Mutational and Structural Studies Aimed at Characterizing the Monomer of HIV-1 Protease and Its Precursor*

Rieko Ishima†, Dennis A. Torchia‡, and John M. Louis¶1

Received for publication, February 13, 2007, and in revised form, April 3, 2007 Published, JBC Papers in Press, April 4, 2007 DOI 10.1074/jbc.M701304200

From the †Department of Structural Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, the ‡Molecular Structural Biology Unit, NIDCR, National Institutes of Health, Bethesda, Maryland 20892-4307, and the ¶Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0520

An experimental protocol for folding the mature human immunodeficiency virus-1 (HIV-1) protease is presented that facilitates NMR studies at a low protein concentration of ~20 μM. Under these conditions, NMR spectra show that the mature protease lacking its terminal β-sheet residues 1–4 and 96–99 (PR5–95) exhibits a stable monomer fold spanning the region 10–90 that is similar to that of the single subunit of the wild-type dimer and the dimer bearing a D25N mutation (PRD25N). Urea-induced unfolding monitored both by changes in 1H-15N heteronuclear single quantum correlation spectra and by protein fluorescence indicates that although PR5–95 monomer displays a transition profile similar to that of the PRD25N dimer (50% unfolded (U50) = ~1.9 M), extending the protease with 4 residues (SFNF) of its N-terminally flanking sequence in the Gag-Pol precursor (SFNFPRD25N) decreases the stability of the fold (U50 = ~1.5 M). Assigned backbone chemical shifts were used to elucidate differences in the stability of the PRT26A (U50 = 2.5 M) and SFNFPRD25N monomers and compared with PRD25N/T26A monomer. Discernible differences in the backbone chemical shifts were observed for N-terminal protease residues 3–6 of SFNFPRD25N that may relate to the increase in the equilibrium dissociation constant (Kd) and the very low catalytic activity of the protease prior to its autoprocessing at its N terminus from the Gag-Pol precursor.

The HIV-1 genome encodes a protease as part of the large Gag-Pol precursor. Like all retroviral proteases, HIV-1 protease is active only as a homodimer. In addition to catalyzing its own autoprocessing at its N terminus, the protease is responsible for cleaving the Gag and Gag-Pol polypeptides at specific sites to produce the mature structural (matrix, capsid, and nucleocapsid) and functional proteins (reverse transcriptase, RNase H, and integrase) crucial for virus maturation and propagation (1, 2). This indispensible role played by the protease in the viral replication cycle motivated extensive studies of structure-based drug design of active site inhibitors of the mature protease (3, 4). Various protease inhibitors are currently being used and developed for the treatment of HIV/AIDS. However, effective long term treatment of AIDS patients has been hampered by the rapid emergence of drug-resistant variants that are less susceptible to inhibition even under highly active antiviral therapy (5, 6). Recent second generation inhibitors of the PR, designed to overcome resistance (i) by reducing the size of their hydrophobic groups so that mutation of the active site residues will have less effect on inhibitor affinity and (ii) by optimizing favorable polar interactions with main chain atoms and to conserved residues, have been shown to curtail drug resistance (7).

The initial steps in the maturation of the protease (Gag-Pol precursor) involve the folding and dimerization of the protease domain when it is in the form of the large Gag-Pol precursor. Previous studies using a mini-precursor, in which the protease is flanked by sequences corresponding to the cleavage sites at its termini, showed that the processing at the N terminus of the protease, which is concomitant with the appearance of catalytic activity, precedes the C-terminal cleavage (8–10). Further examination of protease precursor containing the flanking transframe region sequences (Fig. 1) by kinetics and NMR revealed that the very low catalytic activity of the protease precursor prior to the cleavage at the N terminus is due to a much higher dimer dissociation constant (Kd) as compared with the mature protease (11). Upon intramolecular cleavage at the N terminus, the protease forms a stable dimer and exhibits a very low Kd (<10 × 10−5 M) in 50 mM sodium acetate, pH 5, at 25 °C (8, 13). This suggests that inhibition of the protease function either by preventing or by disrupting dimer formation prior to its maturation provides an attractive avenue for inhibitor design.

For the mature wild-type protease, protein folding appears to be concomitant with dimerization at a concentration of 5 nM and above (11, 14–16). Therefore, a systematic understanding of the folding and dimerization events and the properties of the protease both as a monomer and as a dimer in its precursor and mature forms is essential for the rational design of inhibitors of dimerization. There have been several studies reporting the development of dimerization inhibitors toward the mature HIV-1 protease based on the terminal interface conformation, but to date, the structures of these complexes have not been verified. The low Kd of the mature protease had also precluded

* This work was supported by the National Institutes of Health Intramural Research programs of NIDDK and NIDCR. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 To whom correspondence should be addressed: NIDDK, National Institutes of Health, Bldg. 5, Rm. E3-29, LCP, Bethesda, MD 20892-0520. Tel.: 301-594-3122; Fax: 301-480-4001; E-mail: johnl@intra.niddk.nih.gov.

2 The abbreviations used are: HIV-1, human immunodeficiency virus, type 1; HSQC, heteronuclear single quantum correlation; PR, protease; Nle, norleucine; TFR, transframe region; TFP, transframe peptide; U50, 50% unfolded.
Mutational and Structural Studies of HIV-1 Protease

Despite our extensive studies to characterize the monomer, the propensity for these monomers to aggregate at 0.2–0.3 mM concentrations had precluded further studies to determine the stability of the monomer fold as compared with the dimer, and importantly, to use the monomer to reliably access the binding of potential inhibitors of dimerization. Herein to facilitate these studies, we initially optimized conditions for protein folding and for acquiring $^1$H-$^{15}$N HSQC spectra at $\sim 20 \mu M$ concentration using the high sensitivity cryogenic NMR probe. Optimizing the NMR experiments at low protein concentration notably allowed studies of the urea-induced unfolding of the monomer and dimer by monitoring protein fluorescence under similar equilibrium conditions. Under these optimized conditions, the minimal sequence required for monomer folding was accessed by successive terminal deletions. Subsequently, urea-induced unfolding was monitored both by fluorescence and by NMR, and backbone chemical shifts of the monomers were assigned to characterize the effect of the mutation on the stability of the fold and the monomer–dimer equilibrium. Finally, $^1$H-$^{15}$N HSQC spectra of the monomer and the dimer were acquired in the presence of excess putative dimerization inhibitors of the mature protease. Contrary to the results that suggest that these peptides dissociate the dimer (6, 24), there was no change in the chemical shifts of residues at the termini or the active site. These observations indicate the lack of interaction between either the monomer or the dimer with the inhibitor under the conditions tested.

**EXPERIMENTAL PROCEDURES**

**Protease Constructs**—The mature PR (11) domain in all constructs, optimized for NMR and kinetic studies, bears 5 mutations, Q7K, L33I, L63I to minimize autoproteolysis and C67A detailed studies of the protease monomer until it was observed that subtle mutations of conserved residues involved in interface contacts increase the $K_d$ by several orders of magnitude, thus allowing NMR studies of the monomer to be performed at 0.2–0.3 mM concentrations. The conserved intra-monomer contact between Asp-29 and Arg-87 residues and the active site inter-monomer (interface) contact of Asp-29 and Thr-26 residues play a key role in stabilizing dimer formation with T26A exhibiting the largest increase in $K_d$ (Table 1 and Fig. 2) (17–19). In addition, interface contacts between the terminal 1–4 and 96–99 residues are also critical for dimerization (20). The inter-subunit $\beta$-sheet contacts between the C-terminal strands (96–99/96′–99′), the prime indicates the second subunit) have a larger influence on dimerization than those between the N- and C-terminal $\beta$-strands (1–4/96′–99′ (17)). Thus, the construct bearing the 4-residue deletion at the C terminus of the mature protease, which does not significantly dimerize even up to 1 mM concentration, facilitated the determination of the first three-dimensional structure of the protease monomer by NMR (Fig. 2) (21).

Sequences flanking the protease domain also influence dimer formation, although the exact molecular mechanism of their action on dimerization is not understood. Unlike the flanking C-terminal residues (reverse transcriptase region), which do not appear to affect the catalytic activity or the $K_d$ in fusion with the protease (9, 22, 23), the native transframe region, which flanks the N terminus of the protease in the Gag-Pol precursor, drastically influences the dimer stability of the protease indicated by the very low catalytic activity of the protease prior to its maturation at its N terminus (Fig. 1) (10, 11). Even a two to four residue extension at the N terminus of the protease significantly impairs dimerization similar to the 4-residue N-terminal deletion construct PR$_{5-99}$ (21). Although the N-terminal extensions destabilize the dimer, our NMR investigation of the protease precursor flanked by 56 amino acids of the native transframe region at the N terminus revealed that although the transframe region is mainly unstructured, the protease domain adopts a monomer fold that is similar to that of the mature protease monomer PR$_{5-95}$ (21).

**TABLE 1**

| Construct | Approximate $K_d$ values of protease constructs estimated from $^1$H-$^{15}$N HSQC spectra |
|-----------|----------------------------------------------------------------------------------|
| PR$_{22}$ | $0.5 \mu M$ |
| PR$_{26}$ | $>1000$ |
| PR$_{262}$ | $>1000$ |
| PR$_{26}$ | $>1000$ |
| PR$_{26}$ | $>500$ |
| PR$_{99}$ | $>1000$ |
| PR$_{95}$ | $500$ |
| PR$_{95}$ | $>1000$ |
| PR$_{95}$ | m/u/ua |
| PR$_{95}$ | m/u/ua |
| PR$_{95}$ | m/u/ua |
| PR$_{95}$ | m/u/ua |
| PR$_{95}$ | m/u/ua |
| PR$_{95}$ | m/u/ua |

All constructs bear five mutations, three mutations, Q7K, L33I, and L63I, that restrict autoproteolysis and two mutations, C65A and C95A, to avoid Cys-thiol oxidation. The site of mutation(s) is subscripted. $^\text{m}$ denote four C-terminal residues SFNF of p6pol fused to the N-terminus of PR$_{22}$; $^\text{m}$ denote monomer, unfolded, unfolded aggregate, respectively. At $\sim 20 \mu M$ concentration, PR and PR$_{22}$ are mostly dimeric, and PR$_{262}$ are mostly monomeric. Further deletions into the protease (PR$_{5-99}$, PR$_{5-99}$) lead to destabilization of the monomer fold and aggregation. Addition of DMP323, a potent active site inhibitor, promotes ternary complex formation (dimer + DMP323) of PR$_{262}$, PR$_{26}$, PR$_{95}$, PR$_{99}$, and PR$_{22}$, but not of PR$_{95}$ and PR$_{5-99}$ (21).
and C95A to prevent cysteine-thiol oxidation. Plasmid DNA (pET11a, Novagen, Madison, WI) encoding PR (11) was used with the appropriate oligonucleotide primers to generate the constructs PRD25N and PRD25N/T26A. The PR encoding plasmid was sequentially extended one codon at a time to produce SFNFPR. SFNFPR template was then used to introduce a D25N mutation. Construction of PR 5–99 and PR 1–95 has been described before (11, 17). A stop codon was engineered into constructs PR5–99 to produce PR5–90 and PR5–95 and in PR1–95 to produce PR1–90. All constructs were generated using the QuikChange mutagenesis protocol (Stratagene, La Jolla, CA).

To express PR10–90, the region encoding residues 10–90 was amplified using PR template and appropriate forward and reverse primers and cloned into pET11a vector between NdeI and BamH1 sites. A Gly residue was included at the N terminus of PR10–90 to allow nearly complete excision of the N-terminal methionine and attain N-terminal homogeneity. All constructs were verified both by DNA sequencing and by mass spectrometry.

**Protein Preparation**—*Escherichia coli* BL21(DE3) were grown in minimal media containing 15N ammonium chloride with or without 13C glucose as the sole nitrogen and carbon sources, respectively, at 37 °C and induced for expression. Proteins were purified from inclusion bodies using an established protocol as described previously involving size-exclusion chromatography under denaturing conditions followed by reverse-phase high pressure liquid chromatography. Peak fractions (~0.5 mg/ml) were stored in aliquots at −70 °C. Proteins were dialyzed extensively against 30 mM formic acid and concentrated to ~2 mg/ml using Millipore YM-10 Centriprep and Centrifor concentrators (Millipore, Beford, MA) and stored at 4 °C. Samples were folded for NMR, kinetics, and fluorescence studies freshly when needed using the protocol described in Fig. 3. Concentrations of PR and its mutants are given for a monomer, unless noted otherwise.

**Fluorescence and Kinetics**—Intrinsic fluorescence was recorded using FluoroMax-3 fitted with a Peltier temperature controller (HORIBA Jobin Yvon, Edison, NJ) over a period of 30 min at 20 °C with an excitation and emission wavelength of 285 and 350 nm, respectively, and a bandwidth of 2 nm. Protein maintained in 30 mM formic acid was diluted to a volume of 1.5 ml in 5 mM sodium acetate, pH 6, and the resulting solution was then 2-fold diluted with either 0.1 M sodium acetate, pH 5 (solution A: folded) or 10 mM urea, 0.1 M sodium acetate, pH 5 (solution B: unfolded) to achieve a final protein concentration of ~10 μM in 3 ml. Solution A was transferred into a 4-ml cuvette and allowed to equilibrate in the instrument for 5–10 min, and the fluorescence reading was acquired over a period of 30 min. Normally a stable reading was achieved in about 15 min. For urea denaturation studies, fluorescence was measured in incre-
ments of 0.25 M urea by removing solution A from the cuvette and adding an equal volume of solution B up to a final concentration of 4.75 M urea.

Protease activity was monitored using the substrate Lys-Ala-Arg-Val-Nle-(4-nitrophenylalanine)-Glu-Ala-Nle-NH2 (California Peptide Research, Napa, CA) in 50 mM sodium acetate buffer, pH 5, at 25 °C. Proteins were folded in the same manner as prepared either for fluorescence or for NMR experiments. Protein concentrations were determined both spectrophotometrically (absorbance at 280 nm) and by Bio-Rad assay (Bio-Rad Laboratories). Specific cleavages giving rise to products due to the autoprocessing of active precursor proteases and autoproteolysis of active mature proteases were accessed both by SDS-PAGE and by mass spectrometry.

NMR Spectroscopy—Samples for acquiring NMR spectra were prepared by mixing 40–45 μl of protein with 105–110 μl of 5 mM sodium acetate buffer, pH 6, and subsequently either with 150 μl of 0.1 M acetate buffer, pH 5, or with the same buffer containing urea. All 1H-15N-correlation spectra were recorded using 20–25 mM protein (in monomer) in 50 mM acetate buffer at pH 5 in 95% H2O/5% D2O and a sample volume of ~321 μl in a 5-mm Shigemi tube (Shigemi, Inc., Allison Park, PA). Putative dimerization inhibitors (California Peptide Research) were either mixed with 5 mM acetate buffer (pH adjusted to 6) prior to the protein folding step or mixed with 0.1 M acetate buffer (pH adjusted to 5) when the pH is set (Fig. 3, Quench protocol). Backbone chemical shifts of Ca, and N were determined using HNCA experiments (33) for 0.2–0.3 mM proteins in the same buffer condition as above. Spectra were acquired on DMX500 spectrometers with a CryoProbe (Bruker Instruments, Billerica, MA) at 20 °C. NMR data were processed and analyzed using the nmrPipe, nmrDraw, and PIPP software (34, 35).

RESULTS AND DISCUSSION
Rationale and Experimental Design—The wild-type mature protease catalyzes its own cleavage rapidly (termed autoproteolysis), leading to severe loss in catalytic activity even at low mM concentrations. Thus, solution NMR studies of the uninhibited dimeric protease are feasible for a limited period of time only by using an optimized PR construct, to limit autoproteolysis, and by conducting experiments at pH 5.8, which is above the optimal pH for catalytic activity. The optimized PR construct for expression bears the mutations Q7K, L33I, L63I, C67A, and C95A, the first 3 to limit autoproteolysis and the latter 2 to avoid cysteine-thiol oxidation, leading to protein aggregation (11, 25, 26). Because these mutations do not discernibly affect either the Kd or the kinetics of PR as compared with the wild-type protease, PR has been used as a pseudo wild-type for successive mutational studies and for comparison of the fold, stability, and catalytic properties of the mature protease and its precursor forms (11, 26, 27). An active site D25N mutation, to abolish catalytic activity, was introduced into PR to permit long term studies of either the free or the substrate-bound protease dimer in solution (28). This mutation, however, increases the Kd by >50-fold as compared with PR, which is ~1000-fold less than other interface and non-interface mutations (Table 1).

Studies of protease monomers at 0.2–0.3 mM concentration require suppressing both the autoproteolysis in mutants, where there is significant dimer formation in NMR experimental conditions, as well as the propensity of the folded monomers to aggregate. In addition, previous monomer studies required a starting protein amount of ~2 mg, two dialysis steps, and concentrating the protein to ~0.3 mM; together these steps consume about 6 h of total preparation time (Fig. 3, Dialysis protocol).

To minimize both aggregation and autoproteolysis, we have performed NMR studies of mutant proteases at ~20 μM in monomer concentration using high sensitivity cryogenic probes. We have also devised a simpler folding scheme that is accomplished within 5 min (quench protocol) with reproducible monomer folding efficiency of >90% and negligible signal loss, even 3 months after sample preparation. Thus, 1 mg of protein is sufficient for at least seven NMR samples at a final monomer concentration of ~20 μM. Specific extension, substitution, and deletion mutations were created in PR (Table 1) to access folding and stability of these mutants monitored both by NMR and by fluorescence measurements. Autoproteolysis was almost non-existent in all of the mutants studied because of the high Kd (>2 orders of magnitude as compared with PR) except for PR5–99, which exhibits minor amounts of degradation products even at 20 μM concentration. The autoproteolysis observed for PR5–99 is consistent with our earlier studies showing that it exhibits significant dimer formation and catalytic activity assayed using the chromogenic substrate (11). Additionally, a D25N mutation was introduced in some constructs to totally abolish catalytic activity to allow acquiring data for extended periods of time. Protein folding, stabilities, and Kd values were accessed and compared for three substitution, six deletion, and one extension mutants (Table 1, constructs marked with an asterisk).

PR Lacking the Terminal Residues 1–4 and 96–99 Exhibits a Stable Monomer Fold—In previous studies, we had shown that a 4-residue deletion at either the N terminus (PR5–99) or the C terminus (PR1–95) of the protease shifts the monomer-dimer equilibrium toward the monomer. In 1H-15N HSQC spectra recorded using 0.2–0.3 mM protein, PR5–99 exhibited signals characteristic of a mixture of monomers and dimers, whereas PR1–95 was almost exclusively monomer (17). However, previous assessment of the mutant bearing 4-residue deletions at both termini (PR5–95) of the mature PR was deemed unreliable due to a large decrease in the observed signal intensities of the HSQC spectrum immediately after protein preparation using the dialysis protocol (18). Now we have re-examined PR5–95 that was prepared using the quench protocol at a final concentration of ~20 μM, thus avoiding the step for concentrating the protein after protein folding as in the case of the dialysis protocol. Comparison of the 1H-15N HSQC spectra of PR5–99 and PR5–95 (Fig. 4, A and B) indicates that PR5–99 exhibits a monomer fold similar to that of PR5–99 with negligible aggregation. These two spectra are nearly identical to that of other monomer constructs described previously (17, 18). Although peaks that correspond to the dimer were not observed in the PR5–99 spectrum at ~20 μM, the minor signals in the center of the spectrum of PR5–99 (Fig. 4A, 8.0–8.5 ppm for 1H and 108–128 ppm for 15N) most likely arise from unstructured fragments of the protease due to autoproteolysis and a very small fraction of
unfolded protein. Subjecting PR5–99 to SDS-PAGE analysis after acquiring the HSQC spectrum reveals fragments of the protease resulting from autoproteolysis. Based on the observations that no catalytic activity was detectable for PR5–95 at 1–2 μM protein using a sensitive chromogenic substrate (29) and no fragments resulting from autoproteolysis were detectable by SDS-PAGE after acquiring NMR spectra at 20 μM concentration, very minor signals observed in the PR5–95 spectrum corresponding to the narrow region of 8–8.5 ppm for 1H and 108–128 ppm for 15N most likely represent the unfolded fraction of the protein (<5%). Although an increase in the Kd of some mutants (e.g. PR87K, PRT26A; Table 1) is offset by enhancing the dimer interface contacts through the addition of an inhibitor, the Kd for PR5–95 is expected to be the same or higher than PR1–95, which does not significantly dimerize at concentrations up to 1 mM even in the presence of the potent inhibitor DMP323 (21, 30).

In contrast to PR5–99 and PR5–95, the 1H–15N HSQC spectrum recorded using PR10–90 (Fig. 4C) exhibited weak signals that are characteristic of an unstructured protein. The loss in signal intensities suggests that the majority of the protein undergoes aggregation even at a low protein concentration of 20 μM. Analysis of two other constructs, PR1–90 and PR5–90, showed significant portions of both proteins existing both in aggregated and in folded monomer forms. These results indicate that the region spanning 5–95 of PR is sufficient to maintain a stable fold in 50 mM sodium acetate buffer, pH 5, and that further deletion of residues spanning the regions 5–9 and/or 91–95 significantly reduces the stability of the monomer fold. The observation that nearly all of the PR5–95 exhibits a monomer fold at these low concentrations, unlike at higher concentrations (>0.2 mM), similar to the mutants PR1–95 or PRT26A, suggests that the aggregation of PR5–95 at a higher concentration is most likely caused by the instability of the fold rather than by the misfolded protein that induces aggregation. Apparently, although the regions corresponding to residues 1–9 and 91–99 of the monomer are unstructured, they contribute to maintaining the monomer fold at a higher concentration, supported by the observation that a major fraction of the T26A monomer (PRT26A) and other monomers exhibit a monomer fold at ~0.2 mM concentration, whereas at the same concentration, a majority of the PR5–95 aggregates.

Comparison of Urea-induced Unfolding of Protease Monomers and Dimers—Our recent mutational and NMR studies showing a stable fold for the protease monomer are consistent with a dimer dissociation model in which dimer dissociation and PR unfolding can occur independently. PR dimer contains 4 tryptophanyl residues, Trp-6 and 6 located near the dimer interface and Trp-42 and 42 located externally at the base of the flap region (Fig. 2). Urea-induced unfolding as monitored by associated protein fluorescence change was used to determine the stability of the monomer fold. Folded monomers were prepared at a final concentration of ~0.05 μM by the quench protocol of protein folding using the same protein stock solutions, maintained in 25–50 mM formic acid, that were used to make the NMR samples.

The addition of urea to a final concentration of 5M to the folded monomer, which lacks a defined structure of the terminal residues 1–9 and 91–99 (21), causes a 33–37% quenching of the fluorescence emission as compared with PR or PRT26A dimeric proteases, which show approximately a 50% decrease (13, 15). Plots of the intrinsic fluorescence monitored as a function of increasing urea concentration are shown in Fig. 5. The observed midpoints (U50) for the transition from the folded to the unfolded state of the monomers PR5–95, PRT26A, PRT26A and the dimer PRT25N are 1.98 ± 0.14,
In earlier studies, we showed that the D25N mutation increases the $K_d$ by $\sim$2 orders of magnitude as compared with PR (28) (Table 1). At $\sim$20 $\mu$M monomer concentration, achieved using the same quench protocol of protein folding as used for the monomers above, the majority of PR$_{D25N}$ is dimeric with $\sim$10% of the signal intensity corresponding to the folded monomer (28). The average intensities of four signals of residues Gly-16, Gly-52, Ala-67, and Gly-68 in both PRD25N dimer and PRD25N/T26A monomer are plotted in Fig. 7 as a function of increasing urea concentration (11) have shown that denaturation curves obtained by monitoring enzymatic activity as a function of urea concentration superimpose well onto those obtained by monitoring protein fluorescence. The similarity in the denaturation profiles of the dimer PR$_{D25N}$ and monomer PR$_{S-95}$ suggests that the region spanning 5–95 is sufficient to maintain a native-like stable fold and that the core unfolding of the monomer is not drastically influenced by the terminal residues.

Comparison of the profiles also indicate that the midpoint for the transition from a folded to an unfolded state of the PR$_{T26A}$ monomer (Fig. 5B, $U_{50} = 2.51$ M) is distinctly higher than all other mutants examined ($U_{50}$ range = 1.5–2 M). This increased stability of PR$_{T26A}$ mutant is also clearly evident when comparing the $^1$H–$^{15}$N HSQC spectra at 2 M urea. Signals corresponding to a random-coil region observed at $^1$H 8–8.5 ppm are more significant in the spectra of PR$_{D25N}$ and PR$_{S-95}$ (Fig. 6, D and F), whereas folded monomer signals are still strongly observed in the spectrum for PR$_{T26A}$ (Fig. 6E).

In contrast to PR$_{T26A}$, SFNFPR$_{D25N}$ exhibits a slightly lower stability against urea as compared with PR$_{S-95}$ and PR$_{D25N}$. If the N-terminal residues (SFNF plus 1–9 of PR) of SFNFPR$_{D25N}$ were entirely flexibly similar to that of PR$_{1-95}$, the urea denaturation profile would be similar to that of PR$_{S-95}$ and PR$_{D25N}$. This difference in the urea-induced unfolding profile suggests that the SFNF may influence the monomer stability. The denaturation curves exhibit different shapes, with SFNFPR$_{D25N}$ showing the sharpest transition and PR$_{T26A}$ and PR$_{D25N/T26A}$ showing very broad transitions, possibly indicative of the presence of multiple intermediate species. Additional experiments will be required to characterize these processes in more detail.

**Urea-induced Unfolding of PR$_{D25N}$ Dimer**—Although fluorescence measurements permitted the urea-induced unfolding profiles of the protease monomers to be compared with that of PR$_{D25N}$ dimer, further NMR measurements were required to determine the monomer and dimer populations throughout the transition from the folded to the unfolded state. These populations were obtained from $^1$H–$^{15}$N HSQC spectra of PR$_{D25N}$ recorded at various urea concentrations.

In earlier studies, we showed that the D25N mutation increases the $K_d$ by $\sim$2 orders of magnitude as compared with PR (28) (Table 1). At $\sim$20 $\mu$M monomer concentration, achieved using the same quench protocol of protein folding as used for the monomers above, the majority of PR$_{D25N}$ is dimeric with $\sim$10% of the signal intensity corresponding to the folded monomer (28). The average intensities of four signals of residues Gly-16, Gly-52, Ala-67, and Gly-68 in both PR$_{D25N}$ dimer and PR$_{D25N}$ monomer are plotted in Fig. 7 as a function of increasing urea concentration. The initial ratio of the average dimer signal to the average monomer signal is given by the volume (or area) ratio of the peaks (Fig. 7).

The midpoint of the denaturation curve for PR$_{D25N}$ dimer as measured by NMR (Fig. 7, upper curve) is $\sim$1.8 M urea, in excellent agreement with the fluorescence results (Fig. 5A). The denaturation curves for the monomer and dimer (Fig. 7) are very similar in shape, suggesting that, for this construct, the dissociation of the dimer is not significantly more sensitive to urea than the denaturation of the protein fold. Thus, no
increase in the relative amount of folded monomer (lower curve) at the low urea concentration range is observed, as predicted if the dimer were to dissociate to a stably folded monomer at these urea concentrations. The decrease in signal observed for the dimer must correspond to unfolding (with or without concomitant dissociation) rather than to conversion to folded monomer since this species does not increase at low urea concentrations. The results are consistent either with unfolding and dissociation having approximately equal sensitivities to urea or with unfolding of the monomer being more urea-sensitive than dissociation.

Conformational Characteristics of the Monomers Accessed from Chemical Shifts—To relate the structural features to differences observed in the denaturation profile among the monomer proteases, we assigned and compared the backbone chemical shifts of SFNFPD25N, PRD25N, and PRT26A monomers. For these experiments, samples at a concentration of 0.3 mM were prepared using the dialysis protocol. The results, as expected, showed changes in chemical shifts around the mutation sites. Comparison of Cα and 15N chemical shifts of PRD25N/PRT26A with that of PRT26A indicates significant differences for amides around residue 25 (Fig. 8A). In addition to
these changes, residues 10–12 and 84–95 also exhibit small differences in C/H and 15N chemical shifts between PRD25N/T26A and PRT26A. The comparison of PRD25N/T26A and SFNFPRD25N chemical shifts also indicates small changes in shifts at residues 84–95. The differences in shifts in this region could be either due to the effect of small differences in pH or due to conformational change. Overall, the difference in chemical shifts between PRD25N/T26A and PRD26T is relatively small. The higher stability against urea of PRT26A as compared with PRD25N/T26A cannot be clearly explained by structural differences as manifested by differences in backbone chemical shifts. Thus, the significantly higher stability of PRD26A may be related to alteration in side chain interactions rather than in backbone changes. Based on the proximity of some hydrophobic residues, e.g. Leu-24 and Leu-90, to T26A, we postulate that the methyl group of the substituted Ala-26 enhances hydrophobic contacts. The altered packing around Leu-24 may in turn affect the side chain orientation of residue 11, which is consistent with the observed backbone chemical shift change of this residue. The increase in stability attained through the T26A mutation is offset by the D25N mutation.

In earlier studies (11, 21), we had shown that even two to four residues flanking the N terminus of the protease impairs dimer formation and that the hydrolytic cleavage at the N terminus of the protease is crucial for formation of a native-like stable dimer with catalytic activity and a $K_d$ in the low nM range. However, the molecular mechanism by which the flanking region sequences influence the $K_d$ is not understood. For this reason, the C/H and 15N chemical shifts of SFNFPRD25N and PRD25N/T26A are compared to determine the differences in the secondary structure. Overall, the difference in chemical shifts in the secondary structure between PRD25N/T26A and PRD25N is not significant.

FIGURE 7. Urea denaturation profile of the PRD25N dimer (open circles) and PRD25N monomer (closed circles) monitored by acquiring 1H-15N HSQC spectra. Signals of 4 residues, Gly-16, Gly-52, Ala-67, and Gly-68, were averaged at various concentrations of urea to plot the curve.

FIGURE 8. Difference in 13C/H and 15N chemical shifts (B) between SFNFPRD25N and PRD25N/T26A (in red) and between PRD26A and PRD25N/T26A (in black). Differences in chemical shifts between PRD25N/T26A and PRD25N/T26A are observed in the active site region and in the helical region. Differences in the chemical shifts between SFNFPRD25N and PRD25N/T26A are observed for the N-terminal protease residues 3–6 in addition to the active site region. The secondary structure as determined for PRD25N is indicated on the top. The arrows and cylinder indicate 3-sheet and a-helical structures, respectively. A direct comparison of the SFNFPRD25N and PRD25N constructs is not feasible because under the conditions of the NMR experiments, SFNFPRD25N is a monomer, whereas PRD25N is a dimer. For this reason, a T26A mutation was introduced in PRD25N to increase the $K_d$ to enable comparison of the SFNFPRD25N monomer with the mature protease monomer having the same D25N mutation. PRD26A monomer serves as a control for PRD25N/T26A monomer.

Mutational and Structural Studies of HIV-1 Protease
Mutational and Structural Studies of HIV-1 Protease

understanding of the effect of the SFNF sequence on $K_d$ and stability of the fold.

Qualitative Assessment of Putative Dimerization Inhibitors on the Mature Protease Dimer and Monomer by NMR—The terminal $\beta$-sheet interface has been the target for the development of dimerization inhibitors of the protease. Thus, the new quench protocol of protein folding developed for NMR studies also facilitated the evaluation of the interaction of some previously described dimerization inhibitors of the protease. Our ability to acquire $^1$H-15N HSQC spectra at the low concentration of $\sim 20 \mu M$ also permits achieving higher inhibitor-to-protein ratios, particularly with inhibitors exhibiting poor solubility. Putative dimerization inhibitors, peptide 24 (inhibitor-to-protein ratio = 15:1), reported $K_d$ using Ac-TLNF-NH$_2$ (inhibitor-to-protein ratio = 130:1), produced a change in the spectrum of the dimeric PR$_{126A}$. Furthermore, neither peptide 24 (inhibitor-to-protein ratio = 15:1), Ac-TLNF-NH$_2$ tetrapeptide (inhibitor-to-protein ratio = 65:1), nor peptide 52 (inhibitor-to-protein ratio = 30:1), produced a change in the spectrum of the PR$_{126A}$ monomer. These observations indicate that these peptides do not interact with either the dimer or the monomer under the conditions studied.

Concluding Remarks—Here we describe a simple and fast folding scheme (quench protocol) together with a cryogenic probe that enables $^1$H-15N HSQC spectra of the protease to be acquired at $\sim 20 \mu M$ concentration, $\sim 10$ times lower than the concentration used previously. This development minimizes protein aggregation to the extent that these monomer samples can be stored for more than 3 months without significant loss in signal intensity. In addition, this methodology allows NMR experiments to be directly compared with intrinsic fluorescence resonance energy transfer spectroscopy.

Acknowledgments—We thank J. M. Sayer, E. Matei, and A. M. Gronenborn for helpful discussions, F. Delaglio and D. Garrett for data processing software, and S. M. Lynch and P. Houle for expert technical assistance.

REFERENCES
1. Oroszlan, S., and Luftig, R. B. (1990) Curr. Top. Microbiol. Immunol. 157, 153–185
2. Louis, J. M., Weber, I. T., Tozser, J., Clore, G. M., and Gronenborn, A. M. (2000) Adv. Pharmacol. 49, 111–146
3. Erickson, J. W., and Burt, S. K. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 7970–7974
4. Wondrak, E. M., Nashed, N. T., Haber, M. T., Kimmel, A. R., and Louis, J. M. (1999) J. Biol. Chem. 274, 4477–4481
5. Louis, J. M., Wondrak, E. M., Kimmel, A. R., Wingfield, P. T., and Nashed, N. T. (1999) J. Biol. Chem. 274, 23437–23442
6. Louis, J. M., Clore, G. M., and Gronenborn, A. M. (1999) Nat. Struct. Biol. 6, 868–875
7. Louis, J. M., Dyda, F., Nashed, N. T., Kimmel, A. R., and Davies, D. R. (1998) Biochemistry 37, 2105–2110
8. Louis, J. M., Dyda, F., Kimmel, A. R., and Davies, D. R. (1998) Biochemistry 37, 2105–2110
9. Wondrak, E. M., and Louis, J. M. (1996) Biochemistry 35, 12957–12962
10. Zhang, Z. Y., Poorman, R. A., Maggiora, L. L., Heinrikson, R. L., and Kezdy,
F. J. (1991) J. Biol. Chem. 266, 15591–15594
15. Grant, S. K., Deckman, I. C., Culp, J. S., Minnich, M. D., Brooks, I. S., Hensley, P., Debouck, C., and Meek, T. D. (1992) Biochemistry 31, 9491–9501
16. Todd, M. J., Semo, N., and Freire, E. (1998) J. Mol. Biol. 283, 475–488
17. Ishima, R., Ghirlando, R., Tozser, J., Gronenborn, A. M., Torchia, D. A., and Louis, J. M. (2001) J. Biol. Chem. 276, 49110–49116
18. Louis, J. M., Ishima, R., Nesheiwat, I., Pannell, L. K., Lynch, S. M., Torchia, D. A., and Gronenborn, A. M. (2003) J. Biol. Chem. 278, 6085–6092
19. Strisovsky, K., Tessmer, U., Langner, J., Konvalinka, J., and Krausslich, H. G. (2000) Protein Sci. 9, 1631–1641
20. Weber, I. T. (1990) J. Biol. Chem. 265, 10492–10496
21. Ishima, R., Torchia, D. A., Lynch, S. M., Gronenborn, A. M., and Louis, J. M. (2003) J. Biol. Chem. 278, 43311–43319
22. Cherry, E., Liang, C., Rong, L., Quan, Y., Inouye, P., Li, X., Morin, N., Kotler, M., and Wainberg, M. A. (1998) J. Mol. Biol. 284, 43–56
23. Sluis-Cremer, N., Arion, D., Abram, M. E., and Parniak, M. A. (2004) Int. J. Biochem. Cell Biol. 36, 1836–1847
24. Sluis-Cremer, N., and Tachedjian, G. (2002) Eur. J. Biochem. 269, 5103–5111
25. Rose, J. R., Salto, R., and Craik, C. S. (1993) J. Biol. Chem. 268, 11939–11945
26. Mildner, A. M., Rothrock, D. J., Leone, J. W., Bannow, C. A., Lull, J. M., Reardon, I. M., Sarcich, J. L., Howe, W. J., Tomich, C. S., and Smith, C. W. (1994) Biochemistry 33, 9405–9413
27. Szeltner, Z., and Polgar, L. (1996) J. Biol. Chem. 271, 5458–5463
28. Katoh, E., Louis, J. M., Yamazaki, T., Gronenborn, A. M., Torchia, D. A., and Ishima, R. (2003) Protein Sci. 12, 1376–1385
29. Richards, A. D., Phylip, L. H., Farmerie, W. G., Scarborough, P. E., Alvarez, A., Dunn, B. M., Hiel, P. H., Konvalinka, J., Strop, P., and Pavlickova, L. (1990) J. Biol. Chem. 265, 7733–7736
30. Lam, P. Y., Ru, Y., Jadhav, P. K., Aldrich, P. E., DeLuca, G. V., Eyermann, C. J., Chang, C. H., Emmett, G., Holler, E. R., Daneker, W. F., Li, L., Con- falone, P. N., McHugh, R. J., Han, Q., Li, R., Markwalder, J. A., Seitz, S. P., Sharpe, T. R., Bacher, L. T., Rayner, M. M., Klabe, R. M., Shum, L., Winslow, D. L., Kornhauser, D. M., Jackson, D. A., Erickson-Viitaneu, S., and Hodge, C. N. (1996) J. Med. Chem. 39, 3514–3525
31. Bowman, M. J., and Chmielewski, J. (2002) Biopolymers 66, 126–133
32. Park, S. H., and Raines, R. T. (2000) Nat. Biotechnol. 18, 847–851
33. Kay, L. E., Ikura, M., Tschudin, R., and Bax, A. (2006) J. Magn. Reson. 89, 496–514
34. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (2001) J. Magn. Reson. 95, 214–220
35. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
36. 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

Mutational and Structural Studies of HIV-1 Protease

JUNE 8, 2007 • VOLUME 282 • NUMBER 23
JOURNAL OF BIOLOGICAL CHEMISTRY 17199