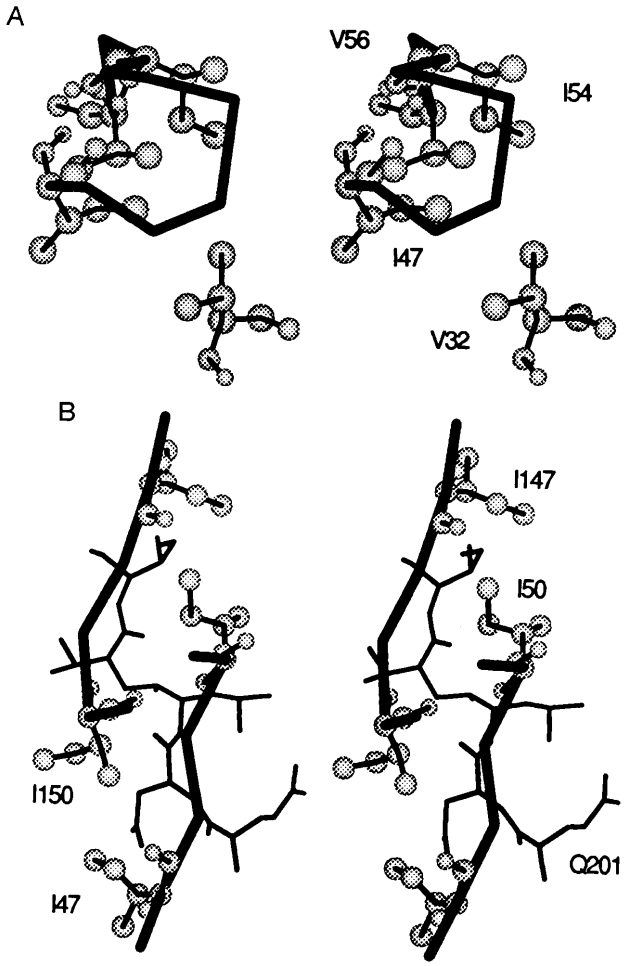

FIG. 4. Hydrophobic interaction in the flap region. (*A*) Stereoview of the interactions between residues Ile-47, Ile-54, and Val-56 within the flap and the underlying residue in the hydrophobic core Val-32. This figure was made using the structure of the PR bound to the inhibitor A-77003 (57). (*B*) Stereoview showing the interactions of Ile-47 (and Ile-47 from the other subunit designated I147) and Ile-50 (and Ile-50 from the other subunit designated I150) with the bound inhibitor U-85548e. Q201 represents the P3 amino acid glutamine of the inhibitor. (P1 represents the substrate amino acid to the N-terminal side of the scissle bond, P2 is next to the P1 residue continuing toward the N terminus, with P3 being next to P2, etc. P1′ represents the substrate amino acid to the C terminus of the scissle bond, with P2′ being next to P1′, etc. S1 represents the pocket in the PR that surrounds the P1 side chain, S2 surrounds the P2 side chain, etc.) This figure was made using the coordinates of the structure of the PR bound to U-85548e (15).

substitution to valine at either 47 or 54 showed wild-type or nearly wild-type activity (Table 1; see also ref. 44) but no activity when the long aliphatic side chain was replaced with alanine (Table 1). Val-56 could not be changed to either of the longer aliphatic amino acids (leucine or isoleucine) but was active when substituted with the isosteric but more polar amino acid threonine, the wild-type amino acid found in the HIV-2 PR. This is in contrast to positions 47 and 54, which showed good activity when substituted to valine but little or no activity when substituted to threonine (Table 1). These results indicate both tight and highly organized packing of these residues in the flap. The inability of Ile-47 and Val-56 to accept leucine may be due to the internally oriented, close interaction between the aliphatic side chains of these two residues (Fig. 4*A*). By contrast, the side chain of Ile-54 is much closer to the surface, perhaps allowing for repacking to accommodate leucine.

Ile-50 can participate in both the S1/S1′ and S2/S2′ subsites, although in an asymmetric way (Fig. 4*B*). The δ carbon can be involved in hydrophobic packing within the flap or extend out toward solvent, with the γ carbons making contacts with bound inhibitors in either orientation. Thus, while a valine substitution would reduce some of the internal packing, it would not preclude continued interaction with substrate, consistent with the near wild-type level of activity that was observed with the Ile-50 → Val mutant (Table 1). However, the loss of the internal packing of the δ carbon may provide more flexibility to the remaining atoms of the side chain. This explanation could account for the appearance of Ile-50 → Val as a substitution associated with at least partial resistance to some PR inhibitors (51), and our observation that this substitution selectively affects K_m but not k_{cat} (Table 2). By contrast, the δ carbon of Ile-47 is in contact with the P2/P2′ side chains of the inhibitor/substrate. Substitution to valine has been reported to have a modest effect (2- to 3-fold) on both K_m and k_{cat} (44). This mutation has also been reported to affect sensitivity to some PR inhibitors (51). In the simian immunodeficiency virus SIVmac and HIV-2 PRs, position 47 is naturally occupied by valine, but in these cases the position 32 residue is isoleucine rather than the valine found in the HIV-1 PR; this may represent a compensatory mutation since in each case the δ carbon of the isoleucine (Ile-47 of HIV-1 and Ile-32 of SIVmac and HIV-2) is in approximately the same position (52, 53).

Glycines in the Turn. Because of their positions in the turn of the flap (Fig. 3), the glycines at positions 49, 51, and 52 are severely constrained in their ability to accommodate a side chain (A. Wlodawer and A. Gustchina, personal communication). Thus it is not surprising that these positions are nearly intolerant of any substitutions (Table 1). A similar conclusion has been drawn for the glycine at position 69 in the Rous sarcoma virus PR (equivalent to position 52 in the HIV-1 PR) based on its sensitivity to substitution with leucine (37). The one exception to the intolerance of these glycines is seen in the active HIV-1 PR with the conservative substitution alanine at position 51. The addition of a methyl group on the top of the flap is unlikely to allow for a new direct interaction with substrate. This substitution selectively affects k_{cat} (Table 2). It is possible that the addition of the methyl group leads to changes in ϕ-ψ angles in the flap that compromise its ability to participate in catalysis. By analogy to the work of Baca and Kent (40), if the Gly-51 → Ala substitution caused a repositioning of the main chain nitrogens of Ile-50, then the coordinated water molecule under the flap might be either more difficult to trap or improperly positioned, leading to a reduced k_{cat}, although other explanations are also possible.

The exception to the highly sensitive glycines in the flap is Gly-48, which is adjacent to the turn within the flap. We have previously shown that a glutamic acid substitution at this position restricts the specificity of the PR, as seen by limited cleavage of the processing site within p66 RT to generate the p51 subunit (23). We also showed that this restriction was due to an unfavorable interaction with the P3′ amino acid, which selectively affected K_m (54). A side chain at position 48 could potentially interact with either the P1/P1′ or P3/P3′ side chains. Position 48 is very tolerant of substitution when assayed using the bacterial expression assay (Table 1), and a similar tolerance to substitution has been seen with purified recombinant PR (46, 54). Selection with the PR inhibitor saquinavir results in the appearance of a Gly-48 → Val substitution (55, 56), again indicating the potential impact of a side chain at position 48 on ligand binding specificity.

We thank George Rose, Mark Moody, Alla Gustchina, Alex Wlodawer, John Erickson, Scott Michael, and Steve Pettit for helpful discussions. This work was supported by National Institutes of Health Grants R01 AI25321, National Research Service Award T32 GM07092 (Genetics; M.M.), and National Research Service Award F32 AI09308 (W.S.).

1. Oroszlan, S. & Luftig, R. B. (1990) *Curr. Top. Microbiol. Immunol.* **157,** 153–185.
2. Pettit, S. C., Michael, S. F. & Swanstrom, R. (1993) *Perspect. Drug Discovery Des.* **1,** 69–83.
3. Meek, T. D., Dayton, B. D., Metcalf, B. W., Dreyer, G. B., Strickler, J. E., Gorniak, J. G., Rosenberg, M., Moore, M. L., Magaard, V. W. & Debouck, C. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1841–1845.
4. Miller, M., Jaskolski, M., Rao, J. K., Leis, J. & Wlodawer, A. (1989) *Nature (London)* **337,** 576–579.
5. Navia, M. A., Fitzgerald, P. M., McKeever, B. M., Leu, C. T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L. & Springer, J. P. (1989) *Nature (London)* **337,** 615–620.
6. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J. & Kent, S. B. (1989) *Science* **245,** 616–621.
7. Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J., Danley, D. E., Geoghegan, K. F., Hawrylik, S. J., Lee, S. E., Scheld, K. G. & Hobart, P. M. (1989) *Nature (London)* **342,** 299–302.
8. Toh, H., Ono, M., Saigo, K. & Miyata, T. (1985) *Nature (London)* **315,** 361 (lett.).
9. Davies, D. R. (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19,** 189–215.
10. Tang, J., James, M. N., Hsu, I. N., Jenkins, J. A. & Blundell, T. L. (1978) *Nature (London)* **271,** 618–621.
11. Rao, J., Erickson, J. & Wlodawer, A. (1991) *Biochemistry* **30,** 4663–4671.
12. Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. & Wlodawer, A. (1989) *Science* **246,** 1149–1152.
13. Hutchison, C. A., III, Nordeen, S. K., Vogt, K. & Edgell, M. H. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 710–714.
14. Tsui, P. l., Sweet, R. W., Sathe, G. & Rosenberg, M. (1992) *J. Biol. Chem.* **267,** 9361–9367.
15. Slepak, V. Z., Wilkie, T. M. & Simon, M. I. (1993) *J. Biol. Chem.* **268,** 1414–1423.
16. Wrenn, C. K. & Katzenellenbogen, B. S. (1993) *J. Biol. Chem.* **268,** 24089–24098.
17. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S. & Miller, J. H. (1994) *J. Mol. Biol.* **240,** 421–433.
18. Schultz, D. W. & Yarus, M. (1994) *J. Mol. Biol.* **235,** 1381–1394.
19. Liang, J., Moye-Rowley, S. & Maurer, R. A. (1995) *J. Biol. Chem.* **270,** 25520–25525.
20. Uozumi, N., Gassmann, W., Cao, Y. & Schroeder, J. I. (1995) *J. Biol. Chem.* **270,** 24276–24281.
21. Hill-Eubanks, D., Burstein, E. S., Spalding, T. A., Brauner-Osborne, H. & Brann, M. R. (1996) *J. Biol. Chem.* **271,** 3058–3065.
22. Ohnuma, S., Nakazawa, T., Hemmi, H., Hallberg, A. M., Koyama, T., Ogura, K. & Nishino, T. (1996) *J. Biol. Chem.* **271,** 10087–10095.
23. Loeb, D. D., Swanstrom, R., Everitt, L., Manchester, M., Stamper, S. E. & Hutchison, C. A., III. (1989) *Nature (London)* **340,** 397–400.
24. Loeb, D. D., Hutchison, C. A., III, Edgell, M. H., Farmerie, W. G. & Swanstrom, R. (1989) *J. Virol.* **63,** 111–121.
25. Ratner, L., Fisher, A., Jagodzinski, L. L., Mitsuya, H., Liou, R. S., Gallo, R. C. & Wong, S. F. (1987) *AIDS Res. Hum. Retroviruses* **3,** 57–69.
26. Hutchison, C. A., III, Swanstrom, R. & Loeb, D. D. (1991) *Methods Enzymol.* **202,** 356–390.
27. Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987) *Methods Enzymol.* **154,** 367–382.
28. Manchester, M., Everitt, L., Loeb, D. D., Hutchison, C. A., III., & Swanstrom, R. (1994) *J. Biol. Chem.* **269,** 7689–7695.
29. Kaplan, A. H., Manchester, M. & Swanstrom, R. (1994) *Methods Enzymol.* **241,** 58–69.
30. Pettit, S. C., Moody, M. D., Wehbie, R. S., Kaplan, A. H., Nantermet, P. V., Klein, C. A. & Swanstrom, R. (1994) *J. Virol.* **68,** 8017–8027.
31. Williams, J. W. & Morrison, J. F. (1979) *Methods Enzymol.* **63,** 437–467.
32. Jordan, S. P., Zugay, J., Darke, P. L. & Kuo, L. C. (1992) *J. Biol. Chem.* **267,** 20028–20032.
33. Kageyama, S., Mimoto, T., Murakawa, Y., Nomizu, M., Ford, H., Jr., Shirasaka, T., Gulnik, S., Erickson, J., Takada, K., Hayashi, H., Broder, S., Kiso, Y. & Mitsuya, H. (1993) *Antimicrob. Agents Chemother.* **37,** 810–817.
34. Gulnik, S. V., Suvorov, L. I., Liu, B., Yu, B., Anderson, B., Mitsuya, H. & Erickson, J. W. (1995) *Biochemistry* **34,** 9282–9287.
35. Mous, J., Heimer, E. P. & Le Grice, S. F. (1988) *J. Virol.* **62,** 1433–1436.
36. Farmerie, W. G., Loeb, D. D., Casavant, N. C., Hutchison, C. A., III, Edgell, M. H. & Swanstrom, R. (1987) *Science* **236,** 305–308.
37. Grinde, B., Cameron, C. E., Leis, J., Weber, I. T., Wlodawer, A., Burstein, H. & Skalka, A. M. (1992) *J. Biol. Chem.* **267,** 9491–9498.
38. Grinde, B., Cameron, C. E., Leis, J., Weber, I. T., Wlodawer, A., Burstein, H. & Skalka, A. M. (1992) *J. Biol. Chem.* **267,** 9491–9498.
39. Cameron, C. E., Ridky, T. W., Shulenin, S., Leis, J., Weber, I. T., Copeland, T., Wlodawer, A., Burstein, H., Bizub, B. D. & Skalka, A. M. (1994) *J. Biol. Chem.* **269,** 11170–11177.
40. Baca, M. & Kent, S. B. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11638–11642.
41. Wlodawer, A. & Erickson, J. W. (1993) *Annu. Rev. Biochem.* **62,** 543–585.
42. Jaskolski, M., Tomasselli, A. G., Sawyer, T. K., Staples, D. G., Heinrikson, R. L., Schneider, J., Kent, S. B. & Wlodawer, A. (1991) *Biochemistry* **30,** 1600–1609.
43. Rodriguez, E. J., Debouck, C., Deckman, I. C., Abu, S. H., Raushel, F. M. & Meek, T. D. (1993) *Biochemistry* **32,** 3557–3563.
44. Sardana, V. V., Schlabach, A. J., Graham, P., Bush, B. L., Condra, J. H., Culberson, J. C., Gotlib, L., Graham, D. J., Kohl, N. E., LaFemina, R. L., Schneider, C. L., Wolanski, B. S., Wolfgang, J. A. & Emini, E. A. (1994) *Biochemistry* **33,** 2004–2010.
45. Vacca, J. P., Dorsey, B. D., Schleif, W. A., Levin, R. B., McDaniel, S. L., Darke, P. L., Zugay, J., Quintero, J. C., Blahy, O. M. & Roth, E. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 4096–4100.
46. Lin, Y., Lin, X., Hong, L., Foundling, S., Heinrikson, R. L., Thaisrivongs, S., Leelamanit, W., Raterman, D., Shah, M., Dunn, B. M. & Tang, J. (1995) *Biochemistry* **34,** 1143–1152.
47. Griffiths, J. T., Phylip, L. H., Konvalinka, J., Strop, P., Gustchina, A., Wlodawer, A., Davenport, R. J., Briggs, R., Dunn, B. M. & Kay, J. (1992) *Biochemistry* **31,** 5193–5200.
48. Collins, J. R., Burt, S. K. & Erickson, J. W. (1995) *Nat. Struct. Biol.* **2,** 334–338.
49. Ho, D. D., Toyoshima, T., Mo, H., Kempf, D. J., Norbeck, D., Chen, C. M., Wideburg, N. E., Burt, S. K., Erickson, J. W. & Singh, M. K. (1994) *J. Virol.* **68,** 2016–2020.
50. Kaplan, A. H., Michael, S. F., Wehbie, R. S., Knigge, M. F., Paul, D. A., Everitt, L., Kempf, D. J., Norbeck, D. W., Erickson, J. W. & Swanstrom, R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5597–5601.
51. Partaledis, J. A., Yamaguchi, K., Tisdale, M., Blair, E. E., Falcione, C., Maschera, B., Myers, R. E., Pazhanisamy, S., Futer, O., Cullinan, A. B., Stuver, C. M., Byrn, R. A. & Livingston, D. J. (1995) *J. Virol.* **69,** 5228–5235.
52. Zhao, B., Winborne, E., Minnich, M. D., Culp, J. S., Debouck, C. & Abdel, M. S. (1993) *Biochemistry* **32,** 13054–13060.
53. Priestle, J. P., Fassler, A., Rosel, J., Tintelnot, B. M., Strop, P. & Grutter, M. G. (1995) *Structure (London)* **3,** 381–389.
54. Moody, M. D., Pettit, S. C., Shao, W., Everitt, L., Loeb, D. D., Hutchison, C. A., III, & Swanstrom, R. (1995) *Virology* **207,** 475–485.
55. Jacobsen, H., Yasargil, K., Winslow, D. L., Craig, J. C., Krohn, A., Duncan, I. B. & Mous, J. (1995) *Virology* **206,** 527–534.
56. Eberle, J., Bechowsky, B., Rose, D., Hauser, U., von der Helm, K., Gurtler, L. & Nitschko, H. (1995) *AIDS Res. Hum. Retroviruses* **11,** 671–676.
57. Hosur, M. V., Bhat, T. N., Kempf, D. J., Baldwin, E. T., Lui, B., Gulnik, S., Wideburg, N. E., Norbeck, D. W., Appelt, K. & Erickson, J. W. (1994) *J. Am. Chem. Soc.* **116,** 847–855.