Mechanism of Inhibition of the Retroviral Protease by a Rous Sarcoma Virus Peptide Substrate Representing the Cleavage Site between the *gag* p2 and p10 Proteins*

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

From the Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106-4935

Irene T. Weber

*From the Department of Pharmacology, Jefferson Cancer Institute, Thomas Jefferson University,
Philadelphia, Pennsylvania 19107*

Terry D. Copeland and Alexander Wlodawer

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

The activity of the avian myeloblastosis virus (AMV) or the human immunodeficiency virus type 1 (HIV-1) protease on peptide substrates which represent cleavage sites found in the *gag* and *gag-pol* polyproteins of Rous sarcoma virus (RSV) and HIV-1 has been analyzed. Each protease efficiently processed cleavage site substrates found in their cognate polyprotein precursors. Additionally, in some instances heterologous activity was detected. The catalytic efficiency of the RSV protease on cognate substrates varied by as much as 30-fold. The least efficiently processed substrate, p2-p10, represents the cleavage site between the RSV p2 and p10 proteins. This peptide was inhibitory to the AMV as well as the HIV-1 and HIV-2 protease cleavage of other substrate peptides with K_i values in the 5–20 μM range. Molecular modeling of the RSV protease with the p2-p10 peptide docked in the substrate binding pocket and analysis of a series of single-amino acid-substituted p2-p10 peptide analogues suggested that this peptide is inhibitory because of the potential of a serine residue in the P1' position to interact with one of the catalytic aspartic acid residues. To open the binding pocket and allow rotational freedom for the serine in P1', there is a further requirement for either a glycine or a polar residue in P2' and/or a large amino acid residue in P3'. The amino acid residues in P1–P4 provide interactions for tight binding of the peptide in the substrate binding pocket.

Translation of avian retrovirus genomic RNA results in the production of two polyprotein precursors, Pr76[gag] and Pr180[gag-pol]. Pr76[gag] encodes the structural proteins required for virus particle assembly including the MA,[1] p2, p10, CA, NC, and PR proteins, listed in order from the amino-terminal end of the precursor. MA is responsible for the attachment of polyprotein precursors to the cell membrane (1) and forms the matrix between the lipoprotein envelope and the inner capsid core of the virus particle (2). It can also bind weakly to viral RNA (3, 4). The p2 and p10 proteins have no known function. CA forms the inner capsid core of the particle (5), and NC is required for packaging of viral specific RNA in the nucleoprotein core (6, 7). Protease is responsible for the processing of the Pr76[gag] and Pr180[gag-pol] polyprotein precursors (8). Pr180[gag-pol] contains reverse transcriptase and integrase, in addition to *gag*-encoded structural proteins.

The protease-dependent process of cleavage of the *gag* and *pol* polyproteins releases the mature forms of the structural and enzymatic proteins required for retrovirus replication (5). This maturation process appears to be regulated both spatially and temporally. Activation of the protease appears to occur in nascent virus particles with processing occurring normally either coincident with virus particle assembly at the cell membrane or soon after budding of virus particles. Loss of protease activity results in the production of immature, noninfectious virus (9–11). Since protease activity is required for the production of infectious virus, this enzyme is a target for the design of antiviral agents.

We have compared the activity of the AMV and HIV-1 PRs on peptide substrates representing cleavage sites found in the RSV and HIV-1 polyprotein precursors. We have identified a peptide corresponding to a cleavage site present in the amino-terminal region of the RSV polyprotein precursor, which acts both as a weak substrate for the AMV PR and an effective competitive inhibitor of AMV, HIV-1, and HIV-2 PRs. In this report, we propose a mechanism for inhibition of the AMV PR by this RSV *gag* p2-p10 cleavage-site peptide.

EXPERIMENTAL PROCEDURES

Purification of Retroviral Proteases—AMV PR, purified from virus as described (12), was obtained from Molecular Genetic Resources,

Inhibition of RSV and HIV-1 Proteases by RSV Peptide Substrate

Tampa, FL. Note that the RSV and AMV PRs, which differ in primary structure by 2 amino acids, are biochemically indistinguishable and have been used interchangeably in these studies. HIV-1 PR expressed in bacteria was a generous gift of Dr. Joe Giam, Case Western Reserve University. Purified HIV-2 PR, also expressed in bacteria, was a generous gift of Dr. Stuart F. J. LeGrice, Case Western Reserve University.

Peptides—The peptides used in this study were synthesized and purified as described by Copeland and Oroszlan (13). Lyophilized peptides were solubilized in 1 mM 2-mercaptoethanol, and peptide concentrations were determined by amino acid composition analysis. Specificity of cleavage was established by isolation of cleavage products by reverse-phase HPLC and determination of their amino acid compositions (14) or direct amino-terminal analysis of product peptides without separation (15).

Assay of Protease Activity—The reaction mixture contained 100 mM sodium phosphate, pH 5.9, 2.4 M sodium chloride (1 M sodium chloride was used for HIV-1 and HIV-2 PR reactions), 10–400 μM peptide as indicated, and 0.5–5 μg/ml protease. Reaction volumes were 25 μl. Reactions were initiated by the addition of protease and stopped by the addition of 175 μl of 0.5 M sodium borate, pH 8.5. Twenty μl of 0.05% (w/v) fluorescamine was then added. The time of incubation was varied as indicated. HIV-1 PR was never incubated more than 8 min because of the instability of this enzyme, presumably the result of autodegradation. This is not a problem with HIV-2 PR since the labile site present in this enzyme was mutated and can no longer be processed.[2] Fluorescamine reacts specifically with primary amines, and its relative fluorescence intensity was determined on a Perkin-Elmer Cetus Instruments LS-5B spectrofluorometer. The excitation wavelength was 386 nm, and the emission wavelength was 477 nm. Both excitation and emission slits were set to 10 nm. Relative fluorescence intensity was converted to nmol of product using a standard curve described by the following equation: nmol of product = relative fluorescence intensity/313. The standard curve was obtained using a hexapeptide with a free amino terminus (17). The concentration of this peptide was determined by amino acid composition analysis. The peptides used in this study were designed with prolines at their amino termini. Since prolines do not react with fluorescamine, the relative fluorescence intensity represents only the newly formed amino terminus present on the carboxyl-terminal peptide fragment. Peptides containing lysine residues did not cause background problems since the fluorescamine reactions were carried out at pH 8.5, 2 pH units below the pK_a of the ϵ-amino group.

Steady-state Kinetic Analysis—Kinetic constants were determined using the assay described above. Concentrations of peptides ranged from 0.25 to 4 times the K_m value. No more than 20% of the substrate was allowed to be consumed during the course of any given experiment. Initial velocity data used to calculate kinetic constants were obtained from at least three experiments performed in duplicate. Kinetic constants were determined by a nonlinear fit of the data to the Michaelis-Menten equation using the NFIT program (17). Correlation coefficients of the fit were greater than 0.98. The standard deviation of the constants reported was no greater than 20%.

Determination of K_i for the RSV p2-p10 Peptide Substrate—Dixon analysis was used to determine the K_i of the RSV p2-p10 peptide substrate with all proteases studied (18). K_i values for the other peptide substrate/inhibitors were determined from IC_{50} values obtained from dose-response curves using the following equation (16)

$$K_i = \frac{IC_{50} - 0.5[E_t]}{1 + [S]/K_m} \qquad \text{(Eq. 1)}$$

where $[E_t]$ and $[S]$ are the protease and substrate concentrations, respectively, used in the assay.

Molecular Modeling of the RSV p2-p10 Peptide in the Active Site of the RSV PR—Molecular modeling used the structure of the RSV protease dimer with modeled flaps and substrate as described previously (19) except that the substrate residues were altered to those of the p2-p10 peptide. The side chains of the p2-p10 peptide were initially positioned to correspond to those in co-crystal structures of HIV-1 PR with inhibitors (20–24). The side chain of serine at P1' was then rotated toward the Asp-37. This moved the serine hydroxyl group to within hydrogen bonding distance of the side chain oxygen of the catalytic Asp-37.

RESULTS

Protease Activity on Homologous Peptide Substrates—Peptide substrates of 9–12 amino acids in length, which represent cleavage sites in the *gag* and *gag-pol* polyprotein precursors of the RSV and of the HIV-1, were synthesized. These peptides include the 7 amino acids, P4-P3', that form the minimum sequence required for efficient and specific cleavage by the retroviral protease (25, 26). These residues interact with a unique array of amino acids in the protease which form the subsites S4-S3' of the substrate binding pocket. The P4-P3' and S4-S3' nomenclature is that established by Schecter and Berger (27).

The ability of the AMV or the HIV-1 PR to cleave peptide substrates derived from their cognate polyprotein precursors is presented in Tables I and II, respectively. Each protease was able to specifically process its respective homologous set of peptide substrates. The catalytic efficiency (k_{cat}/K_m) of the AMV PR for the RSV peptide substrates varied over a 6-fold range for the peptides representing the CA-NCa, CA-NCb, NC-PR, PR-RT, and IN-p4 cleavage sites. A peptide representing the p10-CA cleavage site was cleaved as efficiently as the NC-PR peptide. Since this peptide has proline in P1', which is not reactive with fluorescamine, cleavage was followed by separation of the products by HPLC as described previously (17). Two other peptide substrates, MA-p2 and p2-p10, were found to be 20–30-fold less efficiently cleaved than the most active RSV substrate. The low efficiency of cleavage of the MA-p2 peptide could be explained by a much lower

TABLE I

Activity of the AMV PR on RSV gag and pol peptide substrates

Kinetic values were obtained from Michaelis-Menten plots of the initial velocity *versus* substrate concentration. The activity was determined using the fluorescamine assay as described under "Experimental Procedures." The arrowhead indicates a cleavage site.

Substrate	Sequence[a]	K_m	k_{cat}	k_{cat}/K_m
		μM	min^{-1}	$\mu M^{-1}\ min^{-1}$
MA-p2	(P)TSCY ▼ HCGT(RR)	648	20	0.031
p2-p10	PYVG ▼ SGLY(RR)	—[b]	—[b]	0.02
CA-NCa	(P)AAAM ▼ SSAI(RR)	51	11	0.22
CA-NCb	PLIM ▼ ANVN(RR)	17	11	0.65
NC-PR	PPAVS ▼ LAMTM(RR)	47	21	0.45
PR-RT	(P)ATVL ▼ TVAL(RR)	127	20	0.16
IN-p4	PLFA ▼ GISE(RR)	171	25	0.14

[a] The parentheses indicate amino acids not part of the precursor polypeptide. The nomenclature of substrate positions (in brackets) is illustrated with the p2-p10 peptide sequence as follows: P[P4] Y[P3] V[P2] G[P1] ▼ S[P1'] G[P2'] L[P3'].
[b] Dash denotes too low of a turnover for reliable steady state kinetic analysis.

TABLE II

Activity of the HIV-1 PR on selected HIV-1 gag and pol peptide substrates

Kinetic values were obtained from Michaelis-Menten plots of the initial velocity *versus* substrate concentration. The activity was determined using the fluorescamine assay as described under "Experimental Procedures." The arrowhead indicates a cleavage site.

Substrate	Sequence[a]	K_m	k_{cat}	k_{cat}/K_m
		μM	min^{-1}	$\mu M^{-1}\ min^{-1}$
CA-NCa	(P)ARVL ▼ AEAM(RR)	52.0	256.8	4.94
CA-NCb	(P)ATIM ▼ MQRE(RR)	127.1	894.1	7.03
NC-p6a	PGNF ▼ LQSR(R)	39.6	27.0	0.68
NC-p6b	PRQAN ▼ FLGK(RR)	28.4	23.2	0.82
RT-IN	(P)RKIL ▼ FLDG(RR)	37.0	170.1	4.60

[a] The parentheses indicate amino acids not part of the precursor polypeptide.

binding affinity for protease compared with the other substrates; its K_m for the AMV PR was 648 μM. In the case of the least efficient substrate, p2-p10, the rate of cleavage was so slow that it was only possible to determine the combined k_{cat}/K_m value from the slope of the initial velocity plot.

The HIV-1 PR specifically and rapidly cleaved peptides representing *gag* and *pol* polyprotein precursor cleavage sites, as reported previously (25, 29–31). Our steady-state kinetic analysis of the HIV-1 PR with five of the nine HIV-1 cleavage-site peptide substrates is presented in Table II. A 10-fold range in catalytic efficiencies was observed for these peptides. Additionally, the catalytic efficiencies were 10–20-fold higher than those for the AMV PR toward its homologous substrates, in agreement with other reports (25, 29–31). The increased catalytic efficiency of the HIV-1 PR was caused by increased rates of catalysis; homologous substrates were bound with affinities similar to those of the AMV PR with its cognate substrates (Tables I and II). In disagreement with the reports of others (25, 29, 30) but in agreement with the experiments of Toszer *et al.* (32), we find that the K_m for the HIV-1 PR for peptide substrates is in the μM and not the mM range.

Protease Activity on Heterologous Peptide Substrates—Despite several reports that retroviral proteases are specific for their own polyprotein sequences (5), we analyzed the AMV and HIV-1 PRs for activity on heterologous substrates. The AMV PR could cleave three of the five HIV-1 peptide substrates. However, the catalytic efficiency was so low that a steady-state kinetic analysis was possible for only one substrate, RT-IN. The K_m was estimated at 98 μM and the k_{cat} at 1.3 min^{-1}. This level of activity was considerably lower than that observed for the AMV PR on its homologous substrates. In contrast, the HIV-1 PR exhibited specific processing of all RSV peptide substrates with the exception of p2-p10 (Table III). Interestingly, the catalytic efficiency (k_{cat}/K_m) of the HIV-1 PR for each of the RSV substrates showed the same variation as observed with the AMV PR, which indicates that these differences are intrinsic to the substrates and not the protease. Nevertheless, the heterologous activity of the HIV-1 PR was 10-fold or more lower than that observed for its comparable cognate substrates.

Inhibition of Protease Activity by the RSV p2-p10 Peptide Substrate—To determine if the reduced catalytic efficiency of the AMV PR on the RSV p2-p10 peptide substrate was a result of a reduced affinity or a reduced rate of catalysis, we measured the effect of adding increasing concentrations of this peptide to a processing reaction that contained the RSV NC-PR peptide substrate. As shown in Table I, the NC-PR peptide is one of the most rapidly cleaved substrates for the AMV PR. When added to a reaction mixture, the RSV p2-

TABLE III
Activity of the HIV-1 PR on RSV gag and pol peptide substrates

Kinetic values were obtained from Michaelis-Menten plots of the initial velocity *versus* substrate concentration. The activity was determined using the fluorescamine assay as described under "Experimental Procedures." The arrowhead indicates a cleavage site.

Substrate	Sequence[a]	K_m	k_{cat}	k_{cat}/K_m
		μM	min^{-1}	$\mu M^{-1} min^{-1}$
MA-p2	(P)TSCY HCGT(RR)	250	31.2	0.12
p2-p10	PYVG SGLY(RR)	ND[b]	ND[b]	
CA-NCa	(P)AAAM SSAI(RR)	15	13.5	0.9
CA-NCb	PLIM ANVN(RR)	25	35.3	1.4
NC-PR	PPAVS LAMTM(RR)	16	43.8	2.7
PR-RT	(P)ATVL TVAL(RR)	224	29.9	0.13
IN-p4	PLFA GISE(RR)	167	12.3	0.07

[a] The parentheses indicate amino acids not part of the precursor polypeptide.
[b] ND, cleavage not detected.

p10 peptide inhibited the AMV PR cleavage of the RSV NC-PR peptide substrate (Fig. 1) with a K_i of 20 μM (Table IV). The inhibition was competitive, based on the Dixon plot shown in Fig. 2, where the extrapolated lines for the three concentrations of NC-PR peptide substrate tested intersected above the x axis (18).

A control inhibition study using the RSV p10-CA peptide substrate is presented in Fig. 1. This peptide binds with similar affinity and is processed by the AMV PR at about the same rate as the RSV NC-PR peptide substrate. However, cleavage of this peptide cannot be followed by the fluoresca-
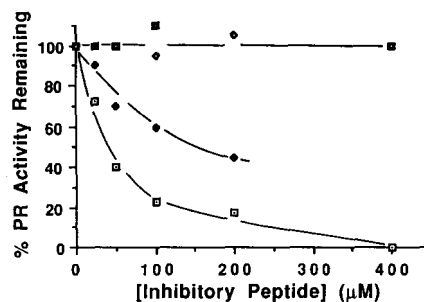

FIG. 1. **Inhibition of the AMV PR by the p2-p10 cleavage-site peptide substrate.** The NC-PR peptide substrate (200 μM) was incubated with the AMV PR (50 ng) and different concentrations of the RSV p2-p10 (□), RSV p10-CA (◆), RSV MA-p2 (■), or Moloney leukemia virus CA-NC (◇) peptide substrates. The extent of NC-PR cleavage was determined as described under "Experimental Procedures."

TABLE IV
Inhibition of AMV PR cleavage of the RSV NC-PR peptide substrate by wild type or RSV p2-p10 peptide analogues

Inhibitory peptide[a]	K_i
	μM
PYVG SGLY(RR)	20
PYVG TGLY(RR)	12.5
PYVG SSLY(RR)	22.7
PYVL SSLY(RR)	5
PYVG SNLY(RR)	20.1
PYVG SGYY(RR)	8.9
PYVG SGAY(RR)	27.1
Ac-SGLY(RR)	>200
PYVL SGLY(RR)	6
PYVW SGLY(RR)	3.3
PFVW SGLY(RR)	0.5
PHFVW SGLY(RR)	0.5

[a] The RSV p2-p10 peptide was modified in the P4-P3' positions as indicated by the underlined letter(s). Parentheses indicate amino acids not present in the RSV precursor polypeptide. The arrowhead indicates a cleavage site.
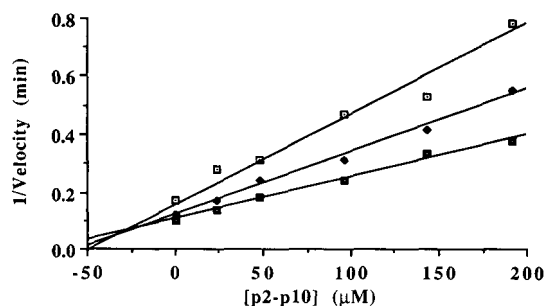

FIG. 2. **Dixon analysis of the p2-p10 peptide inhibition of the AMV PR.** The AMV PR (50 ng) was incubated as described under "Experimental Procedures" with three concentrations of the NC-PR peptide substrate as follows: 50 μM (□), 100 μM (◆), and 200 μM (■), and p2-p10 peptide in concentrations indicated.

mine assay since proline is present in the P1′ position. As expected, the addition of this peptide to a reaction mixture containing the NC-PR substrate resulted in 50% inhibition of NC-PR cleavage, as measured by fluorescence, when equimolar amounts of the two peptide substrates were present.

Also included in Fig. 1 are the results when the RSV MA-p2 peptide substrate was added to the reaction. This substrate was cleaved by the AMV PR at a very low efficiency because of its high K_m; less than 5% cleavage was observed under the conditions used in this experiment. The addition of this peptide had little effect on the cleavage of the NC-PR substrate. Similar results were obtained with the Moloney leukemia virus CA-NC peptide substrate (Fig. 1).

Cleavage of the RSV p2-p10 peptide by the HIV-1 PR was not detected (Table III). We therefore tested whether the RSV p2-p10 peptide could inhibit the cleavage of a HIV peptide (NC-p6a) by either the HIV-1 or HIV-2 PRs. These results are presented in Table V. The RSV p2-p10 peptide was found to be an effective inhibitor of these enzymes, with K_i values of 18 and 5 μM for the HIV-1 and HIV-2 PRs, respectively. Similar K_i values were obtained with the HIV-1 PR using the RSV NC-PR peptide substrate (Table V). Therefore, the RSV p2-p10 peptide is an inhibitor of all three of the retroviral proteases tested.

Modeling of the RSV p2-p10 Peptide Substrate in the Active Site of the AMV PR—To evaluate the potential interactions between residues of the RSV p2-p10 peptide substrate and the enzyme and to understand the basis of its inhibitory activity, we examined a model of the RSV PR (19) with the p2-p10 peptide substrate (Fig. 3A). This model is derived from the structure of an inhibitor bound to the HIV-1 PR (20) (see Fig. 3B). In the RSV model it was possible to rotate the side chain of the P1′ serine in the p2-p10 peptide to allow formation of a hydrogen bond between its side chain OH and the side chain δ oxygen of one of the two catalytic aspartates (Fig. 3A). We propose that formation of such an interaction would disrupt the electrostatic charge distribution required for catalysis.

Analysis of RSV p2-p10 Peptide Substrates Modified at the P1′ Position—To investigate the putative inhibitory role of the serine in the P1′ position, we synthesized a series of p2-p10 peptide analogues in which alanine, leucine, or threonine replaced serine. The first two substitutions lack a hydroxyl group, whereas the threonine, which has a hydroxyl, contains a methyl group extension on the β carbon. As shown in Table VI, substitution of the P1′ serine with an alanine or leucine allowed cleavage of the p2-p10 analogue by the AMV PR with an efficiency comparable to that of the NC-PR peptide substrate. Concomitant with this activation, both of these P1′-modified p2-p10 peptides exhibited decreased effectiveness as competitive inhibitors. In contrast, the substitution of threonine did not increase the rate of cleavage of the p2-p10 peptide. Instead, it improved its ability to act as an inhibitor (Table IV). These results are consistent with the hypothesis

TABLE V
Inhibition of the HIV-1 or 2 PRs by the wild-type and RSV p2-p10 peptide analogues

PR	Substrate	Inhibitory peptide[a]	K_i
			μM
		▼	
HIV-1	RSV NC-PR	PYVG SGLY(RR)	8.3
		PYVL SGLY(RR)	13.3
	HIV NC-p6a	PYVG SGLY(RR)	18
HIV-2	HIV NC-p6a	PYVG SGLY(RR)	5

[a] The RSV p2-p10 peptide was modified at the P1 position as indicated by the underlined letter. The arrowhead indicates a cleavage site.

that the inhibitory properties of the p2-p10 peptide are related to the presence of a hydroxyl group on a residue which occupies the P1′ position.

Analysis of RSV p2-p10 Peptide Substrates Modified at the P1 and P2′ Positions—Examination of the various RSV *gag* and *pol* cleavage sites indicates that the CA-NCa and PR-RT have serine and threonine, respectively, in the P1′ position and that NC-PR has serine in P1, yet these peptide substrates are efficiently cleaved by the AMV PR (Table I). The p2-p10 sequence differs from these by the presence of glycine residues in P1 and P2′. Glycine, which lacks a side chain, may provide greater conformational flexibility for the P1′ serine and facilitate formation of the hydrogen bond between the serine OH and the nearby catalytic aspartic acid residue. If this hypothesis is correct, then substitution of small nonpolar amino acids, such as alanine or valine, in P2′ might hinder formation of the putative Asp-Ser interaction by binding more tightly to the relatively hydrophobic S2′ enzyme subsite and thereby restrict the rotational freedom of the P1′ serine. This is predicted to activate the p2-p10 peptide for cleavage. Our results (Table VI) show that introduction of either alanine or valine in P2′ resulted in cleavage of the p2-p10 peptide analogues by the AMV PR, with efficiencies comparable to that observed with the NC-PR peptide substrate (Table VI). Concomitantly, there was a decrease in effectiveness to act as inhibitors (data not shown). In contrast, the substitution in P2′ of more polar amino acid residues, such as serine or asparagine, maintained the inhibitory properties of the analogues (Table IV) and did not activate them for cleavage. The polar residues in P2′ would be predicted to "weaken" the interaction between the P2′ position of the substrate and the nonpolar S2′ protease subsite and thus result in more rotational freedom for the serine in P1′ to form the hydrogen bond with Asp-37.

We also tested whether the glycine residue in P1 was required for the inhibitory properties of the p2-p10 peptide. In contrast to the alanine or valine substitutions in P2′, the substitution of leucine at P1 produced an effective inhibitor with a K_i of 6 μM (Table IV). This result is not totally surprising since the S1 enzyme subsite is known to accommodate large amino acid residues readily (19). A possible steric role for the glycine in P1, however, could not be excluded, since this peptide contained the glycine in P2′. This possibility was eliminated by the peptide in Table IV in which the glycine residues in P1 and P2′ were replaced with leucine and serine, respectively. This peptide proved to be an inhibitor with a K_i of 5 μM.

To determine if the serine at P1′ and the glycine at P2′ were both necessary and sufficient to confer upon a peptide the inhibitory properties observed for the p2-p10 peptide, we constructed a RSV CA-NCa peptide analogue with a glycine residue in P2′. The wild type CA-NCa peptide contains methionine, serine, and serine residues, in the P1, P1′, and P2′ positions, respectively, and is efficiently cleaved by the AMV PR (Table I). Substituting glycine for serine in P2′ resulted in a 90% reduction in its rate of cleavage. However, mixing this analogue with the NC-PR substrate had little effect upon cleavage of the latter (Table VII). Similar results were obtained with a peptide corresponding to the sequence PASVL-SGGERR which is found in the amino terminus of the HIV-1 and Moloney leukemia virus MA proteins (data not shown). Taken together, these results indicate that the serine and glycine residues at P1′ and P2′ positions, respectively, of the wild type p2-p10 peptide are necessary but not sufficient for its inhibitory activity.

Analysis of RSV p2-p10 Peptide Substrates Modified at the P1–P4 Positions—An additional element that appears to influence the inhibitory properties of the p2-p10 peptide is its
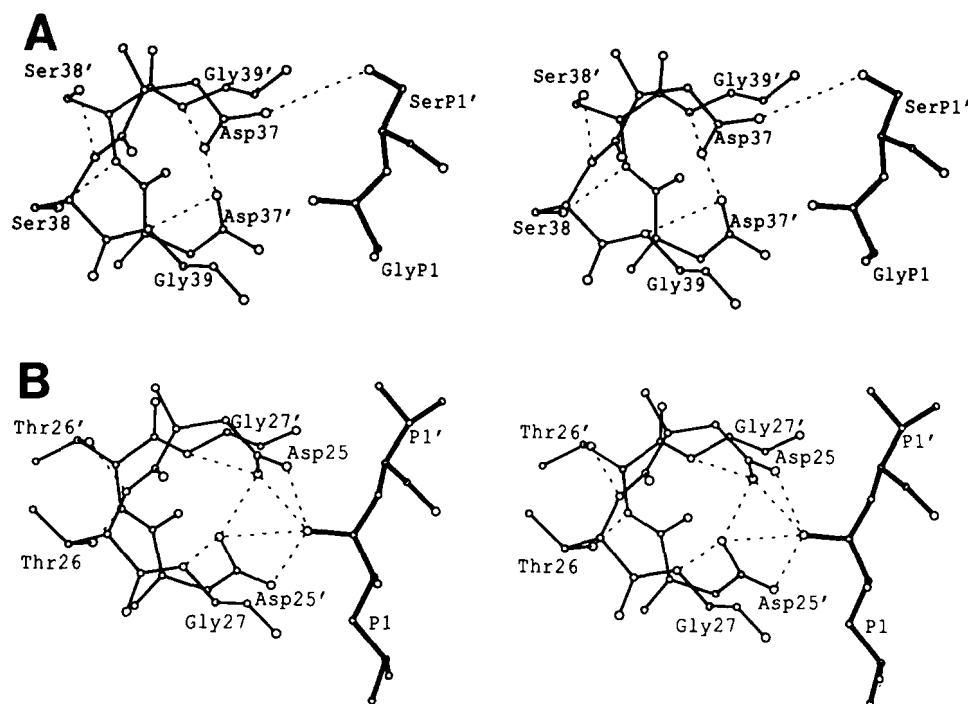
FIG. 3. *Panel A*, stereo view of P1-P1' of the p2-p10 peptide modeled in the active site of the RSV PR. The p2-p10 peptide has Gly-[CONH]-Ser at P1-P1' and is denoted by the *thick lines*. The RSV PR active site residues Asp-37, Ser-38, Gly-39, and 37'-39' from the other subunit (*thin lines*) are shown in an orientation similar to that shown in *panel B*. The P1' serine OH side chain has been rotated, as described under "Experimental Procedures," so that a hydrogen bond (*dashed line*) is possible with the δ oxygen of Asp-37. *Panel B*, stereo view of P1-P1' and the catalytic triplets of HIV-1 PR from the crystal structure with inhibitor U85548e (2). P1-P1' of the inhibitor are Leu-ψ[CH(OH)CH2]-Val, and residues Asp-25, Thr-26, and Gly-27 and 25'-27' from the two subunits in the protease dimer are shown. Hydrogen bond interactions are indicated by *dashed lines*.

TABLE VI

Activity of the AMV PR on the RSV p2-p10 peptide with amino acid substitutions in the P1' and P2' positions

Sequence[a]	K_m	k_{cat}	k_{cat}/K_m
	μM	min^{-1}	$\mu M^{-1}\ min^{-1}$
PYVG SGLY(RR)			0.02
PYVG AGLY(RR)	19.2	3.4	0.18
PYVG LGLY(RR)	22.9	7.5	0.33
PYVG SALY(RR)	13.5	9.8	0.73
PYVG SVLY(RR)	9.8	11.5	1.17

[a] The RSV p2-p10 peptide sequence was modified at the P1' and P2' positions as indicated by the underlined letters. All of these substrates were cleaved at the correct site as established by amino-terminal analysis of the products (data not shown). The arrowhead indicates a cleavage site.

TABLE VII

Inhibition of AMV PR cleavage of the RSV NC-PR peptide substrate by RSV p2-p10 and CA-NCa hybrid peptide analogues

	Inhibitory peptide[a]	K_i
		μM
p2-NC	PYVG SSAI(RR)	56.0
	PYVL SSAI(RR)	33.3
	PYVG SGAI(RR)	20.5
CA-p10	(P)AAAM SGLY(RR)	>200
	(P)AAAM SGAI(RR)	>200
	(P)AAAM SGLI(RR)	>200

[a] Hybrid peptide analogues combining the amino- and carboxyl-terminal portions of the RSV p2-p10 and CA-NCa cleavage sites were prepared as indicated under "Experimental Procedures." Amino acids which differ from the parental hybrid are underlined. None of the peptide analogues was cleaved efficiently by the AMV PR. The arrowhead indicates a cleavage site.

affinity for protease mediated by the P1-P4 amino acids. This is inferred from the fact that the carboxyl-terminal portion of the p2-p10 peptide, which contains the P1' serine, has a K_i of more than 200 μM (Table IV). To investigate the role of the P1-P4 amino acids in binding the peptide to the protease, leucine, tryptophan, phenylalanine, or histidine was introduced in the P1, P3, or P4 position of the p2-p10 substrate as indicated in Table IV. Substitution of leucine or tryptophan in P1 resulted in a 3- and 6-fold, respectively, decrease in the K_i of the p2-p10 analogue. Furthermore, combining the substitutions of phenylalanine in P3 with the tryptophan in P1 produced a 40-fold decrease in the K_i. The additional substitution of histidine for proline in P4 had little effect (Table IV). These results directly parallel those from independent experiments in which the same amino acids were introduced into comparable positions of the NC-PR substrate (19). Steady-state kinetic analysis of these substrates indicated that substitution of leucine or tryptophan into P1 or phenylalanine into P3 resulted in a 2.4-, 4.4-, and 6.3-fold decrease, respectively, in the K_m of these NC-PR analogues for the AMV PR (19). Previous studies have shown that the enzyme subsites act independently in recognition of substrate (19). Therefore, combining the effect of substituting tryptophan in P1 with phenylalanine in P3, one calculates a 28-fold additive decrease in the K_m for the NC-PR peptide. Substitution of histidine in P4 produced a 1.5-fold increase in the K_m for the NC-PR peptide (data not shown). Thus, amino acids in the P1-P4 positions predicted to improve the binding affinity of the peptide to protease showed increased effectiveness of p2-p10 analogues to act as inhibitors.

The importance of the P1-P4 substrate positions as a docking element is further illustrated by analysis of the chimeric peptides which combined the amino- and carboxyl-terminal portions of the CA-NCa and p2-p10 peptides (Table VII). The chimeric p2-NC peptides, which had the P1-P4

residues of the p2-p10 peptide, were effective inhibitors of the AMV PR, whereas the CA-p10 peptides were not. A docking role for the P1 substrate position is also observed with the HIV-1 PR. When leucine is substituted into the P1 position of the RSV NC-PR and p2-p10 peptides, respectively, it resulted in an increase in the K_m for the former (19) and K_i for the latter (Table V).

Analysis of RSV p2-p10 Peptide Substrates Modified at the P2' and P3' Positions—Although there is a good correlation between decreasing the K_i of p2-p10 peptide analogues and decreasing the K_m of comparable NC-PR peptide analogues for amino acid substitutions in P1–P4, this relationship does not hold for placing substitutions in P1'–P3'. For instance, substitution of serine in P2' or tyrosine in P3' of the NC-PR peptide resulted in increases of greater than 7- and 1.2-fold, respectively, in the K_m with the AMV protease.[3] Yet these same substitutions in the p2-p10 peptide produced inhibitors with decreased K_i values (Table IV). Thus, the P2' and P3' amino acids do not serve the same role of the P1–P4 amino acids for binding the peptide to the protease substrate binding pocket. Instead, their function is more related to providing rotational freedom for the serine in P1'. In the case of P3', which is adjacent to P1' in the substrate binding pocket, the presence of larger amino acids may prevent tight interactions between the protease and the peptide, thereby providing additional rotational freedom for the serine in P1'. Results with the substitution of tyrosine for leucine at P3', which decreased the K_i of the p2-p10 analogue by more than 3-fold, are consistent with this interpretation. Introduction of a smaller amino acid, alanine, at the same position resulted in a small increase in the K_i (Table IV).

DISCUSSION

We have studied the activity of the AMV and HIV-1 PRs on peptide substrates which represent cleavage sites in the RSV and HIV-1 *gag* and *gag-pol* polyprotein precursors. The kinetic properties of the HIV-1 peptide substrates reported here are similar to those published previously (32). Analysis of the catalytic efficiencies of the AMV protease on its cognate peptide substrates indicates that, for the most part, there are only small kinetic differences between the various cleavage-site peptide substrates with the exception of those representing the MA-p2 and p2-p10 cleavage sites. In all cases, the sites of cleavage of the peptide substrates were independently verified. Strop *et al.* (28) have reported different results for cleavage of larger peptides containing some of the RSV cleavage sites. These results may not be directly comparable, since these larger peptides have solubility problems. Additionally, a steady-state kinetic analysis was not completed, and the sites of cleavage were not established in these studies.

In the case of RSV, it has been suggested that maturation of the *gag* and *gag-pol* polyprotein precursors occurs with an ordered release of mature proteins (5). Our study suggests that the slow processing observed *in vivo* of the amino-terminal proteins of the *gag* precursor may be related in part to the low efficiency with which the MA-p2 and p2-p10 cleavage sites are processed by protease. This hypothesis is further supported by the observation that substitution of the 7 amino acids of the slow MA-p2 cleavage site with those of the fast NC-PR cleavage site *in vivo* resulted in a greater than 10-fold increase in the rate of processing at this site.[4] However, apparent differences in catalytic efficiency of protease *in vivo* with the other cleavage sites in the *gag* and *gag-pol*

polyproteins, suggested by the order of appearance of intermediates, would be influenced more by the three-dimensional folding of these polyproteins and the accessibility of the different cleavage sites.

A comparison of the activity of the retroviral proteases on heterologous substrates suggests that the AMV enzyme is more stringent than HIV-1 in recognition of cleavage sites. The AMV PR has barely detectable activity on the HIV-1 peptide substrates tested. In contrast, the HIV-1 PR cleaved all of the RSV peptide substrates with efficiencies equivalent to those of AMV PR. The only exception was the RSV p2-p10 substrate which neither the HIV-1 nor the HIV-2 PRs were able to cleave. This peptide is cleaved by the AMV PR, but at a low efficiency because of a low catalytic rate. However, it binds reasonably well to the protease and acts as a competitive inhibitor of the AMV as well as the HIV-1 and HIV-2 PRs.

Analysis of several co-crystal structures of HIV-1 PR with different inhibitors (20–24) has shown that peptidic inhibitors bind in an extended β-conformation between the active site and the two symmetrically related flexible flaps. There is a series of β-sheet-like hydrogen bonds between the carbonyl and amide groups of the inhibitor and protease residues that connect the flaps, the inhibitor, and residues 27–29 (HIV-1) at the active site which provide binding energy for a peptide or part of a protein to bind to protease. In addition, each side chain of the inhibitor lies in a corresponding subsite formed by amino acid residues of the enzyme. These alternate on either side of the main chain of the inhibitor and provide specific interactions with substrate side chains. The inhibitor side chains protrude into the subsites approximately perpendicular to the plane of the β-sheet formed by the substrate backbone. This means that the interactions of the side chains are largely independent of the hydrogen bond interactions to the main chain. Computer modeling and comparison with kinetic measurements have suggested that substrate peptides bind to HIV-1 (31) and RSV (19) PRs in a manner resembling the inhibitors in the crystal structures of HIV-1 PR.

Many of the more effective inhibitors of retroviral proteases are transition state analogues and contain a nonhydrolyzable isostere instead of the scissile bond between P1 and P1'. Peptidic compounds with hydroxyethylene have a hydroxyl group (OH) in the place of the carbonyl oxygen (C=O) of the scissile peptide bond, as illustrated by the co-crystal structures of HIV-1 PR with these inhibitors (20–23). This is consistent with the reaction mechanism deduced by Hyland *et al.* (33). In these inhibitors, which include acetyl-pepstatin (22), the hydroxyl group interacts with the carboxylates of the two catalytic aspartic acids (Asp-25 in HIV-1 PR) and may form several hydrogen bonds, as shown in Fig. 3B for the HIV-1 PR inhibitor structure of Jaskolski *et al.* (20).

We propose that the inhibitory p2-p10 peptide acts by formation of a hydrogen bond between the hydroxyl side chain of the serine at P1' and the carboxylate oxygen of the closest catalytic aspartate. This is shown in the structural model of the complex between the RSV PR and the p2-p10 peptide in Fig. 3A. This interaction can be made by a simple rotation of the serine side chain which is a minor adjustment of the conformation. We propose that this hydrogen bond interaction alters the protonation state of the 2 catalytic aspartates. These aspartates participate in a conserved network of hydrogen bond interactions that connect the two halves of the active site (Asp-Ser-Gly from each subunit of the RSV protease dimer). This network must stabilize the correct conformation and electronic state necessary for proteolysis to occur. Any interference with these interactions would be expected to reduce the efficiency of catalysis.

Amino acid substitutions of the p2-p10 peptide have shown

that its inhibitory effect depends upon tight binding of the peptide to the substrate binding pocket mediated by residues at P4-P1. Substitution of residues in the P1-P4 positions that are predicted to increase the affinity of the p2-p10 peptide for the protease produce more effective inhibitors. Combining the substitution of tryptophan and phenylalanine in P1 and P3, respectively, resulted in an additive effect on the K_i producing a nM inhibitor. This result confirms the relative independence of the different enzyme subsites in recognition of substrate described previously (19).

The final element for inhibition is the requirement for a glycine or small polar residue, such as serine or asparagine, at P2' and/or a larger amino acid residue at P3'. A small hydrophobic residue at P3' reduces the inhibitory effect. We propose that this is due partly to a tight fit of the binding site around the smaller residues, which would restrict the rotation of serine at P1'. In contrast, large hydrophobic residues at P3' would tend to expand the binding sites and have the opposite effect on the rotation of the P1' serine. The side chain of P3' is adjacent to P1' because of the alternating β-like conformation, so the size of the P3' side chains may directly effect the size of the S1' subsite. The effect of the P2' residue is more complex, since small hydrophobic residues such as valine and alanine position the p2-p10 peptide for substrate cleavage by providing tight binding to the S2' enzyme subsite. This is proposed to restrict the rotation of the P1' serine. In contrast, substitution of glycine, serine, or asparagine at P2' resulted in inhibitors. Glycine at P2' is most effective as an inhibitor, even when P3' contains small residues. This is because glycine lacks a side chain and allows easier rotation of serine at P1'. Serine and asparagine at P2' also form inhibitory p2-p10 analogues but are more effective if a large hydrophobic residue is present at P3'. The polar side chains of serine and asparagine are predicted to bend toward the polar environment formed by the main chain amide and carbonyl groups. This would leave the S2' subsite of protease partly vacant, providing the flexibility necessary for the P1' serine hydroxyl to rotate into the conformation required for inhibition.

The mechanism of inhibition of the HIV-1 PR by the p2-p10 peptide is more complicated. In contrast to results with the AMV PR, none of the p2-p10 peptide analogues, including leucine or alanine in P1' or alanine in P2' positions, was found to activate this peptide as a substrate. This indicates that there are some distinct, but subtle, differences between the HIV-1 and the AMV PRs. Structural studies with HIV-1 PR and p2-p10 analogues may clarify this issue.

Is there any biological significance to having an inhibitory sequence in *cis* in the RSV Pr76^{gag}? It is known that the viral precursor polypeptide is not normally processed until the virus buds from the cell surface. One can speculate on putative mechanisms that would delay activation of the protease. One possibility would be that the p2-p10 inhibitory sequence on one of the precursors binds to the active site of a protease precursor homodimer. Upon reaching the site of viral assembly at the cell membrane, cleavage and/or displacement of the "inhibitory" sequence may lead to activation of protease. Alternatively, one can visualize regulating protease precursor homodimer formation directly through different conformations or aggregation states of the associated precursor molecules or kinetically by the low turnover number of the protease. Since the RSV PR is relatively inefficient, only a weak

inhibitory effect may be needed. Once the polyproteins are enclosed in the viral particles, however, the higher concentration of protease will increase the effectiveness of cleavage and subsequent maturation. The possibility that the p2-p10 sequence is involved in control of protease activation can be tested by introducing mutations into this sequence and determining the effects on processing of precursor polypeptides *in vivo*.

REFERENCES

1. Wills, J., Craven, R., Weldon, R., Nelle, T., and Erdie, C. (1991) *J. Virol.* **65**, 3804–3812
2. Pepinsky, R., and Vogt, V. (1984) *J. Virol.* **52**, 145–153
3. Leis, J., McGinnis, J., and Green, R. (1978) *Virology* **84**, 87–98
4. Steeg, C., and Vogt, V. (1990) *J. Virol.* **64**, 847–855
5. Dickson, C., Eisenman, R., Fan, H., Hunter, E., and Teich, N. (1984) in *RNA Tumor Viruses* (Weiss, R., Teich, N., Varmus, H., and Coffin, J., eds) vol. 1, pp. 513–648 Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
6. Fu, X., Katz, R. A., Skalka, A. M., and Leis, J. (1988) *J. Biol. Chem.* **263**, 2140–2145
7. Gorelick, R., Henderson, L., Hanser, J., and Rein, A. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 8420–8424
8. Skalka, A. M. (1989) *Cell* **56**, 911–913
9. Kohl, N., Emini, E., Schleif, W., and Davis, L. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 4185–4189
10. Crawford, S., and Goff, S. (1985) *J. Virol.* **53**, 899–907
11. Gottlinger, H., Sodroski, J., and Haseltine, W. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 5781–5785
12. Alexander, F., Leis, J., Soltis, D., Crowl, R., Danho, W., Poonian, M., Pan, Y.-C., and Skalka, A. M. (1987) *J. Virol.* **61**, 534–542
13. Copeland, T. D., and Oroszlan, S. (1988) *Gene Anal. Tech.* **5**, 109–115
14. Kotler, M., Danho, W., Katz, R., Leis, J., and Skalka, A. M. (1989) *J. Biol. Chem.* **264**, 3428–3435
15. Allen, G. (1983) in *Laboratory Techniques in Biochemistry and Molecular Biology*, Vol. 9, pp. 135–139, Elsevier Science Publishing Co., Amsterdam
16. Maibaum, J., and Rich, D. (1988) *J. Med. Chem.* **31**, 625–629
17. Grinde, B., Cameron, C., Leis, J., Weber, I., Wlodawer, A., Burstein, H., Bizub, D., and Skalka, A. M. (1992) *J. Biol. Chem.* **267**, 9481–9490
18. Dixon, M. (1953) *Biochem. J.* **55**, 170–171
19. Grinde, B., Cameron, C., Leis, J., Weber, I., Wlodawer, A., Burstein, H., and Skalka, A. M. (1992) *J. Biol. Chem.* **267**, 9491–9498
20. Jaskolski, M., Tomasselli, A., Sawyer, T. Staples, D., Heinrikson, R., Schneider, J., Kent, S., and Wlodawer, A. (1991) *Biochemistry* **30**, 1600–1609
21. Erickson, J., Neidhart, D. J., VanDrie, J., Kempt, D. J., Wang, X. C., Norbeck, D. W., Plattner, J. J., Rittenhouse, J. W., Turon, M., Wideburg, N., Kohlbrenner, W. E., Simmer, R., Helfrich, R., Poul, D. A., and Knigge, M. (1990) *Science* **249**, 527–533
22. Fitzgerald, P. M. D., McKeever, B. M., VanMiddlesworth, J. F., Springer, J. P., Heimbach, J. C., Leu, C.-T., Herber, W. K., Dixon, R. A. F., and Darke, P. L. (1990) *J. Biol. Chem.* **265**, 14209–14219
23. Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Kent, S. B. H., and Wlodawer, A. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 8805–8809
24. Miller, M., Schneider, J., Sathyanarayana, B., Toth, M., Marshall, G., Clawson, L., Selk, L., Kent, S., and Wlodawer, A. (1989) *Science* **246**, 1149–1152
25. 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
26. Kotler, M., Katz, R., Danho, W., Leis, J., and Skalka, A. M. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 4185–4189
27. Schecter, I., and Berger A. (1967) *Biochem. Biophys. Res. Commun.* **27**, 157–162
28. Strop, P., Konvalinka, J., Stys, D., Pavlickova, L., Blaha, I., Velek, J., Travnicek, M., Kostka, V., and Sedlacek, J. (1991) *Biochemistry* **30**, 3437–3443
29. Tomasselli, A., Hui, J., Sawyer, T., Staples, D., Bannow, C., Reardon, I., Howe, W., DeCamp, D., Craik, C., and Heinrikson, R. (1990) *J. Biol. Chem.* **265**, 14675–14683
30. Hyland, L., Tomaszek, T., Jr., Roberts, G., Carr, S., Magaard, V., Bryan, H., Fakhoury, S., Moore, M., Minnich, M., Culp, J., DesJarlais, R., and Meek, T. (1991) *Biochemistry* **30**, 8441–8453
31. Tozser, J., Gustchina, A., Weber, I., Blaha, I, Wondrak, E., and Oroszlan, S. (1991) *FEBS Lett.* **279**, 356–360
32. Tozser, J., Blaha, I., Copeland, T., Wondrak, E., and Oroszlan, S. (1991) *FEBS Lett.* **281**, 77–80
33. Hyland, L. J., Tomaszek, T. A., Jr., and Meek, T. D. (1991) *Biochemistry* **30**, 8454–8463
34. Leis, J., Baltimore, D., Bishop, J. M., Coffin, J., Fleissner, E., Goff, S. P., Oroszlan, S., Robinson, H., Skalka, A. M., Temin, H. M., and Vogt, V. (1988) *J. Virol.* **62**, 1808–1809