A Side Chain at Position 48 of the Human Immunodeficiency Virus Type-1 Protease Flap Provides an Additional Specificity Determinant

MARK D. MOODY,* STEVE C. PETTIT,* WEI SHAO,* LORRAINE EVERITT,* DANIEL D. LOEB,† and RONALD SWANSTROM**†‡

*Linssenberger Comprehensive Cancer Center, †Department of Microbiology and Immunology, and the **Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Received November 15, 1994; accepted December 29, 1994

Substitution of glycine with glutamic acid at position 48 of the human immunodeficiency virus protease resulted in an enzyme with reduced activity on one of the protease processing sites in the viral Pol polyprotein precursor. Cleavage at this site was restored by a second-site substitution in the substrate replacing an aspartic acid with either glycine or asparagine. These results suggest that the glutamic acid side chain in the mutant protease has an unfavorable charge-charge interaction with this position in the substrate. Cleavage of a processing site in the viral Gag polyprotein precursor with the mutant enzyme was enhanced, and this enhancement was dependent on the presence of an arginine residue in the substrate, again suggesting a charge-charge interaction. The potential for such interactions was confirmed using molecular modeling. The effect of the position 48 substitution was attributed to a 10-fold increase in $K_m$ for the processing site in Pol. These results indicate that the addition of a side chain at position 48 can alter the specificity of the HIV-1 protease to substrate in a sequence-specific manner and that compensatory changes can be made in the substrate.


INTRODUCTION

All retroviruses encode a protease (PR) that is responsible for the proteolytic processing of the Gag and Gag–Pol polyprotein precursors during the assembly of virus particles (reviewed in Oroszlan and Luftig, 1990). The viral protease cleaves a diverse set of generally hydrophobic sequences to generate the mature proteins products found in the virion. Structural analysis of the human immunodeficiency virus type 1 (HIV-1) protease bound to a variety of inhibitors has shown that substrate in the active site is bound in an extended conformation and that the side chains of seven consecutive amino acids of the substrate interact with pockets (called sub-sites) in the protease (reviewed in Wlodawer and Erickson, 1993). Previous biochemical analysis demonstrated that substantial enzyme activity is seen with substrates that are at least seven amino acids in length (Darke et al., 1988; Toczzer et al., 1991b). The HIV-1 protease interacts with four amino acids upstream of the scissile bond and three amino acids downstream, and these interactions must provide the specificity needed for the correct selection of the limited number of cleavage sites used in the processing of the polyprotein precursors.

Retroviral proteases have been the object of mutational analysis to explore the basis of specificity. The retroviral protease is a symmetric homodimer and a member of the aspartic proteinase family of enzymes (Oroszlan and Luftig, 1990; Wlodawer and Erickson, 1993). Mutations have been introduced into either the Rous sarcoma virus (RSV) protease (Cameron et al., 1994; Grinde et al., 1992a) or the avian myeloblastosis-associated virus protease (Konvalinka et al., 1992) in an effort to determine the basis for differences in specificity compared to the protease of HIV-1. These results have shown that substitutions of amino acids in the subsites can affect both the binding and catalytic efficiency of the protease. Altered specificity to protease inhibitors has been demonstrated with the HIV-1 protease through the testing of protease mutants carrying designed mutations (Sardana et al., 1994) or through the selection of virus variants with reduced sensitivity to growth inhibition by the inhibitor (Efland et al., 1994; Ho et al., 1994; Kaplan et al., 1994; Otto et al., 1993).

We have previously identified a mutant of the protease of HIV-1 that has greatly reduced cleavage at one of the processing sites within the Pol precursor (Loeb et al., 1989b). The mutant has a substitution at position 48 of the protease which changes a glycine to glutamic acid.
(G48E). Position 48 is within the flap of the retroviral protease, a section of beta sheet that moves to expose the substrate binding cleft of the enzyme then closes down to make both backbone and side chain contacts with the bound substrate (Miller et al., 1989). In this report we describe experiments designed to determine the basis for the restricted specificity seen with this mutant. We found that the glutamate side chain in the mutant enzyme interacts with a specific acidic amino acid in the substrate and that the restricted phenotype was largely reversed by mutating this amino acid in the substrate. Furthermore, the mutant protease showed enhanced cleavage of a Gag substrate containing a basic amino acid.

MATERIALS AND METHODS

Plasmid construction and mutagenesis

All plasmids were derivatives of the vector pBl20 (International Biotechnologies). HIV-1 DNA used in the plasmid constructions was obtained from an infectious clone of the HXB2 isolate (Ratner et al., 1987). The features and construction of the plasmid pART2 utilized for the expression of the pol gene in Escherichia coli, have been previously described (Farmerie et al., 1997; Loeb et al., 1999a, 1999b). Briefly, pART2 contains the HIV-1 pol gene (PR-RT-IN domains) inserted downstream of the iac promoter of pBl20. The plasmid pART2ΔPR is a derivative of pART2 in which the coding regions of HIV-1 RT-IN were placed under transcriptional control of the bacteriophage T7 promoter. A eukaryotic ribosomal recognition sequence (Kozak, 1989) and an initiating methionine were placed at the start of the RT coding domain by mutagenesis. pGagS contains the HIV-1 gag gene under transcriptional control of the bacteriophage T7 promoter (Pettit et al., 1994).

Site-directed mutagenesis of pART2, pART2ΔPR, and pGagS was performed as described (Loeb et al., 1999a; Pettit et al., 1991) using the method of Kunkel et al. (1991) with the modification of Bebenek and Kunkel (1989). All mutations were identified by DNA sequence analysis (Sanger et al., 1977) utilizing modified T7 DNA polymerase (U.S. Biochemicals).

Assay for the processing of Pol in E. coli

E. coli strain JM101 (containing either pART2 or pART2 with mutations) was induced for the expression of the pol gene from the iac promoter as described (Pettit et al., 1991). Ninety minutes postinduction, the processed Pol products were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and detected by Western blot (Loeb et al., 1999a) with anti-RT monoclonal antibody (Du Pont).

Expression and purification of the HIV-1 protease from E. coli

HIV-1 protease was expressed in E. coli BL21 (DE3) cells by inserting the coding domain of the protease into a bacteriophage T7 expression system (Studier and Moffat, 1986). The DNA sequence of the protease was as reported (Ratner et al., 1987) except for the addition of an initiating methionine codon. Following induction of expression, protease was purified from isolated inclusion bodies described by Pettit et al. (1994). The protease purified by this procedure was greater than 90% pure as judged by SDS–PAGE (Laemmli, 1970) and was stable for several months when stored at −70°C.

The synthesis and proteolytic processing of Gag and Pol precursors in vitro

Capped mRNA was synthesized in vitro from linearized pGagS or pART2ΔPR DNA as described (Pettit et al., 1994). HIV-1 gag or pol mRNA (1.0 µg) was translated in a rabbit reticulocyte lysate system (RRL) (Promega Corp.), in the presence of 50 µCi [35S]cysteine (1000 Ci/mM) (Du Pont NEN), according to the manufacturer’s instruction.

The Gag and Pol precursors were processed in vitro by HIV-1 PR in 50-µl reactions containing 5 µl of the RRL translation mixture and 1.0 to 5.0 µg of purified recombinant protease as described (Pettit et al., 1994). Aliquots of 5 µl were taken at various times, and the products were separated by SDS–PAGE (Laemmli, 1970). Relative quantitation of labeled protein products was performed by densitometry analysis of autoradiographs on a Molecular Dynamics model 300 densitometer or by analysis of the radioactivity in the gel using a Molecular Dynamics model 400 phosphoimager.

Calculation of cleavage rates for processing sites in the Gag or Pol precursors

Cleavage of the Gag or Pol substrate in the in vitro processing reactions was monitored by the method of Pettit et al. (1994). Briefly, the relative rate of cleavage for an individual site was calculated from the first order rate equation where the substrate concentration is represented by the percentage of the processing site remaining uncleaved in the substrate and processing intermediates. Time points representing 20 to 80% cleavage were used in the estimation of rates. The maximum amount of assay time utilized for the rate calculations was 1 hr, during which time the protease retained at least 80% of its initial activity (Pettit et al., 1994). For the comparison of the rates of cleavage between the different processing sites, and with wild-type and mutant proteases, the rate of cleavage of a secondary processing site was used as an internal standard.
Analysis of enzyme kinetics

The kinetic parameters of $K_m$ and $V_{max}$ were determined for the wild-type and mutant proteases using substrates AETFnFVG and AETFnFVG (where nF is nitrophenyalalanine) in a reaction buffer modified from Tozser et al. (1991a) consisting of 0.25 M potassium phosphate (pH 6.5), 7.5% glycerol, 5 mM DTT, 1 mM EDTA, 0.1% PEG, and 2 M NaCl. For the substrate KARLMANSnF the reaction conditions were 100 mM sodium acetate (pH 5.5), 200 mM NaCl, 1 mM EDTA, and 0.1% PEG. The reactions were carried out at 37°C with cleavages limited to less than 20% of the substrate. The reactions were stopped by the addition of TFA, and the cleavage products were detected by gel electrophoresis after elution from a PepRPC-FPLC column (Pharmacia). Active site titration was done using a tight binding inhibitor and determining initial rates of cleavage at variable concentrations of enzyme and different fixed concentrations of the inhibitor (Williams and Morrison, 1979). The enzyme and inhibitor were preincubated on ice for 10 min prior to the start of the reaction. Titration of active sites was calculated by fitting of initial rates in the presence and absence of inhibitor to equation 6 as described by Grant et al. (1991). The fraction of active enzyme was then used to correct $V_{max}$ to $K_{cat}$.

RESULTS

Reduced cleavage at the RT p51/p13 cleavage site

The HIV-1 protease cleaves the Gag and Gag/Pol polyprotein precursors during virion maturation, and these processing events can be recapitulated in experimental systems. Expression of the pol gene in bacteria leads to cleavage of the Pol polyprotein by the pol-encoded viral protease (Farmerie et al., 1987; Figs. 1A and 1B, lane 1). Cleavage of Pol can be monitored by the release of the mature viral proteins (Loeb et al., 1989a), the protease p11 PR, two forms of reverse transcriptase p66/p51 RT (Fig. 1B, lane 1), and the C-terminal inteinase domain p32 IN. Cleavage of the p66 form of RT to the p51 form is limited to one-half of the molecules since only one of the p51/p13 cleavage sites in the p66 homodimer is accessible; the other is involved in forming the structure of the RNase H domain. We assume that the limited cleavage of RT we observe in the bacterial expression and processing system is due to dimerization of RT, although we have not tested this assumption directly. This bacterial expression and protease assay system has been used to characterize a large number of different HIV-1 protease mutants (Loeb et al., 1989b). One of these mutants, which has the wild-type glycine at position 48 replaced with glutamic acid (G48E), shows greatly reduced cleavage at the site within the RT domain that generates the p51 form of RT, leaving predominately the larger form of RT p66 (Loeb et al., 1989b; Fig. 1B, lane 2). This mutant protease is not grossly defective for activity at the other sites in the Pol protein since cleavages occur to remove the protease from the Pol precursor and to release the C-terminal IN domain.

Position 48 is within the flap region of the protease (Fig. 1C). The addition of a side chain at this position could result in new interactions with the substrate, and this interaction would likely involve the P3 and P3' amino acid side chains of the substrate. (P3 designates the substrate amino acid that lies three positions N-terminal to the cleavage site, while P3' designates the substrate amino acid that lies three positions C-terminal to the cleavage site). This can be seen in Fig. 1D where the G48 position is juxtaposed to the P3-equivalent position of the inhibitor MVT-101, Thr200 (Miller et al., 1989). The possibility of altered interaction with the substrate was explored by examining the activity of wild-type and mutant protease with substrate in several different experimental settings.

Reversion of the G48E phenotype

One possible explanation for the reduced cleavage at the RT p51/p13 cleavage site by the G48E mutant protease is the creation of an unfavorable interaction between the enzyme and the P3 and/or P3' amino acids of the substrate. To test this possibility we mutated the P4 through P3' amino acids at the RT p51/p13 cleavage site and determined the effect on protease cleavage with the mutant and wild-type enzymes. Only substitutions at the P3' site were able to revert the phenotype and restore wild-type levels of cleavage (see below) while substitutions at other positions did not revert the phenotype (data not shown).

The P3' position of the RT p51/p13 cleavage site is aspartic acid in the wild-type sequence (Mizrahi et al., 1989; Graves et al., 1990). Changing the P3' amino acid to glutamic acid reduced even the residual cleavage at this site by the G48E mutant protease (Fig. 1B, lane 3) without affecting cleavage by the wild-type enzyme (not shown). Changing the P3' position of this cleavage site to either glycine or asparagine restored cleavage at this site by the G48E mutant protease in the bacterial expression system (Fig. 1B, lanes 4 and 5) and had no apparent effect on cleavage by the wild-type enzyme (Fig. 1B, lanes 7 and 8). Thus, the altered specificity of the G48E mutant appears to be the result, at least in part, of an unfavorable interaction between the glutamic acid at position 48 within the flap and the aspartic acid at the P3' position of the p51/p13 cleavage site. The fact that either glycine or asparagine can restore cleavage suggests that the basis for this unfavorable interaction is primarily due to charge repulsion of the negatively charged amino acids in substrate and enzyme rather than unfavorable con-
tacts, since aspartic acid and asparagine occupy approximately the same volume. Restored cleavage of a substrate with a P3' glycine represents a situation where the protease position 48 amino acid and the P3' amino acid of the substrate have been exchanged.

The phenotypes for cleavage by the G46E mutant protease at the p51/p13 RT cleavage site for different P3' substitutions are shown in Table 1 and can be summarized as follows. While substitution with asparagine or glycine restored cleavage with the mutant protease, substitution with glutamic acid, valine, tyrosine, or histidine completely blocked cleavage, and lysine, serine, or the wild-type aspartic acid permitted only a low level of cleavage. The wild-type protease was tolerant to substrate P3' substitution of asparagine or glycine (Fig. 1B, lanes 7 and 8), as well as glutamic acid, serine, or alanine; like...
TABLE 1
Cleavage Phenotype of Mutant and Wild-Type Proteases on Substrates With Different P3’ Substrate Substitutions

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Wild type-like</th>
<th>Partial</th>
<th>Over</th>
<th>No cleavage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>Asp, Asn, Gly, Glu, Ser, Ala</td>
<td>Lys</td>
<td>Val, His</td>
<td></td>
</tr>
<tr>
<td>G48E mutant</td>
<td>Gly, Asn, Asp, Ser, Lys, Glu, Val, Tyr, His</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The observed phenotypes were wild type-like, cleavage of approximately one-half of the p66 form of RT to p61; partial, cleavage of significantly less than one-half of the p66 form of RT to p51; over, cleavage of over 60% of the p66 form of RT to p51; and no cleavage, the p66 form of RT persisted without any generation of the p51 form.

* The identity of the amino acid at the P3’ position that gives the indicated phenotype. The three letter abbreviations for amino acids are used.

the mutant protease, the wild-type protease was intolerant of P3’ substrate substitution of valine or histidine. Substitution with lysine resulted in over-processing of the p66 form of RT with the wild-type protease to generate predominantly the p51 form of RT (not shown). We interpret the overcleavage to represent disruption of the structure of the p51/p13 cleavage site in both subunits of RT making the cleavage site in each subunit susceptible to the protease. An example of the over-processing phenotype is seen with the P2 threonine to isoleucine mutant (Fig. 1B, lane 9).

The P3 side chain of the substrate has interactions that are similar (but not identical) to those made by the P3’ side chain with the dimer protease. We attempted to identify phenotypic revertants at the P3 position of the p51/p13 processing site but obtained less clear results, including several examples of over-processing with the wild-type enzyme. Given the limitations of the bacterial expression/protease assay system we turned to other experimental approaches to confirm and extend these results.

Cleavage of an RT/IN substrate in vitro

To confirm the validity of the bacterial expression/processing assay for the HIV-1 protease we examined the cleavage of a Pol-derived substrate in vitro using purified mutant and wild-type protease. A radiolabeled RT/IN substrate was generated by in vitro transcription followed by translation using a reticulocyte lysate system. The resulting labeled RT/IN substrate, containing the p51/p13 cleavage site as well as the RT p66/IN p32 cleavage site, was cleaved with either wild-type protease or mutant G48E protease (Fig. 2A). In order to compare the relative rates of cleavage, the amount of enzyme in each reaction was adjusted so that cleavage of the RT/IN site occurred at similar rates with the two enzymes, allowing this site to serve as an internal standard. Cleavage by the wild-type protease generated IN, p66 RT, and p51 RT, and the cleavage reaction was essentially complete by the 20-min time point (Fig. 2B). Cleavage by the G48E mutant protease similarly generated IN and p66 RT, but only trace amounts of p51 RT were seen even at the 1- and 6-hr time points (Fig. 2B). These results confirm that the mutant protease displays a site-specific reduction in its ability to cleave the RT p51/p13 processing site.

Mutant RT/IN substrates were constructed which had either a P3’ glycine or asparagine at the p51/p13 cleavage site, and these substrates were cleaved in vitro with the wild-type and mutant proteases. The mutant protease cleaved both mutant substrates more efficiently than it cleaved the wild-type substrate. For the wild-type protease, the substrate with the P3’ glycine substitution was cleaved as well as the wild-type substrate, but the substrate with the P3’ asparagine substitution was cleaved at a reduced rate. Thus, improved cleavage in vitro by the mutant enzyme of substrates with these P3’ substitutions was consistent with the patterns of cleavage observed using the bacterial expression/processing system. However, in each case cleavage by the mutant protease of the mutant substrate was less extensive than was cleavage of these substrates by the wild-type protease (data not shown).

Altered cleavage of Gag in vitro

Since the G48E mutant protease had an altered pattern of cleavage with the RT/IN polypeptide substrate we examined the pattern of cleavage of the Gag polyprotein substrate. There are five major cleavage sites in Gag, with four of the five cleavages occurring readily in vitro. The fifth cleavage occurs slowly and therefore is difficult to detect (Fig. 3A). A radiolabeled Gag substrate was generated using in vitro transcription-translation, and the labeled Gag substrate was cleaved with purified wild-type and mutant proteases. Cleavage of Gag with the wild-type protease occurs sequentially, generating defined intermediates, with the initial cleavage occurring at the p2/NC site, followed by cleavage at the MA/CA and NC/p6 sites at rates similar to each other, and slow cleavage at the CA/p2 site (Erickson-Vitalen et al., 1989; Krausslich et al., 1988; Partin et al., 1990; Tritt et al., 1991). The amount of each enzyme was adjusted to give similar rates of cleavage at the p2/NC site, and the rates of cleavage at the other sites compared. Cleavage with the mutant enzyme resulted in a small but significant acceleration in the appearance of the p24 CA protein relative to its p25 CA/p2 precursor (as seen by the altered ratio of p24/p25), suggesting accelerated cleavage of the CA/p2 site (Fig. 3B). Cleavage of a peptide substrate
confirmed the enhanced cleavage by the mutant enzyme at the CA/p2 site (see below).

The CA/p2 processing site has an arginine at the P3 position raising the possibility that a favorable charge interaction with the mutant enzyme may be the basis for the apparent increase in the rate of cleavage. We examined the effect of the P3 amino acid on cleavage at this site by mutating it to a series of either charged or polar amino acids (Fig. 4). To determine the relative effect of each amino acid, the rates were corrected to

FIG. 2. Processing of an RT7N substrate with P3' substitutions by wild type and mutant proteases. (A) The left panel shows the result of cleavage with the wild-type protease (PR), and the right panel shows the result of cleavage with the mutant protease (G48E). The cleavage reaction was run for the indicated times, and the labeled substrate and products were separated by polyacrylamide gel electrophoresis followed by autoradiography. The positions of the RT7N substrate and products are indicated on the left, and the positions of the size markers are shown in the center. (B) The amount of radioactivity in each protein species was quantitated and used to calculate the extent of cleavage at each site as a function of time of the cleavage reaction. The circles represent the amount of uncleaved substrate remaining for the site that generates the p51 RT product. The squares represent the amount of uncleaved substrate remaining for the site that generates the p66 RT product. The open symbols are for cleavage by the G48E mutant protease. The filled symbols are for cleavage by the wild-type protease.

FIG. 3. Processing of the Gag precursor in vitro with wild type and mutant proteases. (A) Cartoon of the Gag coding domain under the control of a T7 promoter. The Gag precursor (Pr55 Gag) is cleaved to generate the mature products p17 MA, p7 NC, p6, the processing intermediate p25, and the products p24 CA and p2. The sequence of the cleavage site between p24 CA and p2 is shown at the bottom, with the slash indicating the cleavage site and the P3 arginine underlined. (B) Extent of cleavage at three of the cleavage sites in Gag by the wild-type (closed circles) and G48E mutant (open circles) protease.
give the rate of cleavage of an alanine-containing substrate a relative cleavage rate of 1. With the wild-type protease, a Gag substrate with the wild-type arginine at the P3 position of the CA/p2 site was cleaved at a rate nearly three times faster than a substrate with alanine at P3. The rate of cleavage increased even more with a P3 glutamine substitution (to six times the rate with alanine). Asparagine, lysine, and aspartic acid were nearly equivalent to alanine, while the substrate with a P3 glutamic acid was cleaved significantly slower. Cleavage by the G48E mutant protease was improved nearly twofold for substrates with the basic amino acids arginine or lysine at P3, compared to cleavage with the wild-type protease, and decreased twofold on substrates containing acidic amino acids. This observation suggests that position 48 of the HIV-1 protease can function as a determinant of specificity and that the negative charge of glutamic acid at position 48 of the mutant protease alters the rate of cleavage by interacting with the P3 side chain of the substrate.

**Catalytic properties of the mutant protease**

We next examined the cleavage of defined peptide substrates to provide a more detailed characterization of the interaction between enzyme and substrate. Two substrates were used, one peptide representing the wild-type RT p51/p13 cleavage site (AETF/YVGD) and the other peptide representing the phenotypic revertant mutant sequence (AETF/YVGG). For each peptide nitrophenylalanine replaced the P1 phenylalanine to provide a detectable chromophore. We measured the kinetic parameters of $K_m$ and $k_{cat}$ for the two enzymes using these two peptide substrates (Table 2). The $k_{cat}$ values were similar for both proteases with both substrates. The $K_m$ value for the wild-type protease was over 10-fold lower with the wild-type substrate than was the $K_m$ for the G48E mutant protease with this substrate. This observation is consistent with the initial observation that the G48E mutant protease cleaves the p51/p13 cleavage site less well than the wild-type protease (Figs. 1B and 2). In contrast, both enzymes cleaved the mutant substrate equivalently, again consistent with the interpretation that this mutant substrate sequence compensates for the mutation in the G48E protease. However, the absolute value of $K_m$ for the wild-type enzyme was significantly higher with the mutant substrate than with the wild-type substrate. It is possible that the higher $K_m$ is the result of increased entropy in the unbound form of the glycine-containing mutant substrate. The absence of a similar increase in $K_m$ for the G48E mutant protease with the mutant substrate implies that the mutant substrate relieves an unfavorable interaction with the wild-type (P3' aspartic acid) substrate.

Catalytic constants were also determined for the wild-type and mutant enzymes using a peptide representative of the CA/p2 cleavage site. As shown above, cleavage at this site in the context of the Gag precursor appears

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Wild type $K_m$</th>
<th>Wild type $k_{cat}$</th>
<th>G48E mutant $K_m$</th>
<th>G48E mutant $k_{cat}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AETF/YVGD</td>
<td>0.01</td>
<td>2.1</td>
<td>0.1</td>
<td>3.2</td>
</tr>
<tr>
<td>AETF/YVGG</td>
<td>0.08</td>
<td>2.0</td>
<td>0.07</td>
<td>2.0</td>
</tr>
<tr>
<td>KARVL/AEGAMSNF</td>
<td>0.35</td>
<td>5.6</td>
<td>0.16</td>
<td>4.8</td>
</tr>
</tbody>
</table>

Note. All values are the average from duplicate experiments. The variation in values was less than 20% of the average. The units of $K_m$ are mM, and the units of $k_{cat}$ are sec$^{-1}$. 
to be accelerated using the G48E mutant protease (Fig. 3B). The $K_m$ of the mutant enzyme using this peptide substrate was two-fold lower than for the wild-type enzyme (Table 2). The $k_{cat}$ values for the two enzymes were very similar on this substrate. The lower $K_m$ may be due to a favorable interaction between the position 48 glutamic acid in the mutant enzyme and the P3 arginine in the substrate, and this result with a peptide substrate is consistent with the observation of accelerated cleavage at this site within the Gag substrate (Fig. 3B).

**DISCUSSION**

In these experiments we have confirmed the original observation that the G48E mutant of the HIV-1 protease has altered substrate specificity (Loeb et al., 1989b) and extended our understanding of this mutant by identifying substrate residues that are involved in binding by the mutant enzyme. These results suggest that a side chain at position 48 of the protease can participate in substrate interaction. This potential specificity determinant is unused in HIV-1; naturally occurring coding changes at this position have not been observed in any HIV-1 isolate (Fontenot et al., 1992). Whatever balance is needed between the protease and its substrate in substrate recognition and cleavage is stably maintained without an interaction involving a position 48 side chain.

Potential interactions between a side chain at position 48 of the protease and the P3 and P3' side chains of the substrate are readily interpretable based on known structures of the HIV-1 protease complexed with various substrate-based inhibitors. Figure 5 shows a model of an hypothesized interaction between a glutamic acid side chain at position 48 of the flap with either a P3 arginine or a P3' aspartic acid. The models were generated based on the structure of the protease complexed with the inhibitor MV1-101 (Miller et al., 1989). In generating the models it was necessary to move the phenylalanine side chain at position 53 slightly, but no other adjustments were made, and the modeled P3/P3' side chains and position 48 glutamic acid side chain all assume reasonable conformations. These models offer a structural basis for interpreting the data which suggest an interaction, either positive or negative, between these side chains. We did not find that arginine and lysine behaved equivalently as substrate substitutions, suggesting chain length plays a role in determining the quality of the charge-charge interaction.

A glycine is conserved at position 48 (or its equivalent position) in the proteases of the primate lentiviruses. In other retroviruses the amino acids isoleucine, glutamine, leucine, and histidine can occupy this position (Rao et al., 1991). Cross-referencing the type of amino acid at position 48 (or its equivalent position) with the amino acids found at positions P3 and P3' in the cleavage sites for a variety of retroviruses (Orozloslan and Luftig, 1990) does not reveal a simple correlation to suggest a dominant interaction (not shown), although for all retroviruses these two substrate positions are rather heterogeneous in the range of allowed amino acids (Petit et al., 1991). Since the other amino acids found at this position of the flap are uncharged, their interaction with substrate may be through van der Waals forces, and the effect of a
side chain at this position may be to enclose the S3/S3' subsites, which in HIV-1 are somewhat open due to the glycine at position 48.

The role of a glycine at position 48 has been examined with the protease from RSV, which normally has a histidine (at position 65 in the RSV protease). Substitution of glycine for histidine at position 65 in the RSV protease results in an inactive enzyme (Grinde et al., 1992a). However, when the H65G substitution is included with two other substitutions in the core of the enzyme, activity is restored. In this case, the contribution of the glycine to the double mutation in the core was to reduce the activity of the enzyme on a variety of both RSV and HIV-1 cleavage sites (Grinde et al., 1992a). The glycine substitution did enhance the effect of the two core mutations in cleaving certain P1 and P3 substituted substrates (Grinde et al., 1992b). In this triple mutant, the absence of the side chain in the flap may either enhance the backbone hydrogen bonds made between the flap and the substrate or allow an altered conformation of the substrate, the enzyme core, or both, to affect the rate of cleavage. In the case of HIV-1 protease, position 48 is tolerant to many different substitutions (Loeb et al., 1989b; L. Everitt and R. Swanstrom, unpublished observation), although potential changes in substrate specificity have not been examined for most of these active mutants. One substitution at this position appears likely to have important biological significance. A change at position 48 from glycine to valine has been seen after selection for growth of HIV-1 in the presence of the protease inhibitor Ro 31-8959 (Jacobsen et al., 1994) and also with the inhibitor A-77003 (T. Smith, D. Iribarce, and R. Swanstrom, unpublished observation), suggesting that substitutions at position 48 may contribute to the evolution of drug resistance. One reason position 48 in the flap can accept substitutions is that the glycine already has phi/psi angles that are compatible with a side chain, unlike the other glycines in the flap (A. Wlodawer, personal communication).

A glutamic acid at position 48 can influence the K_m of the protease, lowering it with one substrate and increasing it with another substrate (Table 2). The flap of the protease has been hypothesized to move as much as 15 Å to allow substrate to enter the substrate binding cleft in the core of the enzyme (Gustchina and Weber, 1990). This suggests that interaction between the flap and substrate probably does not occur until after the substrate has entered the enzyme core and the flaps have moved over the substrate binding cleft. For the flap to still have an effect on K_m implies that substrate that has attained this level of interaction with the enzyme is not irreversibly positioned to be cleaved and can still leave the enzyme uncleaved. Thus, flap closure is not a final determinant of cleavage.

Cleavage at the p51/p13 site in RT is regulated by at least two different mechanisms. One mechanism is the sequence of the cleavage site, as discussed above. The second mechanism is inferred from the structures of RT (Jacobo-Molina et al., 1993; Kohlstaedt et al., 1992) and RNase H (Davies et al., 1991). It is likely that one of the two p51/p13 cleavage sites in the p66 RT homodimer is inaccessible as part of the folded RNase H domain while the other site is in a more open conformation and available to cleavage, both by the HIV-1 protease and by other proteases (Lowe et al., 1988). The structure of RNase H offers an explanation for the observation that some of the substitutions made in the processing site result in overprocessing. The P3 glutamic acid in the folded RNase H structure forms a hydrogen bond with a downstream threonine. Thus, the P3 substitutions may be disruptive to structure and therefore result in overprocessing, as we observed in the bacterial expression assay. In contrast, the P3' aspartic acid in the folded RNase H structure does not appear to be involved in any hydrogen bonds with other amino acids and, therefore, may be less sensitive to disruption of structure after substitution (Davies et al., 1991), although overprocessing was observed with a lysine substitution at this position.

In these experiments we have examined a variant of the HIV-1 protease with altered substrate specificity, and examples of phenotypic reversal of the altered specificity. The basis for the change in specificity appears to be charge-charge interaction between the side chain at position 48 and the P3 or P3' side chains of the substrate. Given the tolerance of position 48 to substitution it may be possible to manipulate the specificity further with other substitutions at this position. In the case of neutral endopeptidase-24.11 it has been possible to predict a change in specificity that involves a charge-charge interaction with substrate by reversing the positions of the protease and substrate side chains involved in the interaction (Beaumont et al., 1992); similarly, charge reversal has been used to make compensatory changes in tissue-type plasminogen activator and its inhibitor PAI-1 (Madsen et al., 1990). We do not know if it will be possible to generate a replication competent virus carrying a G48E protease substitution. It is possible that such a virus may require (or evolve) compensatory mutations in processing sites to attain full infectivity.

ACKNOWLEDGMENTS

We gratefully acknowledge the help of Dr. Alla Gustchina for her efforts with the molecular modeling, for providing Fig. 5, and for her helpful discussion and also Dr. Charles Carter for assistance with Fig. 1C. This work was supported by Public Health Service Grant RO1AI0521. W.S. was supported by N.I.H. Training Grant T32 CA0915. S.C.P. was supported by N.I.H. N.R.S.A. Award F32-AI084000.

REFERENCES


regulates sequential proteolytic processing and is required to produce fully infectious virions. J. Virol. 68, 8017–8027.


