Backbone Dynamics of the N-Terminal Domain of the HIV-1 Capsid Protein and Comparison with the G94D Mutant Conferring Cyclosporin Resistance/Dependence†

Ramón Campos-Olivas and Michael F. Summers*

Howard Hughes Medical Institute and Department of Chemistry and Biochemistry, University of Maryland Baltimore County, Baltimore, Maryland 21250

Received April 30, 1999; Revised Manuscript Received June 10, 1999

ABSTRACT: Nuclear magnetic resonance (NMR) ^{15}N relaxation methods have been used to characterize the backbone dynamics of the N-terminal core domain of the HIV-1 capsid protein (CA^{151}). The domain, which has an unusually flat, triangular shape, tumbles in solution at 28 °C with an effective rotational correlation time of 11.5 ns. Relaxation data for backbone amides in the domain's seven α-helices are indicative of fully anisotropic rotational diffusion. The principal axes of the rotational diffusion tensor calculated from the NMR data are aligned to within 12−23° of the principal axes of the inertial tensor, with the axis of fastest rotational diffusion coincident with that of minimal inertia, and vice versa. Large variations in the ^{15}N−^{1}H nuclear Overhauser effects for individual amino acids correlate with the degree of convergence in the previously calculated NMR structure. In particular, the partially disordered residues Val86−Arg97 that contain the human cyclophilin A (CypA) packaging signal have ^{15}N heteronuclear NOEs and transversal relaxation rates consistent with a high degree of dynamic conformational averaging. The N-terminal domain of a CA mutant (G94D) that confers both resistance to and dependence on cyclosporin A analogues was also analyzed. Our results indicate that this mutation does not influence the conformation or dynamics of CA^{151}, and therefore probably affects the function of the protein by modifying essential intermolecular CA−CA interactions.

The RNA genome of human immunodeficiency virus type-1 (HIV-1)[1] encodes a 55 kDa Gag polyprotein that plays a central role in virus assembly (for reviews, see refs *1*, *2*). Approximately 2000 copies of Gag self-associate specifically at the cell membrane, capture 2 copies of the unspliced genomic viral RNA, and bud to form an immature virion. After budding, Gag is proteolytically processed by the viral protease into the three major structural proteins: matrix (MA), capsid (CA), and nucleocapsid (NC) (*3*−*5*). Major morphological changes occur during this maturation process, in which the capsid proteins condense to form the cone-shaped core particle in the center of the virion, the nucleocapsid proteins form a ribonucleoprotein complex with the two RNA molecules that are encapsidated in the viral core, and the matrix proteins remain associated with the inner face of the viral membrane (for a review, see ref *6*). During the subsequent infection stage, the capsid core disassembles in the cytosol of the newly infected host cell, releasing a reverse transcription/preintegration complex that is responsible for the synthesis of proviral DNA and its subsequent nuclear localization and chromosomal integration.

There is currently great interest in understanding the nature of the intermolecular CA−CA interactions responsible for capsid particle assembly and disassembly. Structural studies of the intact HIV-1 CA protein have been hampered by its tendency to aggregate in solution, forming tubes, spheres, and other noncrystalline structures (*7*, *8*). NMR-detected limited proteolysis and biochemical experiments demonstrated that HIV-1 CA contains an independently folded N-terminal domain (here denoted CA^{151}), and high-resolution structural information has now been obtained by NMR (*9*) and X-ray crystallographic methods (*10*−*12*) for both the N- and C-terminal domains. The N-terminal domain comprises a highly helical folded monomer with an unusually flat, triangular overall shape. The C-terminal domain, which is also highly helical, forms dimers and exhibits a dimerization constant equal to that of the intact CA protein (K_d = 16 μM). Very recently, the crystal structure of the complete HIV-1 CA protein bound to an antibody fragment that recognizes the C-terminal domain has been reported (*13*). The structure exhibits an unusual head-to-tail dimer interface that is significantly different from the predicted models (*12*).

HIV-1 CA is also responsible for packaging the cellular human protein cyclophilin A (CypA), a human peptidyl-prolyl isomerase that is essential for HIV-1 infectivity (*14*−*16*). A single mutation in CA (P90A), or the addition of the CypA inhibitor cyclosporin A (CsA), or CsA analogues, to

† This research was supported by NIH Grant AI30917. R.C.-O. acknowledges a postdoctoral fellowship from Fundación Ramón Areces.

* Address correspondence to this author. Telephone: (410)-455-2527. Fax: (410)-455-1174. E-mail: summers@hhmi.umbc.edu.

[1] Abbreviations: CA, HIV-1 capsid protein; CA^{151}, N-terminal core domain of the HIV-1 capsid protein; CA^{151}(G94D), CA^{151} with substitution of the native residue Gly in position 94 by Asp; CypA, human cyclophilin A; CsA, cyclosporin A; **D**, molecular rotational diffusion tensor; **I**, molecular inertial tensor; NMR, nuclear magnetic resonance; HIV-1, human immunodeficiency virus type-1; HSQC, heteronuclear single quantum correlation; PDB, protein data bank; R_1 (T_1), longitudinal relaxation rate (time); R_2 (T_2), transversal relaxation rate (time); rms, root-mean-square; SIV, simian immunodeficiency virus; XNOE, heteronuclear ^{15}N{^{1}H} nuclear Overhauser effect; wt, wild type.

10.1021/bi990991x CCC: $18.00
Published on Web 07/20/1999

infected cells, prevents the incorporation of CypA into otherwise indistinguishable virions and inhibits infectivity by blocking HIV-1 replication (*16*−*18*). Although CypA packaging is essential for HIV-1 replication, its functional role has yet to be unambiguously established. Approximately 200 CypA molecules are packaged in each virion via intermolecular interactions with residues Pro85−Pro93 of CA^{151} (*10*, *19*). Although the conformation of these residues is well-defined in the X-ray structure of the CypA−CA^{151} complex (*10*) [and in complexes of CypA with CA peptides and of CA with an antibody fragment (*11*, *20*)], the CypA binding residues are poorly defined in the NMR structure of the free domain (*9*).

Long-term treatment of virions with CsA analogues leads to infective revertant mutants that are both CsA resistant and also CsA dependent (*21*). Interestingly, these mutations occur in a type-II tight turn that is located within the conformationally disordered loop and is immediately adjacent to the CypA binding site (*9*). The mutations do not affect the affinity of CA for CypA, indicating that CsA resistance is not due to a packaging defect of CypA (*22*). It is thus possible that these mutations disrupt the tight turn, resulting in local or possibly global conformational changes, and possibly altering CA−CA interactions.

To gain insights into the dynamic behavior of the HIV-1 CA N-terminal domain, including the residues responsible for CypA packaging, we have performed quantitative ^{15}N NMR relaxation studies of the backbone amide groups of CA^{151}. In addition, the NMR chemical shifts, interproton NOEs, and ^{15}N relaxation properties of the HIV-1 CA revertant mutant, G94D, have been determined in order to assess the effects of this single amino acid substitution on the local and global properties of the domain.

MATERIALS AND METHODS

Protein Preparation. The U-^{15}N-labeled CA^{151} protein and the expression plasmid pET32a-CA^{151}(G94D), produced as described (*19*), were provided by Dr. Sundquist (University of Utah). pET32a-CA^{151}(G94D) was used to transform *E. coli* BL21(DE3) pLysS. Cells were grown in M9 minimal medium containing $^{15}NH_4Cl$ as the only nitrogen source to produce [U-^{15}N]-CA^{151}(G94D). The purification protocol was slightly modified with respect to that published for the wild-type protein (*9*) and consisted of NH_4SO_4 (20−40%) precipitation, followed by two FPLC fractionations by ionic exchange: cationic (Q, Pharmacia) in 25 mM Tris-HCl, pH 8.0, and anionic (SP, Pharmacia) in 25 mM potassium phosphate, pH 7.2. Protein-containing chromatographic fractions were pooled and concentrated, and the buffer was exchanged with acetate buffer (CD_3COO^-), 25 mM, pH 5.5, by membrane ultrafiltration (Centricom). NMR samples were 5% D_2O/95% H_2O and 0.7 and 0.5 mM in protein concentration [determined assuming $\epsilon_{280} = 25\,320\ cm^{-1}\ M^{-1}$ (*23*)] for CA^{151} and CA^{151}(G94D), respectively.

NMR Data Collection. All NMR spectra were recorded at 28 °C with a GE Omega PSG 600 spectrometer equipped with an inverse broadband 5 mm probehead and a shielded *z*-gradient coil. 2D ^{15}N HSQC, 3D TOCSY−HSQC (τ_m = 60 ms), and 3D NOESY−HSQC (τ_m = 100 ms) were recorded on the mutant protein sample for the assignment of ^{15}N and ^1HN amide and $^1H\alpha$ backbone resonances (*24*, *25*). A ^{15}N spectral window of 1228 Hz centered at 118.3 ppm and a ^1H spectral width of 9090 Hz centered at the H_2O signal (4.76 ppm) were used for all experiments. The size of the acquired 3D matrixes was $1024^*(t_3) \times 32^*(t_2) \times 128^*(t_1)$, where the asterisk denotes complex points, and 16 scans were recorded per point. $^{15}N-\{^1H\}$ steady-state heteronuclear NOE (XNOE) data were obtained for wild-type and mutant proteins using a relaxation delay of 7 s, yielding data sets of $1024^*(t_2) \times 90^*(t_1)$ after accumulation of 128 scans per point.

Longitudinal (R_1) and transversal (R_2) relaxation data for the backbone ^{15}N nuclei of the wild-type protein were recorded as $1024^*(t_2) \times 28^*(t_1)$ data sets with 32 scans per point, using a recovery delay between scans of 2 s. Eight different values for the relaxation time were recorded in an interleaved manner for ^{15}N T_1 ($=1/R_1$; T_1 delays = 60.2, 100.3, 220.6, 401.1, 802.2, 952.6, and 1203.2 ms) and T_2 ($=1/R_2$; T_2 delays = 15.5, 31.0, 46.6, 62.1, 77.6, 93.1, 108.6, and 124.2 ms) relaxation experiments. All the above experiments used conventional pulse sequences (*26*, *27*) and "flip-back" methods (*28*) for water suppression.

NMR Data Processing and Analysis. The NMR data were processed and analyzed with the NmrPipe (*29*) and NMRView (*30*) software packages, respectively. Heteronuclear dimensions were extended by linear prediction and zero-filled prior to Fourier transformation. The XNOE value for a given residue was calculated as the intensity ratio (I/I_0) of the $^{15}N-^1H$ correlation peak in the presence (I) and absence (I_0) of 3 s proton saturation, and errors were estimated from the base line noise in the two spectra. Relaxation rates were determined by nonlinear fits of the time dependence of the peak intensities, and errors were obtained from the covariance matrix.

The analysis of the overall tumbling of CA^{151} from the T_1/T_2 ratios of N−H groups without significant internal motion contributions was carried out with the program Quadric Diffusion 1.11 from Prof. A. G. Palmer, Columbia University (*31*). With this approach, a local diffusion constant for the *i*th residue (D_i) is calculated from its T_1/T_2 value using the "model free" expression for the isotropic spectral density function (*35*). The local diffusion constant can be related to a local diffusion correlation time, τ_i, via the relationship: $\tau_i = [(6D_i)^{-1}]$. The local diffusion constants are then represented in a convenient quadratic form for obtaining the diffusion tensor by minimization techniques (*31*, *33*). The inertial tensor parameters ($\{D_{iso}\}$, $\{D_\perp, D_\parallel, \phi, \theta\}$, and $\{D_{zz}, D_{yy}, D_{xx}, \phi, \theta, \psi\}$) for isotropic, axially symmetric, and fully anisotropic models were fitted to the experimental data via a χ^2 minimization (*31*). The χ^2 function is defined as the sum (over all evaluated residues) of the squared difference between experimental and model-calculated D_i values, weighted by the squared experimental uncertainty in D_i (see eq 16 in ref *31*). The value of the χ^2 function obtained for two different models cannot be used to identify the most appropriate model, as the inclusion of more adjustable parameters (i.e., the smaller number of degrees of freedom) will generally result in lower χ^2 values. Instead, the F statistic, as defined in eq 17 of (*31*), takes into account the different number of degrees of freedom of the two models, and is used to test the improvement in the statistical fit. In practice, a value of $F > 4.5$ indicates that the improved fit to the more highly parametrized model is statistically significant at a confidence level greater than 99%.

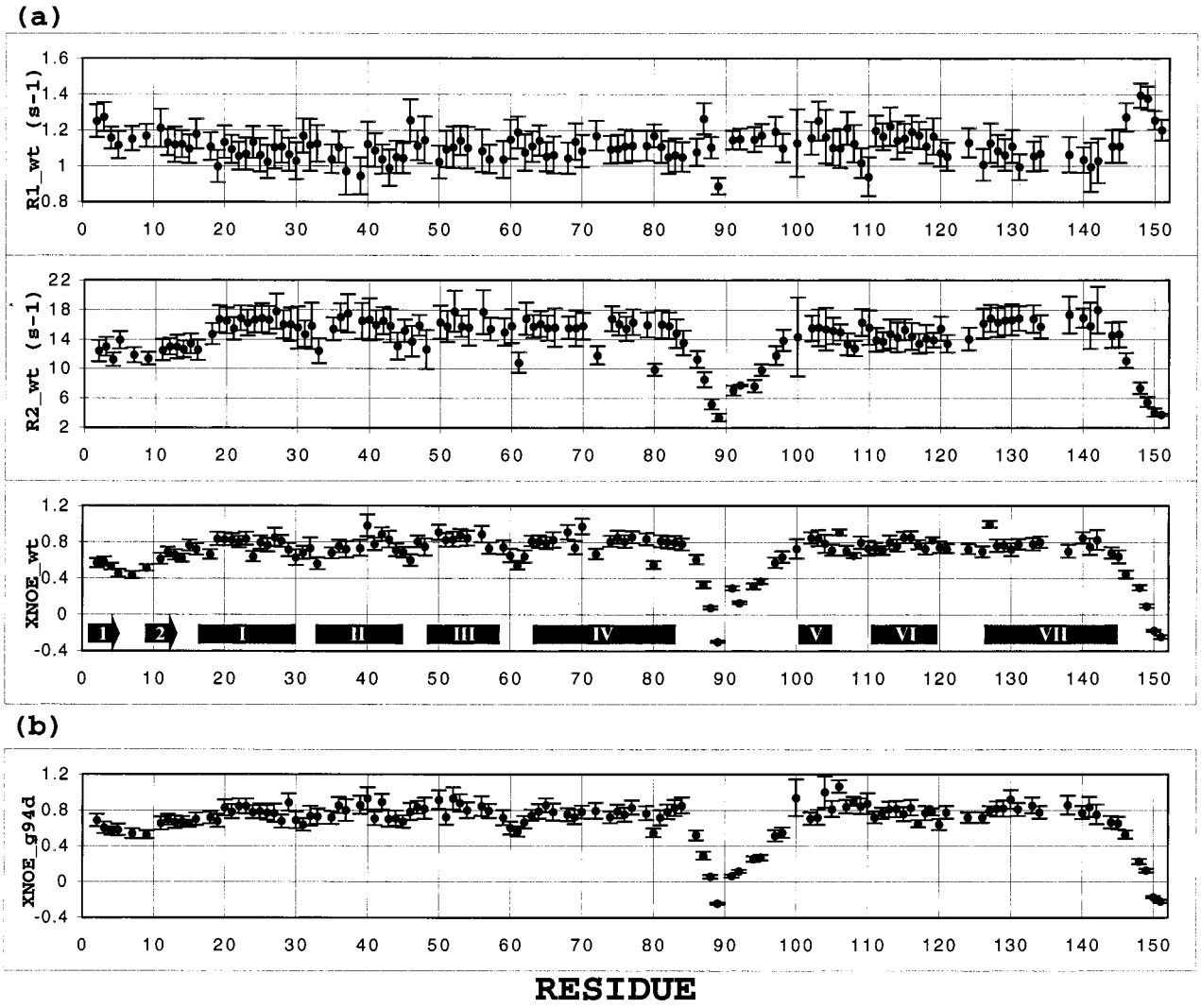

FIGURE 1: ^{15}N backbone amide relaxation parameters measured for (a) CA^{151} and (b) CA^{151}(G94D) versus the protein sequence. (a) From top to bottom: longitudinal relaxation rates (R_1), transverse relaxation rates (R_2), and ^{15}N{^{1}H} NOE (XNOE) for the wild-type CA^{151} are shown. The errors are on average 8%, 12%, and 9%, respectively. Only data for residues sufficiently well resolved in the HSQC spectra (121 in total) are displayed. The secondary structure elements present in CA^{151} [β-strands 1(1−5) and 2(9−13); α-helices I(16−30), II(33−45), III(48−58), IV(63−83), V(100−105), VI(110−119), and VII(126−145)] are represented schematically in the XNOE_wt panel. (b) ^{15}N{^{1}H} NOE for the mutant protein CA^{151}(G94D). The average error is 11%.

RESULTS AND DISCUSSION

NMR Relaxation Data. A total of 138 ^{15}N−^{1}H backbone correlation signals are expected in the HSQC spectrum (151 residues − 13 prolines). The N−H signals of 17 of these residues (6, 8, 10, 55, 58, 67, 71, 73, 78, 96, 101, 132, 135, 136, 137, 139, and 143) were either too weak or severely overlapping for reliable quantification of their intensities. Relaxation data were therefore obtained for 121 backbone amides, which constitute 88% of all backbone amide sites in the protein. Backbone ^{15}N relaxation data obtained for the wild-type CA^{151} protein, including longitudinal (R_1) and transversal (R_2) relaxation rates and heteronuclear ^{15}N{^{1}H} NOE (XNOE), are plotted in Figure 1a, along with XNOE data obtained for CA^{151}(G94D) (Figure 1b). Four residues (40, 70, 106, and 127) in CA^{151} exhibit XNOE values greater than the theoretical maximum of 0.834, indicating that their errors are underestimated and/or they are affected by chemical exchange with the solvent (*28*). These residues were not considered in the analysis of the protein internal dynamics and rotational tumbling.

Internal Dynamics. As described elsewhere (*26*), NMR relaxation rates and XNOEs are influenced by fluctuations of the backbone N−H vector on the time scale of 10^{-3}−10^{-12} s (ms−ps), and therefore by the internal dynamics of the individual NH vector as well as by the overall rotational diffusion of the molecule. For proteins the size of CA^{151}, the ^{15}N XNOE is very sensitive to variations in NH bond vector internal dynamics. The maximum theoretical XNOE of 0.834 (at 60 MHz ^{15}N frequency) qualitatively reflects highly restricted internal motion, whereas values smaller than 0.65 are indicative of substantial internal motion. As observed for a number of proteins, the N- and C-terminal segments of the protein exhibit significantly smaller XNOE values compared to most other residues (Figure 1a). In particular, residues 146 to 151 (the C-terminal tail) exhibit very small XNOE values (average = 0.17), indicative of rapid internal motion. The α-helices exhibit the largest XNOE values (average for $\alpha_I = 0.77$, $\alpha_{II} = 0.75$, $\alpha_{III} = 0.83$, $\alpha_{IV} = 0.80$, $\alpha_V = 0.78$, $\alpha_{VI} = 0.78$, $\alpha_{VII} = 0.77$), whereas the two β-strands of the N-terminal β-hairpin ($\beta_1 = 0.53$, $\beta_2 = 0.62$)

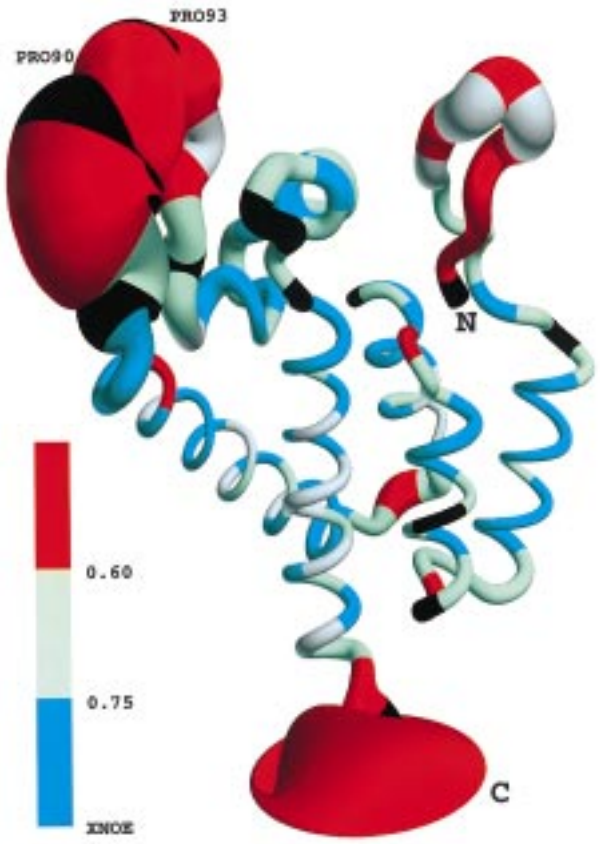

FIGURE 2: Schematic representation of the solution structure of CA^{151}. To illustrate the definition of the different regions, the radius of the tube representing the mean backbone conformation is drawn proportional to the rms deviation of backbone coordinates between the 50 conformers used to represent the solution structure. In addition, residues are grouped and color-coded according to the XNOE value as given in the legend. Proline residues are indicated in black (those closer to the mutation site are explicitly labeled) and those for which relaxation data are not available in gray. This figure was prepared with the program MOLMOL (*44*).

and some tight turns (e.g., residues 59−62; average XNOE = 0.64) exhibit intermediate values.

A representation of the CA^{151} NMR structure is provided in Figure 2, in which the radius of the backbone is proportional to the uncertainty (global heavy atom rms deviation) of the NMR-derived structure (*9*) (PDB entries 1gds and 1gdy), and the residues are color-coded according to the XNOE value. This representation reveals a high correlation between the degree of convergence in the NMR structure and the ^{15}N XNOE data. Regions of the NMR structure that exhibit poor convergence due to a paucity of intermediate- and long-range $^1H-^1H$ NOEs also exhibit relatively low XNOE values, indicative of high dynamic mobility. Thus, the NMR relaxation data indicate that the regions of poor convergence are due to high internal mobility (*9*).

The CypA binding loop, which exists between helices IV and V, exhibits a strikingly high mobility with a minimum XNOE value of −0.3 at G89. The same trends are observed in the R_2 data. Although residues in exposed loops are generally found to be more flexible than those in regular secondary structures, loops of this size (<15 residues) do not normally exhibit XNOEs as low as those observed here for the CA^{151} CypA binding loop.

Rotational Diffusion Analysis of CA^{151}. CA^{151} is a highly asymmetric molecule with an overall structure that resembles an arrowhead (leading edge length of ∼31 Å, trailing edge length of ∼39 Å, and a thickness of ∼16 Å) (*9*). Indeed, the principal components of the inertial tensor of CA^{151}, calculated using the program pdbinertia by A. G. Palmer (*31*), have relative magnitudes of 1.00:0.76:0.46 (average values over the 50 NMR structures, with a standard deviation of ∼3%). This indicates that the protein should exhibit fully anisotropic rotational diffusion. The orientation of the inertial tensor in the molecular frame of reference is shown in two different views in Figure 3a,b, with the principal components of the tensor defined in order of increasing value, i.e., $I_{xx} > I_{yy} > I_{zz}$. The major and minor axes of the inertia tensor are oriented orthogonal to the plane of the arrowhead and in the direction of the arrowhead, respectively.

For a rigid protein tumbling anisotropically, the ^{15}N R_2/R_1 ($=T_1/T_2$) ratios are dependent on the orientation of the NH bond vectors relative to the principal components of the diffusion tensor, **D**. The R_2/R_1 ratio of a given NH will be maximum when the NH bond vector is parallel to the major principal component of the diffusion tensor (D_{zz} in our convention) and perpendicular to the minor principal component (D_{yy}). Conversely, the observed R_2/R_1 will be smallest for NH vectors that are perpendicular to the D_{zz} direction and parallel to that of D_{yy}. The rotational diffusion tensor can thus be determined from the relaxation rate ratios and the relative orientations of the corresponding NH vectors determined in the three-dimensional structure of the protein (*31*−*34*).

A first step in the analysis of rotational diffusion from the relaxation rates involves the identification and elimination from the analysis of those residues for which internal motion contributes significantly to the observed R_2/R_1 values. Of the 121 amide sites available, 50 exhibited small XNOE values (<0.65), 4 exhibited XNOE values greater than the theoretical upper limit, and 1 exhibited a significantly longer T_2 (decreased R_2 without concomitant decrease in R_1). The remaining 66 residues were identified as having "rigid" N−H bond vectors since internal motions do not appear to contribute significantly to the NMR relaxation parameters. The T_1/T_2 ratios for the rigid residues selected using the above criteria are shown in Figure 3c. All but 12 of these residues (3 N-caps: 16, 48, 110; and 15, 47, 59, 84, 107, 109, 120, 121, and 124) are located in helical segments of the protein (indicated with the colored bars).

Since the NH bond vectors within a given α-helix are nearly parallel to the helix axis (with a rms deviation of ∼15°), the T_1/T_2 values for the residues of a given helix should be generally uniform and thus provide redundant angular information. Residues of different helices may exhibit different T_1/T_2 values, depending on the helix orientation relative to the rotational diffusion tensor. The observed average values of T_1/T_2 for the different helices in the protein have characteristic values of $\alpha_I = 15.0$, $\alpha_{II} = 15.3$, $\alpha_{III} = 14.6$, $\alpha_{IV} = 14.4$, $\alpha_V = 13.2$, $\alpha_{VI} = 12.6$, and $\alpha_{VII} = 15.8$ (see bars, color-coded according to the helix number in Figure 3c). Neglecting hydration effects, the diffusion tensor should be collinear with the inertial tensor of a given molecule, with the axis of fastest diffusion (D_{zz}) corresponding with that of minimal inertia (I_{zz}), and vice versa (i.e., D_{yy} corresponds to I_{xx}). Helix VII (purple) exhibits the largest

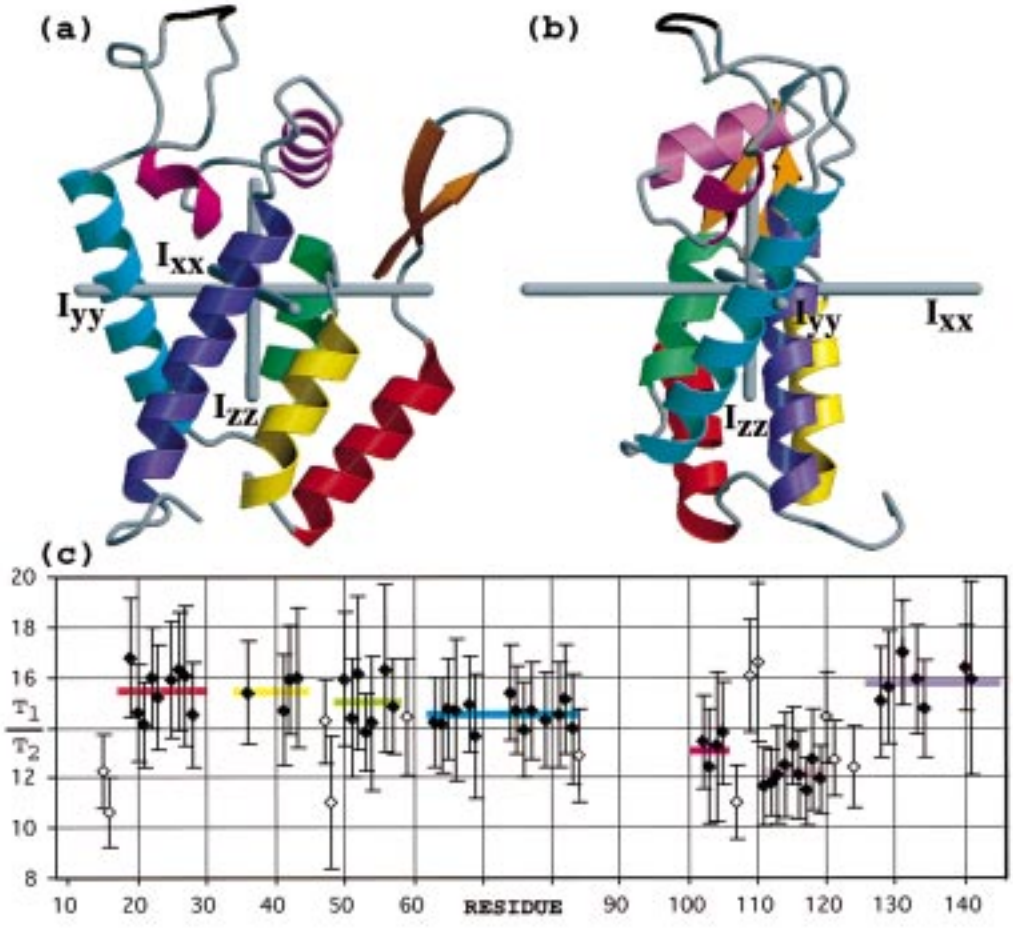

FIGURE 3: (a) Ribbon representation of one of the conformers describing the solution structure of CA[151]. The β-strands are colored in orange; α-helices in red (I), yellow (II), green (III), cyan (IV), fuchsia (V), pink (VI), and purple (VII); the rest in gray; and segment 92−95 is shown in black. The inertial tensor **I** is represented as the three mutually perpendicular gray axes, that are of relative magnitude 1.00: 0.77:0.47 for I_{xx}:I_{yy}:I_{zz}, respectively. (b) Orthogonal view of (a) illustrating the asymmetry of the molecule. (c) Plot of the T_1/T_2 values of the rigid NHs selected for diffusion analysis versus the protein sequence. The position of the α-helices is indicated by the colored bars, where N-cap residues 16, 33, 48, and 110 have been excluded from helices I, II, II, and VI, respectively. The bars are drawn at the average T_1/T_2 value of each particular helix and are colored as in (a). Data for nonhelical residues are represented with the white symbol.

average T_1/T_2 value, and its helix axis is most closely aligned with the minor principal component of the inertia tensor I_{zz} and is perpendicular to the major component I_{xx}, as expected. In contrast, helix VI (pink) has the lowest average T_1/T_2 value and is nearly perpendicular to the minor principal axis of the inertia tensor I_{zz} and nearly parallel to the major component, I_{xx}. The other helices exhibit intermediate orientations and, as expected, exhibit intermediate T_1/T_2 ratios. The actual angles between helix VII, the closest to the orientation of the minor principal axis of the inertia tensor, and the helix axes of helices α_I through α_{VI} are 25, 20, 31, 33, 46, and 53°, respectively. These angles are in excellent qualitative agreement with the observed variations in the average T_1/T_2 ratios for the different helices.

Rotational diffusion tensors were calculated quantitatively for isotropic, axially symmetric, and fully anisotropic diffusion models using the program [Quadric Diffusion] (*31*). All 66 "rigid" NH vectors (see above) were considered in the initial analysis. Statistical significance tests indicate that the data fit better to anisotropic models compared to the isotropic rotation model (Table 1). Calculations assuming an axially symmetric model afforded reasonably good fits for all residues except Ser16, Thr48, Thr107, and Ser109. Elimination of these residues from the calculations resulted in a significant statistical improvement of the fitting, in which χ^2 dropped by almost 50% (Table 1). Ser16, Thr48, and Thr107 have relatively small T_1/T_2 values compared to the other "rigid" residues, which is indicative of flexibility. However, the measured XNOE values are relative large, and this could be due to saturation transfer from H_2O in the reference NOE experiment (*28*). Chemical exchange may contribute to the XNOEs measured for these residues since their N−H protons are exposed to solvent. Exchange might also be facilitated by the side chain hydroxyl group on these residues.

Our data clearly indicate that the criteria used to identify "rigid" amides is not sufficiently stringent. Amides with relatively small chemical or conformational exchange dynamics that are not identified by the commonly used protocols may be identified during the course of quantitative rotational diffusion calculations if the molecule is sufficiently anisotropic. To minimize the chances that relatively small exchange contributions may affect the fitting of the relaxation data, only the α-helical residues were included in the subsequent calculations (a total of 54 residues). This conservative approach was considered to be the most reliable since, as discussed above, the NH data for a given helix should be highly redundant. The results of the fitting to the different rotational diffusion models are summarized in Table 1.

Table 1: Rotational Diffusion Analysis of CA^{151} [a]

model	D_{iso} ($\times 10^{-7}$ s^{-1})	$2D_{zz}/(D_{xx}+D_{yy})$	D_{xx}/D_{yy}	θ (deg)	ϕ (deg)	ψ (deg)	χ^2	F
isotropic	1.41 ± 0.01						27.8	
	(1.42 ± 0.01)						(41.5)	
axial (1)	1.44 ± 0.02	0.72 ± 0.06		95.1 ± 9.9	8 ± 12		9.7	31.3
	(1.45 ± 0.02)	(0.70 ± 0.06)		(80.1 ± 6.7)	(185 ± 9)		(18.5)	(25.7)
axial (2)	1.45 ± 0.02	1.28 ± 0.08		12.3 ± 8.6	81 ± 56		10.9	25.7
	(1.46 ± 0.02)	(1.29 ± 0.07)		(17.7 ± 8.5)	(54 ± 23)		(20.6)	(21.1)
anisotropic	1.45 ± 0.02	1.24 ± 0.08	1.26 ± 0.14	24.0 ± 33.6	84 ± 78	-73 ± 61	7.4	7.3/11.4
	(1.46 ± 0.02)	(1.27 ± 0.07)	(1.29 ± 0.13)	(20.6 ± 34.7)	(67 ± 58)	(-60 ± 48)	(15.2)	(6.6/10.7)

[a] The fit was done using a single structure that was closest to the mean structure of the 50 conformer NMR ensemble and T_1/T_2 values of either the 54 helical rigid NH vectors or the 66 rigid vectors (results in parentheses). Values are reported for the optimal fit, and the errors are determined from the covariance matrix (isotropic and axially symmetric models) or from Monte Carlo simulations (anisotropic model). The Euler angles describe the orientation of the diffusion tensor in the frame of reference of the inertial tensor. χ^2 is a measure of the quality of the fit. F evaluates the statistical significance of the improvement over the previous, simpler model, with a smaller number of parameters, and in the case of the anisotropic model, the two values reported are with respect to the two axial solutions.

Elimination of the nonhelical residues results in a significant decrease of the target function, but the diffusion tensor parameters obtained with both datasets are very similar. It should also be noted that similar results for **D** were obtained using different conformers of the NMR ensemble, or when **D** was obtained by averaging the results of individual fits (data not shown).

Two solutions are obtained when fitting the data to an axially symmetric tensor, and in both cases the fits represent a statistically significant improvement over the isotropic model (Table 1). In one solution, the anisotropy parameter $[2D_{zz}/(D_{xx} + D_{yy})]$ is greater than unity (axial solution 2) which corresponds to a prolate ellipsoidal tensor. The second solution describes an oblate ellipsoidal tensor with an anisotropy parameter less than unity (axial solution 1). The predicted orientation dependence of D_i for the optimized prolate and oblate tensors is plotted with the experimental D_i values in Figures 4a and 4b, respectively. These plots reveal that the NH vectors of the helices provide a good sampling of the angular space about the two principal axes of the two axial solutions. When the Euler angles θ and ϕ reported in Table 1 are used to find the orientations of the two identified principal axes of the diffusion tensors in the molecule, the intuitively expected result is obtained. That is, for the prolate ellipsoidal model, the axis about which rotational diffusion is most rapid is nearly collinear ($\sim$12°) with the axis of minimal inertia (I_{zz} in Figure 3a,b); the tensor elements have values of $D_{||} = 1.71 \times 10^7$ s^{-1} ($\tau_{||} = 9.8$ ns) and $D_{\perp} = 1.33 \times 10^7$ s^{-1} ($\tau_{\perp} = 12.5$ ns). Likewise, for the oblate ellipsoid model solution, the axis about which diffusion is slowest is nearly collinear ($\sim$23°) with the major principal component of the inertial tensor (I_{xx} in Figure 3a,b); the tensor elements in this solution have magnitudes of $D_{||} = 1.14 \times 10^7$ s^{-1} ($\tau_{||} = 14.7$ ns) and $D_{\perp} = 1.59 \times 10^7$ s^{-1} ($\tau_{\perp} = 10.5$ ns).

Recent attempts to fit NMR relaxation data obtained for a fully anisotropic diffusing molecule to axially symmetric models resulted in two equally valid solutions (*36*, *37*), consistent with the above findings. It is therefore not surprising that, upon fitting the relaxation data of CA^{151} to a fully anisotropic model, a statistically significant improvement in the fit is obtained (Table 1). Predicted and experimentally measured D_i values for the fully anisotropic diffusion model are plotted as a function of residue number in Figure 4c. The principal components of the fully asymmetric diffusion tensor are $D_{zz} = 1.67 \times 10^7$ s^{-1}, $D_{xx} = 1.50 \times 10^7$ s^{-1}, and $D_{yy} = 1.19 \times 10^7$ s^{-1}, which correspond to rotational correlation times of $\tau_{zz} = 9.97$ ns, $\tau_{xx} = 11.13$ ns, and $\tau_{yy} = 14.02$ ns, respectively. These values are in very good agreement with the $\tau_{||}$ and $\tau_{\perp}$ values obtained from the two axially symmetric solutions. In addition, the orientations of the major and minor axes of the optimal fully asymmetric diffusion tensor are nearly coincident with the orientations of the two principal axes identified in the axially symmetric analysis (Table 1). The three principal components of the diffusion tensor ($D_{zz} > D_{xx} > D_{yy}$) also have orientations similar to those of the three principal axis of the inertial tensor: D_{zz} forms an angle of 22° with I_{zz}, D_{yy} forms an angle of 12° with I_{xx}, and D_{xx} and I_{yy} are separated by 23° (compare Figure 3a,b with Figure 5a,b). Thus, the axis of minimal inertia approximately coincides with that of maximum diffusion, and that of maximum inertia with that of slowest diffusion. As previously noted, the correspondence between inertial and diffusion tensors is expected since the hydrodynamic properties are determined mainly by overall molecular shape (*32*, *38*).

The effective isotropic rotational diffusion constant ($D_{iso} = 1.46 \times 10^7$ s^{-1}, $\tau_{iso} = 11.48$ ns) is larger than expected for a protein of this mass, at 28 °C. This may be due to the unusually broad and flat shape of the molecule. In fact, hydrodynamic calculations support the intuitive expectation that a compact, spherical protein would tumble in solution more rapidly than a wide and flatter protein of the same mass (*38*). However, it is also possible that a small amount of protein aggregation could contribute to the larger-than-expected rotational correlation time. In this regard, we previously noted that proton T_2 values decrease at CA^{151} concentrations above 1.3 mM, consistent with protein self-aggregation (*9*).

NMR Characterization of CA^{151}(G94D) and Comparison with CA^{151}. To characterize the structural differences between CA^{151} and CA^{151}(G94D), the ^{1}HN, ^{15}N, and ^{1}Hα chemical shifts of the two proteins were compared. Except for a few residues, the ^{1}H$-^{15}$N HSQC spectra of CA^{151}(G94D) and CA^{151} were essentially identical, which facilitated assignment of the backbone ^{1}HN and ^{15}N signals. ^{1}Hα assignments were obtained from the ^{15}N-edited 3D TOCSY−HSQC and 3D NOESY−HSQC spectra. Differences between the NMR chemical shifts observed for CA^{151}(G94D) and CA^{151} are generally very small and are plotted in Figure 6. The only significant variations occur for residues Ile91−Arg97 (^{15}N), His84−Glu98 (^{1}HN), and Ala92−Gln95 (^{1}Hα), all of which

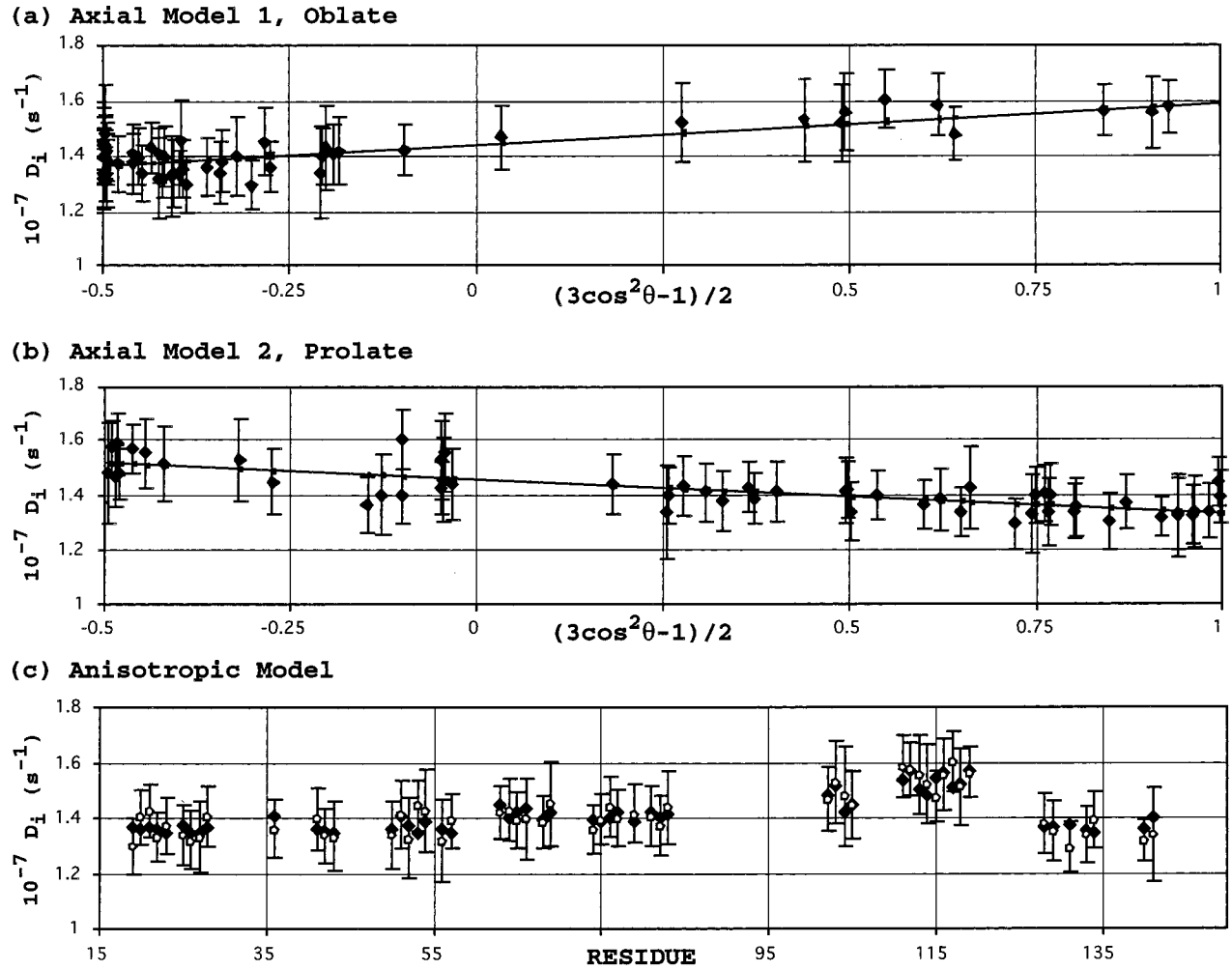

FIGURE 4: Fit of the experimental data of helical and rigid NHs to the different diffusion models. For the axially symmetric models, the local diffusion coefficient D_i for each residue i considered is represented versus $Y^2_0(\theta) = (3\cos^2\theta - 1)/2$, for (a) the oblate approximation and (b) the prolate approximation, where θ_i is the angle between the principal component of **D**, $D_{zz} = D_{||}$, and the ith NH bond vector. The straight lines represent, in each case, the dependence for the best-fit axial tensor. The fit to the fully anisotropic model is shown in (c) where the experimental D_i values (white symbols in the center of the error bars) and the values of the optimal fit (black symbols) are represented for the seven helices, from left to right along the protein sequence.

are in the vicinity of the mutation site. As expected, the mutated residue consistently displays the largest shift difference. Very small perturbations occur for residues Ile115−Met118 (^{1}HN and ^{1}Hα), which constitute the central part of helix VI that is proximal to the CypA binding loop (*9*). Thus, the chemical shift comparisons indicate that the G94D mutation results in localized and relatively insignificant structural perturbations.

To examine the structural consequences of the mutation more precisely, ^{1}H−^{1}H homonuclear NOE data obtained for the mutant and wild-type proteins were compared. In agreement with the chemical shift data, the NOE patterns of the two proteins were essentially indistinguishable. ^{1}H to amide ^{1}H NOEs for residues of the CypA binding loop of CA^{151} and CA^{151}(G94D) exhibited similar intensities, and the pattern and intensities of NOEs in helix VI were also identical. The only exceptions were the absence of detectable NOEs for the 86CγH$_3$−98HN and 91Cγ_2H$_3$−95HN proton pairs, which were very weak in the NOE data obtained for CA^{151}. In CA^{151}, residues Ala92-Pro93-Gly94-Gln95 form a type-II tight turn as evidenced by the following NOEs: 92CβH$_3$−95HN, 93Hα−94HN, 93Hα−95HN, and 94HN−95HN. For CA^{151}(G94D), 93Hα−95HN is not observed, but the other three are clearly present. This suggests that the mutation of G for D in position 3 of the turn at most destabilizes the turn, but does not completely disrupt it. In fact, this is the expected effect according to a statistical analysis of the presence of the amino acids in the different types of β-turns (*39*). Thus, APGQ has a large statistical preference to adopt either a type-II or a type-I′ β-turn, whereas the mutant APDQ has a smaller but still relatively significant tendency to adopt type-II or type-I turns.

We previously demonstrated that the G89−P90 peptide bond in CA^{151} exists in a conformational equilibrium of the trans and cis Pro−peptide bonds. It is well-known that proline readily adopts both cis and trans peptide bond conformations in small unstructured peptides (*40*), and the observed equilibrium in CA^{151} is consistent with the present observation of high internal dynamics for these residues. In the case of the G94D mutant, a small subset of additional signals were also observed that correspond to the cis-Pro90 conformation. The estimated cis/trans population for Pro90 is indistinguishable from that observed in the wild-type protein (*9*). It thus appears that the conformational equilibrium associated with the cis−trans isomerization of P90 is not fundamentally altered in the CA^{151}(G94D) mutant.

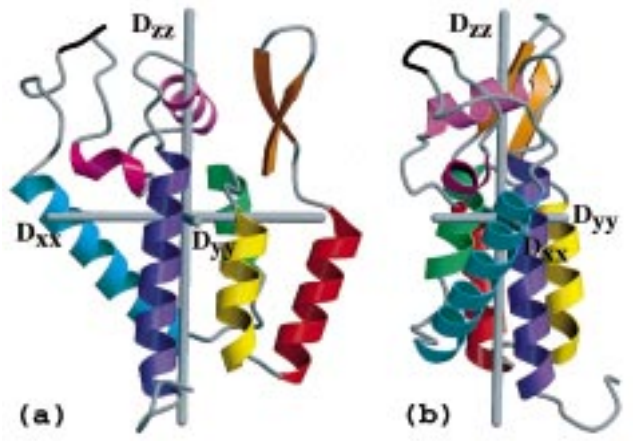

FIGURE 5: Schematic representation of CA^{151} aligned in the fully anisotropic rotational diffusion tensor. The tensor **D** is represented by the three mutually perpendicular gray axes, in the direction of the three principal components of the diffusion tensor. The length of each axis is drawn proportionally to $10^{D_{ii}}$, where $D_{ii} = D_{xx}$, D_{yy}, D_{zz} are the magnitudes of the principal components. The color code is as in Figure 3. The two orthogonal views in (a) and (b) are to be compared with Figure 3a and Figure 3b, respectively. In Figures 3 and 5, the Cartesian systems representing the inertia and diffusion tensors are displayed in the same orientation, so that the change in orientation of the molecule between the two figures is determined by the relative orientations of **I** and **D**.

As for the wild-type protein, quantitative XNOE data were obtained for 120 out of the 137 backbone $^{15}N-^{1}H$ groups. These values of XNOE for CA^{151}(G94D) are shown in Figure 1b. Of these, only residue Gly106 exhibited an XNOE value greater than the theoretical maximum of 0.834. The heteronuclear NOE values measured for CA^{151} and CA^{151}(G94D) are nearly identical, with the average difference (0.01) falling well within the experimental error. Residues 89 and 108 appear to be slightly less mobile in the mutant (on average, XNOE increases by ~0.14) whereas the mutated residue at position 94 and residue 95 exhibit slightly increased mobility (on average, XNOE decreases by ~0.13). The variations are compatible with local effects resulting from the change of residue at position 94, and are small compared to the relatively large mobility of the loop. Thus, it does not appear that the phenotype observed for the G94D mutant is due to large differences in the internal dynamics of the domain.

CONCLUSIONS

The N-terminal domain of the HIV-1 capsid protein provides an ideal system for evaluating the methods used to determine anisotropic rotational diffusion. The domain has an unusually asymmetric shape and is comprised mainly of α-helices. Since the N−H vectors of a given helix are essentially collinear with the helix axis, the relaxation data obtained for amides of a given helix provide redundant information regarding the angle of the helix relative to the diffusion tensor. Our studies indicate that the relaxation data obtained for backbone amides in α-helices are internally consistent, and thus can be used to reliably evaluate rotational diffusion.

However, it is clear that the standard criteria used to identify "rigid" amides is not sufficiently stringent, at least

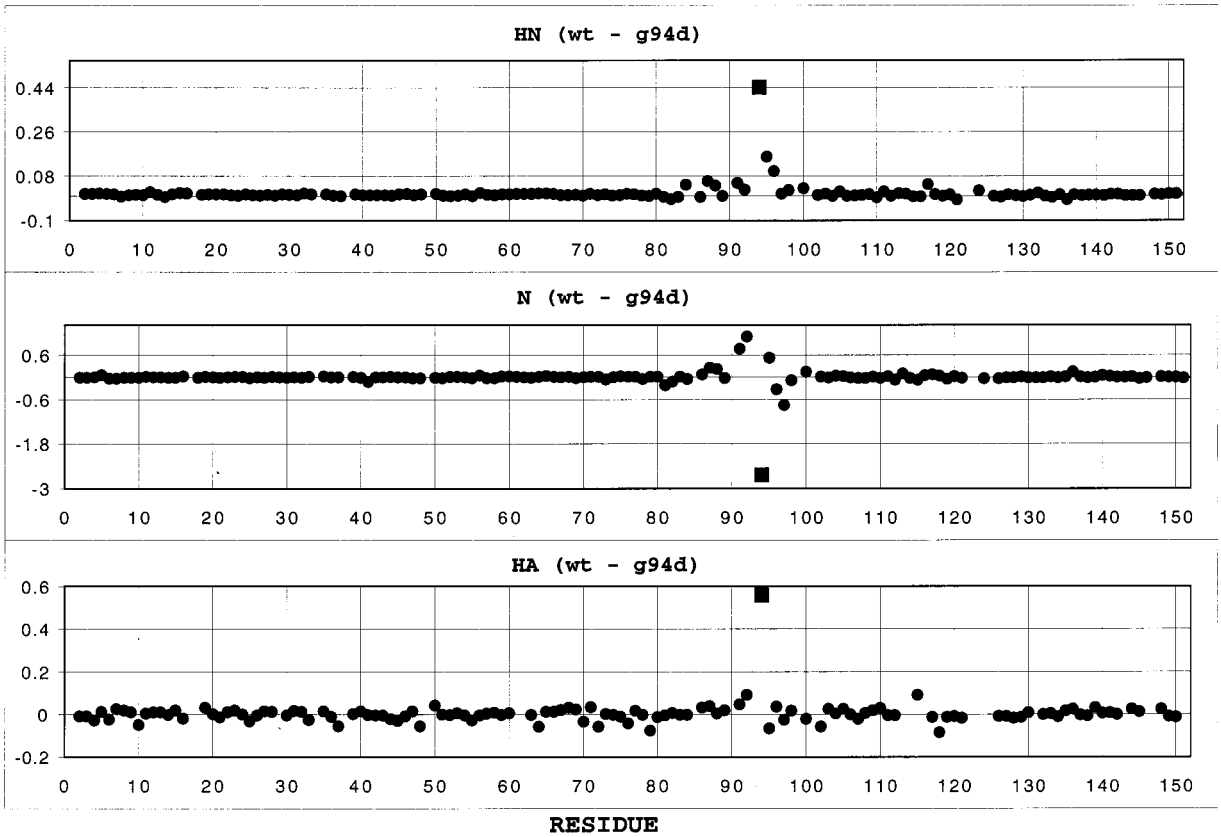

FIGURE 6: Chemical shifts differences of ^{1}HN, ^{15}N, and $^{1}H\alpha$ backbone nuclei between the wild-type protein (CA^{151}) and the mutant CA^{151}-(G94D). Errors in the measurement of the chemical shifts, taken as the spectral resolution in the relevant dimension, are approximately ±0.06, ±0.02, and ±0.16 ppm, for $^{1}H\alpha$, ^{1}HN, and ^{1}N, respectively. The data for the mutated residue at position 94 are indicated with the square symbol.

in this case. For example, backbone amides of N-terminal helix capping residues Ser16 and Thr48 were identified as "rigid" by comparison of individual relaxation parameters with average values, but the T_1/T_2 ratios for these residues fall outside the limits established by the analysis of the highly redundant data obtained for the α-helical amides (Figure 4). Our studies suggest that initial rotational diffusion analyses should probably involve only amide groups in regions of well-defined secondary structure (i.e., α-helices or β-sheets), and that backbone amides with relatively minor chemical or conformational exchange dynamics that are not identified by the commonly used protocols may be subsequently identified during the course of the rotational diffusion calculations. Of course, this procedure will only work if the molecule is sufficiently anisotropic and the orientations of the amides in regions of well-defined secondary structure adequately sample the limits of vector space about the rotational diffusion axes.

The present studies also reveal that regions of CA^{151} that were poorly defined in the NMR structure, including the cyclophilin A binding loop, exhibit a high degree of conformational mobility. This conformational freedom explains the observation of an equilibrium distribution of cis- and trans-Pro90 backbone conformations for the protein in solution, and also confirms that Pro90 is freely accessible for binding to cyclophilin A.

Although the HIV-1 capsid protein is the functional target of CypA, the exact physiological role of CypA is not well understood. Substitution of the simian immunodeficiency virus (SIV) CA gene by the HIV-1 CA gene results in chimeric virions that package CypA and are sensitive to inhibition by CsA, even though native SIV particles do not package or require CypA (*41*). Even the transfer of only the CypA binding site of HIV-1 CA (residues 86−93) to the corresponding position in SIV confers the ability to package CypA and the sensitivity to CsA. In addition, the transfer of residues 86−90 required the presence of CsA for efficient virus replication (*42*), coincident with the phenotype displayed by the spontaneous CsA resistant/dependent mutations. Although it is tempting to speculate that the G94D mutant might alter the structure of the type-II turn, or possibly exert a longer-range conformational change, we were unable to detect significant differences in chemical shifts, homonuclear $^1H-^1H$ NOEs, or heteronuclear $^{15}N-\{^1H\}$ NOEs upon comparison of the wild-type and CA^{151}-(G94D) mutant domains. The mutation also does not significantly affect the dynamic properties of the partially ordered CypA binding loop, or the cis/trans X−Pro peptide bond equilibrium exhibited by the CypA binding residue, Pro90.

Since the G94D mutation does not detectably alter the structure or dynamics of the domain, it appears likely that the mutation affects intermolecular CA−CA interactions that take place as the capsid molecules condense during capsid maturation. The partially ordered CypA binding loop, including the type-II tight turn, may function by binding to an adjacent capsid protein in the assembling capsid core particle. Residues of the type-II turn may form specific intermolecular interactions, and CypA may be subsequently required to break these intermolecular interactions and facilitate capsid disassembly during the infectivity stage of the virus replication cycle. In this model, the function of the revertant mutants would be to partially destabilize the intermolecular interactions, to the extent that CypA is no longer required for infectivity. The recent preparation of capsid-like particles in vitro (*43*) offers the exciting prospect that these and other issues regarding the mechanism of capsid assembly/disassembly and the structural nature of the capsid core can be addressed directly.

ACKNOWLEDGMENT

We are very grateful to Dr. Sanghee Yoo and Dr. W. I. Sundquist (University of Utah) for providing the CA^{151} sample and the plasmid with the CA^{151}(G94D) gene, and to Dr. M. E. Zawrotny (HHMI−UMBC) for technical support.

REFERENCES

1. Gelderblom, H. R. (1991) *AIDS 5*, 617−637.
2. Wills, J., and Craven, R. (1991) *AIDS 5*, 639−654.
3. Mervis, R. J., Ahmad, N., Lillehoj, E. P., Raum, M. G., Salazar, F. H. R., Chan, H. W., and Venkatesan, S. (1988) *J. Virol. 62*, 3993−4002.
4. Henderson, L. E., Bowers, M. A., Sowder, R. C., II, 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.
5. Wiegers, K., Rutter, G., Kottler, H., Tessmer, U., Hohenberg, H., and Krausslich, H.-G. (1998) *J. Virol. 72*, 2846−2854.
6. Turner, B. G., and Summers, M. F. (1999) *J. Mol. Biol. 285*, 1−32.
7. Ehrlich, L. S., Agresta, B. E., and Carter, C. A. (1992) *J. Virol. 66*, 4874−4883.
8. Rosé, S., Hensley, P., O'Shannessy, D. J., Culp, J., Debouck, C., and Chaiken, I. (1992) *Proteins: Struct., Funct., Genet. 13*, 112−119.
9. Gitti, R. K., Lee, B. M., Walker, J., Summers, M. F., Yoo, S., and Sundquist, W. I. (1996) *Science 273*, 231−235.
10. Gamble, T. R., Vajdos, F., Yoo, S., Worthylake, D. K., Houseweart, S. M., Sundquist, W. I., and Hill, C. P. (1996) *Cell 87*, 1285−1294.
11. Momany, C., Kovari, L. C., Prongay, A. J., Keller, W., Gitti, R. K., Lee, B. M., Gorbalenya, A. E., Tong, L., McClure, J., Ehrlich, L. S., Summers, M. F., Carter, C., and Rossmann, M. G. (1996) *Nat. Struct. Biol. 9*, 763−770.
12. Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Korthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997) *Science 278*, 849−853.
13. Berthet-Colominas, C., Monaco, S., Novelli, A., Sibai, G., Mallet, F., and Cusack, S. (1999) *EMBO J. 18*, 1124−1136.
14. Luban, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V., and Goff, S. P. (1993) *Cell 73*, 1067−1078.
15. Franke, K. E., En Hui Yuan, H., and Luban, J. (1994) *Nature 24*, 359−362.
16. Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C. T., Sodroski, J., and Gottlinger, H. G. (1994) *Nature 372*, 363−365.
17. Steinkasserer, A., Harrison, R., Billich, A., Hammerschmid, F., Werner, G., Wolff, B., Peichl, P., Palfi, G., Schnitzel, W., Mlynar, E., and Rosenwirth, B. (1995) *J. Virol. 69*, 814−824.
18. Braaten, D., Franke, E. K., and Luban, J. (1996) *J. Gen. Virol.*, 3551−3560.
19. Yoo, S., Myszka, D. G., Yeh, C., McMurray, M., Hill, C. P., and Sundquist, W. I. (1997) *J. Mol. Biol. 269*, 780−795.
20. Zhao, Y., Chen, Y., Schutkowsky, M., Fischer, G., and Ke, H. (1997) *Structure 5*, 139−146.
21. Aberham, C., Weber, S., and Phares, W. (1996) *J. Virol. 70*, 3536−3544.
22. Braaten, D., Aberham, C., Franke, E. K., Yin, L., Phares, W., and Luban, J. (1996) *J. Virol. 70*, 5170−5176.
23. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 189*, 319−326.
24. Clore, G. M., and Gronenborn, A. M. (1991) *Science 252*, 1390−1399.

25. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
26. Kay, L. E., Torchia, D. A., and Bax, A. (1989) *Biochemistry 28*, 8972−8979.
27. Kay, L. E., Nicholson, L. K., Delaglio, F., Bax, A., and Torchia, D. A. (1992) *J. Magn. Reson. 97*, 359−375.
28. Grzesiek, S., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 12593−12594.
29. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
30. Johnson, B. A., and Blevins, R. A. (1994) *J. Biomol. NMR 4*, 603−614.
31. Lee, L. K., Rance, M., Chazin, W. J., and Palmer, A. G. I. (1997) *J. Biomol. NMR 9*, 287−298.
32. Tjandra, N., Feller, S. E., Pastor, R. W., and Bax, A. (1995) *J. Am. Chem. Soc. 117*, 12562−12566.
33. Brüschweiler, R., Liao, X., and Wright, P. E. (1995) *Science 268*, 886−889.
34. Tjandra, N., Wingfield, P., Stahl, S., and Bax, A. (1996) *J. Biomol. NMR 8*, 273−284.
35. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559.
36. Blackledge, M., Cordier, F., Dosset, P., and Marion, D. (1998) *J. Am. Chem. Soc. 120*, 4538−4539.
37. Cordier, F., Caffrey, M., Brutscher, B., Cusanovich, M. A., Marion, D., and Blackledge, M. (1998) *J. Mol. Biol. 281*, 341−361.
38. Garcia de la Torre, P., and Bloomfield, V. A. (1981) *Q. Rev. Biophys. 14*, 81−139.
39. Hutchinson, E. G., and Thornton, J. M. (1994) *Protein Sci. 3*, 2207−2216.
40. Grathwohl, C., and Wuthrich, K. (1976) *Biopolymers 15*, 2025−2041.
41. Dorfman, T., and Göttlinger, H. (1996) *J. Virol. 70*, 5751−5757.
42. Bukousky, A. A., Weiman, A., Accola, M. A., and Gottlinger, H. G. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 10943−10948.
43. Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T., and Sundquist, W. I. (1999) *Science 283*, 80−83.
44. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.

BI990991X