Hydrophilic Peptides Derived from the Transframe Region of Gag-Pol Inhibit the HIV-1 Protease†

John M. Louis,*,‡,§ Fred Dyda,|| Nashaat T. Nashed,⊥ Alan R. Kimmel,§ and David R. Davies||

Laboratories of Chemical Physics, Molecular Biology, Bioorganic Chemistry, and Cellular and Developmental Biology, National Institute of Diabetes, Digestive and Kidney Diseases, The National Institutes of Health, Bethesda, Maryland 20892

Received August 19, 1997; Revised Manuscript Received December 2, 1997

ABSTRACT: The HIV-1 transframe region (TFR) is between the structural and functional domains of the Gag-Pol polyprotein, flanked by the nucleocapsid and the protease domains at its N and C termini, respectively. Transframe octapeptide (TFP) Phe-Leu-Arg-Glu-Asp-Leu-Ala-Phe, the N terminus of TFR, and its analogues are competitive inhibitors of the action of the mature HIV-1 protease. The smallest, most potent analogues are tripeptides: Glu-Asp-Leu and Glu-Asp-Phe with K_i values of ∼50 and ∼20 μM, respectively. Substitution of the acidic amino acids in the TFP by neutral amino acids and D or retro-D configurations of Glu-Asp-Leu results in an >40-fold increase in K_i. Protease inhibition by Glu-Asp-Leu is dependent on a protonated form of a group with a pK_a of 3.8; unlike other inhibitors of HIV-1 protease which are highly hydrophobic, Glu-Asp-Leu is extremely soluble in water, and its binding affinity decreases with increasing NaCl concentration. However, Glu-Asp-Leu is a poor inhibitor (K_i ∼ 7.5 mM) of the mammalian aspartic acid protease pepsin. X-ray crystallographic studies at pH 4.2 show that the interactions of Glu at P2 and Leu at P1 of Glu-Asp-Leu with residues of the active site of HIV-1 protease are similar to those of other product−enzyme complexes. It was not feasible to understand the interaction of intact TFP with HIV-1 protease under conditions of crystal growth due to its hydrolysis giving rise to two products. The sequence-specific, selective inhibition of the HIV-1 protease by the viral TFP suggests a role for TFP in regulating protease function during HIV-1 replication.

Retroviruses encode their structural and functional proteins through translation of polycistronic messenger RNAs into precursor proteins (*1*). The Gag polyprotein consists of the structural proteins in the arrangement MA−CA−p2−NC−p1−p6[1] (see Figure 1; *2*−*4*). The Gag-Pol polyprotein is translated via a mechanism in which a −1 frame shift of an adenosine residue changes the open reading frame from the Gag to the Pol at a frequency of 5−10% (*5*, *6*). This ribosomal frame shift site corresponds to the second codon within p1 leading to the synthesis of the transframe region (TFR or p6*) which links the structural Gag domain to the functional Pol domain [Figure 1 (Genbank accession number HIVHXB2CG)]. Thus, the structure of Gag-Pol is MA−CA−p2−NC−TFR−PR−RT−IN (*4*, *7*).

The homodimeric HIV-1 protease (PR) is a member of the family of aspartic acid proteases (*1*, *8*) and is responsible for its own maturation from Gag-Pol (*1*, *9*, *10*) and processing of the viral polyproteins for producing mature structural and functional proteins (*1*). The precise steps of assembly and maturation of the polyproteins are crucial in the life cycle of HIV-1. Premature activation of the protease or partial inhibition of its activity during virus maturation leads to defects in viral assembly and consequently to the formation of noninfectious aberrant particles (*11*, *12*).

In an effort to understand the detailed mechanism of the autoprocessing of the viral protease from the Gag-Pol polyprotein, we had examined the kinetics of maturation of the protease from a model polyprotein (*13*) that contains only the two native protease cleavage sites that flank the protease domain. The mature protease (PR) is liberated from the polyprotein by a two-step mechanism (*9*). The initial and faster step involves the hydrolyses of the peptide bonds at the $p6^{pol}$−PR junction via an intramolecular mechanism (see Figure 1). This cleavage at the N terminus of PR is accompanied by a large increase in protease activity. The flanking C-terminal sequences, which do not appear to influence the catalytic activity of the protease precursor, are cleaved subsequently via an intermolecular process (*10*).

Although no specific in vivo structural or functional role has been ascribed to the HIV-1 TFR, the N-terminal flanking sequence of PR, a major deletion within the TFR leads to enhanced processing of the viral Gag-Pol, as compared to that of the wild-type Gag-Pol, by the HIV-1 protease (*14*).

† This research was supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to J.M.L. and D.R.D.).

* Corresponding author at Laboratory of Chemical Physics, NIDDK, Bldg. 5, Room 411, NIH, Bethesda, MD 20892. Telephone: (301) 594-3122. Fax: (301) 496-0825. E-mail: jmlouis@helix.nih.gov.

‡ Laboratory of Chemical Physics.

§ Laboratory of Cellular and Developmental Biology.

|| Laboratory of Molecular Biology.

⊥ Laboratory of Bioorganic Chemistry.

[1] Abbreviations: HIV-1, human immunodeficiency virus type 1; TFR, transframe region; TFP, transframe peptide Phe-Leu-Arg-Glu-Asp-Leu-Ala-Phe; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; PR, HIV-1 protease; Phe(NO_2), 4-nitrophenylalanine; Nle, norleucine; RP-HPLC, reverse-phase high-performance liquid chromatography; DMSO, dimethyl sulfoxide. The nomenclature of HIV-1 proteins is according to Leis et al. (*4*): MA, matrix; CA, capsid; NC, nucleocapsid; PR, protease; RT, reverse transcriptase; IN, integrase.

S0006-2960(97)02059-X
Published on Web 02/07/1998

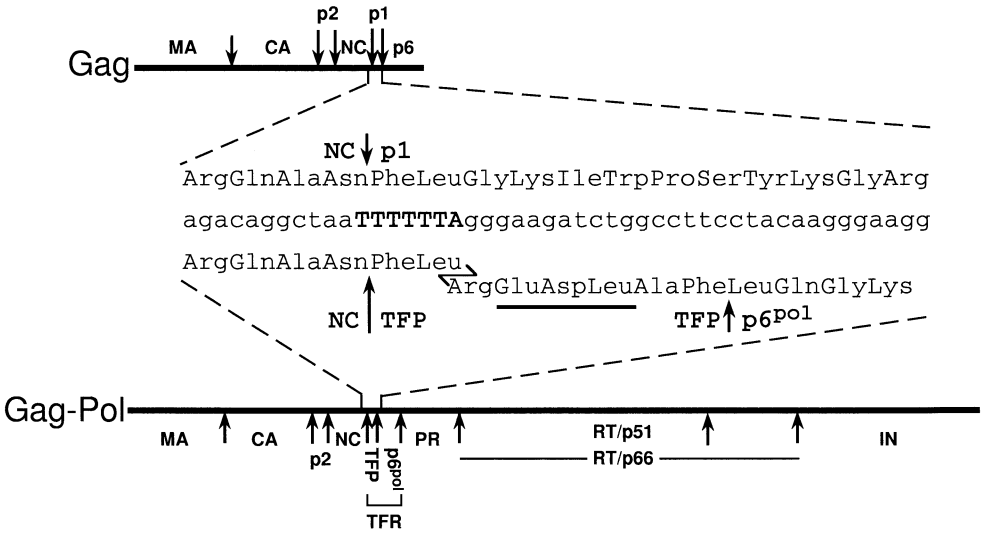

FIGURE 1: Nucleotide and protein sequence surrounding the translation frame shift site in HIV-1. The Gag and Gag-Pol polyproteins of HIV-1 are indicated by bold lines. Arrows indicate the specific sites of cleavage by the viral PR. The DNA and encoded amino acids spanning the NC−p1 junction in the Gag polyprotein and NC−TFP and TFP−p6^{pol} junctions in the Gag-Pol polyprotein are aligned. Bold letters in the DNA sequence denote the conserved signal sequence required for ribosomal frame shifting. TFP denotes the eight conserved amino acids flanking NC. The remainder of the transframe region (TFR) constitutes p6^{pol}. The N-terminal dipeptide sequence of TFP is common with p1. The second amino acid of TFP is either Leu or Phe at a ratio of ∼7:3, which is dependent upon variation in the frame-shifting mechanism (*5*, *6*) and does not alter the remaining sequence of TFP. The underlined amino acid sequence represents the tripeptide core of TFP that inhibits the PR. The coding sequence for the tripeptide Glu-Asp-Leu corresponds to a region in the viral RNA between conserved regions of the signal sequence and the RNA secondary structure (*5*, *6*).

These results imply that TFR or a sequence within this region has a regulatory function for the autoprocessing of the protease from the Gag-Pol precursor.

The HIV-1 TFR consists of two domains (see Figure 1), an N-terminal transframe peptide (TFP) Phe-Leu-Arg-Glu-Asp-Leu-Ala-Phe and the remaining 48 amino acids (p6^{pol}), which are separated by a protease cleavage site (*7*, *15*; J. M. Louis and E. M. Wondrak, manuscript in preparation). While the p6^{pol} domain is highly variable among HIV-1 isolates (*16*, *17*), the N-terminal TFP domain, which shares two amino acids that are common with p1, is well conserved (*16*).

In this study, we have focused on the N-terminal domain of TFR and shown that peptides derived from TFP and its analogues are sequence-specific, selective inhibitors for the action of the mature HIV-1 protease. We have studied the mechanism of inhibition and structure−activity relationship of TFP and its derivatives and show that the hydrophilic tripeptides Glu-Asp-Leu and Glu-Asp-Phe are the most potent inhibitors of the enzyme.

To gain further insight into the sequence-specific interactions of the TFP with the HIV-1 protease, we cocrystallized the protease in the presence of TFP. However, analysis of the crystal with the octapeptide TFP was not possible due to its hydrolysis to form a protease−pentapeptide Phe-Leu-Arg-Glu-Asp complex and a free tripeptide Leu-Ala-Phe. Next, we analyzed the crystal structure of Glu-Asp-Leu, the hydrophilic tripeptide analogue of TFP, complexed to the protease. From these studies, we conclude that the intact TFP may bind to the enzyme, leading to its hydrolysis being different in conformation from that of the tripeptide Glu-Asp-Leu which we describe at 2 Å resolution. We have compared the latter complex to other known protease−product complexes. The inhibitory octapeptide TFP spans the sequence of two half-substrates (NC−TFP and TFP−p6^{pol}; see Figure 1) of the viral protease and is the smallest of the known proteolytic products of the Gag and Gag-Pol polyproteins. Direct evidence that TFP is formed in vivo comes from the sequence determination of peptides isolated from HIV (*3*, *7*). It is conceivable that TFP plays a role in regulating the action of the protease during virus replication.

MATERIALS AND METHODS

Enzymes. HIV-1 protease was purified (*13*; J. M. Louis, unpublished), and its active site concentration was determined as previously described (*10*). The purified enzyme was stored as a 5−8 μM stock solution in 50 mM sodium acetate at pH 5.0, 1 mM DTT, 1 mM EDTA, and 0.05% reduced Triton X-100 at −80 °C. A stock solution of 1−2 mg/mL porcine pepsin (Sigma Chemical Co., St. Louis, MO) was prepared in 50 mM formic acid at pH 3.0.

Synthetic Peptides. With the exception of dipeptides in Table 1, all peptides were synthesized and purified by California Peptide Research Inc. (Napa, CA). In all cases, they were at least 98% homogeneous as determined by chromatography, and their peptide content ranged from 74 to 79%. All peptides were analyzed for their amino acid composition, and their mass spectra showed the expected molecular ions. The dipeptides in Table 1 were purchased from Sigma Chemical Co. Stock solutions of all peptides were prepared in water except for peptides **1** and **2** (see Table 1) which were dissolved in 50 mM sodium bicarbonate. The solution concentration of peptides containing phenylalanine was determined spectrophotometrically.

Kinetics and Inhibition. HIV-1 protease activity was measured at 25 °C using substrate **I**, Lys-Ala-Arg-Val-Nle-Phe(NO_2)-Glu-Ala-Nle-NH_2, as previously described (*9*, *10*). In a typical assay, 2.5 μL of the enzyme was added to 97.5 μL of buffer (see the figure legends) and 2.5 mM DTT in the absence or presence of varying concentrations of the inhibitor. The reaction was initiated by the addition of 10

Table 1: Inhibition of the Action of HIV-1 Protease by TFP Phe-Leu-Arg-Glu-Asp-Leu-Ala-Phe and Its Analogues[a]

	peptide	K_i (μM)
1	Phe-Leu-Arg-Glu-Asp-Leu-Ala-Phe	98 ± 10
2	Gly-Glu-Asp-Leu-Ala-Phe	64 ± 13
3	Glu-Asp-Leu-Ala	160 ± 20
4	Glu-Asp-Leu	50 ± 9
5	Arg-Glu-Asp	3360 ± 400
6	Asp-Leu-Ala-Phe	270 ± 25
7	Asp-Glu-Leu	230 ± 20
8	Glu-Glu-Leu	310 ± 25
9	Asp-Asp-Leu	140 ± 6
10	Glu-Asn-Leu	3070 ± 210
11	Glu-Lys-Phe	2010 ± 380
12	Glu-Asp-Phe	25 ± 3
13	Asp-Asp-Phe	43 ± 7
14	Phe-Asp-Asp-Phe	55 ± 14
15	Asp-Asp	970 ± 140
16	Asp-Glu	5400 ± 380
17	Asp-Gln	>1000
18	Glu	>50,000
19	Asp	>50,000
20	Phe-Leu-Arg-Gln-Asn-Leu-Ala-Phe	>2000
21	D-Glu-D-Asp-D-Leu	>2000
22	D-Leu-D-Asp-D-Glu	>2000

[a] K_i was obtained from a plot of the apparent K_m vs [**I**] for peptides **1**, **2**, and **4** as shown in Figure 2. For the others, K_i was obtained from a plot of $1/V$ vs [**I**] at a saturating concentration of the substrate. Protease was assayed in the presence of the inhibitor in 50 mM sodium formate at pH 4.25 and 2.5 mM DTT. The final enzyme and substrate concentrations were 150 nM and 390 μM, respectively. Peptides **13** and **14** were assayed at pH 4.4 in the presence of 0.08 and 0.02% DMSO, respectively. All amino acids are of the L configuration unless otherwise noted.

μL of a 4 mM solution of substrate **I** in water and monitored by following the decrease in absorption at 310 nm ($\Delta\epsilon$ = 1800) in a 100 μL spectrophotometric cell. Pepsin was assayed using substrate **II**, Phe-Gly-His-Phe(NO_2)-Phe-Ala-Phe-OCH_3 (*18*). This reaction was monitored by following the decrease in absorbance at 269 nm in a final volume of 110 μL of 50 mM formate buffer (pH 3.0) containing 150 nM and 160 μM enzyme and substrate, respectively, at 25 °C. Concentrations of substrate solutions were determined spectrophotometrically at 280 nm (ϵ = 12 000; *9*, *10*).

X-ray Crystallography. HIV-1 protease (100 μL, 5−7.5 mg/mL) in 50 mM sodium formate at pH 4.2 and 2.5 mM DTT was mixed with 6 μL of a 150 mM solution of tripeptide Glu-Asp-Leu in water or with 25 μL of TFP in 50 mM $NaHCO_3$. The enzyme−inhibitor complex (3 μL) was then mixed with an equal volume of buffer A at pH 4.0 to give a final composition of 60 mM sodium citrate, 132 mM NaH_2PO_4, 5 mM DTT, and 1.1 M sodium formate. This mixture was placed in an airtight vessel over 1 mL of buffer A. Repeated macroseeding was used to obtain diffraction quality single crystals. The crystal used for data collection was cryoprotected in 15% glycerol and subsequently flash-frozen in liquid propane. X-ray diffraction data were collected at 95 K on a Raxis-IIC imaging plate detector mounted on an RU-200 rotating anode generator operated at 50 kV and 100 mA using double-mirror focused CuKα radiation. The data were integrated and internally scaled with the HKL suite (*19*) of programs. The single crystal used for X-ray data collection was about 0.3 mm in size, belonged to the orthorombic space group *P*21212, with unit cell parameters of a = 58.47 Å, b = 86.44 Å, and c = 45.81 Å, containing one HIV-1 protease dimer in its asymmetric unit (Table 3). Because the structure was nearly isomorphous with respect to the crystal structure of the HIV-1 protease−pepstatin complex, it was possible to start refinement with the HIV-1 protease coordinates from Protein Data Bank entry 5HVP (*20*). Rigid body refinement, simulated annealing, and energy minimization followed by individual restrained *B* factor refinement were carried out with X-PLOR (*21*). The model was displayed, and the tripeptide was built into the difference density with O (*22*). The coordinates have been deposited with the Protein Data Bank with the assigned identification code 1a30.

Table 2: K_i's Obtained from a Plot of $1/V$ vs [**I**] at a Saturating Concentration of the Substrate[a]

peptide	[NaCl] (M)	K_i (μM)
Glu-Asp-Leu	0	52 ± 10
	1	435 ± 65
	2	1001 ± 170
Glu-Asp-Phe	0	15 ± 2
	1	410 ± 25
Asp-Leu-Ala-Phe	0	270 ± 25
	1	3260 ± 390

[a] The enzyme was assayed in the presence of the inhibitor in 50 mM sodium formate at pH 4.25 and 2.5 mM DTT. The final enzyme and substrate concentrations were 150 nM and 390 μM, respectively.

Table 3: X-ray Data Collection and Refinement Statistics[a]

number of reflections ($I > 0.0$)	65424
number of unique reflections	14407
completeness (%)	
(40.0−2.0 Å)	88.7
(2.06−2.0 Å)	72.5
number of unique reflections in refinement [$F > 2\sigma(F)$, 40.0−2.0 Å]	13987
number of non-hydrogen atoms in refinement	1791
R factor (%)	18.9
free R factor (%)	22.7

[a] R factor: $\Sigma|F_{obs}-F_{calc}|/\Sigma F_{obs}$. The free R factor is reported on a randomly selected 5% of unique reflections excluded from refinement.

RESULTS AND DISCUSSION

Inhibition of HIV-1 Protease by Transframe Peptide TFP and Its Analogues. The octapeptide Phe-Leu-Arg-Glu-Asp-Leu-Ala-Phe (TFP, see Figure 1; peptide **1** in Table 1) and peptides corresponding to truncated forms (peptides **3** and **4**) are competitive inhibitors for HIV-1 protease-catalyzed hydrolysis of substrate (see Materials and Methods, Table 1, and Figure 2). Control experiments show that the rate of HIV-1 protease-catalyzed reaction initiated by the addition of the enzyme (E) to a mixture of inhibitor [I (peptide **1** or **4**)] and substrate (S) is equal to that initiated by the addition

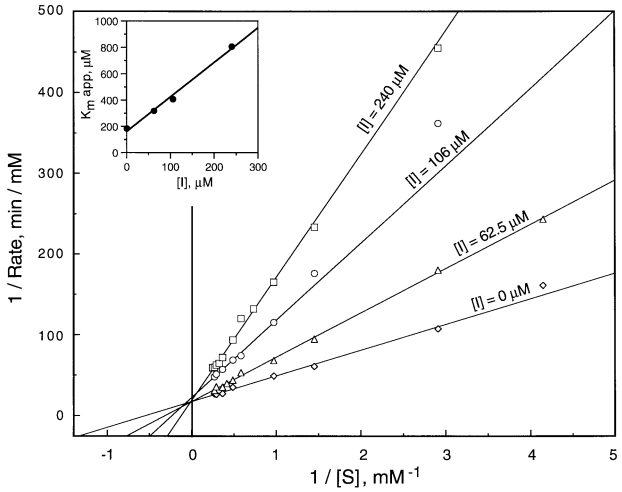

FIGURE 2: Lineweaver−Burk plots of HIV-1 protease-catalyzed hydrolysis of substrate **I** at various concentrations of the inhibitor tripeptide Glu-Asp-Leu in 50 mM sodium formate at pH 4.25 containing 2.5 mM DTT and 150 nM HIV-1 protease at 25 °C. The lines are based on values of K_m and k_{cat} calculated from nonlinear curve fitting of the Michaelis−Menten equation to the data. The inset is a plot of the apparent K_m vs inhibitor concentration.

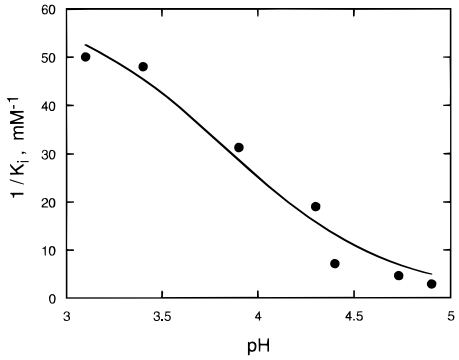

FIGURE 3: Plot of the inhibition constant (K_i) by Glu-Asp-Leu for HIV-1 protease-catalyzed hydrolysis of substrate **I** vs pH. K_i's were obtained from plots of $1/V$ vs [**I**] at a saturating concentration of the substrate (390 μM). Reactions were carried out between pH values of 3 and 5 in 50 mM sodium formate or 100 mM sodium acetate buffers containing 2.5 mM DTT and 150 nM HIV-1 protease at 25 °C. The line is a calculated curve with a pK_a of 3.8 and a limiting K_i value of 20 μM.

of S to a preincubated mixture (10 min) of E and I. This indicates a fast equilibrium between the EI complex and the free E and I. The measured K_i's for compounds **1**−**22** are summarized in Table 1. The shortest and most potent native sequence that produces inhibition that is comparable to that of TFP is the tripeptide Glu-Asp-Leu (peptide **4**). The potency of this tripeptide is comparable to that of pepstatin A and its analogues, the best known inhibitors of aspartic proteases, assayed under similar conditions (*18*, *23*, *24*). While Glu-Asp-Leu is also a competitive inhibitor for the action of the mammalian cellular aspartic acid protease pepsin, the K_i value for this inhibition is about 150-fold higher (∼7.5 mM at pH 3.0) than that for the inhibition of HIV-1 protease.

Inhibition of the action of the mature HIV-1 protease by TFP and its derivatives is sequence-specific. Substitution of the two acidic amino acids (Glu and Asp) of TFP (Table 1, peptide **20**) with neutral residues (Gln and Asn) increases the K_i by a factor of >40. A single substitution of either of these two acidic amino acids in the tripeptide Glu-Asp-Leu results in an increase in K_i. An interchange or duplication of the acidic residues of the tripeptide results in a 3−6-fold increase in K_i (peptides **7**−**9**). A substitution of the charged Asp to a neutral Asn (peptide **10**) or major substitutions (peptide **11**) in Glu-Asp-Leu increases K_i by a factor of >40. Incorporation of native Arg at the N terminus but deletion of Leu at the C terminus to produce the tripeptide Arg-Glu-Asp (peptide **5**) gives an inhibitor 50-fold less effective than Glu-Asp-Leu (peptide **4**) and suggests that a hydrophobic amino acid flanking the carboxyl terminus of the charged Asp is critical. The only change that shows a modest reduction in K_i substitutes Phe for Leu in peptide **4** to produce Glu-Asp-Phe (peptide **12**). The spatial arrangement of the two acidic residues within the tripeptide is critial for the inhibitory activity. D (peptide **21**) and retro-D (peptide **22**) configurations of Glu-Asp-Leu and dipeptides (peptides **15**−**17**) are very poor inhibitors; glutamic or aspartic acid do not inhibit HIV-1 protease up to 50 mM.

Although Glu-Asp-Leu is a competitive inhibitor for the HIV-1 protease, its interaction with the enzyme is different from that of other known competitive inhibitors such as pepstatin. The binding of Glu-Asp-Leu to the HIV-1 protease requires the protonated form of an ionizable group with a pK_a of 3.8 (Figure 3). Glu-Asp-Leu is a highly charged molecule that displays good solubility in water (>112 mg/mL). Other competitive inhibitors of HIV-1 protease that are derived from peptide sequences of known substrates are generally hydrophobic and have poor water solubility (*18*, *25*). These latter inhibitors have been shown to bind to the hydrophobic cavity of the enzyme's active site, and their binding affinity is enhanced with increasing NaCl concentrations (*18*, *25*). However, for inhibitors Glu-Asp-Leu (peptide **4**), Asp-Leu-Ala-Phe (peptide **6**), and Glu-Asp-Phe (peptide **12**), increasing NaCl concentrations decrease their affinity for the protease (Table 2). These results suggest that the interactions between the enzyme and inhibitor are different from those of the other competitive inhibitors and are mostly hydrophilic in nature, even though a hydrophobic residue at the C terminus may be required (see Table 1).

X-ray Crystallographic Analysis of HIV-1 Protease in Complex with TFP and the Tripeptide Glu-Asp-Leu. To define the details of the interactions between HIV-1 protease and TFP, we first cocrystallized TFP (8-fold excess) with the mature HIV-1 protease at pH 4.2 (see Materials and Methods). The crystals were in the space group *P*21212. The electron density map, however, did not correspond to a bound, intact octapeptide but to a smaller peptide, suggesting that TFP may have been hydrolyzed under the conditions of crystal growth. This interpretation was assessed by enzyme assays, performed under conditions similar to that used for crystallization. The parent peptide TFP and potential product peaks were fractionated by reverse-phase HPLC and subjected to mass spectrometric analyses. Results clearly indicated that TFP is hydrolyzed between Asp and Leu, yielding the pentapeptide Phe-Leu-Arg-Glu-Asp and the tripeptide Leu-Ala-Phe. Although data suggest that the pentapeptide of TFP remains bound to the protease while the tripeptide is released, it was difficult to define precise interactions due to orientational disorder.

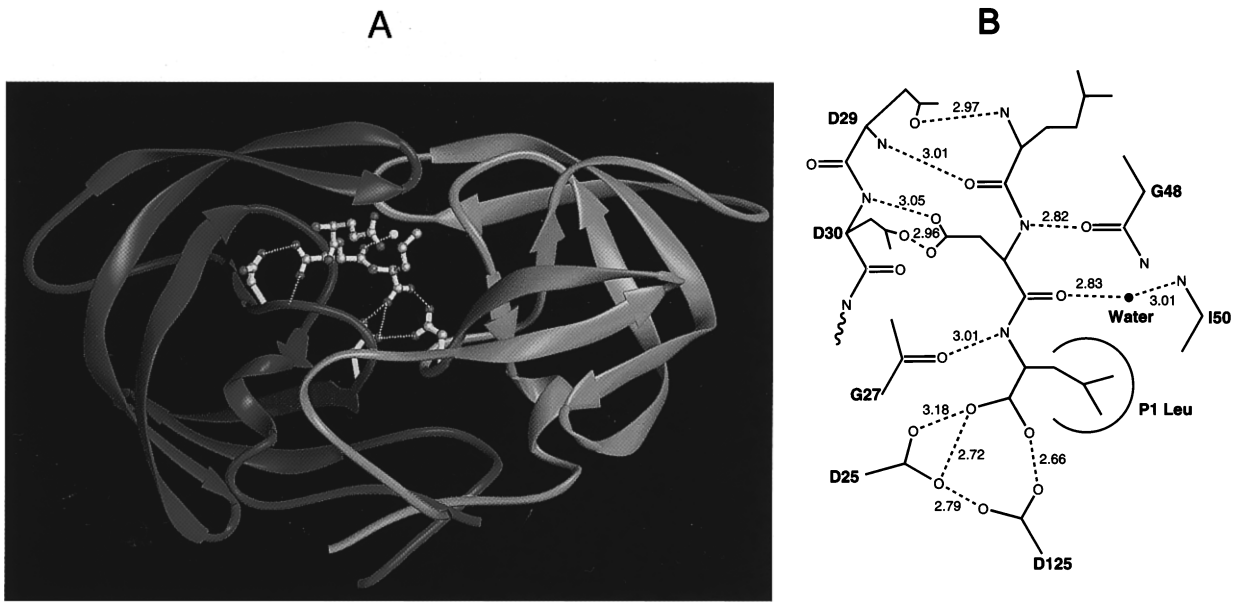

FIGURE 4: (A) Crystal structure of the HIV-1 protease bound to the tripeptide inhibitor Glu-Asp-Leu. One protease monomer is shown in green and the other in orange. The tripetide is represented as a ball-and-stick model. The Asp 25 side chains of both monomers of protease and Asp 30 of one monomer are shown in yellow. Hydrogen bond interactions between the protein and the inhibitor are shown in dotted lines. The figure was prepared using the program RIBBONS 2.5 (*31*). (B) Hydrogen bond interactions between the tripeptide inhibitor Glu-Asp-Leu and the HIV-1 protease.

Next, we investigated the interaction of the smaller potent analogue Glu-Asp-Leu with the protease by X-ray crystallography. At the end of the refinement, clear difference electron density appeared spanning the active site of the enzyme. In one half-site, this difference electron density could unambiguously be interpreted as density representing the tripeptide Glu-Asp-Leu, while density in the other half-site was lower, and possibly the sum of several different and lower occupancy binding modes of the tripeptide. This observed difference between the two half-sites is very likely the result of the different packing environment in the crystal lattice. The bulk of the interactions between the enzyme and the tripeptide are hydrophilic (Figure 4A,B), with the exception of the Leu side chain at P1 occupying a hydrophobic pocket formed by Gly 49 and Ile 50 from one of the flap β-strands from one monomer and Leu 23, Pro 81, and Ile 84 from the other monomer. All of the tripeptide main chain nitrogen and carbonyl oxygen atoms participate in a hydrogen bonding network with the enzyme. The N-terminal Glu of the inhibitor is 3.0 Å from OD of Asp 29, while its carbonyl oxygen is hydrogen bonded to the main chain nitrogen of Asp 29. The main chain nitrogen of the Asp residue of the tripeptide is hydrogen bonded to the carbonyl oxygen of Gly 48, while its carbonyl oxygen is one of the hydrogen bonds coordinating the flap-bridging water molecule. The nitrogen of the Leu is 3.0 Å from the carbonyl oxygen of Gly 27.

The terminal carboxylate group of the Leu residue interacts with both active site aspartic acid side chains. One of the carboxylate oxygens is hydrogen bonded to both oxygens of Asp 25 in one monomer, while the other oxygen is hydrogen bonded only to OD1 of Asp 125 in the other monomer. OD2 of the latter aspartic acid is hydrogen bonded to OD2 of Asp 25, a hydrogen bond which is a standard feature of HIV-1 protease. The side chain of the Glu of the inhibitor appears to be disordered and is presumed to have multiple conformations, interacting with both the enzyme and the solvent. The side chain of the Asp residue of the tripeptide participates in two hydrogen bond interactions with the enzyme. OD1 is 3.0 Å from OD1 of Asp 30, while OD2 is 3.1 Å from the main chain nitrogen of the same residue. These two hydrogen bonds appear to provide a unique sequence-specific interaction between the tripeptide inhibitor and the enzyme. The existence of the hydrogen bond between the two usually identically charged side chains is consistent with the pH dependence of K_i for the tripeptide.

The observed interactions between the terminal carboxylate at Leu and the active site aspartic acid residues indicate that the inhibitor is bound to the enzyme in a fashion similar to that of the product of the hydrolytic reaction. Recently, Rose and colleagues have reported structures of HIV and SIV protease product complexes (*26*). A comparison of the coordinates of the tetrapeptide Ac-SLNF, a p17/p24 cleavage site analogue, complexed to the HIV-1 protease shows that the interactions of Asn at P2 and Phe at P1 with the protease are nearly identical with that of the Asp at P2 and Leu at P1 of the hydrophilic peptide Glu-Asp-Leu as observed here. In both structures, the C-terminal carboxylate group of the peptide makes the same triangular arrangement with the enzyme carboxylates as shown in Figure 4. A single substitution of Asp to Asn in the tripeptide Glu-Asp-Leu increases K_i by >60 times (see Table 2). It is not clear from the structural information why these interactions contribute to the higher affinity of Glu-Asp-Leu to the protease as compared to that of Glu-Asn-Leu. Furthermore, since hydrolysis of TFP results only in two products, cleavage occurring between Asp and Leu suggests that during catalysis TFP binds in a conformation different from that of the tripeptide Glu-Asp-Leu.

TFP and Maturation of Gag-Pol. In vitro kinetic studies of the maturation of Gag-Pol have been restricted by its structural complexity (*9*), and thus, very little is known about the mechanism of regulation of autoprocessing of the protease from the Gag-Pol polyprotein and the pathway for

releasing TFP. However, comparisons of the kinetic parameters of HIV-1 protease-catalyzed hydrolysis of peptide substrates that correspond to the NC−TFP (RQAN*FLRE) and TFP−p6^{pol} (DLAF*LQGK) junctions (see Figure 1; J. M. Louis and E. M. Wondrak, manuscript in preparation) and the neighboring p2−NC (*27*), NC−p1 (*2*), and p6^{pol}−PR (*27*) sites indicate that the rate of cleavage of NC−TFP is at least 10^5 times slower than that of other substrates. Sequence determination of HIV-1 proteins also suggests the presence of TFP (*3*, *7*). Thus, TFP and/or NC−TFP in vivo may play a pH-dependent role in regulating the maturation process as well as the activity of the mature protease. Inhibition of HIV-1 protease by TFP and hydrolysis by the protease of the CA−p2 site, a critical regulatory step for the sequential processing of Gag precursor (*28*), are both optimal at lower pH. The CA−p2 cleavage site has a Glu at P2′ and is likely to be responsible for enhancing substrate binding in a manner similar to that of its analogue Lys-Ala-Arg-Val-Nle-Phe(NO_2)-Glu-Ala-Nle-NH_2 which has a pH optimum of 4.0 for hydrolysis (*29*).

It was shown recently that, in the presence of a potent protease inhibitor (1 μM), Gag-Pol is cleaved in vivo in the vicinity of the N terminus of TFR, leading to the accumulation of a Pol intermediate, TFR−PR−RT−IN (*30*). We suggest that this intermediate is produced by cleavage of the TFP−p6^{pol} site in one of the early steps in the maturation of the protease from Gag-Pol. The other intermediate product of this initial cleavage is presumably MA−CA−p2−NC−TFP (see Figure 1). The insensitivity of the TFP−p6^{pol} cleavage to inhibition in vivo may be the result of a conformational difference of the protease precursor as compared to that of the mature enzyme. It is critical to determine if inhibitors designed for the mature enzyme have equivalent affinity to the protease precursor prior to this cleavage. Inhibition of the protease by TFP and its analogues also suggests the possibility of designing an alternative class of hydrophilic protease inhibitors that may target both the mature protease and intermediates in its maturation as analogues of a putative natural regulatory protein.

ACKNOWLEDGMENT

We are thankful to Drs. H. M. Fales and R. Dugky for mass spectral analyses, Dr. J. M. Sayer for discussions and critical review of the manuscript, and Drs. A. M. Gronenborn and G. M. Clore for support.

REFERENCES

1. Oroszlan, S., and Luftig, R. B. (1990) *Curr. Top. Microbiol. Immunol. 157*, 153−185.
2. Wondrak, E. M., Louis, J. M., de Rocquigny, H., Chermann, J. C., and Roques, B. P. (1993) *FEBS Lett. 333*, 21−24.
3. Henderson, L. E., Bowers, M. A., Sowder, R. C., Serabyn, S. A., Johnson, D. G., Bess, J. W., Jr., Arthur, L. O., Bryant, D. K., and Fenselau, C. (1992) *J. Virol. 66*, 1856−1865.
4. Leis, J., Baltimore, D., Bishop, J. M., Coffin, J., Fleissner, E., Goff, S. P., Oroszlan, S., Robinson, H., Skalka, A. M., and Temin, H. M. (1988) *J. Virol. 62*, 1808−1809.
5. Hatfield, D. L., Levin, J. G., Rein, A., and Oroszlan, S. (1992) *Adv. Virus Res. 41*, 193−239.
6. Jacks, T., Power, M. D., Masiarz, F. R., and Varmus, H. E. (1988) *Nature 331*, 280−283.
7. Gorelick, R. J., and Henderson, L. E. (1994) in *Human retroviruses and AIDS* (Myers, G., Korber, B., Wain-Hobson, S., Jeang, K.-T., Henderson, L., and Pavlakis, G., Eds.) Part III: Analyses, pp 2−5, The Los Alamos National Laboratory, Los Alamos, NM (http://hiv-web.lanl.gov).
8. Wlodawer, A., and Erickson, J. W. (1993) *Annu. Rev. Biochem. 62*, 543−585.
9. Louis, J. M., Nashed, N. T., Parris, K. D., Kimmel, A. R., and Jerina, D. M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 7970−7974.
10. Wondrak, E. M., Nashed, N. T., Haber, M. T., Jerina, D. M., and Louis, J. M. (1996) *J. Biol. Chem. 271*, 4477−4481.
11. Kaplan, A. H., Zack, J. A., Knigge, M., Paul, D. A., Kempf, D. J., Norbeck, D. W., and Swanstrom, R. (1993) *J. Virol. 67*, 4050−4055.
12. Krausslich, H. G. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 3213−3217.
13. Louis, J. M., McDonald, R. A., Nashed, N. T., Wondrak, E. M., Jerina, D. M., Oroszlan, S., and Mora, P. T. (1991) *Eur. J. Biochem. 199*, 361−369.
14. Partin, K., Zybarth, G., Ehrlich, L., DeCrombrugghe, M., Wimmer, E., and Carter, C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 4776−4780.
15. Phylip, L. H., Mills, J. S., Parten, B. F., Dunn, B. M., and Kay, J. (1992) *FEBS Lett. 314*, 449−454.
16. In *Human retroviruses and AIDS* (1994) Sequence Alignment: parts I and II (Myers, G., Korber, B., Wain-Hobson, S., Jeang, K.-T., Henderson, L., and Pavlakis, G., Eds.) The Los Alamos National Laboratory, Los Alamos, NM (http://hiv-web.lanl.gov/).
17. Candotti, D., Chappey, C., Rosenheim, M., M'Pele, P., Huraux, J. M., and Agut, H. (1994) *C. R. Acad. Sci., Ser. III 317*, 183−189.
18. Tropea, J. D., Nashed, N. T., Louis, J. M., Sayer, J. M., and Jerina, D. M. (1992) *Bioorg. Chem. 20*, 67−76.
19. Otwinowski, Z., and Minor, W. (1997) in *Methods in Enzymology* (Carter, C. W., and Sweet, R. M., Eds.) Vol. 276, pp 301−326, Academic Press, San Diego.
20. Fitzgerald, P. M., McKeever, B. M., VanMiddlesworth, J. F., Springer, J. P., Heimbach, J. C., Leu, C. T., Herber, W. K., Dixon, R. A., and Darke, P. L. (1990) *J. Biol. Chem. 265*, 14209−14219.
21. Brünger, A. T. (1992) *X-PLOR Version 3.1, A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
22. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A 47*, 110−119.
23. Huff, J. R. (1991) *J. Med. Chem. 34*, 2305−2314.
24. Rich, D. H. (1985) *J. Med. Chem. 28*, 263−273.
25. Wondrak, E. M., Louis, J. M., and Oroszlan, S. (1991) *FEBS Lett. 280*, 344−346.
26. Rose, R. B., Craik, C. S., Douglas, N. L., and Stroud, R. M. (1996) *Biochemistry 35*, 12933−12944.
27. Tozser, J., Blaha, I., Copeland, T. D., Wondrak, E. M., and Oroszlan, S. (1991) *FEBS Lett. 281*, 77−80.
28. Pettit, S. C., Moody, M. D., Wehbie, R. S., Kaplan, A. H., Nantermet, P. V., Klein, C. A., and Swanstrom, R. (1994) *J. Virol. 68*, 8017−8027.
29. Polgar, L., Szeltner, Z., and Boros, I. (1994) *Biochemistry 33*, 9351−9357.
30. Lindhofer, H., von der Helm, K., and Nitschko, H. (1995) *Virology 214*, 624−627.
31. Carson, M. (1991) *J. Appl. Crystallogr. 24*, 958−961.

BI972059X