

# Programming the Rous Sarcoma Virus Protease to Cleave New Substrate Sequences\*

(Received for publication, February 9, 1996)

Todd W. Ridky<sup>‡</sup>, Diane Bizub-Bender<sup>§</sup>, Craig E. Cameron<sup>¶</sup>, Irene T. Weber<sup>||</sup>,  
Alexander Wlodawer<sup>\*\*</sup>, Terry Copeland<sup>\*\*</sup>, Anna Marie Skalka<sup>§</sup>, and Jonathan Leis<sup>‡‡</sup>

From the Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106, the <sup>§</sup>Fox Chase Cancer Center, Institute for Cancer Research, Philadelphia, Pennsylvania 19111, the <sup>||</sup>Jefferson Cancer Center, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, and the <sup>\*\*</sup>Macromolecular Structural Laboratory, NCI-Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, Maryland 21702

The Rous sarcoma virus protease displays a high degree of specificity and catalyzes the cleavage of only a limited number of amino acid sequences. This specificity is governed by interactions between side chains of eight substrate amino acids and eight corresponding subsite pockets within the homodimeric enzyme. We have examined these complex interactions in order to learn how to introduce changes into the retroviral protease (PR) that direct it to cleave new substrates. Mutant enzymes with altered substrate specificity and wild-type or greater catalytic rates have been constructed previously by substituting single key amino acids in each of the eight enzyme subsites with those residues found in structurally related positions of human immunodeficiency virus (HIV)-1 PR. These individual amino acid substitutions have now been combined into one enzyme, resulting in a highly active mutant Rous sarcoma virus (RSV) protease that displays many characteristics associated with the HIV-1 enzyme. The hybrid protease is capable of catalyzing the cleavage of a set of HIV-1 viral polyprotein substrates that are not recognized by the wild-type RSV enzyme. Additionally, the modified PR is inhibited completely by the HIV-1 PR-specific inhibitor KNI-272 at concentrations where wild-type RSV PR is unaffected. These results indicate that the major determinants that dictate RSV and HIV-1 PR substrate specificity have been identified. Since the viral protease is a homodimer, the rational design of enzymes with altered specificity also requires a thorough understanding of the importance of enzyme symmetry in substrate selection. We demonstrate here that the enzyme homodimer acts symmetrically in substrate selection with each enzyme subunit being capable of recognizing both halves of a peptide substrate equally.

Rational design of HIV-1<sup>1</sup> protease inhibitors as therapeutic agents for AIDS will require a thorough understanding of the molecular mechanisms that govern the complex interactions between the enzyme and its substrates. This necessitates identification of key amino acids that determine substrate specificity. To identify these critical residues, we have exploited differences in structure and specificity between the related proteases from Rous sarcoma virus (RSV) and HIV-1. Although these two enzymes share a 30% amino acid identity and common overall topology (1, 2), they possess markedly different substrate specificities. HIV-1 protease catalyzes the cleavage not only of its own Gag and Gag-Pol polyprotein sequences but also those of the native RSV protease (PR) substrates (3, 4). In contrast, RSV PR has a more limited substrate range, cleaving its own, but not HIV-1 PR, polyprotein sequences. To gain insight into the basis for these differences, we identified protease residues located within 10 Å of a bound substrate positioned by analogy to x-ray crystal structures of HIV-1 PR-inhibitor complexes (5). Alignment of the two enzyme structures revealed that many amino acid residues in the RSV substrate binding pockets are identical to those in the structurally equivalent positions of HIV-1 protease. However, a small number of structurally equivalent residues differ between the two proteins (Fig. 1). We hypothesized that these amino acid differences contribute to the difference in substrate preference between RSV and HIV-1 proteases. To test this idea, a number of RSV protease mutants were constructed by site-directed mutagenesis that replaced one or two of these RSV residues with the structurally equivalent HIV-1 residues (3–6). Many of these constructs were active and displayed partially altered specificity for substrate selection at one or two of the eight enzyme subsites. Consistent with this result, Sedlacek *et al.* (7) also showed that some of these changes allow for partial HIV-1-like specificity with modified peptide substrates. Recently, several of these residues have been shown to be substituted in viral mutants that arise during the development of drug resistance to HIV-1 protease inhibitors in clinical trials with AIDS patients (8). For a review of the resistant phenotypes, see Ref. 9. Changes in substrate preference obtained when single amino acid substitutions were introduced into RSV PR were sufficient to allow for some catalytic activity on peptides that represented one or two HIV-1 polyprotein cleavage sites (3–6). However, it became clear that because subsites were acting somewhat independently in substrate selection,

\* This work was supported in part by United States Public Health Service Grants CA52047 (to J. L.), CA58166 (to I. W.), and CA06937, CA47486, and RR05539 (to A. M. S.); an appropriation from the Commonwealth of Pennsylvania (to A. M. S.); and NCI, DHHS, Contract N01 C074101 with ABL (to A. W. and T. C.). 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.

<sup>‡</sup> Medical Scientist Trainee. Supported by Grant GM07250 from the National Institutes of Health.

<sup>¶</sup> Recipient of Predoctoral Fellowship GM13628 from the National Institutes of Health. Present address: Dept. of Chemistry, Pennsylvania State University, 152 Davey Laboratory, University Park, PA 16802.

<sup>‡‡</sup> To whom all correspondence should be addressed: Dept. of Biochemistry, Case Western Reserve University School of Medicine, 2119 Abington Rd., Cleveland, OH 44106-4935. Tel.: 216-368-3360; Fax: 216-368-4544; E-mail: jxl8@biocserver.bioc.cwru.edu.

<sup>1</sup> The abbreviations used are: HIV, human immunodeficiency virus; RSV, Rous sarcoma virus; AMV, avian myeloblastosis virus; PR, retroviral protease; NC, retroviral nucleocapsid protein; CA, capsid protein; PAGE, polyacrylamide gel electrophoresis; NTA, nitrilotriacetic acid; kb, kilobase pair(s).

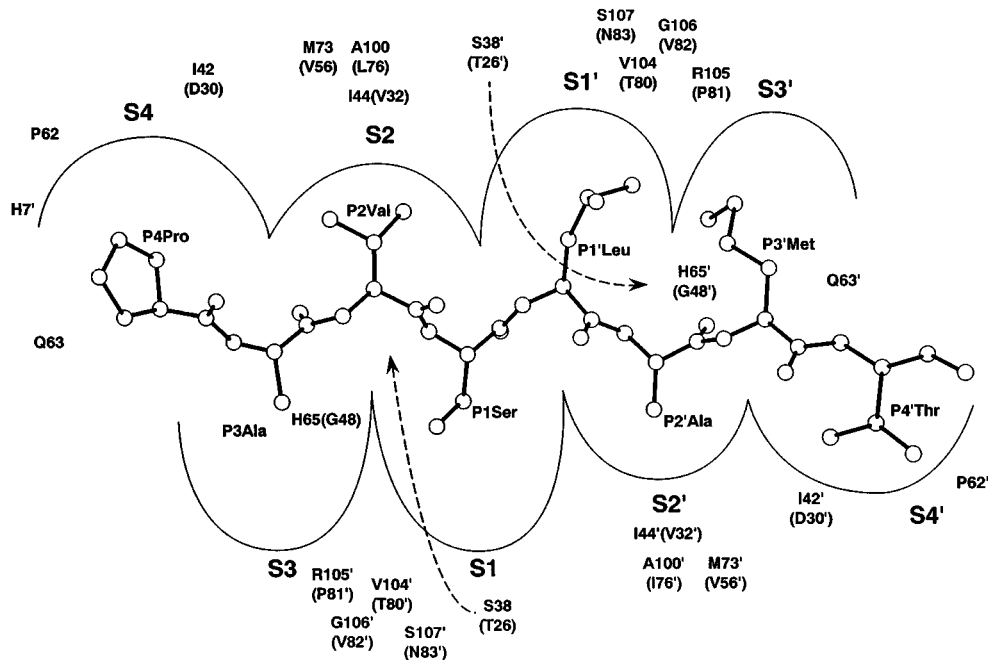


FIG. 1. Schematic representation of the RSV NC-PR substrate, PAVSLAMT, from P4 to P4' in the S4-S4' subsites of PR. The relative size of each subsite is indicated approximately by the area enclosed by the curved line around each substrate side chain. Protease residues forming the subsites are shown for those that differ between the RSV and HIV-1 PRs. RSV PR residues are shown outside the parentheses, whereas the HIV-1 PR residues are shown within the parentheses. Most of the residues contribute to more than one adjacent subsite and this is indicated by the relative positions of the labels.

multiple amino acid substitutions would be required to affect a complete change in enzyme specificity. In this report, multiple amino acid changes in enzyme subsites have been combined into one construct and shown to impart substantial HIV-1-like behavior upon the RSV PR. Moreover, a covalently linked dimer PR was used to demonstrate that the symmetric subunits of the enzyme recognize both halves of a substrate equally.

#### EXPERIMENTAL PROCEDURES

**Protease Assay**—PR activity was assayed in a volume of 25  $\mu$ l of 100 mM sodium phosphate, pH 5.9, 2.4 M sodium chloride, 0–320  $\mu$ M peptide substrate, and 50–500 ng of PR. Reactions were initiated by the addition of protease, incubated at 37  $^{\circ}$ C for 3–15 min, and stopped by the addition of 300  $\mu$ l of 0.5 M sodium borate, pH 8.5. Twenty  $\mu$ l of 0.05% (w/v) of fluorescamine was added. Relative fluorescence intensity was measured on a Perkin-Elmer LS-50B luminescence spectrophotometer using an excitation wavelength of 386 nm and an emission wavelength of 477 nm. The concentration of peptide substrate was determined by amino acid composition analysis. Each activity measurement represents the mean of at least three independent experiments. In each case, the standard error for all experiments did not exceed 20% of the value reported. Kinetic constants were determined using the assay described above. Initial rate data from substrate saturation curves were fit to the Michaelis-Menten equation using the NFIT program (5). Correlation coefficients of the fit were greater than 0.98, and the standard deviation of the constants reported was <20%.

**Peptides**—Peptide substrates 9–12 amino acids in length were synthesized, based on cleavage sites in Gag and Gag-Pol polyprotein precursors from RSV and HIV-1. These peptides include the 8 amino acids, P4-P4', required for efficient and specific cleavage by the retroviral protease. These residues interact with a unique array of amino acids in the protease which form the S4-S4' subsites of the substrate binding pocket. Peptides were synthesized with an amino-terminal proline residue to prevent uncleaved substrate from reacting with fluorescamine. Peptides were solubilized in 1 mM dithioerythritol, and their concentrations were determined by quantitative amino acid composition analysis.

**Plasmid Construction and Mutagenesis**—The PR I42D,I44V was constructed by our previously published mutagenesis procedure (4). Briefly, oligodeoxynucleotide 5'-GACTCTGGAGCGGACACCGTCA-TATCAGAGGAGGAT-3' was annealed to *Pvu*I linearized PR I44V and *Bgl*II/*Acc*I gapped PR. PR I42D,I44V clones were screened for the presence of a new *Asp*I site. PR I42D,I44V,M73V,A100L was con-

structed by ligating a dephosphorylated 2.6-kb *Eco*RI/*Acc*I fragment from PR I42D,I44V to a 0.6-kb *Eco*RI/*Acc*I fragment from PR M73V, A100L (4). Clones were screened by digestion with *Asp*I (I42D,I44V), *Sfu*I (M73V), and *Spe*I (A100L). Finally, PR I42D,I44V,M73V,A100L, V104T,R105P,G106V, S107N was constructed by a three-way ligation with a 2.3-kb *Bgl*II/*Bss*HII fragment from PR, a 0.6-kb *Bgl*II/*Alu*I fragment from PR I42D,I44V, and a 0.3-kb *Alu*I/*Bss*HII fragment from PR V104T,R105P, G106V,S107N (4). Clones were screened for the presence of an *Asp*I site (I42D,I44V), *Sfu*I and *Spe*I sites (M73V,A100L), and an *Age*I site (V104T,R105P,G106V,S107N). The protease gene from the above construct was also amplified using PCR with an upstream primer that included a *Bam*HI site followed by a nucleotide sequence coding for a bovine factor Xa cleavage site. The downstream primer included a *Hind*III site at the carboxyl end of the protease sequence. This product was then cloned into the expression plasmid p6HRT (S. F. J. LeGrice, Case Western Reserve University, Cleveland, OH) between the *Bam*HI and *Hind*III sites and replaced the HIV reverse transcriptase gene with the mutant RSV protease gene. The S38T mutation was then added to this construct using standard PCR mutagenesis procedures (4). Mutations were confirmed by sequencing. The final RSV(S9) has the following substitutions: PR S38T, I42D, I44V, M73V, A100L, V104T, R105P, G106V, and S107N. The I42D,I44V mutations were also introduced into the upstream or amino-terminal subunit of the dimer by ligating together the 220-base pair *Stu*I-*Eco*RI fragment from the wild-type RSV covalently linked PRGGGGPR expression plasmid, the 2599-base pair *Eco*RI-*Acc*I fragment from the I42D,I44V monomer expression plasmid, and the 820-base pair *Stu*I-*Acc*I fragment from the linked dimer expression plasmid. The I42D,I44V mutation was introduced into the downstream subunit by using PCR techniques to amplify the entire I42D,I44V gene from the monomer expression plasmid. The upstream oligodeoxynucleotide for amplification contained a *Stu*I restriction site (5'-CCC GCCG CCG GAG GCCTAG CCG ATGACAATGGAACATAAAGAT-3'). The downstream oligodeoxynucleotide annealed at position 700 on the monomer expression plasmid (5'-AGCTGTGACCGTCTC-3'). This PCR amplified product was digested with *Bss*HII and *Stu*I and ligated into the wild-type linked-dimer plasmid between the *Bss*HII and *Stu*I sites.

**Purification of AMV and HIV-1 Proteases**—AMV protease was purified directly from virus particles as described previously (10). Intact AMV was obtained from Molecular Genetic Resources (Tampa, FL). HIV-1 protease was expressed in *Escherichia coli* and refolded from the inclusion body fraction as described previously (11).

**Purification of Wild-type and Mutant Linked-dimer RSV Proteases**—Bacterially expressed RSV wild-type and mutant linked-dimer, as well

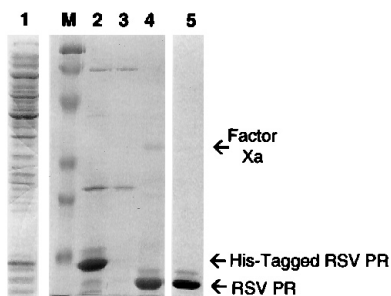
as the I42D,I44V nonlinked homodimer proteases, were prepared from *E. coli* MC1061 transformed with the temperature-sensitive  $\lambda$ -cI repressor plasmid, pRK248cIts, as described previously (12).

**Purification of Soluble Bacterially Expressed Wild-type and Mutant RSV Proteases**—RSV wild-type and RSV(S9) histidine-tagged proteases were expressed in *E. coli* M15 pDM1.1. Cells were grown at 30 °C in 4 liters of YT media, pH 7.5, to  $A_{600} = 0.6$ . Isopropyl- $\beta$ -D-thiogalactopyranoside was added to 0.5 mM and protein induced in cells for 2.5 h. Bacterial cells were pelleted, washed in 10 mM Tris, pH 8.0, 10 mM EDTA, and suspended in 10 mM HEPES, pH 8.3. Cells were lysed by addition of lysozyme to 67  $\mu$ g/ml and viscosity was reduced by incubation with DNase I to 33  $\mu$ g/ml in the presence of 4 mM  $MgCl_2$  at room temperature for 30 min. Cell debris was pelleted by centrifugation, and the clarified supernatant was passed over a 1-ml Ni-NTA (Qiagen, Chatsworth, CA) column equilibrated with 10 mM HEPES, pH 8.3, and washed with 10 mM HEPES, pH 8.3, 30 mM imidazole, 10% glycerol. Protease was then eluted with 10 mM HEPES, pH 8.3, 250 mM imidazole. Imidazole was removed by dialysis against 2 liters of 10 mM HEPES, pH 8.3, 1 mM  $\beta$ -mercaptoethanol at 4 °C overnight. The histidine tag was removed from the fusion protein by treatment with 100  $\mu$ g of bovine factor Xa (Hematological Technologies, Essex Junction, VT) in 10 mM HEPES, pH 8.3, 0.1 M NaCl, and 1 mM  $CaCl_2$ . Cleavage was monitored by separation of the proteins by SDS-PAGE and was usually complete after 6-h incubation at room temperature. At this point the

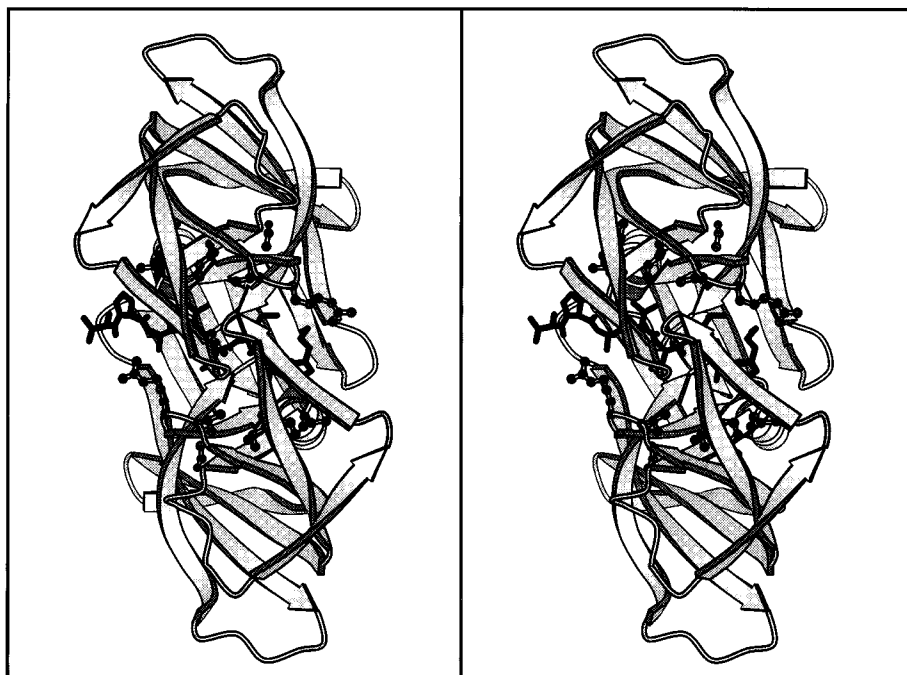
RSV protease was maximally active and possessed the native sequence at its amino terminus. The native PR was purified by passing the digested enzyme preparation through a fresh Ni-NTA column, equilibrated with 10 mM HEPES, pH 8.3. The cleaved PR now eluted from the affinity resin with 30 mM imidazole. This eluted protease was contaminated with factor Xa, which was subsequently removed by treatment with immobilized benzamidine (Pharmacia Biotech Inc., Uppsala, Sweden). Alternatively, biotinylated Factor Xa (Boehringer Mannheim) was removed with immobilized streptavidin. The final yield of active purified PR is about 2 mg. The RSV protease is greater than 95% pure as judged by SDS-PAGE (Fig. 2). Note that viral AMV and bacterially expressed RSV protease differ in primary structure by only two amino acids and are biochemically indistinguishable.

## RESULTS AND DISCUSSION

**Bacterial Expression and Purification of Soluble RSV PR**—We sought to develop a method of purifying active bacterially expressed RSV PR that could be used for substrate specificity and structural studies. The earlier method of purifying PR from inclusion bodies yielded pure enzyme, but had a relatively low specific activity. While this activity was sufficient for specificity studies, these preparations did not crystallize. To avoid denaturation and refolding of PR, we developed a method to purify recombinant PR from the soluble fraction of bacterial cell extracts. A polyhistidine sequence was added to the amino terminus of the PR to allow for efficient and rapid purification of enzyme (Fig. 2, lane 2). The specific activity of the histidine-tagged fusion PR was 5% that of AMV PR isolated directly from virions. However, when the polyhistidine sequence was removed by treatment with factor Xa (Fig. 2, lane 4), leaving the native  $NH_2$ -terminal sequence, active RSV protease was obtained with a specific activity 75–100% that of the enzyme purified from virus. Factor Xa was removed from the PR preparation by affinity chromatography using immobilized benzamidine as described under “Experimental Procedures.” The resultant protein is greater than 95% pure as analyzed by SDS-PAGE (Fig. 2, lane 5). One contaminating protein band was detected that migrates slower than PR; it may have resulted from limited cleavage of PR at an alternate internal factor Xa site or could represent a bacterial protein. The presence of this band does not affect protease activity, and these preparations are of a sufficient purity to form crystals suitable for structural studies. A second construct was made which



**FIG. 2. Bacterial expression and purification of RSV PR.** RSV PR (wild-type) was expressed and purified as described under “Experimental Procedures.” Soluble extract from induced cells (lane 1) was passed over a Ni-NTA column. PR was eluted with a 250 mM imidazole wash (lane 2). The  $NH_2$ -terminal histidine tag was removed by treatment with bovine factor Xa. After treatment, the PR was applied to the Ni-NTA column which was washed with buffer (lane 3) and then 30 mM imidazole wash (lane 4). Factor Xa was then removed by treatment with immobilized benzamidine (lane 5). Protein molecular mass markers 143, 97, 50, 35, 30, and 22 kDa are shown (lane 1). The molecular mass of native RSV PR monomer is 13.5 kDa.



**FIG. 3. Stereo representation of the RSV PR containing nine amino acid substitutions that influence the selection of substrate and catalytic rate.** Amino acids important for substrate selection and catalytic rate by RSV protease are shown by ball and stick structures in a crystal structure model of a homodimeric enzyme. These include RSV PR positions 38, 42, 44, 73, 100, 104, 105, 106, and 107. The RSV NC-PR peptide substrate is shown in the active site with thick lines. The figure was prepared with MOLSCRIPT (15).

TABLE I  
Comparison of steady state kinetic activity by AMV, HIV-1, and RSV(S9) proteases on peptide substrates

Varying concentrations of peptide substrates were incubated with different PRs, and the extent of cleavage was determined using the fluorescence assay described under "Experimental Procedures."

Protease	Substrate <sup>a</sup>	$K_m$	$k_{cat}$	$k_{cat}/K_m$
		$\mu\text{M}$	$\text{min}^{-1}$	$\text{min}^{-1} \mu\text{M}^{-1}$
AMV	RSV NC-PR	67	25	0.4
HIV-1	RSV NC-PR	6	14	2.3
RSV (S9)	RSV NC-PR	6	43	7.2
HIV-1	HIV RT-IN	32	161	5.0
	HIV CA-NC <sub>b</sub>	30	207	6.9
	HIV CA-NC <sub>a</sub>	44	50	1.1
	HIV NC-p6a	100	33	0.3
RSV (S9)	HIV RT-IN	111	492	4.4
	HIV CA-NC <sub>b</sub>	132	179	1.4
	HIV CA-NC <sub>a</sub>	366	98	0.3
	HIV NC-p6a	283	17	0.1

<sup>a</sup> The amino acid sequences of the peptide substrates are as follows: RSV NC-PR, **PPAVS-LAMTMR**R; HIV-1 RT-IN, **PARVL-FLDGRR**; HIV-1 CA-NC<sub>b</sub>, **PATIM-MQRERR**; HIV-1 CA-NC<sub>a</sub>, **PARVL-AEAMRR**; HIV-1 NC-p6a, **PGNF-LQSR**R. These amino acid sequences are presented in one letter notation with the amino terminus to the left. Specific PR cleavage sites are indicated by hyphens with the natural sequence in bold letters. Nonboldface Arg residues were added to the substrate to improve solubility, and a Pro residue was added to the amino terminus to prevent the starting substrate from reacting with fluorescamine. Neither modification affects PR activity.

expresses an altered RSV PR, referred to as RSV PR(S9). This PR has nine substituted residues, including S38T, I42D, I44V, M73V, A100L, V104T, R105P, G106V, and S107N. The RSV(S9) PR was expressed, purified from the soluble fraction, and activated as described for the wild type PR.

**Changing the Specificity of the RSV PR**—We have combined nine mutations into the RSV PR gene that introduce specific amino acid substitutions into the substrate binding pocket. The position of these residues relative to the RSV NC-PR peptide substrate is shown in Fig. 3. The S38T substitution was added to the PR, since we have confirmed that this mutation, by itself, increased the rate of catalytic activity of PR about 2-fold. When combined into a single enzyme, these multiple substitutions allow the protease to cleave four HIV-1 peptide substrates at initial rates more than 100-fold greater than wild-type RSV PR. Steady state kinetic data obtained from substrate saturation curves show that the RSV(S9) protease has  $K_m$  values determined with the HIV-1 PR substrates that are 2–8-fold higher than the corresponding values for HIV-1 PR. The  $K_m$  values for HIV-1 PR ranged from 30 to 100  $\mu\text{M}$ , while those for RSV(S9) PR ranged between 111 and 366  $\mu\text{M}$  for the same substrates (Table I). Note that the  $K_m$  values for AMV PR on the HIV-1 PR substrates could not be measured, because these peptides are not cleaved to any significant extent. Sites of cleavage were established by  $\text{NH}_2$ -terminal amino acid analysis of the product peptides (Ref. 4 and data not shown). The overall catalytic efficiencies ( $k_{cat}/K_m$ ) of RSV(S9) PR are 1.1–4.9-fold lower than those determined with the HIV-1 PR. However, substrates that are cleaved most efficiently by HIV-1 protease are also those that are cleaved most efficiently by the RSV(S9) mutant. With the RSV NC-PR reference substrate, the RSV PR(S9) enzyme resembles the HIV-1 PR more than the parental RSV enzyme. Its specific activity on this substrate is approximately 2.5 times greater than the wild-type enzyme. Its overall catalytic efficiency ( $k_{cat}/K_m$ ) is about 10 times higher due primarily to a sharp decrease in the  $K_m$  for the NC-PR substrate. These results demonstrate that the modest specificity changes introduced by each of the single mutations can be combined to produce a highly active enzyme with a signifi-

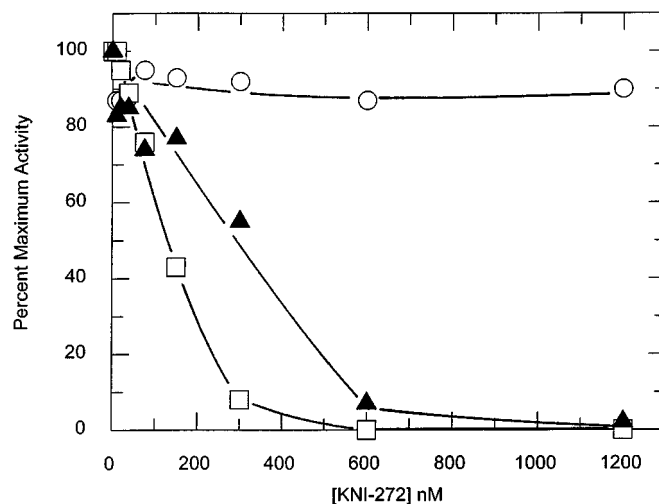


FIG. 4. Inhibition of AMV, RSV(S9), and HIV-1 PRs by the HIV-1 PR inhibitor KNI-272. Ten ng of AMV (○), RSV(S9) (▲), or HIV-1 PR (□) was incubated with 100  $\mu\text{M}$  NC-PR peptide and various concentrations of KNI-272. Activity is expressed as a percentage of activity in the absence of inhibitor.

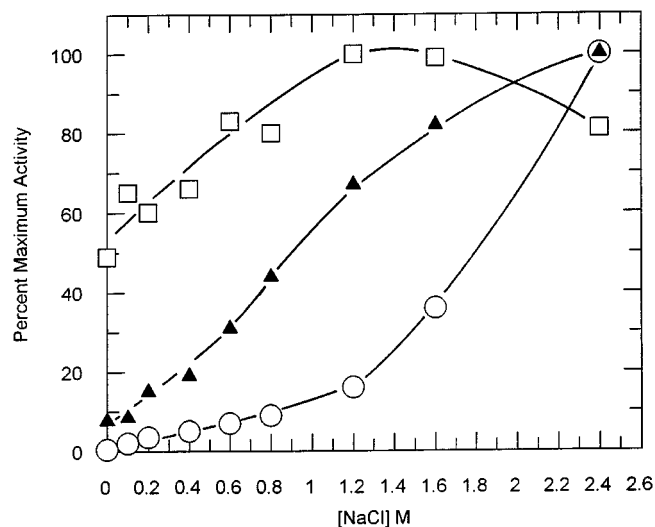


FIG. 5. Effect of salt on the activity of AMV, RSV(S9), and HIV-1 PRs. PR activity was measured as described in the legend to Fig. 4, except that the amount of the NaCl in the assay was varied as indicated. □, HIV-1 PR; ▲, RSV(S9) PR; ○, AMV PR.

cantly altered substrate preference. Furthermore, by combining all nine amino acid substitutions into one enzyme, cleavage of the HIV-1 RT-IN, CA-NC<sub>a</sub>, and NC-p6a substrates was considerably greater than that observed with enzymes containing the same mutations as either single or double substitutions (4). This is to be expected, as many of these residues contribute to more than one enzyme subsite and all have to be changed in order to affect a complete change in specificity.

There are two additional indicators that the RSV(S9) PR has substantial HIV-1 PR character. First, it is inhibited effectively by the previously described nanomolar HIV-1 PR inhibitor, KNI-272 (14). At inhibitor concentrations which completely block activity of both HIV-1 PR and RSV(S9) PR, wild-type AMV PR is unaffected (Fig. 4). Second, the salt dependence of RSV(S9) PR is closer to that observed with HIV-1 PR than that with RSV PR (Fig. 5). For instance, in the presence of 1 M NaCl, HIV-1 PR has 100%, RSV(S9) PR has greater than 60%, and the AMV PR has less than 20% of their respective maximal activities with the NC-PR peptide substrate.

While RSV(S9) PR displays substantial HIV-1 PR-like sub-

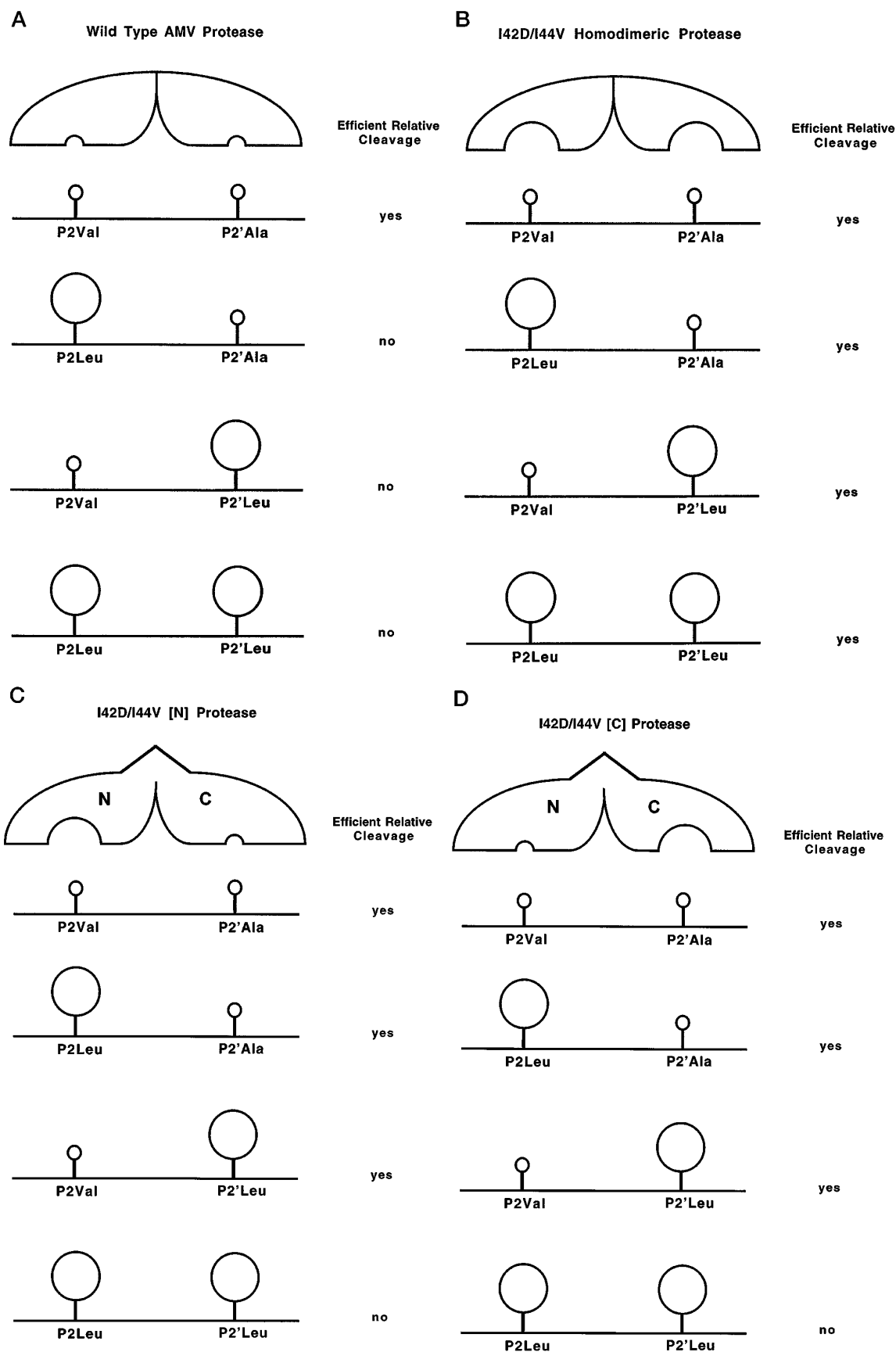


FIG. 6. Schematic representation of the effects of asymmetric S2 subsite PR mutations on cleavage of NC-PR peptide substrates with leucine substitutions in the P2 and/or P2' positions. The noncovalently linked PR is indicated at the top of the figure by the presence of two separated subunits (A and B). The covalently linked PR is indicated by subunits connected with the triangle at the top (C and D). The wild-type S2 and S2' subsites are represented by the small half-circles. S2 and S2' subsites with the I42D, I44V specificity altering substitutions are represented by the large half-circles. A diagrammatic representation of the NC-PR peptide substrate is depicted below with wild type residues in P2 and P2' denoted by small circles and Leu substitutions by the large circles.

strate specificity and kinetics, complete conversion to HIV-1 PR specificity has not been reached. This is seen by differences in activity between HIV-1 and RSV(S9) proteases with HIV-1 substrates ( $k_{cat}/K_m$  values in Table I), differences in effectiveness of the KNI-272 inhibitor (Fig. 4), and salt dependence for activity (Fig. 5). One of several possible reasons for this is that the nine amino acids substituted into the RSV PR influence substrate amino acid selection primarily in six of the eight enzyme subsites, S3 to S3'. These substitutions have limited influence on substrate selection in the S4 and S4' subsites. An additional mutation that deletes RSV PR residues 61–63, at the base of the enzyme flaps, alters preference for amino acids interacting with the S4 and S4' subsites to resemble that of HIV-1 PR (4). These residues are unique to the RSV enzyme, which has larger flaps than the HIV-1 PR. Unfortunately, when this deletion was combined with other RSV PR mutations, it produced an inactive enzyme. It seems likely that the removal of these residues caused a conformational change which was not tolerated in the context of the other mutations. Additional changes in the RSV(S9) PR will probably have to be made in order to accommodate the S4 and S4' deletions.

Analysis of RSV(S9) PR activities provides some insight into protease substrate recognition. One can explain the varied steady state kinetic data with different HIV-1 substrates in Table I by the fact that the peptide substrates each have markedly different amino acid sequences, and the (S9) mutations do not affect cleavage at each site equally. Thus, strong interactions between the enzyme and the CA-NCb, CA-NCA, and NC-p6a substrates may depend not only on differences in the S3 to S3' subsites, but also on the S4 and S4' subsites that were not altered in the RSV PR(S9). In contrast, RT-IN and inhibitor KNI-272 interactions do not seem to depend on changes in the S4 and S4' subsites.

**Asymmetry Is Introduced in the PR Homodimer by Substrate Binding**—The active retroviral PR is a homodimer. Therefore, each of the amino acid substitutions that has been shown to influence selection has been, in effect, a double substitution that alters specificity of at least two subsites formed by the two subunits on opposite sides of the enzyme. In order to examine the effect of substitutions in individual subunits, a catalytically active protease dimer covalently linked by four glycine residues between the carboxyl terminus of one subunit and the amino terminus of the second subunit was prepared as described previously (12). This construct was then used to introduce the double substitution, I42D,I44V, into the separate PR subunits designated N or C for NH<sub>2</sub>-terminal half or COOH-terminal half as depicted in Fig. 6, A–D. These substitutions were chosen because they confer a change in specificity for substrate amino acids at the S2 and S2' subsites. A homodimeric, noncovalently linked enzyme with the same amino acid substitutions in both subunits is capable of cleaving efficiently RSV NC-PR peptide substrates containing Leu substituted for either P2 Val or P2' Ala, whereas the wild-type enzyme cleaves these substrates poorly (Fig. 7). The I42D,I44V(N) mutation introduces the appropriate changes into the coding sequences of the upstream subunit in the plasmid that expresses the linked-dimer. As expected, the resultant enzyme is capable of cleaving the NC-PR peptide substrate with the large Leu in P2 (Fig. 7), because the corresponding S2 subsite is made larger by the mutations. Additionally, the I42D,I44V(N) dimer also cleaves a peptide with Leu at the P2' position efficiently. These results suggest that a subsite containing a single mutation can accommodate either the P2 or the P2' amino acids allowing for the cleavage of a substrate bound in two different orientations, as depicted in Fig. 6, A–D. To verify that the enzyme is functionally symmetric, the same I42D,I44V substitutions were intro-

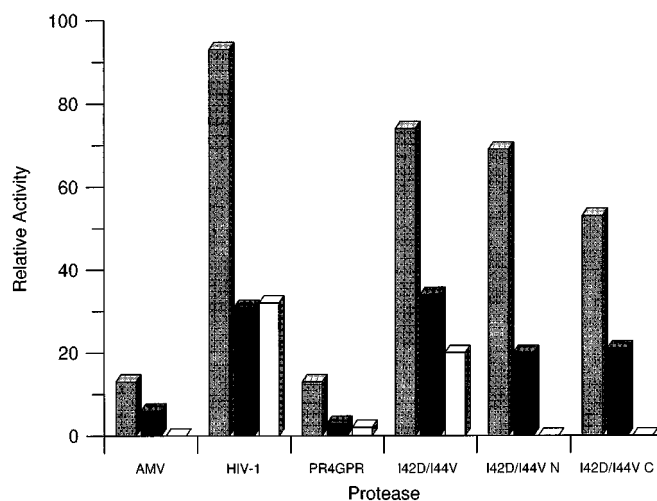


FIG. 7. Effects of S2 subsite substitutions in different subunits of the PR dimer on cleavage of NC-PR substrates with amino acid substitutions in the P2 and/or P2' positions. Changes in protease substrate preference caused by substitutions in the S2 enzyme subsite in either one or both subunits was determined. Enzyme activity was measured using a RSV NC-PR peptide substrate with Leu substituted in P2 (PPALS-LAMTMRR) (gray boxes), in P2' (PPAVS-LLMT-MRR) (black boxes), or in both P2 and P2' (PPALS-LLAMTMRR) (white boxes). Activity is expressed as a percentage relative to the initial rate of cleavage of the wild-type NC-PR peptide substrate and was measured using the fluorescamine assay (4). PRGGGPR is a wild-type RSV PR that has the two subunits of the homodimer linked with four Gly residues. RSV PR (I42D,I44V) is a noncovalently linked protease homodimer with substitutions in the S2 and S2' subsites. RSV PRGGGGPR (I42D,I44V(N)) and RSV PRGGGGPR (I42D,I44V(C)) are covalently linked homodimers with asymmetric substitutions in either the amino or carboxyl subunits as indicated by the N and C, respectively. The activity data are summarized in Fig. 6, A–D.

duced into the carboxyl-terminal half of the linked-dimer. The enzyme produced from this clone also cleaves both the P2 Leu and the P2' Leu-modified substrates with efficiencies equal to those seen with the I42D,I44V(N) mutant (Fig. 7). In contrast, a substrate that includes Leu in both the P2 and P2' positions is cleaved efficiently only by the noncovalently linked PR that contains the 42 and 44 substitutions in both S2 and S2' simultaneously. Taken together, these results indicate that the RSV protease homodimer is symmetric when not bound to a substrate and that each enzyme subunit is capable of binding equally to either half of an asymmetric substrate. This is consistent with structural data which show that substrate-based inhibitors can occur in two different conformations when bound to the HIV-1 PR (13).

The results presented here demonstrate that amino acids can be substituted at key residues in most of the enzyme subsites that alter RSV PR specificity. The present construct combined nine separate amino acid substitutions and produced an RSV PR that cleaves peptide substrates representing HIV-1 gag and pol gene cleavage sites. Furthermore, by using a covalently linked PR dimer, we have demonstrated that key amino acid residues can be changed in separate subunits to produce an asymmetrically substituted enzyme. This strategy can be used to create a PR, targeted to a new protein sequence, in which each enzyme subunit is custom designed to bind efficiently to one-half of the new substrate sequence. The resulting enzyme will prefer one substrate orientation over the reverse orientation. Finally, examination of structural differences between similar yet catalytically unique enzymes has advanced our understanding in a way that would not have been possible if each had been examined individually. This general approach can be extended to many other protein families to

identify key enzyme residues that mediate specific protein-substrate and protein-protein interactions.

*Acknowledgments*—We thank C. Giam for a HIV-1 PR expression clone and purification procedure, Ramothea Webster for technical assistance, E. Houts from Molecular Genetic Resources for AMV protease, and S. F. J. LeGrice for the expression plasmid p6HRT and *E. coli* strain M15 pDM1.1. KNI-272 was a generous gift from Dr. John Erickson, NCI-Frederick Cancer Center.

#### REFERENCES

- Jaskolski, M., Miller, M., Rao, J. K. M., Leis, J., and Wlodawer, A. (1990) *Biochemistry* **29**, 5889–5898
- Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., and Kent, S. B. H. (1989) *Science* **245**, 616–621
- Cameron, C., Grinde, B., Jacques, P., Jentoft, J., Leis, J., Wlodawer, A., and Weber, I. (1993) *J. Biol. Chem.* **268**, 11711–11720
- Cameron, C., Ridky, T., Shulenin, S., Leis, J., Weber, I., Copeland, T., Wlodawer, A., Burstein, H., Bizub-Bender, D., and Skalka, A. (1994) *J. Biol. Chem.* **269**, 11170–11177
- Grinde, B., Cameron, C., Leis, J., Weber, I., Wlodawer, A., Burstein, H., Bizub, D., and Skalka, A. (1992) *J. Biol. Chem.* **267**, 9481–9490
- Grinde, B., Cameron, C., Leis, J., Weber, I., Wlodawer, A., Burstein, H., and Skalka, A. (1992) *J. Biol. Chem.* **267**, 9491–9498
- Sedlacek, J., Fabry, M., Coward, J., Horejsi, M., Strop, P., and Luftig, R. (1993) *Virology* **192**, 667–672
- Condra, J., Schleif, W., Blahy, O., Gabryelski, L., Graham, D., Quintero, J., Rhodes, A., Robbins, H., Roth, E., Shivaprakash, M., Titus, D., Yang, T., Teppler, H., Squires, K., Deutsch, P., and Emmini, E. (1995) *Nature* **374**, 569–571
- Ridky, T., and Leis, J. (1995) *J. Biol. Chem.* **270**, 29621–29623
- Alexander, F., Leis, J., Soltis, D., Crowl, R., Danho, W., Poonian, M., Pan, Y., and Skalka, A. (1987) *J. Virol.* **61**, 534–542
- Ridky, T. W., Cameron, C. E., Cameron, J., Leis, J., Copeland, T., Wlodawer, A., Weber, I., and Harrison, R. (1996) *J. Biol. Chem.* **271**, 4709–4717
- Bizub, D., Weber, I., Cameron, C., Leis, J., and Skalka, A. (1991) *J. Biol. Chem.* **266**, 4951–4958
- Fitzgerald, P., McKeever, B., VanMiddlesworth, J., Springer, J., Heimbach, J., Leu, C., Herber, W., Dixon, R., and Darke, P. (1991) *J. Biol. Chem.* **265**, 14209–14219
- Gulnik, S. V., Suvorov, L., Liu, B., Yu, B., Anderson, B., Mitsuya, H., and Erickson, J. W. (1995) *Biochemistry* **34**, 9282–9287
- Kraulis, P. (1991) *J. Appl. Crystallogr.* **24**, 946–950

## **Programming the Rous Sarcoma Virus Protease to Cleave New Substrate Sequences**

Todd W. Ridky, Diane Bizub-Bender, Craig E. Cameron, Irene T. Weber, Alexander Wlodawer, Terry Copeland, Anna Marie Skalka and Jonathan Leis

*J. Biol. Chem.* 1996, 271:10538-10544.

doi: 10.1074/jbc.271.18.10538

---

Access the most updated version of this article at <http://www.jbc.org/content/271/18/10538>

### 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 15 references, 10 of which can be accessed free at <http://www.jbc.org/content/271/18/10538.full.html#ref-list-1>