Molecular Model of Equine Infectious Anemia Virus Proteinase and Kinetic Measurements for Peptide Substrates with Single Amino Acid Substitutions†

Irene T. Weber,*,‡ József Tözsér,§,∥ Jin Wu,‡ Deirdre Friedman,∥ and Stephen Oroszlan*,∥

Department of Pharmacology, Jefferson Cancer Institute, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, Department of Biochemistry, Medical University of Debrecen, P.O. Box 6, H-4012 Debrecen, Hungary, and Laboratory of Molecular Virology and Carcinogenesis, ABL-Basic Research Program, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, Maryland 21702

Received October 12, 1992; Revised Manuscript Received January 15, 1993

ABSTRACT: A molecular model has been built of the equine infectious anemia virus (EIAV) proteinase on the basis of the crystal structures of the related Rous sarcoma virus (RSV) and human immunodeficiency virus (HIV) proteinases. The 104 residue long EIAV proteinase has 30 identical and 11 similar amino acids compared to those in HIV-1 proteinase and 25 identical and 18 similar amino acids compared to RSV proteinase. The overall structure is predicted to be close to that of HIV-1 proteinase. Two regions show differences: there are 6 additional residues leading to the tip of the flap, which is predicted to be involved in interactions with substrate, and there is a single residue deletion in the β b′ strand at a position equivalent to residue 60 in HIV-1 proteinase. The conformation of the residues leading to the flap was modeled by analogy to the corresponding region of RSV proteinase. The peptide substrate, VSQNYPIVQ, was modeled by analogy to the inhibitors in the cocrystal structures of HIV-1 proteinase, and the residues forming the substrate binding sites of EIAV proteinase were identified. EIAV proteinase showed several nonconservative substitutions in these residues compared to HIV-1 proteinase: Thr 30 instead of Asp in subsites S2, S2′, S4, and S4′, Ile 54 instead of Gly 48 in subsites S1, S1′, S3, and S3′, Arg 79 instead of Thr 74 in S4 and S4′, and Ile 85 instead of Thr 80 in subsites S1 and S1′. The presence of additional residues in the flaps which lie over the substrate at S4 and S4′ and the substitutions in the individual subsites can be correlated with differences in specificity of HIV and EIAV proteinases for various substrate peptides.

Retroviral proteinases (PR)[1] are part of the viral Gag or Gag–Pol polyproteins of immature virions which they cleave into functional domains after activation by an unknown mechanism. The structure and function of retroviral proteinases have been the subject of several reviews (Skalka, 1989; Oroszlan & Luftig, 1990; Fitzgerald & Springer, 1991). The recent observation that the PR of equine infectious anemia virus (EIAV) was able to process the nucleocapsid protein within intact capsids has suggested an important role for PR action in the early phase of viral infection (Roberts & Oroszlan, 1989; Roberts et al., 1991). EIAV is a lentivirus with strong homology to the immunodeficiency viruses (Stephens et al., 1986) and is the only lentivirus for which capsids have been prepared. Therefore, it is an important model for HIV. The retroviral PR is a potential target for chemotherapy (Krausslich et al., 1989); therefore, better understanding of the substrate specificity of retroviral proteinases may help to design more potent inhibitors. Crystal structures are available of Rous sarcoma virus (RSV) PR (Miller et al., 1989a) and of HIV-1 PR alone (Navia et al., 1989; Wlodawer et al., 1989; Lapatto et al., 1989) and complexed with different inhibitors (Miller et al., 1989; Erickson et al., 1990; Fitzgerald et al., 1990; Swain et al., 1990; Jaskolski et al., 1991). Both the experimental structures and molecular modeling of the PR structures have greatly facilitated studies of the substrate specificity, as well as the rational design of proteinase inhibitors. The retroviral proteinases are enzymatically active as dimers of two chemically identical subunits that are structurally very similar. Each subunit contains the characteristic active site triplet, Asp-Thr-Gly, of aspartic proteinases and has a similar fold to one domain of the pepsin-like proteinases (Tang et al., 1978). The cocrystal structures of HIV-1 PR with different inhibitors show that peptide-based inhibitors are bound by a series of hydrogen bond interactions between the NH and C=O groups of the inhibitor and the PR, and each inhibitor side chain lies in successive subsites formed by PR residues.

The structural alignment of the amino acid sequences was initially shown in Weber (1989) for several retroviral proteinases. Comparison of the crystal structures of HIV-1 and RSV proteinases, combined with the alignment of amino acid sequences, suggested that all retroviral proteinases shared a conserved structural core (Weber, 1990a), which facilitates molecular modeling of one PR on the basis of the available crystal structure of another PR. Previously, HIV-1 PR was modeled by analogy to the RSV PR crystal structure (Weber et al., 1989), and a model of the HIV-2 PR was built from the coordinates of HIV-1 PR (Gustchina & Weber, 1991; Tomasselli et al., 1990). In addition, the interactions of HIV-1, HIV-2, and RSV PRs with substrates were modeled on the basis of the crystal structures of HIV-1 PR–inhibitor complexes and correlated with kinetic measurements of substrate cleavage (Tözsér et al., 1991, 1992; Grinde et al., 1992).

† Research sponsored in part by the National Cancer Institute, DHHS, under Contract No. NO1-CO-74101 with ABL and by the Hungarian Science and Research Fund (OTKA F 5104 to J.T.). The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

‡ Thomas Jefferson University.

§ Medical University of Debrecen.

∥ NCI–Frederick Cancer Research and Development Center.

[1] Abbreviations: EIAV, equine infectious anemia virus; HIV-1 and HIV-2, human immunodeficiency virus type-1 and type-2; RSV, Rous sarcoma virus; PR, retroviral proteinase; RP-HPLC, reversed-phase high-performance chromatography.

0006-2960/93/0432-3354$04.00/0

Here we describe the molecular modeling of the dimer of EIAV PR in a complex with a peptide substrate, on the basis of the crystal structures of HIV-1 PR with inhibitor (Miller et al., 1989b; Swain et al., 1990; Jaskolski et al., 1991). This model was used to interpret kinetic data obtained with oligopeptide substrates and inhibitors [see also Tözsér et al. (1993)].

EXPERIMENTAL PROCEDURES

Molecular Modeling. The amino acid sequence of EIAV PR was substituted for that of HIV-1 PR using the refined crystal structure of HIV-1 PR with inhibitor JG365 (Swain et al., 1990). The computer graphics program FRODO (Jones, 1985) was used on an Evans and Sutherland computer graphics system PS390 and subsequently on an ESV10 to build the molecular model. Residues that were identical in EIAV proteinase and HIV-1 proteinase were not altered. Other residues were superimposed on the atoms of the HIV-1 proteinase residue, and the side chains were adjusted if necessary by rotating torsion angles to remove any undesirably close contacts with adjacent atoms. The main chain atoms were not altered, except for substitutions of Gly and at the two regions of insertion and deletion. The backbone conformation was adjusted to make negative φ torsion angles when necessitated by a change from Gly in HIV-1 PR to another amino acid in EIAV PR. The FRODO idealization routine, REFI, was run to maintain reasonable geometry. Most of the structure of EIAV PR was modeled by analogy to the crystal structure of HIV-1 proteinase with inhibitor. The region leading to and including the flaps, residues 36–52, was modeled by analogy to RSV PR, as described in Grinde et al. (1992). One subunit was built in this manner, as described previously for HIV-2 PR (Gustchina & Weber, 1991) and for the flaps of RSV PR (Grinde et al., 1992). The second subunit was modeled by superimposing the first subunit on the second in the dimeric crystal structure. The hydrogen bond interaction that connected the two flaps in the dimer was reintroduced by altering the carbonyl of residue 56′ to resemble that of HIV-1 proteinase Ile 50′ (a prime indicates the second subunit). To ensure that the interactions between the two flaps in the dimr were maintained, the main chain atoms of residues 54′–58′ were repositioned to correspond to the equivalent atoms in HIV-1 PR. Both the JG365 inhibitor (acetyl-S-L-N-F-ψ(CH(OH)CH_2N)-P-I-V-O-methyl) and the U85548e inhibitor (V-S-Q-N-L-ψ(CH(OH)CH_2)-V-I-V) have similarities with the desired substrate peptide, and their cocrystal structures with HIV-1 PR (Swain et al., 1990; Jaskolski et al., 1991) were used to model the substrate peptide. The inhibitor from the starting model of the cocrystal structure was mutated to a substrate with the amino acid sequence of the HIV-1 MA-CA cleavage site peptide, V-S-Q-N-Y*P-I-V-Q-NH2 (SP-211, the asterisk indicates the cleavage site), similar to the models for the HIV-1 and HIV-2 proteinases (Tözsér et al., 1991, 1992).

Energy Minimization Procedure. The model coordinates for the EIAV PR dimer with substrate were subjected to energy minimization using the program XPLOR (Brunger, 1990) run on a Convex 240, as described in Kumar and Weber (1992). The hydrogen atoms of potential hydrogen bond donors were added and refined in XPLOR, and standard topologies and parameters were used for the amino acids. The two catalytic aspartates were modeled as a charged aspartic acid and a protonated aspartate to approximate the close interaction between the side chain oxygens of the two aspartates, as described in Sansom et al. (1992). The substrate terminated in $CONH_2$, and this was used in the model rather than the standard carboxyl terminus. The empirical energy function used a 6–12 van der Waals potential and a Coulomb potential with an 8-Å cutoff for the nonbonded terms. The effect of hydrogen bonds is taken into account by appropriate partial charges and van der Waals parameters. The conformational terms included potentials for covalent bonds, bond angles, dihedral angles, and chirality or planarity. The Powell conjugate gradient algorithm (Powell, 1977) was used for 500 cycles of energy minimization until the maximum gradient was less than 0.1 kcal/mol. This new molecular model for EIAV proteinase was then compared with the structures of HIV-1 proteinase and the model of HIV-2 proteinase for analysis and interpretation of kinetic measurements with different peptide substrates.

Proteinase Activity Measurements. EIAV proteinase was purified to apparent homogeneity from concentrated virus using reversed-phase HPLC, as described in the preceding paper (Tözsér et al. 1993). Recombinant HIV-1 proteinase was expressed in *Escherichia coli* and purified as described previously (Louis et al., 1989; Wondrak et al., 1991). Oligopeptides were synthesized by solid-phase peptide synthesis on a Model 430A automated peptide synthesizer (Applied Biosystems, Inc) or a semiautomatic Vega peptide synthesizer (Vega-Fox Biochemicals) using *tert*-butoxycarbonyl chemistry and purified by reversed-phase high-performance chromatography (RP-HPLC) (Copeland & Oroszlan, 1988). Amino acid composition of the peptides was determined by amino acid analysis on either a Durrum D-500 or a Waters Pico-Tag analyzer. The proteinase assays were performed in 0.25 M potassium phosphate buffer, pH 5.6, containing 7.5% glycerol, 5 mM dithiothreitol, 1 mM EDTA, 0.2% Nonidet P-40, and 2 M NaCl. The reaction mixture containing 0.4 mM peptide was incubated at 37 °C for 1 h or for 24 h. The reaction was stopped by the addition of guanidine hydrochloride (6 M final concentration), acidified with trifluoroacetic acid, and subjected to RP-HPLC separation, as described previously (Tözsér et al., 1991, 1992, 1993). Amino acid analysis of the collected peaks was used to confirm the site of cleavage with HIV PR (Bláha et al., 1991; Tözsér et al., 1992). For the EIAV PR, cleavage products were identified by the retention time, which was found to be identical to that obtained with HIV PR. Relative activities were calculated from the molar amounts of peptides cleaved per unit time, by dividing the activity on a given peptide by the activity on the unmodified substrate (SP-211), at less than 20% substrate turnover, as described in Bláha et al. (1991). Measurements were performed in duplicate, and the average values were calculated. The error was less than 10%.

RESULTS AND DISCUSSION

Model Structure of EIAV PR. Structural alignment of the amino acid sequences of several retroviral proteinases showed that EIAV PR resembles both RSV and HIV proteinases (Weber, 1989). HIV-1 PR and RSV PR have about 30% identical amino acids and share a similar conserved three-dimensional structure (Weber, 1990a). They differ in length, and the 25 extra residues in RSV PR are accommodated in extended surface loops in three major regions (the flap and turns between β-strands b and c, and b′ and c′) and three smaller insertions compared to the structure of HIV proteinase. The alignment of the EIAV PR amino acid sequence based on the elements of secondary structure (Figure 1) suggested that it resembles the HIV PR more closely and has similar turns between β-strands b and c, and b′ and c′. The length

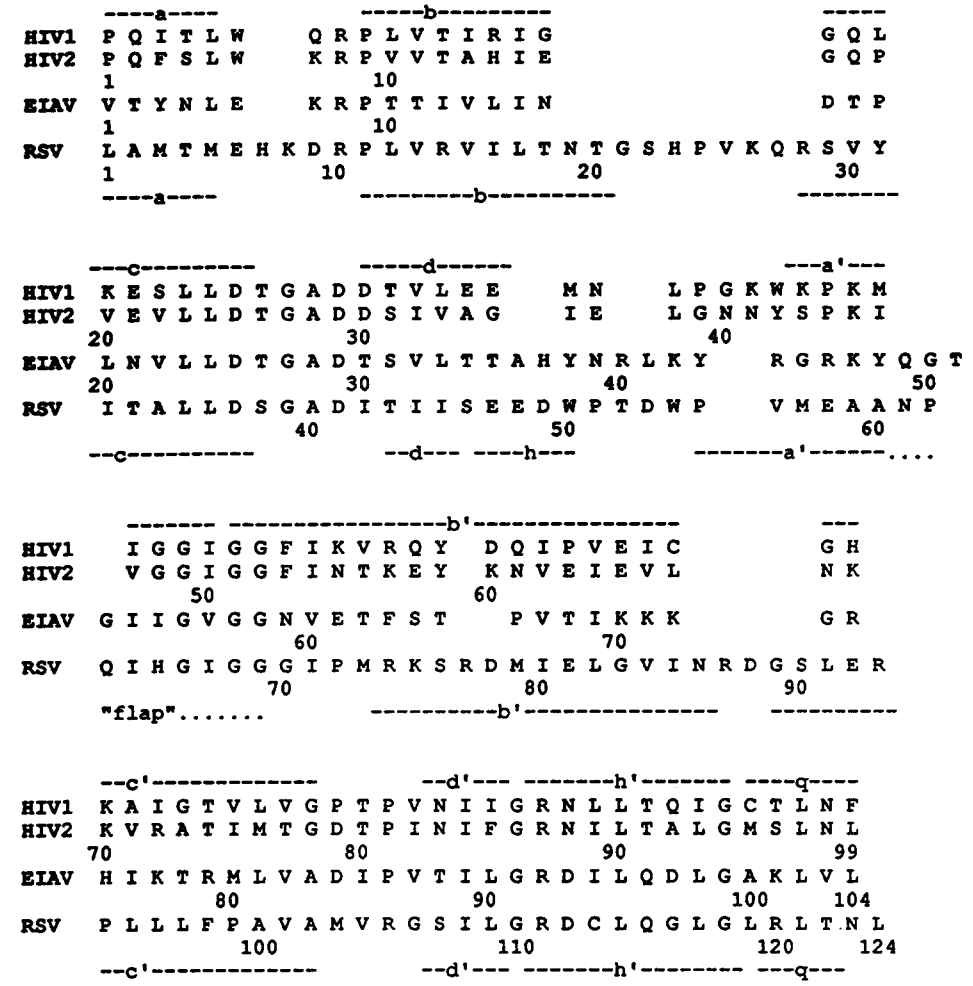

FIGURE 1: Sequence comparison of HIV-1, HIV-2, EIAV, and RSV proteinases based on the structural alignment of Weber (1989). The elements of secondary structure observed for the crystal structures of HIV-1 and RSV PR are indicated by letters a through d and q for β-strands and h for helix. The dotted line indicates the residues in the flap that were not visible in the crystal structure of RSV PR (Miller et al., 1989a).

Table I: Comparison[a] of Amino Acid Sequences of EIAV, HIV, and RSV PRs

	HIV-1	HIV-2	RSV PR
EIAV PR			
no. identical	30	32	25
no. similar	11	13	18
no. identical + similar	41	45	43
no. identical residues in substrate-binding site	15	11	10

[a] The sequence alignment of Figure 1 was used for this analysis. Similar residues: Val/Leu/Ile, Asp/Glu, Asn/Gln, Ser/Thr, Arg/Lys/His, Phe/Tyr/Trp.

and sequence of residues in the flap of EIAV PR suggested that this region is more similar to RSV PR than to HIV-1 PR. Table I compares the amino acid sequence of EIAV PR with those of HIV-1, HIV-2, and RSV proteinases. EIAV PR has 30% identical amino acid residues compared to HIV-1 PR and 25% compared to RSV PR. However, combining the numbers of identical and similar residues, these retroviral proteinases all share about 42% similarity in amino acid sequence with EIAV PR. This supports our prediction that the three-dimensional structures are similar and share a conserved core structure. If the 20 residues that form the inhibitor binding region are examined (not including the insertions in RSV and EIAV proteinases), then HIV-1 and -2 are the most similar with 16 identical residues, followed by HIV-1 and EIAV PRs which share 15 identical residues. The other pairs of proteinases, including HIV-2 compared with EIAV PR, share only 10–11 identical substrate binding residues. Therefore, in many respects, the substrate specificity of EIAV PR is expected to resemble that of HIV-1 PR.

Energy Minimization of EIAV PR Model. The model coordinates were energy-minimized using XPLOR in order to remove any close contacts between atoms. The minimization resulted in coordinates with a root mean square (rms) deviation of 0.73 Å for αC atoms, 0.81 Å for main chain atoms, and 1.06 Å for all atoms compared with the starting model. This is slightly less than observed on minimization of a model for one domain of an ion channel (Kumar & Weber, 1992). Comparison of the minimized and the starting model structures showed that the largest changes in the backbone atoms were located around the residues 16–17 of the β b to c turn, residues 26–30 near the catalytic Asp 25, residues 47–57 of the flaps, residues 64–67 around the single residue deletion relative to HIV PR, and residues 71–74 in the β b′ to c′ turn. The ends of the model substrate also show deviations. The differences near the catalytic aspartates are probably due to the simplified model for the partial charges (Sansom et al., 1992). The other differences are all at surface loops or at regions of deletion or insertion relative to the structure of HIV-1 PR. These changes were smaller and more evenly distributed throughout the structure than those obtained for minimization of the HIV-1 PR–inhibitor crystal structures using the BIOSYM program, Discover, as described by Sansom et al. (1992). The energy-minimized model for the dimer of EIAV PR is

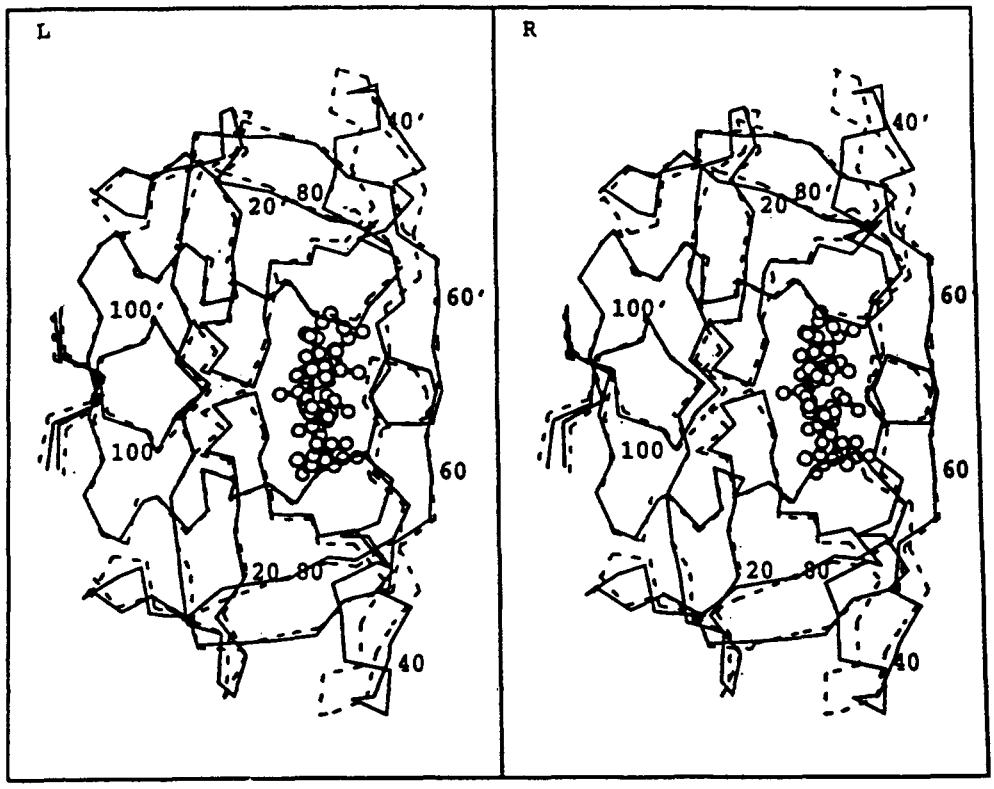

FIGURE 2: Stereoview of superposition of α-carbon atoms of the model structure of the dimer of EIAV PR (continuous lines) with HIV-1 PR crystal structure (Swain et al., 1990) (dashed lines). The inhibitor from the cocrystal structure with HIV-1 PR is shown in a ball-and-stick representation.

compared to the crystal structure of HIV-1 PR complexed with inhibitor in Figure 2.

Accuracy of Model Structure. Analysis of experimental structures of homologous proteins (Chothia & Lesk, 1986) suggests that the EIAV and HIV proteinases with 30% identical residues would be expected to have about 1.45 Å rms deviation for main chain atoms of the conserved core structure. This is consistent with comparison of the crystal structures of RSV and HIV proteinases which have an rms deviation of 1.45 Å for 86 common αC atoms in the subunit, and also show about 30% identical amino acid residues. The energy minimization of the model structure of EIAV PR resulted in an rms deviation of 0.81 Å for main chain atoms, which is less than the total expected deviation of 1.45 Å from the atoms of HIV-1 PR. Previous analysis of model structures has indicated that the structure of the active site and substrate binding residues can be predicted relatively accurately, while the conformation of surface loops is very difficult to predict accurately (Weber, 1990b; Greer, 1990). The structural alignment of the amino acid sequences suggested that the EIAV PR subunit structure should resemble HIV-1 PR, except for two regions: residues 36–52 were predicted to resemble residues 48–63 of RSV PR, and there is a single amino acid deletion of residue 60 in HIV PR. The surface loops and regions around deletions or insertions were examined for favorable interactions with other parts of the protein which may indicate that the conformation was modeled correctly, as described for HIV-2 proteinase (Gustchina & Weber, 1991).

Surface Loops in the Model Structure. The turn between antiparallel β-strands b and c is the same length in HIV and EIAV proteinases; however, the turn in HIV-1 PR is between two glycines, and these are replaced by Asn-Asp in the EIAV PR. The backbone atoms were adjusted to obtain negative φ torsion angles. The final model after energy minimization showed several stabilizing interactions among the side chains in this region. Lys 42, in the region leading to the flap that was modeled by analogy to RSV PR, formed hydrogen bond and ionic interactions with Asn 16 and Asp 17 at the β b to c turn. The side chain of Asn 39 also forms hydrogen bond interactions with the side chains of Asp 17 and Thr 18. Then, the hydroxyl of Thr 88 interacts with the carbonyl oxygen of residue 21 in β c and the His 37 side chain; these hydrogen bonds connect β-strand d′ with β c and the region leading to the flap. Equivalent interactions occur in the crystal structure of RSV PR between the side chains of Asp 49 in the region leading to the flap and Ser 107 in β d′.

The region leading to the flap in EIAV PR differs from that of HIV-1 PR, and it was modeled to resemble the corresponding region in RSV PR (Figure 2). It is possible that this region is not in the correct conformation, especially since surface loops are the most difficult to model accurately (Weber, 1990; Greer, 1990). Several stabilizing interactions are observed for this region in the model structure of EIAV PR. The hydroxyl of Tyr 38 has hydrogen bond interactions with the carbonyl oxygens of residues 42 and 43, and the hydroxyl of Thr 65 near the deletion in β b′. Tyr 38 is in an equivalent position to that of Trp 50 in RSV PR. There are interactions between amino acid side chains which connect the two strands of the flap in the model structure. A salt bridge is formed between Arg 46 and Glu 61 connecting β-strands a′ and b′. The hydroxyl of Thr 51 forms hydrogen bond interactions with the carbonyl oxygen and the hydroxyl of Thr 62. Asn 59, which replaces Phe 53 in HIV-1 PR, forms a hydrogen bond interaction with the carbonyl oxygen of residue 53. This is close to the region of the flap in RSV PR where His 65 was predicted to form a hydrogen bond to the carbonyl oxygen of Gly 69 and stabilize the conformation of the flap (Grinde et al., 1992). The interactions observed in the model of the EIAV PR indicate that this region may be modeled reasonably correctly from residue 46 to 62.

The sequence alignment of Figure 1 shows that there is a single amino acid insertion or deletion next to HIV PR residue 60 in β b′ in the RSV and EIAV PR sequences. In HIV-1 PR, Gln 61 forms a bulge in β b′ next to β c′. HIV-1 PR Tyr 59 OH interacts with the carbonyl oxygen of Leu 38 in the region leading to the flap that differs in conformation among the proteinases. The side chain of Asp 60 interacts with the side chains of Lys 43, Gln 58, and Thr 74 and connects β a′, b′, and c′. In RSV PR there is an additional residue and a different conformation for residues 76–79 compared to HIV-1 PR. The side chain of Asp 78 interacts with Asn 19 and His 23 from the extended loop between β b and c that is absent in HIV-1 PR. In contrast, EIAV PR shows a single amino acid deletion compared to HIV-1 PR. This deletion was modeled by removing the bulge in β b′, so that there are continuous β-sheet hydrogen bonds between β b′ and c′ in this region of EIAV PR. This is expected to be an energetically favorable structure. The amino acid side chains no longer form hydrogen bond interactions with other regions as seen in HIV-1 and RSV PR. This is partly because the longer β b to c loop of RSV PR is absent in EIAV PR, and the residues leading to the flap differ from those of HIV-1 PR.

The surface turn between β-strands b′ and c′ also showed differences in the two experimental structures of HIV-1 and RSV proteinases. This was built to resemble HIV-1 PR in the model of EIAV PR. One unusual feature is that this turn has several basic residues in EIAV PR with Lys-Lys-Lys-Gly-Arg at residues 70–74. In HIV-1 PR, the basic side chain of His 69 may form a salt bridge with the negatively charged C terminus of residue 99′. The corresponding residue in EIAV PR is Arg 74 which can interact in a similar manner with the C terminus at residue 104′. On the other side of the β turn, Glu 65 in HIV-1 PR forms a salt bridge with Arg 14, providing an interaction between β b and b′. Arg 14 is replaced by Leu in EIAV PR, so a different interaction between β b and b′ is observed in the model structure. Lys 71 on β b′ forms hydrogen bond interactions with the hydroxyls of Thr 11 on β b, and Tyr 3 on β a. The model structure of HIV-2 PR also showed conserved interactions between residues in different regions of secondary structure in HIV-1 and -2 proteinases (Gustchina & Weber, 1991). Such interactions may be important for stabilizing the tertiary structure.

The structures of both HIV-1 and RSV proteinases show a similar β-sheet formed by the four termini in the dimer (Weber, 1990a). The model structure of EIAV PR has a similar β-sheet formed by the termini that is more similar to HIV-1 than to the β-sheet in RSV PR. There is an ionic interaction between the negatively charged C terminus and the positively charged N terminus. The C terminus is also stabilized by an ionic interaction with Arg 74′. In addition to the β-sheet hydrogen bonds between the amides and carbonyl oxygens of adjacent β-strands, several side chains form interconnecting hydrogen bonds. The side chain of Asn 4 forms hydrogen bonds with the side chains of Lys 101′ and Glu 6. The hydroxyl of Tyr 3 interacts with Lys 71 and possibly the hydroxyl of Thr 11. Interactions among amino acid side chains of the β-sheet formed by the termini were observed in the crystal structures of HIV-1 and RSV PR (Weber, 1990a). In each case, the exact interactions differ, although the effect is to stabilize the β-sheet and the nearby turn between β b′ and c′. In HIV-1 PR, His 69 from the β b′ to c′ turn interacts with the C terminus. In RSV PR, Asn 86 and Glu 92 from the same turn interact with the N terminus and Asn 123, the residue before the C terminus. Residues Lys 71 and Arg 74 are predicted to make similar interactions in the model of EIAV PR, connecting the β b′ to c′ turn with the C terminus and with Tyr 3 near the N terminus. The C-terminal β-strand q provides a large contribution to the subunit–subunit interaction in the PR dimer, and a tetrapeptide representing the C-terminus of HIV-1 PR was found to inhibit dimer formation (Zhang et al., 1991).

Substrate Binding Site of EIAV PR. The residues forming the substrate binding region of EIAV PR have been located in the model structure, and significant PR–substrate interactions have been described. Analysis of the crystal structures of HIV-1 PR with different inhibitors (Miller et al., 1989b; Erickson et al., 1990; Fitzgerald et al., 1990; Swain et al., 1990; Jaskolski et al., 1991) has shown that the inhibitor is bound in an extended conformation and forms two short β-sheets with residues 27–29 on one side and residues from the antiparallel β-strands of the flaps on the other side (Gustchina & Weber, 1990). The side chain of each inhibitor residue from P4 to P3′ lies in successive subsites S4 to S3′ formed by PR residues. The model built for EIAV PR shows similar interactions. Potential hydrogen bond interactions between EIAV PR and the substrate main chain carbonyl oxygens and amides are illustrated in Figure 3. Most are equivalent to interactions observed in the experimental structures, including the hydrogen bonds linking the conserved water molecule to the flaps and the inhibitor. In addition to the interactions observed in HIV-1 PR, the extra residues in the flaps of EIAV PR provide the possibility of hydrogen bond interactions between Gly 52 and the NH groups of P4 and P5′. It is possible that the NH of P5 also can form hydrogen bond interactions with EIAV PR, although the exact interaction is more difficult to predict. This would extend the size of peptide that is recognized by EIAV PR, and the experimental measurements suggest that the most efficiently cleaved substrate extends from P5 to P4′ or P3′ (Tözsér et al., 1993).

The residues that are predicted to form the substrate binding sites, S4 to S4′, are compared to the equivalent residues in HIV-1 PR in Table II. Only one to three amino acids in each subsite differ in EIAV compared to HIV-1 PR, with the exception of S4 and S4′ where the additional residues of the flaps are predicted to contribute to the subsites. Another interesting feature is that a different set of residues are changed, compared to the differences between HIV-1 and -2 PR subsites (Tözsér et al., 1993). In the subsites of HIV-2 PR, residues 32, 47, 76, and 82 differ from those in HIV-1 PR, and these are all conservative changes among the hydrophobic amino acids, Val/Ile/Leu/Met. In the subsites of EIAV PR, there are changes in residues 30, 48, 50, 74, and 80. Only residue 50 has the conservative Ile/Val change. The other differences are in polar or charged amino acids: Asp/Thr 30, Thr/Arg 74, Thr/Ile 80, and the change Gly/Ile 48 that adds a hydrophobic side chain. This suggests that EIAV PR and HIV-1 PR should show more differences in specificity than HIV-1 and -2 PR.

Correlation of Structure and Activity of EIAV and HIV-1 Proteinases. The differences observed on comparison of the model structure of EIAV PR and the crystal structure of HIV-1 PR have been correlated with measurements of relative activities for peptide substrates. The substrates were chosen as peptides containing single amino acid substitutions in positions P4 to P3′ of the peptide, V-S-Q-N-Y*-P-I-V-Q-NH2, that represents the MA-CA cleavage site of HIV-1. A variety of amino acid substitutions were tested at each position. The relative activities of EIAV and HIV-1 proteinases for each of these peptides are summarized in Table III. Previously,

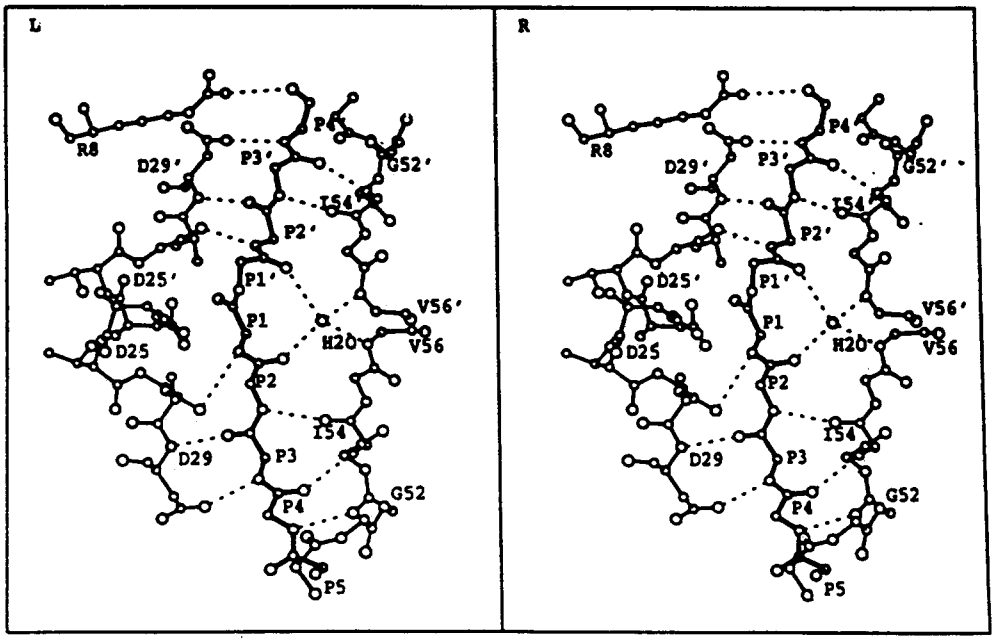

FIGURE 3: Stereoview of the model substrate in binding site of the EIAV enzyme showing the hydrogen bond interactions with the substrate NH and C=O groups. The main chain of the substrate is shown in thick lines for P5 to P4'. EIAV PR residues 8, 25–29, 25'–29', and the main chain atoms of residues 50–56 and 52'–56' from the flaps are shown in thin lines. Hydrogen bonds are indicated by dashed lines.

Table II: Residues[a] Forming the Subsites of EIAV and HIV-1 PRs

subsite	HIV-1/EIAV residues
S4	Arg 8', Asp 29, **Asp/Thr 30 (Gly 50, Thr 51, Gly 52)**, Ile 47/53, **Thr 74/Arg 79**, Leu 76/81
S3	Arg 8', Leu 23', Asp 29, **Gly 48/Ile 54**, Gly 50/55, Pro 81'/86', Val 82'/87'
S2	Ala 28, Asp 29, **Asp/Thr 30**, Val 32, Ile 47/53, **Gly 48/Ile 54**, Gly 49/55, **Ile 50'/Val 56'**, Leu 76/81, Ile 84/89
S1	Arg 8', Leu 23', Asp 25', Asp 25, Gly 27, **Gly** 48/**Ile 54**, Gly 49/55, **Ile 50/Val 56, Thr 80'/Ile 85'**, Pro 81'/86', Val 82'/87', Ile 84'/89'
S1'	Arg 8, Leu 23, Asp 25, Asp 25', Gly 27', **Gly 48'/Ile 54'**, Gly 49'/55', **Ile 50'/Val 56', Thr 80/Ile 85**, Pro 81/86, Val 82/87, Ile 84/89
S2'	Ala 28', Asp 29', **Asp/Thr 30'**, Val 32', Ile 47'/53', **Gly 48'/Ile 54'**, Gly 49'/55', **Ile 50/Val 56**, Leu 76'/81', Leu 84'/89'
S3'	Arg 8, Leu 23, Asp 29', **Gly 48'/Ile 54'**, Gly 50'/55', Pro 81/86, Val 82/87
S4'	Arg 8, Asp 29', **Asp/Thr 30', (Gly 50', Thr 51', Gly 52')**, Ile 47'/53', **Thr 74/Arg 79'**, Leu 76'/81'

[a] Amino acid residues in the second subunit of the dimer are indicated by a prime. The residues that differ in the two PRs are indicated in bold type as HIV-1 PR residue/EIAV PR residue. The EIAV PR residue numbers change after residue 36 due to an insertion relative to HIV-1 PR. The residues in the extra loop of the EIAV PR flap are in parentheses.

Table III: Relative Activities Obtained with EIAV and HIV-1[a] Proteinases for Substrates Having Single Amino Acid Substitutions from Position P4 to P3'[b]

substitutions	Ser P4	Gln P3	Asn P2	Tyr P1	Pro P1'	Ile P2'	Val P3'
Ala	1.2 (0.3)	0.14 (0.34)	5.1 (0.53)	0 (<0.01)	0 (0)[c]	0.06 (0.17)[c]	0.24 (0.34)
Leu	1.5 (0.02)	<0.01 (0.40)	0.74 (0.06)	0.09 (0.21)	<0.01 (0)	0.24 (0.44)[c]	5.2 (3.2)
Val	1.4 (0.05)	<0.01 (0.71)	2.5 (0.17)	0 (0)	<0.01 (0)	1.0 (0.93)[c]	
Ile	1.2 (0.03)		1.3 (0.1)	0 (0)			
Phe	1.3 (0.03)	0.23 (0.22)	0.16 (0.03)	1.4 (1.7)	<0.01 (0)	0.06 (0.07)[c]	5.9 (2.4)
Gly	0.12 (0.76)	0.21 (0.18)	0.09 (0.12)	0 (0)		<0.01 (<0.01)[c]	0.10 (0.24)
Ser				0 (0)	0 (0)		0.27 (0.36)
Thr			3.24 (0.02)				0.10 (0.40)
Met	0.65 (0.04)			0.15 (0.23)	<0.01 (0)		
Asp		<0.01 (<0.01)	0.15 (0.13)	0 (0)		<0.01 (<0.01)[c]	<0.01 (0.05)
Glu							0.16 (0.1)
Gln						<0.01 (0.03)	0.73 (0.67)
Lys	0.25 (0.06)	0.08 (0.12)	0 (0)	0 (0)	0 (0)[c]	0 (0)[c]	0.09 (0.1)
Arg	0.14 (0.15)						0.41 (0.57)
Cys			7.3 (0.92)	0.09 (<0.01)			

[a] Numbers in parentheses. [b] Relative activities were calculated by determining the molar amount of peptide cleaved and dividing the activity on a given peptide by the activity obtained with the unmodified substrate Val-Ser-Gln-Asn-Tyr*Pro-Ile-Val-Gln-NH_2 (SP-211) at 0.4 mM concentration. [c] Taken from Bláha et al. (1991).

the activities of HIV-1 and -2 proteinases were compared using a similar series of oligopeptides (Tözsér et al., 1991, 1993).

S4 binding site: One of the significant differences between HIV-1 PR and EIAV PR is the presence of additional residues 49–52 in the region of EIAV PR leading to the tip of the flap (Figures 1 and 4). The exact conformation of these residues is not easy to model; however, it is predicted to provide additional interactions with substrate residues P4 and P5 and to increase the length of substrate peptide that can be efficiently

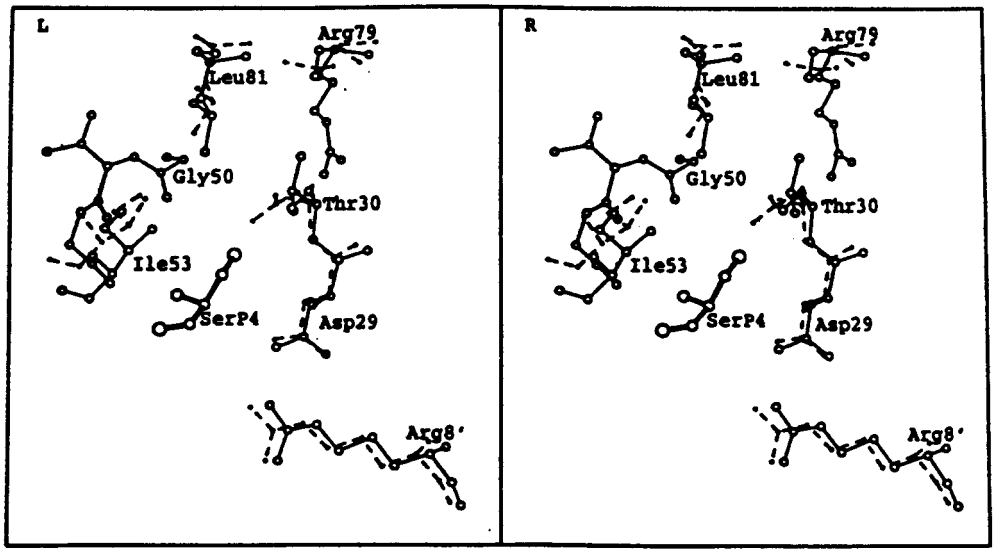

FIGURE 4: Stereoview of the S4 subsite of EIAV PR (thin lines) with Ser at P4 of the peptide substrate (thick lines). The corresponding residues from the crystal structure of HIV-1 PR are shown in dashed lines. Residues 49–52 of EIAV PR have no equivalent in HIV-1 PR. P4 Ser hydroxyl can form a hydrogen bond interaction with the hydroxyl of Thr 30 (not shown).

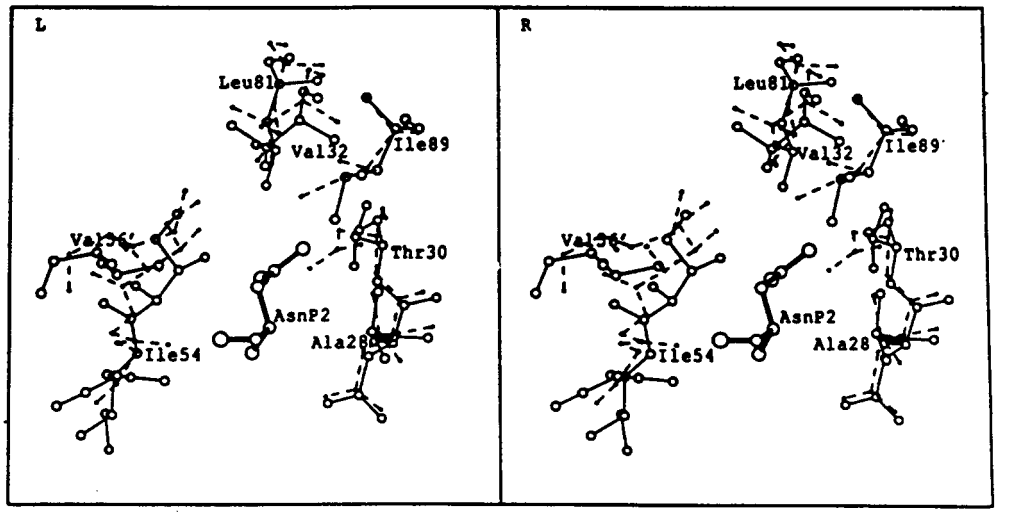

FIGURE 5: Stereoview of the S2 subsite of EIAV and HIV-1 PR with Asn at P2 of the peptide substrate in a representation similar to that of Figure 4.

bound and cleaved. This is in agreement with tests of different lengths of peptide on the basis of the MA-CA cleavage site of HIV-1 (Tözsér et al., 1993). The substitution of Asp 30 to Thr as well as the presence of Arg 79 in this pocket actually turns a negatively charged pocket into a positively charged one (Table II). When the relative activities of EIAV and HIV-1 PR are compared for substitutions of the P4 position, it is clear that hydrophobic residues Leu, Val, Phe, Ile, and Ala form better substrates for EIAV than for HIV-1 PR (Table III). This is partly due to the presence of the additional residues 49–52 in the flaps of EIAV PR, which make subsite S4 more enclosed (Figure 4). EIAV PR also has the more hydrophobic Thr instead of Asp 30 in HIV PR, and hydrophobic residues at P4 can interact with the aliphatic side chain of Arg 79. Gly at P4, which lacks a side chain, is a poor substrate for EIAV PR, while it makes a good substrate for the less enclosed S4 subsite of HIV-1 PR. Similar results have been observed for the substrates of RSV PR, which also has additional flap residues contributing to S4 (Cameron et al., 1993).

S3 and S3' binding sites: The main difference appears to be the Gly 48 to Ile change, which adds a hydrophobic side chain at one side of the subsite. These subsites were predicted to accommodate a variety of side chains, and many of the peptides containing substitutions at the P3' position were hydrolyzed by both HIV and EIAV PR (Table III). The activities of EIAV and HIV-1 PR toward substitutions of P3 were similar, Gln formed the best substrate for both enzymes. The model structure suggested that the side chain of P3 Gln has the possibility of forming hydrogen bond interactions with the side chain of Arg 8' from the PR and with Tyr OH at P1. S3 is more hydrophobic in EIAV PR than in HIV-1 PR due to the Gly 48/Ile 54 change (Table II); however, this does not correlate with the kinetic measurements for Gln or Leu at P3. Results of substrate cleavage by a mutant of RSV PR which has Gly substituted for the wild-type His at this position suggested that this residue primarily affects the S1 and S1' subsites (Grinde et al., 1992). The differences at P3' are more consistent with the model structure. Leu and Phe at P3' form the best substrates for both enzymes. The more hydrophobic S3' site in EIAV compared to HIV-1 PR due to the Gly 48 to Ile 54 change means that the larger hydrophobic Phe and Leu at P3' are better substrates for EIAV than for HIV-1 PR.

S2 and S2' binding sites: Subsite S2 is illustrated in Figure 5. Subsites S2 and S2' show the nonconservative change of

Asp 30 to Thr in EIAV PR (Table II), which will make the site more hydrophobic and less negatively charged. Substitution of Ile 50 to Val 56 introduces a smaller hydrophobic side chain, and substitution of Gly 48 to Ile 54 introduces a larger hydrophobic side chain that is probably directed away from the subsite (Figure 5). These changes suggest that somewhat larger and more hydrophobic side chains may be accommodated at P2 and P2′ in EIAV PR compared to HIV PR. In fact, at P2, the medium to small size hydrophobic residues Ala, Val, Ile, Cys, and Leu are substantially better substrates of EIAV than HIV-1 PR, as is Thr which is partly hydrophobic. The β-branched side chains of Val, Ile, and Thr are accommodated better in the S2 and S2′ subsites of EIAV proteinase due to the presence of the smaller Val 56 instead of the larger Ile 50 of HIV-1 proteinase. Phe is too large to be a good substrate at P2 or P2′ for either enzyme. At P2′, Val and Ile form the best substrates for both enzymes, while the substitution of Ala or Leu results in poorer substrates for EIAV compared to HIV PR. Asymmetrical effects of substitutions at P2 and P2′ were also observed on comparison of HIV-1 and -2 proteinases (Tözsér et al., 1992a). Some of these effects may be due to the sequence of the adjacent residues in the substrate. Results from comparisons of RSV and HIV-1 PR specificities have suggested that the subsites P1′ and P2′ are smaller than the P1 and P2 subsites, respectively.

Subsites S1 and S1′: Pro 81 and Val 82 of HIV-1 PR form a "roof" for subsites S1 and S1′ and have a significant effect on the specificity for P1 and P1′, as described in Grinde et al. (1992). These residues are conserved in EIAV PR. In RSV PR, His 65 is present instead of Gly 48 of HIV-1 PR, and mutation of His 65 to Gly results in inactive proteinase. In EIAV PR, the changes are Gly 48 to Ile 54, Ile 50 to Val 56, and Thr 80 to Ile 85 (Table II), which tend to make the subsites S1 and S1′ slightly smaller and more hydrophobic compared to HIV-1 PR. Both Ile 54 and Val 56 are on the side of the subsite and may have less effect on specificity than the residues forming the "roof", although there may be small changes in the size and shape of residues that can be readily accommodated. The activity measurements are very similar for both proteinases, the large hydrophobic Phe and Tyr form the best substrates at P1, and the medium size hydrophobic Met and Leu also form substrates (Table III). Only Pro at P1′ forms a substrate for both HIV-1 and EIAV proteinases.

CONCLUSIONS

The molecular model for EIAV PR shows many similarities with the structure of HIV-1 PR. One exception is the conformation of the region leading to the flaps which is predicted to resemble RSV more closely than HIV PR, and provides three additional residues in subsites S4 and S4′. Only five residues forming the subsites for substrate binding differ in the two enzymes. These are Asp/Thr 30 which contributes to subsites S4, S2, S2′, and S4′; Gly 48/Ile 54 which forms part of subsites S3, S1, S1′, and S3′; Ile 50/Val 56 which forms part of subsites S2, S1, S1′, and S2′; Thr 74/Arg 79 which contributes to subsites S4 and S4′; and Thr 80/Ile 85 which forms part of subsites S1 and S1′. These differences in the subsites of the two enzymes have been correlated with the relative activities toward peptide substrates with single amino acid substitutions of P4 to P3′. As expected, the activities of EIAV and HIV-1 proteinases are similar for substitutions at most positions. Substantial differences are observed in the relative activities of the two enzymes for peptides with substitutions at the P4 and P2 positions. Hydrophobic residues at P4 form better substrates for EIAV than for HIV-1 proteinases, due to the presence of the additional flap residues 50–52 that contribute to the S4 subsite. Small hydrophobic residues at P2 form good substrates for EIAV PR, while Asn is the best substrate for HIV-1 proteinase. This is predicted to be due to the changes of Asp 30 in HIV-1 to Thr in EIAV PR and Ile 50 to Val 56, which decrease the size and increase the hydrophobicity of the S2 subsite of EIAV PR.

ACKNOWLEDGMENT

We thank Ivo Bláha, Terry D. Copeland, and Patrick Wesdock for the synthesis of the peptide substrates, Cathy Hixson and Suzanne Specht for the amino acid analyses, and John M. Louis and Ewald M. Wondrak for the purified HIV-1 proteinase.

REFERENCES

Bláha, I., Nemec, J., Tözsér, J., & Oroszlan, S. (1991) *Int. J. Pept. Protein Res. 38*, 453–458.
Brunger, A. T. (1992) XPLOR Version 3.0 Manual, Yale University.
Cameron, C. E., Grinde, B., Jacques, P., Jentoft, J., Leis, J., Wlodawer, A., & Weber, I. T. (1993) *J. Biol. Chem.* (in press).
Chothia, C., & Lesk, A. M. (1986) *EMBO J. 5*, 823–826.
Copeland, T., & Oroszlan, S. (1988) *Gene Anal. Tech. 5*, 109–115.
Erickson, J., Neidhart, D. J., VanDrie, J., Kempf, 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., Knigge, M. (1990) *Science 249*, 527–533.
Fitzgerald, P. M. D., & Springer, J. P. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 299–320.
Fitzgerald, P. M. D., McKeever, B. M., VanMiddlesworth, J. F., Springer, J. P., Heimbach, J. C., Leu, C.-D.-T., Herbert, W. K., Dixon, R. A. F., & Darke, P. L. (1990) *J. Biol. Chem. 265*, 14209–14219.
Greer, J. (1990) *Proteins: Struct., Funct., Genet. 7*, 317–334.
Grinde, B., Cameron, C. E., Leis, J., Weber, I., Wlodawer, A., Burstein, H., & Skalka, A.-M. (1992) *J. Biol. Chem. 267*, 9491–9498.
Gustchina, A. & Weber, I. T. (1990) *FEBS Lett. 269*, 269–272.
Gustchina, A., & Weber, I. T. (1991) *Proteins: Struct., Funct., Genet. 10*, 325–339.
Jaskólski, M., Tomasselli, A., Sawyer, T., Staples, D., Heinrikson, R., Schneider, J., Kent, S., & Wlodawer, A. (1991) *Biochemistry 30*, 1600–1609.
Jones, A. T. (1985) *Methods Enzymol. 115*, 157–171.
Kräusslich, H., Oroszlan, S., & Wimmer, E., Eds. (1989) *Viral Proteinases as Targets for Chemotherapy, Current Communications in Molecular Biology*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Kumar, V., & Weber, I. T. (1992) *Biochemistry 31*, 4643–4649.
Lapatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J. R., Whittle, P. J., Danley, D. E., Geoghegan, K. F., Hawrylik, S. J., Lee, S. E., Scheld, K. G., & Hobart, P. M. (1989) *Nature 342*, 299–302.
Louis, J. M., Wondrak, E. M., Copeland, T. D., Smith, C. A. D., Mora, P. T., & Oroszlan, S. (1989) *Biochem. Biophys. Res. Commun. 159*, 420–425.
Miller, M., Jaskólski, M., Mohana Rao, J. K., Leis, J., & Wlodawer, A. (1989a) *Nature 337*, 576–579.
Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S., & Wlodawer, A. (1989b) *Science 246*, 1149–1152.
Navia, M. A., Fitzgerald, P. M. D., McKeever, B. M., Leu, C.-T., Heimbach, J. C., Herber, W. K., Sigal, I. S., Darke, P. L., & Springer, J. P. (1989) *Nature 337*, 615–620.

Oroszlan, S., & Luftig, R. B. (1990) *Curr. Top. Microbiol. Immunol. 157*, 153–185.
Powell, M. (1977) *Math. Programming 12*, 241–254.
Roberts, M. M., & Oroszlan, S. (1989) *Biochem. Biophys. Res. Commun. 160*, 486–494.
Roberts, M. M., Copeland, T. D., & Oroszlan, S. (1991) *Protein Eng. 4*, 695–700.
Sansom, C., Wu, J., & Weber, I. T. (1992) *Protein Eng. 5*, 659–667.
Skalka, A.-M. (1989) *Cell 56*, 911–913.
Stephens, R. M., Casey, J. W., & Rice, N. R. (1986) *Science 231*, 589–594.
Swain, A. L., Miller, M. M., Green, J., Rich, D. H., Kent, S. B. H., & Wlodawer, A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8805–8809.
Tang, J., James, M. N. G., Hsu, I. N., Jenkins, J. A., & Blundell, T. L. (1978) *Nature 271*, 618–621.
Tomasselli, A., Hui, J., Sawyer, T., Staples, D., Bannow, C., Reardon, I., Howe, J., DeCamp, D., Craik, C., & Heinrickson, R. (1990) *J. Biol. Chem. 265*, 14675–14683.
Tözsér, J., Gustchina, A., Weber, I. T., Bláha, I., Wondrak, E. M., & Oroszlan, S. (1991) *FEBS Lett. 279*, 356–360.
Tözsér, J., Weber, I. T., Gustchina, A., Bláha, I., Copeland, T. D., Louis, J. M., & Oroszlan, S. (1992) *Biochemistry 31*, 4793–4800.
Tözsér, J., Friedman, K., Weber, I. T., Bláha, I., & Oroszlan, S. (1993) *Biochemistry* (preceding paper in this issue).
Weber, I. T. (1989) *Gene 85*, 567–569.
Weber, I. T. (1990a) *J. Biol. Chem. 265*, 10492–10496.
Weber, I. T. (1990b) *Proteins: Struct., Funct., Genet. 7*, 172–184.
Weber, I. T., Miller, M., Jaskólski, M., Leis, J., Skalka, A. M., & Wlodawer, A. (1989) *Science 243*, 928–931.
Wlodawer, A., Miller, M., Jaskólski, M., Sathyanarayana B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J., & Kent, S. B. H. (1989) *Science 245*, 616–621.
Wondrak, E. M., Louis, J. M., Mora, P. T., & Oroszlan, S. (1991) *FEBS Lett. 280*, 347–350.
Zhang, Z.-Y., Poorman, R. A., Maggiora, L. L., Heinrikson, R. L., & Kezdy, F. J. (1991) *J. Biol. Chem. 266*, 15591–15594.