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

Robert Powers
University of Nebraska - Lincoln, rpowers3@unl.edu

G. Marius Clore
Laboratory of Chemical Physics, Building 2 National Institute of Diabetes and Digestive and Kidney Diseases National Institutes of Health, Bethesda, MD

Ad Bax
Laboratory of Chemical Physics, Building 2 National Institute of Diabetes and Digestive and Kidney Diseases National Institutes of Health, Bethesda, MD

Daniel S. Garrett
Laboratory of Chemical Physics, Building 2 National Institute of Diabetes and Digestive and Kidney Diseases National Institutes of Health, Bethesda, MD


http://digitalcommons.unl.edu/chemistrypowers/13

This Article is brought to you for free and open access by the Published Research - Department of Chemistry at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Robert Powers Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
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

Robert Powers¹, G. Marius Clore¹,†, Ad Bax¹, Daniel S. Garrett¹, Stephen J. Stahl²
Paul T. Wingfield² and Angela M. Gronenborn¹†

¹Laboratory of Chemical Physics, Building 2
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health, Bethesda, MD 20892, U.S.A.

²Protein Expression Laboratory, Building 6B
Office of the Director, National Institutes of Health
Bethesda, MD 20892, U.S.A.

(Received 13 June 1991; accepted 11 July 1991)

The solution structure of the ribonuclease H domain of HIV-1 reverse transcriptase has been investigated by three-dimensional double and triple resonance heteronuclear magnetic resonance spectroscopy. The domain studied has 138 residues and comprises residues 427 to 560 of the 66 kDa reverse transcriptase with an additional four residues at the N terminus. Initial studies on the wild-type protein were hindered by severe differential line broadening, presumably due to conformational averaging. Mutation of the single tryptophan residue located in a loop at position 113 (position 535 in the reverse transcriptase sequence) to an alanine resulted in much improved spectral properties with no apparent change in structure.

¹H, ¹⁵N and ¹³C backbone resonances were assigned sequentially using a range of three-dimensional double and triple resonance heteronuclear experiments on samples of uniformly (>95%) ¹⁵N and ¹³N/¹³C-labeled protein, and the secondary structure was elucidated from a qualitative analysis of data derived from three-dimensional ¹⁵N- and ¹³C-edited nuclear Overhauser enhancement spectra. The secondary structure comprises three α-helices and five strands arranged in a mixed parallel/antiparallel β-sheet with α1, α2, α3, α4, α5, 1-3x, 3x, -1x topology. The C-terminal region from residue 114 onwards appears to be conformationally disordered in solution as evidenced by an almost complete absence of sequential and medium range nuclear Overhauser effects.

Key words: HIV-1; RNase H domain; reverse transcriptase; solution secondary structure; 3D heteronuclear n.m.r.; double and triple resonance n.m.r.

The ribonuclease H (RNase H) domain of human immunodeficiency virus (HIV-1) reverse transcriptase plays a crucial role in viral replication as evidenced by the failure of mutant provirus, defective for RNase H function, to produce infectious virus particles (Schatz et al., 1989). The RNase H domain catalyzes the cleavage of the RNA portion of a DNA/RNA hybrid, a process that is

† Authors to whom all correspondence should be addressed.

‡ RNase H, the ribonuclease H domain comprising residues 427 to 560 of the 66 kDa reverse transcriptase of HIV-1 together with the four amino acid sequence Met-Asn-Glu-Leu at the N terminus; HIV-1, human immunodeficiency virus-1; AIDS, acquired immunodeficiency syndrome; n.m.r., nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; 3D, three-dimensional; HNCO, amide proton to nitrogen to carbonyl correlation; HN(CO), amide proton to nitrogen to z-carbon correlation; HN(CO)C, amide proton to nitrogen (via carbonyl) to z-carbon correlation; HCA, alpha proton to z-carbon to carbonyl correlation; HCA(CO)N, z-proton to z-carbon (via carbonyl) to nitrogen correlation; c.d., circular dichroism; p.p.m., parts per million; 2D, two-dimensional.
Figure 1. Selected $F_1-F_3$ planes for the 3D HNCO (a), HNCA (b), HN(CO)CA (c), HCACO (d) and HCA(CO)N (e) triple resonance spectra of a sample of uniformly (>95%) $^{15}$N/$^{13}$C-labeled Trp113→Ala RNase H from HIV-1 at 26°C. All spectra were recorded at 600 MHz on a Bruker AM600 spectrometer equipped with a triple resonance probe and modified with additional hardware, as described by Kay et al., (1990a). The HNCO and HNCA experiments were recorded as described by Kay et al. (1990a), the HN(CO)CA experiment as described by Tkura & Bax (1991), and the HCACO and HCA(CO)N experiments as described by Powers et al. (1991). The latter 2 experiments employ a constant time evolution of the $^{13}$C" magnetization in the $F_3$ dimension, which results in in-phase $^{13}$C" signals. The Trp113→Ala mutant of the RNase H domain was purified as described by Becerra et al. (1990) and the sample contained 1-1 mM-Trp113→Ala RNase H in 100 mM-sodium phosphate (pH 5.4), dissolved in 90% H$_2$O/10% $^2$H$_2$O for the HNCO, HNCA and HN(CO)CA experiments and in 99:99:96% H$_2$O for the HCACO and HCA(CO)N experiments. The spectral widths in the $^{15}$N, $^1$H and $^{13}$CO dimensions were 20-16 p.p.m., 31-13 and 64 p.p.m., respectively, with the carrier positions placed at 1185 p.p.m., 56 p.p.m. and 177 p.p.m., respectively. The spectral width in the $^1$H dimension was 13-44 p.p.m. with the carrier at 476 p.p.m. for the HNCO, HNCA and HN(CO)CA experiments. For the HCACO and HCA(CO)N experiments, the $^1$H spectral width was 8-33 p.p.m. with the carrier at 476 p.p.m. For the HNCO, HNCA and HN(CO)CA experiments the number of points acquired in the various dimensions was 32 complex in $F_1$ ($^{15}$N), 64 complex in $F_2$ ($^{13}$CO or $^{12}$C") and 1024 real in $F_3$ ($^1$H). For the HCACO and HCA(CO)N experiments, the number of points acquired was 32 complex in $F_1$ ($^{15}$N), 64 complex in $F_2$ ($^{13}$CO) for the HCACO experiment and 32 complex in $F_3$ ($^{15}$N) for the HCA(CO)N experiment, and 512 real in $F_3$. Zero-filling was used in all dimensions, and in the case of the HCACO and HCA(CO)N experiment, linear prediction by means of the mirror image technique (Zhu & Bax, 1990) was used to further extend the data. The final 3D spectra consisted of 64 x 128 x 1024 data points for the HNCO. HNCA and HN(CO)CA experiments, and 128 x 128 x 512 points for the HCACO and HCA(CO)N experiments. All the spectra in this Figure, as well as in Fig. 2, were processed on a Sun Sparc Workstation using in-house routines for Fourier transformation (Kay et al., 1989) and linear prediction (Zhu & Bax, 1990), together with the commercially available software package NMR2 (New Methods Research, Inc, Syracuse, NY). Analysis of the 3D spectra and peak picking was carried out using the in-house programs CAPP and PIPP (Garrett et al., 1991).
required at several stages during reverse transcription, and displays both endonuclease and 3′→5′ exonuclease activity (Krug & Berger, 1989; Mizraphy, 1989; Schatz et al., 1990). Thus, the RNase H domain presents a potential site for the design of drugs for the treatment of AIDS (Mitsuya et al., 1990). In a recent paper we described the over-expression, purification and physical characterization of the RNase H domain comprising residues 427 to 500 of the 66 kDa reverse transcriptase with an additional four residue sequence at the N terminus (Becerra et al., 1990). In this paper we present initial multi-dimensional heteronuclear studies on the RNase H domain undertaken with the eventual aim of determining its high-resolution three-dimensional structure in solution. Specifically, the backbone 1H, 15N and 13C resonances are assigned in a sequential manner using a combination of 3D double and triple resonance heteronuclear n.m.r. experiments, and the secondary structure is elucidated from a qualitative analysis of NOE connectivities derived from 3D heteronuclear-edited NOESY spectra (for reviews, see Clore & Gronenborn 1991a,b). As the work presented in this paper was being prepared for publication, a 2.4 Å (1 Å = 0.1 nm) resolution X-ray structure of RNase H from a different HIV-1 virus strain was published (Davies et al., 1991). The results from both studies are essentially in agreement, although significant differences are noted at the C terminus.

Initially n.m.r. studies were carried out on the wild-type RNase H domain from strain HXB2 of HIV-1. However, it rapidly became apparent that the assignment procedure was being impaired by suspected conformational exchange processes. Thus, we were unable to see many connectivities in the various 3D triple resonance experiments due to severe line broadening. This was also manifested in the 3D 15N-edited NOESY (Fesik & Zuiderweg, 1988; Marion et al., 1989a,b) and HOHAHA (Marion et al., 1989b; Driscoll et al., 1990a) spectra, as well as in the 2D 1H-13C correlation spectra (Bodenhausen & Ruben, 1980; Bax et al., 1990a). In the latter spectra, a large variation in cross-peak intensities occurred. At a very early stage of this work, we found that, under all conditions tried, we did not observe the 1H-13C correlations for the aromatic ring of the single Trp residue at position 113, while those for the other eight aromatic rings (6 Tyr, 1 Phe and 1 His) were clearly detectable. We therefore postulated that Trp113 was located in a segment of the polypeptide chain that exhibited conformational flexibility, resulting in severe line broadening of adjacent protons through large differences in ring current shifts between the conformers. On the basis of this hypothesis, we proceeded to construct a Trp113→Ala mutant by primer-directed mutagenesis (Oostra et al., 1983). No difference in stability of the native and mutant proteins was found, both of which had a Tm of ~60°C (determined by differential scanning calorimetry), and the location of many of the cross-peaks in the 1H-15N correlation spectra of the two proteins were identical. Further, there was little difference in the e.d. spectrum of the two proteins. This implied that the mutation caused only minimal structural perturbation, which would be entirely consistent with the predicted location of Trp113 on the protein surface based on sequence alignment with the known crystal structure of Escherichia coli RNase H (Yang et al., 1990; Katayanagi et al., 1990). However, the spectra of the mutant enzyme were both qualitatively and quantitatively vastly superior to those of the wild-type protein, so that further detailed study was restricted to the mutant.

The sequential assignment strategy was based on a series of 3D double and triple resonance n.m.r. experiments. In particular, we made use of five triple resonance experiments to establish connectivities along the chain via one- and two-bond heteronuclear couplings. The 3D HNCO, HNCA (Ikura et al., 1990a) and HN(CO)CA (Ikura & Bax, 1991) experiments recorded in water were used to establish NH(i)-15N(i)-13C(i)-13C(i+1) and NH(i)-15N(i)-13C(i+1)-13C(i+2) correlations, respectively. 13C(F2)-NH(F1) planes of these three experiments at a single 15N(F1) frequency of 119.43 ppm are illustrated in Figure 1(a) to (c). As the NH(i)-15N(i+1) intraresidue one-bond coupling is larger than the interresidue two-bond coupling, it is usually the case that the intraresidue correlations in the HNCA experiment are more intense than the interresidue ones (Clore et al., 1990; Kay et al., 1990b). However, in the case of RNase H where complications arise from exchange line broadening this rule is not generally applicable. Thus, for example, the interresidue correlations for Lys9 and He120 seen in Figure 1(a) are actually more intense than the intraresidue ones. Possible ambiguities are resolved by analysis of the HNCACO spectrum, which only displays the interresidue 13C(i+1)-15N(i) correlations. The 3D HCACO and HCA(15N) experiments (Ikura et al., 1990a; Powers et al., 1991) recorded in 2H2O were used to establish CH(i)-13C(i)-15N(i) and CH(i)-13C(i)-13C(i+1) correlations, respectively. These experiments are illustrated in Figure 1(d) to (e), which show a set of 13C(F2)-CH(F1) and 15N(F2)-CH(F1) planes at the same 13C(F1) frequency of 58.87 ppm. In addition, a 3D 15N-edited HOHAHA spectrum recorded in H2O with a DIPSI-2 mixing scheme (Clore et al., 1991) was used to identify 15N(ii)-NH(i)-CH(i) correlations (Fig. 2(a)). Interpretation of these six scalar correlation experiments is sufficient to sequentially assign the backbone 1H, 15N and 13C resonances. Confirmation of the sequential assignment was based on 3D 15N and 13C-edited (Ikura et al., 1990b; Zuiderweg et al., 1990) NOESY spectroscopy to identify through-space (<5 Å) connectivities of the type NH(i)-NH(i+1,2) or NH(i)-NH(i+1,2,3,4). In the case of NH(i)-NH(i+1), NH(i)-NH(i+2) and NH(i)-NH(i+3), which have been extensively used in conventional protein resonance assignment by 2D methods (Withrich, 1986; Clore & Gronenborn, 1987). Examples of selected amide strips through
Figure 2. Amide strips extending from Gln53 to Asp66 taken from the 600 MHz 3D $^{15}$N-edited HOHAHA (a) and NOESY (b) spectra of uniformly (>95\%) $^{15}$N-labeled Trp113→Ala RNase H from HIV-1 recorded at 26°C. The mixing times for the 2 experiments were 30 ms and 100 ms, respectively, and the experiments were recorded as described by Clore et al. (1991) and Driscoll et al. (1990a), respectively. The Figure is composed of narrow strips taken from different $^1H(F_1)$-$^1H(F_2)$ planes of the 3D spectrum, as described by Driscoll et al. (1990a). Asterisks indicate the position of the diagonal peak for each residue and boxes enclose the intraresidue NH-$^1$CH and NH-$^1$CH cross-peaks observed in the NOESY spectrum. Sequential NH-NH(i+1) NOEs are indicated by thick arrows, sequential $^1$CH(i)-NH(i+1) and $^1$CH(i)-NH(i+1) NOEs with thin arrows. Note that while the NH-$^1$CH cross peaks in the HOHAHA spectrum are observed for all the residues in this Figure (with the exception of Gln56, Tyr61 and Asp66), NH-$^1$CH cross-peaks are only seen for Gln65 and Asp66 and these are of very weak intensity. The spectral widths in the $F_1$ ($^1$H), $F_2$ ($^{15}$N) and $F_3$ ($^1$H) dimensions were 11.41 ppm, 2916 ppm and 13.44 ppm with the carrier placed at 1185 ppm in the $^{15}$N dimension and at 476 ppm in the $^1$H dimensions. The total number of points acquired were 128 complex in $F_1$ ($^1$H), 32 complex in $F_2$ ($^{15}$N) and 1024 real in $F_3$ ($^1$H), and the final absorptive part of the 3D spectra, after appropriate zero-filling, consisted of 256 x 64 x 1024 points.
A summary of the sequential scalar and the sequential and medium range (up to i, i+4) NOE connectivities observed for the Trp113→Ala mutant of RNase H domain is shown in Figure 3, and the backbone "H, 15N and 13C assignments are given in Table 1. We were able to confidently obtain backbone resonance assignments for 123 of the 138 residues for the mutant protein compared to 110 tentative assignments for the wild-type. The number of residues for which a complete set of correlations was observed in the triple resonance experiments and the 15N-edited HOHAHA spectrum was 60 for the mutant compared to only 31 for the wild-type. Further NOEs involving 113 NH protons could be assigned in the 3D 15N-edited NOESY spectrum of the mutant compared to only 86 for the wild-type. It is also interesting to note that with the exception of Met1 and Asp76, almost all residues in the mutant protein for which no backbone assignments could be ascertained are located in the C terminus, whereas in the wild-type unassigned residues are spread throughout the sequence.

The secondary structure was deduced from a qualitative analysis of the NOE data involving the backbone NH, CαH and CγH protons (Wüthrich, 1986; Clore & Gronenborn, 1987) derived from the 3D 15N- and 13C-edited NOESY spectra, in conjunction with data on slowly exchanging amide protons (summarized in Fig. 3). The latter was
### Table 1

Backbone $^1H$, $^{13}C$, and $^{15}N$ resonance assignments for the RNase H domain of HIV-1 reverse transcriptase at pH 5.4 and 20°C

<table>
<thead>
<tr>
<th>Residue</th>
<th>Ala12</th>
<th>Val13</th>
<th>Gly14</th>
<th>Ala15</th>
<th>Gly16</th>
<th>Thr17</th>
<th>Thr18</th>
<th>Thr19</th>
<th>Val20</th>
<th>Asp21</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1H$</td>
<td>120.4</td>
<td>58.1</td>
<td>89.2</td>
<td>95.8</td>
<td>96.9</td>
<td>97.3</td>
<td>100.5</td>
<td>101.6</td>
<td>108.9</td>
<td>112.3</td>
</tr>
<tr>
<td>$^{13}C$</td>
<td>53.6</td>
<td>46.8</td>
<td>51.9</td>
<td>54.3</td>
<td>56.7</td>
<td>57.3</td>
<td>57.6</td>
<td>57.8</td>
<td>58.2</td>
<td>58.9</td>
</tr>
<tr>
<td>$^{15}N$</td>
<td>765</td>
<td>767</td>
<td>763</td>
<td>760</td>
<td>756</td>
<td>759</td>
<td>763</td>
<td>764</td>
<td>760</td>
<td>763</td>
</tr>
</tbody>
</table>

### Table 1 (continued)

Chemical shift (p.p.m.)

<table>
<thead>
<tr>
<th>Residue</th>
<th>Ala12</th>
<th>Val13</th>
<th>Gly14</th>
<th>Ala15</th>
<th>Gly16</th>
<th>Thr17</th>
<th>Thr18</th>
<th>Thr19</th>
<th>Val20</th>
<th>Asp21</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1H$</td>
<td>120.4</td>
<td>58.1</td>
<td>89.2</td>
<td>95.8</td>
<td>96.9</td>
<td>97.3</td>
<td>100.5</td>
<td>101.6</td>
<td>108.9</td>
<td>112.3</td>
</tr>
<tr>
<td>$^{13}C$</td>
<td>53.6</td>
<td>46.8</td>
<td>51.9</td>
<td>54.3</td>
<td>56.7</td>
<td>57.3</td>
<td>57.6</td>
<td>57.8</td>
<td>58.2</td>
<td>58.9</td>
</tr>
<tr>
<td>$^{15}N$</td>
<td>765</td>
<td>767</td>
<td>763</td>
<td>760</td>
<td>756</td>
<td>759</td>
<td>763</td>
<td>764</td>
<td>760</td>
<td>763</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Ala12</th>
<th>Val13</th>
<th>Gly14</th>
<th>Ala15</th>
<th>Gly16</th>
<th>Thr17</th>
<th>Thr18</th>
<th>Thr19</th>
<th>Val20</th>
<th>Asp21</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1H$</td>
<td>120.4</td>
<td>58.1</td>
<td>89.2</td>
<td>95.8</td>
<td>96.9</td>
<td>97.3</td>
<td>100.5</td>
<td>101.6</td>
<td>108.9</td>
<td>112.3</td>
</tr>
<tr>
<td>$^{13}C$</td>
<td>53.6</td>
<td>46.8</td>
<td>51.9</td>
<td>54.3</td>
<td>56.7</td>
<td>57.3</td>
<td>57.6</td>
<td>57.8</td>
<td>58.2</td>
<td>58.9</td>
</tr>
<tr>
<td>$^{15}N$</td>
<td>765</td>
<td>767</td>
<td>763</td>
<td>760</td>
<td>756</td>
<td>759</td>
<td>763</td>
<td>764</td>
<td>760</td>
<td>763</td>
</tr>
</tbody>
</table>

### Table 2

Chemical shift (p.p.m.)

<table>
<thead>
<tr>
<th>Residue</th>
<th>Ala12</th>
<th>Val13</th>
<th>Gly14</th>
<th>Ala15</th>
<th>Gly16</th>
<th>Thr17</th>
<th>Thr18</th>
<th>Thr19</th>
<th>Val20</th>
<th>Asp21</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1H$</td>
<td>120.4</td>
<td>58.1</td>
<td>89.2</td>
<td>95.8</td>
<td>96.9</td>
<td>97.3</td>
<td>100.5</td>
<td>101.6</td>
<td>108.9</td>
<td>112.3</td>
</tr>
<tr>
<td>$^{13}C$</td>
<td>53.6</td>
<td>46.8</td>
<td>51.9</td>
<td>54.3</td>
<td>56.7</td>
<td>57.3</td>
<td>57.6</td>
<td>57.8</td>
<td>58.2</td>
<td>58.9</td>
</tr>
<tr>
<td>$^{15}N$</td>
<td>765</td>
<td>767</td>
<td>763</td>
<td>760</td>
<td>756</td>
<td>759</td>
<td>763</td>
<td>764</td>
<td>760</td>
<td>763</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residue</th>
<th>Ala12</th>
<th>Val13</th>
<th>Gly14</th>
<th>Ala15</th>
<th>Gly16</th>
<th>Thr17</th>
<th>Thr18</th>
<th>Thr19</th>
<th>Val20</th>
<th>Asp21</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^1H$</td>
<td>120.4</td>
<td>58.1</td>
<td>89.2</td>
<td>95.8</td>
<td>96.9</td>
<td>97.3</td>
<td>100.5</td>
<td>101.6</td>
<td>108.9</td>
<td>112.3</td>
</tr>
<tr>
<td>$^{13}C$</td>
<td>53.6</td>
<td>46.8</td>
<td>51.9</td>
<td>54.3</td>
<td>56.7</td>
<td>57.3</td>
<td>57.6</td>
<td>57.8</td>
<td>58.2</td>
<td>58.9</td>
</tr>
<tr>
<td>$^{15}N$</td>
<td>765</td>
<td>767</td>
<td>763</td>
<td>760</td>
<td>756</td>
<td>759</td>
<td>763</td>
<td>764</td>
<td>760</td>
<td>763</td>
</tr>
</tbody>
</table>

† $^1H$ and $^{13}C$ chemical shifts are expressed relative to (trimethylsilyl)-propionate-d$_4$ acid, and $^{15}N$ shifts relative to liquid NH$_3$. 

Communications
Figure 4. β-Sheet structure of Trp13→Ala RNase H from HIV-1 as determined from a qualitative analysis of NOE and amide exchange data. The β-strands are indicated on the left by roman numerals and the residue number range. Long-range NOEs are indicated by arrows, and hydrogen bonds derived from the NH/solvent exchange and NOE data are shown as broken lines. The slowly exchanging amide protons are encircled. The 1H chemical shifts of the CαH protons of Val36 and Glu42 are degenerate and indicated by a square box around them.

obtained by recording a series of 15N-1H Overbodenhausen correlation spectra (Bax et al., 1990a) over a period of 80 hours starting within five minutes of dissolving an unexchanged sample of lyophilized protein in 2H2O (Driscoll et al., 1990b). We find three clearly defined helices, which extend from residues 52 to 66 (αA), 78 to 87 (αB) and 93 to 106 (αC), and five β-strands from residues 16 to 24 (strand I), 30 to 38 (strand II), 41 to 48 (strand III), 69 to 75 (strand IV) and 108 to 113 (strand V). These five strands form a mixed parallel-antiparallel β-sheet, which is illustrated in Figure 4. Strands I, II and III are antiparallel and connected via a short loop and β-turn, respectively. Strand I is parallel to strand IV, which in turn is parallel to strand V. The connecting element between strands III and IV is formed by α-helix αA while that between strands IV and V is formed by the other two helices, αB, and αC. Thus, using the notation of Richardson (1981), the topology of the sheet can be described as +1, +1, -3, -1z. It is interesting to note that the slowest exchanging amide protons are located in strands I, IV and V and helices αA and αC, respectively. Further, while there is an extensive network of NOE connectivities between strands I, II and III and between strands IV and V, the number of NOE connectivities between strands I and IV is very limited and their intensities are weak. The latter may be due to fraying at the ends of these two strands, accompanied by conformational mobility. Taken together, these observations suggest that β-strands IV and V and the connecting α-helices αA and αC constitute the most tightly folded portion of the RNase H structure.

The last well-defined residue in solution is Ala113, and the C-terminal region from Val114 to Gly133 appears to be conformationally disordered as evidenced by an almost complete absence of sequential and medium range NOEs (Fig. 3). In the X-ray structure residues 116 to 120 are also disordered and not visible in the electron density map. However, there is a helix (αD) from Gly121 to Ser131 in the X-ray structure, which is absent in solution. This is probably due to greater conformational flexibility in the solution state and implies that this helix is not required to stabilize the overall polypeptide fold.

It should be noted that carboxy-terminal proteolytic cleavage of the RNase H domain results in a polypeptide approximately 3000 Da smaller than the native form (P.T. Wingfield, A. M. Gronenborn, G. M. Clore, unpublished results). This shortened cleavage product no longer exhibits a folded structure, as judged by both e.d. and n.m.r. spectra, which are characteristic of a random coil conformation. The most likely location of the cleavage site is in the sequence Lys105-Lys106-Glu107-Lys108, located at the beginning of β-strand V. Thus, removal of part or all of β-strand V destroys the folded form of RNase H, which is consistent with the above interpretation.

With the exception of the C-terminal helix αD, which is not observed in solution, the secondary structure found in the liquid and crystal states are in agreement. It should be pointed out that the RNase H domains studied by n.m.r. and crystallography were derived from two different HIV-1 strains, HXB2 and BH10, respectively, and differ by several amino acid substitutions (i.e. Ile46→Pro, Asp49→Asn, Glu90→Lys and Val137→Ile). In addition, the sequence of the protein used for the n.m.r. studies contains the engineered Trp13→Ala mutation and the tetrapeptide Met-Asn-Glu-Leu.
appended at the N terminus, while these four residues in the crystallized protein are Tyr-Ala-Ser-Arg. With the exception of the Ile46→Pro change, which is located at the end of strand 111, the differences are conservative, located in loops or at the disordered N and C termini, and clearly do not affect the structure in any significant manner.

Comparison of the secondary structure of HIV-1 RNase H with that of E. coli (Yang et al., 1990; Katayanagi et al., 1990; Yamazaki et al., 1991) is also of interest. There is an almost perfect match of the five β-strands and helices $z_2$, $z_3$ and $z_4$ between HIV-1 and E. coli RNase H. The $z_1$ helix in the HIV-1 RNase H is about two to three residues longer than that of the corresponding helix in the E. coli enzyme, and the $z_2$ helix and the subsequent 12-residue loop that serve to connect helices $z_2$ and $z_3$ in the E. coli protein are replaced by a short loop comprising residues 88 to 92 in the HIV-1 enzyme. In summary we have made use of 3D double and triple resonance heteronuclear n.m.r. spectroscopy to obtain $^1H$, $^{15}N$ and $^{13}C$ backbone sequential assignments for the RNase H domain of HIV-1 reverse transcriptase. The observed secondary structure in solution is consistent with that found in the crystal structure but there appears to be substantially more disorder at the C terminus in the solution structure. Further, extensive line broadening in the n.m.r. spectra of the wild-type protein is strongly suggestive of extensive conformational heterogeneity, which is in part alleviated by the mutation of Trp113 to Ala.

This work was supported by the AIDS Targeted Anti-viral Program of the Office of the Director of the National Institutes of Health. We thank Joshua Kaufman and Ira Palmer for technical assistance with the protein purification, and Joseph Shiloah for fermentations.

References


