ELSEVIER

Biophysical Chemistry 96 (2002) 285–291

Biophysical Chemistry

www.elsevier.com/locate/bpc

Structural properties of the histidine-containing loop in HIV-1 RNase H

Gunther Kern[a,*], Jeff Pelton[b], Susan Marqusee[a], Dorothee Kern[b,1]

[a]*Department of Molecular and Cell Biology, 229 Stanley Hall, University of California, Berkeley, CA 94720, USA*
[b]*Department of Chemistry, Calvin Lab, University of California, Berkeley, CA 94720, USA*

Received 2 January 2001; received in revised form 31 October 2001; accepted 2 November 2001

Abstract

The isolated HIV-1 RNase H domain is inactive. This inactivity has been linked to the lack of structure in the C-terminus of the isolated domain. Thermodynamic stability experiments on the RNase H domain as well as a deletion mutant lacking the C-terminal helix have implied that this region is structured. His539 residing in a loop preceding the C-terminal helix was studied by NMR to determine the stability and conformational properties of this region. The stability of the structural environment of His539 matches that of the entire RNase H domain. Furthermore, His539 is locked into a defined tautometric state in the folded protein and its pK_a is shifted compared to a freely accessible His, suggesting that this region is structured. The data support the view that the overall dynamics rather than the lack of structure in a small portion of the protein render activity of the isolated HIV-1 RNase H.

Keywords: Protein folding; HIV-1 RNase H; His; NMR; Reverse transcriptase

1. Introduction

The heterodimeric (66 kDa and 51 kDa subunits) reverse transcriptase (RT) harbors a DNA-polymerase and an RNase H activity [1–3]. The active protein is produced by asymmetric cleavage of a homodimeric (66 kDa subunit) precursor. The RNase H activity resides in the C-terminal portion of the 66 kDa subunit (Y427-L560) of the heterodimeric RT. This activity is essential for viral infectivity [4–6]. RNase H degrades the RNA portion of the DNA/RNA hybrid, formed after the initiation of minus-strand DNA synthesis, produces the plus-strand RNA primer and removes the RNA template during minus-strand synthesis (for a review, see [7]). Surprisingly, the isolated RNase H domain is inactive [8–10], whereas the structural homologous *E. coli* RNase H is active. There are three striking differences between these homologues [10,11]. Firstly, the HIV protein shows increased overall dynamics [10,12]. Secondly, the basic helix/loop is missing in HIV-1 RNase H

*Corresponding author. AstraZeneca Boston, Infection Discovery, 35 Gatehouse Drive, Waltham, MA 02454, USA. Tel.: +1-781-839-4611.

E-mail address: gunther.kern@astrazeneca.com (G. Kern).

[1] Present address: Department of Biochemistry, Brandeis University, Waltham, MA 02454, USA.

0301-4622/02/$ - see front matter
PII: S0301-4622(02)00019-4

(Thr89 to Asp 102 in *E. coli* RNase H). Thirdly, the C-terminal region of HIV-1 RNase H appears to be unstructured compared to the corresponding region in the *E. coli* homologue [9,11]. The different structural properties are deferred from the lack of electron density in the crystal structure and the failure to obtain structural information for this region by NMR. The C-terminal portion of the protein comprises helix E and the preceding His539-containing loop (Residues 538-560). In the heterodimeric RT, this region diffracts very well. The structural properties of this region in the isolated HIV-1 RNase H domain were, therefore, connected to the inactivity of the isolated HIV-1 RNase H domain. His539, which resides in this loop is not essential for activity, but contributes to the catalytic efficiency (k_{cat}) in both RNase H homologues [13–15]. Replacement of His by Asn renders HIV-1 infectivity [6].

In contrast to all structural data, thermodynamic data clearly show a stabilization of HIV-1 RNase H by its C-terminal helix E and the His-containing loop [16,17]. This implies that this region is not a random coil structure.

To address this contradiction, we used 1D and 2D NMR to determine the stability and structural properties of the region close to His539. Our results suggest that the His539-containing loop is folded in the isolated RNase H domain.

2. Material and methods

2.1. NMR spectroscopy

NMR experiments were recorded on a Bruker DRX-500 spectrometer operating at 25 °C. Two dimensional ^{1}H-^{15}N HMQC experiments [18] optimized for observation of His aromatic ^{2}JNH and ^{3}JN couplings were recorded on uniformly ^{15}N-labeled HIV-1 RNase H in 20 mM sodium phosphate buffer pH 7.5 or 50 mM d_4-sodium acetate buffer pH 5.0 and in the same buffers containing 5 M urea. The experiments were optimized by setting the delay, during which ^{15}N and ^{1}H signals become antiphase, to 22 ms as described previously [19]. The ^{1}H transmitter was set to H O and the ^{15}N carrier was set to 201 ppm. A total of 2048×128 complex points were collected in the t_2 and t_1 dimensions with spectral widths of 6944 and 2102 Hz, respectively. States-TPPI was used [20] for quadrature detection. The data were processed using Felix 97.0 (Biosym Tech.). Typically skewed sinebell window functions were applied in t_1 and t_2 before zero-filling and Fourier transformation. Chemical shifts are referenced to H O (4.77 ppm, 25 °C) and external liquid ammonia (^{15}N) [21]. Uncertainties in chemical shifts are 0.02 ppm for ^{1}H and 0.2 ppm for ^{15}N.

2.2. Stability measurements

The amount of folded His539 was determined from the integrals of the unfolded and folded signal for the Hε1 proton of His539. To isolate the signal for this proton from the amide backbone proton region in the 1D NMR spectrum, all backbone amide protons were exchanged against deuterium by unfolding and refolding HIV-1 RNase H using 5 M deuterated urea and deuterated sodium acetate buffer, pH_{read} 5.5. The equilibrium between the folded and unfolded state was determined from the integrals of the His539 signals for the folded and unfolded form. The overall stability of the protein was obtained by following loss of secondary structure upon addition of urea using far-UV circular dichroism at 222 nm. Thermodynamic stability was determined from the denaturant dependence of the equilibrium between the native and unfolded state, as previously described [22].

2.3. pK_a measurements

All exchangeable protons in HIV-1 RNase H were substituted with deuterium. The protein was incubated in 20 mM d_4-sodium acetate buffer pH 3.0 and adjusted to the indicated pH values using a deuterated 1 M sodium phosphate solution, pH 9 or NaOD above pH 9. Protein precipitate, which was observed between pH 6–7, was removed by centrifugation prior to recording NMR spectra. The signal for the Hε1 proton was followed by ^{1}H-NMR using a 0.9 mM sample and 400 scans.

3. Results

3.1. Thermodynamic stability of His539 against denaturant-induced unfolding

HIV-1 RNase H contains a single His539 in the loop connecting strand 5 and helix E. The signal for the Hε1 in the histidine side-chain is well-detectable in the amide region of a 1D NMR spectrum after all amide backbone and side-chain protons have been exchanged against deuterium. This allowed us to determine whether this His is in an unfolded or folded region of HIV-1 RNase H. If this region is unfolded, a single peak for Hε1 of His539 should be observed in the 1D NMR spectra in the absence and presence of increasing amounts of urea. However, if His539 locates in a folded region of the protein, distinct folded and unfolded His-signals should be observable if both conformations are at slow exchange on the NMR time scale ($<$ms). As shown in Fig. 1a, the signal for His539 initially shifts linearly downfield upon the addition of increasing amounts of urea. This shift does not reflect unfolding. It mirrors the change in solvent upon urea addition and is observable throughout the titration experiment. However, at a concentration above 3 M urea, a second peak is observable. This signal increases in volume with increasing amounts of urea while the other peak decreases. At fully denaturing urea concentrations the second peak sharpens, indicating an increase in the mobility of this region due to complete denaturation of the protein. This supports that His539 is located in a folded region of HIV-1 RNase H. Integration of the peak areas for the unfolded and folded signals yields a thermodynamic stability of the region close to His539 of 4.8 kcal/mol ($c_m = 3.6$ M, $m = 1.33$ kcal/molM. This local stability is identical to the global stability of the protein (Fig. 1b solid line).

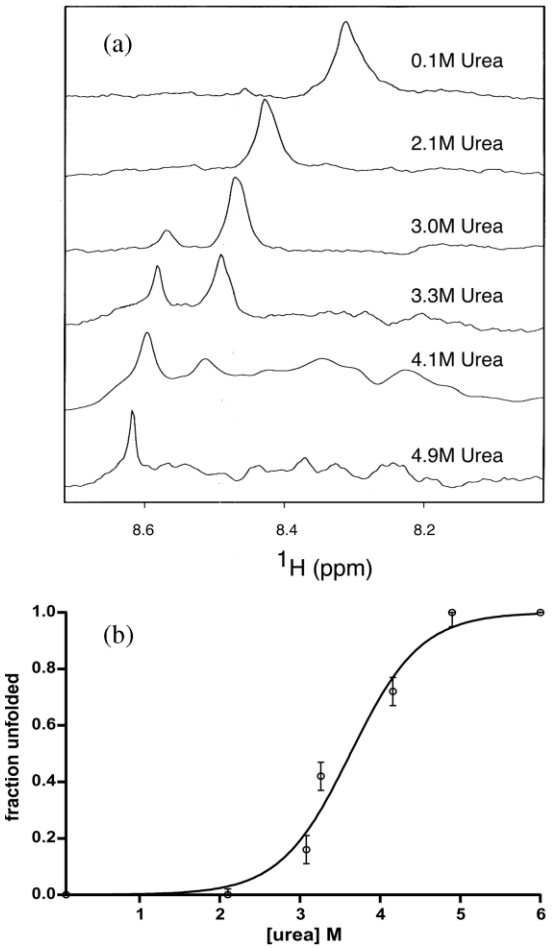

Fig. 1. Thermodynamic stability of His539 in HIV-1 RNase H. (a) ^{1}H NMR spectra for ^{1}Hε signal of His539 in the presence of 0–4.9 M urea in 50 mM d_4 sodium acetate buffer, pH 5.0; (b) Relative integrals of the unfolded ^{1}Hε signal of His539 (circles). The relative error is indicated for each point by a bar. The solid line represents a fit of the global stability of HIV-1 RNase H against urea denaturation followed by far-UV circular dichroism at 222 nm in the same buffer as described for (a).

3.2. pK_a value for His539

A shift of the pK_a value of a histidine side-chain due to protein structure formation can provide for additional structural information of this specific region. While a pK_a-downshift indicates H-bonding or the proximity of His to a base, an increase in the pK_a points to an interaction with an acidic amino acid. The pK_a value for His539 in the folded protein is 6.1. It was determined from the chemical shift dependence of the Hε1 signal on the pH in a range from pH 2.9–10.2 at 25 °C (Fig. 2). The pK_a values for His in a random coil environment were measured for barnase (6.4 and 6.52) [23]. Compared to those values, the pK_a of His539 is decreased. This suggests an

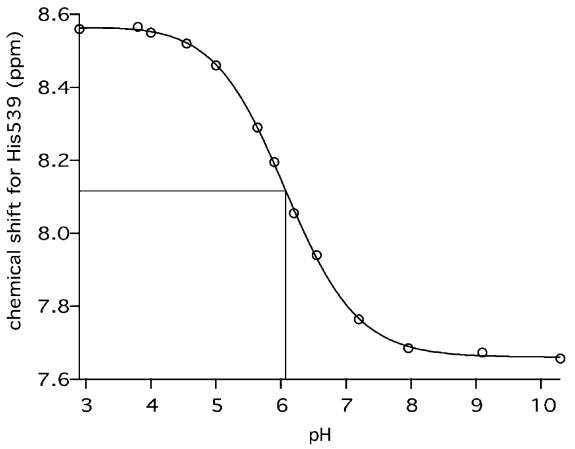

Fig. 2. pK_a determination of His539 in folded HIV-1 RNase H. The chemical shift of the $^1H\varepsilon$ signal of His539 was determined as a function of pH. At the indicated points, a ^{1}H-NMR spectrum was recorded and the chemical shift for this signal determined. A 20 mM d_4-sodium acetate buffer was used which was titrated with either 1 M sodium phosphate, pH 9, or NaOD above pH 9. The error for each data point is smaller than the size of the used symbols.

interaction with a basic amino acid, e.g. formation of a hydrogen bond. In the crystal structure of the reverse transcriptase [24], the closest base is Lys550. Its Nz–atom is 4.66 A distant from the Nε1 of His539. Although this distance would not allow efficient hydrogen bonding, both atoms could be closer in the isolated RNase H domain. Unfortunately, no electron density is observed for this region in the crystal structure for this isolated domain [9]. Interestingly, the pK_a value for the corresponding His124 in the structurally homologous RNase H from *E. coli* is increased to pH 7.1 and was suggested to reflect the proximity of this residue to a glutamic acid [15]. In summary, the decreased pK_a for His539 provides for further evidence that this side-chain is not located in a random coil region of HIV-1 RNase H.

3.3. Detection of the tautomeric state for His539 by 2D NMR

The protonation state [15] and the tautomeric form [25,26] of a His side-chain can be determined from 2D HMQC NMR-experiments using a ^{15}N labeled protein. The chemical shifts and crosspeak patterns for the different tautomeric states of the His side-chain nitrogens have been characterized (for a summary, see Pelton et al. [19]).

In the case of HIV-1 RNase H, the assignment of the His side-chain ^{1}H and ^{15}N resonances is straight-forward, since it contains only one such residue. To obtain information about the protonation state and the tautomeric form of this His sidechain, we recorded HMQC spectra below (pH 5.0) and above (pH 7.8) the pK_a value of His539. For a free histidine, both side-chain nitrogens can be alternatively protonated above the pK_a. Therefore, a histidine located in a random coil part of a protein should show the same properties as a free histidine, where an average crosspeak pattern for both possible tautomeric states should be observed in a 1H-15N HMQC (Fig. 3, right side). However, for His539 of HIV-1 RNase H at pH 7.4, the observed Nε2-Hε2, Nε2-Hδ2 and Nδ?-Hε1 crosspeaks and their intensities (Fig. 3) are in agreement with a tautomeric form where exclusively Nε2 is protonated [19]. The slight downfield ^{15}N chemical shift of Nε2 from the expected value of 167.5 ppm to 172 ppm indicates that this nitrogen serves as a hydrogen bond donor [27,28]. This agrees with the observed decreased pK_a for His539 which also points to a base close to His539.

At pH 5.0, the ^{15}N chemical shifts of the crosspeaks for Nδ1-Hε1, Nδ1-Hδ2, Nε2-Hε1 and Nε2-Hδ2 (Fig. 3) at 172 and 183 ppm show that both side-chain nitrogens are protonated. Obviously the His539 environment within the folded protein is permissive to the protonation of both histidine side-chain nitrogens. However, the number of crosspeaks as well as their chemical shifts clearly indicate impaired rotation of the His sidechain in the folded RNase H domain.

4. Discussion

The reason for the lack of activity of the isolated HIV-1 RNase H domain has not been unambiguously identified. It is not known if the lack of substrate binding or the inability to catalyze confer inactivity. Davies and colleagues [29] associated the lack of activity to a lack of structure in the C-terminal region of RNase H which includes His539.

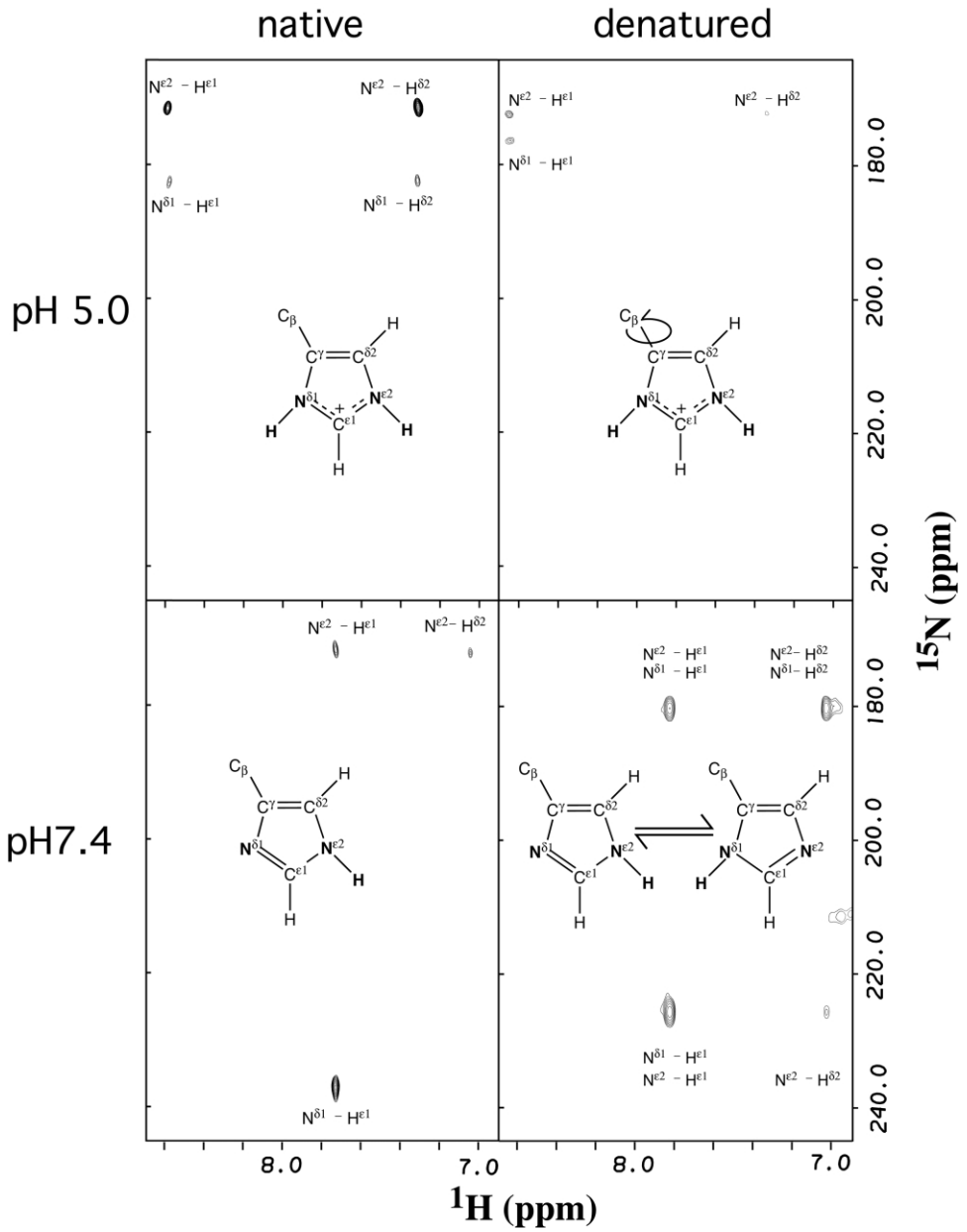

Fig. 3. Tautomeric states of His539 in HIV-1 RNase H. HMQC spectra of ^{15}N labeled HIV-1 RNase H at pH 5.0 (50 mM acetate buffer) and pH 7.4 (20 mM phosphate buffer) in the presence (right) and absence (left) of 5 M urea. The assignment of the respective signals is indicated. At pH 7.4 in the presence of 5 M urea, a double assignment results for three signals due to the exchange between protonation states. The signals in this window at 6.95 ppm (^{1}H dimension) and 212 ppm (^{15}N dimension) result from an impurity.

In contrast to the isolated HIV-1 RNaseH domain, the isolated RNaseH domain of murine leukemia virus (MuLV) retains its activity. However, substrate specificity requires the remainder of the reverse transcriptase protein [30]. Interestingly in HIV-1 RNaseH, insertion of the *E. coli* α_c-helix [31] or addition of an N-terminal His-tag [32] reestablish activity and substrate specificity. It remains to be determined which parts of the protein become stabilized or less dynamic upon this addition.

Deletion of the C-terminal region in *E. coli* RNase H does not completely abolish activity [33]. Therefore, this region is not essential for activity of *E. coli* RNase H.

A characterization of the intramolecular dynamic properties of *E. coli* RNase H [34] and a comparison to the dynamics obtained for HIV-1 RNase H [10] further support this assumption. Although the C-terminal part of HIV-1 RNaseH has much lower order parameters, this difference to the *E. coli* enzyme is also obvious for several other regions in this protein. Therefore, the loss of activity was correlated to the overall increased dynamic property and instability of the isolated HIV-1 domain.

Our findings reveal that the C-terminal region is not unfolded but rather populates a distinct conformation. Our previous studies showed that the C-terminal region contributes approximately 1.1 kcal/mol to the overall stability of the isolated domain [17]. At this stability, the folded form represents 85% of the population and can be detected by the methods we employed in this work. Our stability measurements for His539, however, show that this region has the same stability as the overall protein, suggesting that the region in proximity to this residue is not included in the less stable part of the C-terminal region. His539 is an active site residue and it was therefore suggested that lack of conformation or orientation is responsible for inactivity. Our data support the model where inactivity is not solely due to structural deficiencies of the C-terminal portion but where an overall increased rigidity is required to preserve activity of the isolated RNase H domain.

Acknowledgments

This work was supported by a research grant of the Deutsche Forschungsgemeinschaft (DFG) to G.K. and a fellowship of the Studienstiftung des Deutschen Volkes and BASF to D.K.

References

[1] V.F. diMarzo, T.D. Copeland, A.L. DeVico, et al., Characterization of highly immunogenic p66/p51 as the reverse transcriptase of HTLV-III/LAV, Science 231 (4743) (1986) 1289–1291.

[2] M.M. Lightfoote, J.E. Coligan, T.M. Folks, A.S. Fauci, M.A. Martin, S. Venkatesan, Structural characterization of reverse transcriptase and endonuclease polypeptides of the acquired immunodeficiency syndrome retrovirus, J. Virol. 60 (2) (1986) 771–775.

[3] E.M. Wondrak, J. Lower, R. Kurth, J. Gen. Virol. 67 (1986) 2791–2797.

[4] J. Hansen, T. Schulze, W. Mellert, K. Moelling, Identification and characterization of HIV-specific RNase H by monoclonal antibody, EMBO J. 7 (1) (1988) 239–243.

[5] O. Schatz, F. Cromme, L.F. Grunninger, G.S. Le, Point mutations within the C-terminal domain of HIV1 reverse transcriptase specifically repress RNase H function, FEBS Lett. 257 (1989) 311–314.

[6] M. Tisdale, T. Schulze, B.A. Larder, K. Moelling, Mutations within the RNase H domain of human immunodeficiency virus type 1 reverse transcriptase abolish virus infectivity, J. Gen. Virol. 1 (1991) 59–66.

[7] R.A. Katz, A.M. Skalka, The retroviral enzymes, Annu. Rev. Biochem. 63 (1994) 133–173.

[8] J.L. Keck, S. Marqusee, The putative substrate recognition loop of *Escherichia coli* ribonuclease H is not essential for activity, J. Biol. Chem. 271 (33) (1996) 19883–19887.

[9] J.d. Davies, Z. Hostomska, Z. Hostomsky, S.R. Jordan, D.A. Matthews, Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase, Science 252 (5002) (1991) 88–95.

[10] R. Powers, G.M. Clore, S.J. Stahl, P.T. Wingfield, A. Gronenborn, Analysis of the backbone dynamics of the ribonuclease H domain of the human immunodeficiency virus reverse transcriptase using 15N relaxation measurements, Biochemistry 31 (38) (1992) 9150–9157.

[11] R. Powers, G.M. Clore, A. Bax, et al., Secondary structure of the ribonuclease H domain of the human immunodeficiency virus reverse transcriptase in solution using three-dimensional double and triple resonance heteronuclear magnetic resonance spectroscopy, J. Mol. Biol. 221 (4) (1991) 1081–1090.

[12] A.M. Mandel, M. Akke, A.G.R. Palmer, Backbone dynamics of *Escherichia coli* ribonuclease HI: correla-

tions with structure and function in an active enzyme, J. Mol. Biol. 246 (1) (1995) 144–163.

[13] O. Schatz, F.V. Cromme, L.F. Gruninger, L.G. Sf, Point mutations in conserved amino acid residues within the C-terminal domain of HIV-1 reverse transcriptase specifically repress RNase H function, FEBS Lett. 257 (2) (1989) 311–314.

[14] S. Kanaya, A. Kohara, Y. Miura, et al., Identification of amino acid residues involved in an active site of *escherichia coli* ribonuclease H by site directed mutagenesis, J. Biol. Chem. 265 (1990) 4615–4621.

[15] Y. Oda, M. Yoshida, S. Kanaya, Role of histidine 124 in the catalytic function of ribonuclease HI from *Escherichia coli*, J. Biol. Chem. 268 (1) (1993) 88–92.

[16] E.R. Goedken, S. Marqusee, Folding the ribonuclease H domain of Moloney murine leukemia virus reverse transcriptase, Proteins 33 (1998) 135–143.

[17] G. Kern, T. Handel, S. Marqusee, Characterization of a folding intermediate from HIV -1 ribonuclease H, Prot. Sci. 7 (10) (1998) 2164–2174.

[18] A. Bax, M. Ikura, L.E. Kay, et al., Comparison of different modes of two dimensional reverse-correlation NMR for the study of proteins, J. Magn. Res. 86 (1990) 304–318.

[19] J.G. Pelton, D.A. Torchia, N.D. Meadow, S. Roseman, Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques, Prot. Sci. 2 (1993) 543–558.

[20] D. Marion, M. Ikura, R. Tschudin, R. Bax, Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins, Biochem. Biophys. Res. Commun. 85 (1998) 393–399.

[21] D.S. Wishart, C.G. Bigam, J. Yao, et al., ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, J. Biomol. NMR 6 (2) (1995) 135–140.

[22] M.M. Santoro, D.W. Bolen, Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants, Biochemistry 27 (1988) 8063–8068.

[23] H. Sancho, L. Serrano, A.R. Fersht, Histidine-residues at the N-termini and C-termini of alpha-helices- perturbed Pkas and protein stability, Biochemistry 32 (1992) 2253–2258.

[24] D.W. Rodgers, S.J. Gamblin, B.A. Harris, et al., The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1, Proc. Natl. Acad. Sci. USA 92 (4) (1995) 1222–1226.

[25] W.F. Reynolds, I.R. Peat, M.H. Freedman, J.R. Lyerla Jr, Determination of the tautomeric form of the immidazole ring of L-histidine inj basic solution by carbon-13 magnetic resonance spectroscopy, J. Am. Chem. Soc. 95 (1973) 338–341.

[26] W.W. Bachovin, J.D. Roberts, Nitrogen-15 NMR spectroscopy. The state of histidine in the catalytic triade of α-lytic protease: Implications for the charge-relay mechanism of peptide-bond cleavage by serine proteases, J. Am. Chem. Soc. 100 (1978) 8041–8047.

[27] I.I. Schuster, J.D. Roberts, Nitrogen-15 NMR spectroscopy. Effects of hydrogen bonding and protonation on nitrogen chemical shifts in imidazoles, J. Org. Chem. 44 (1979) 3864–3867.

[28] J.D. Roberts, C. Yu, C. Flanagan, T.R. Birdseye, A nityrogen-15 NMR study of acid-base and tautomeric equilibria of 4-substituted imidazoles and its relevance to the catalytic mechanism of α-lytic protease, J. Am. Chem. Soc. 104 (1982) 3945–3949.

[29] J.F.D. Davies, Z. Hostomska, Z. Hostomsky, S.R. Jordan, D.A. Matthews, Crystal structure of the ribonuclease H domain of HIV-1 reverse transcriptase [see comments], Science 252 (1991) 88–95.

[30] X. Zhan, R.J. Crouch, The isolated RNase H domain of murine leukemia virus reverse transcriptase. Retention of activity with concomittant loss of specificity, J. Biol. Chem. 272 (1997) 22023–22029.

[31] J.L. Keck, S. Marqusee, Substitution of a highly basic helix/loop sequence into the RNase H domain of human immunodeficiency virus reverse transcriptase restores its Mn(2+)-dependent RNase H activity, Proc. Natl. Acad. Sci. USA 92 (7) (1995) 2740–2744.

[32] S.J. Stahl, J.D. Kaufman, T.S. Vikic, R.J. Crouch, P.T. Wingfield, Construction of an enzymatically active ribonuclease H domain of human immunodeficiency virus type 1 reverse transcriptase, Prot. Eng. 7 (9) (1994) 1103–1108.

[33] E.R. Goedken, T.M. Raschke, S. Marqusee, Importance of the C-terminal helix to the stability and enzymatic activity of *Escherichia coli* ribonuclease H, Biochemistry 36 (23) (1997) 7256–7263.

[34] A.M. Mandel, M. Akke, A.R. Palmer, Backbone dynamics of *Escherichia coli* ribonuclease HI: correlations with structure and function in an active enzyme, J. Mol. Biol. 246 (1) (1995) 144–163.