4006 *Biochemistry* **2010,** *49,* 4006–4017
DOI: 10.1021/bi902196e

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

NMR Relaxation Studies of an RNA-Binding Segment of the Rous Sarcoma Virus Gag Polyprotein in Free and Bound States: A Model for Autoinhibition of Assembly[†]

Gwen M. Taylor,[‡] Lixin Ma,[‡] Volker M. Vogt,[§] and Carol Beth Post*,[‡]

[‡]*Department of Medicinal Chemistry and Molecular Pharmacology, Markey Center for Structural Biology, and Purdue Cancer Center, Purdue University, West Lafayette, Indiana 47907, and*
[§]*Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853*

Received December 22, 2009; Revised Manuscript Received March 28, 2010

ABSTRACT: Assembly of retrovirus particles is promoted by interaction of the Gag polyprotein with RNA. Nonspecific RNA association with the nucleocapsid domain (NC) of Gag induces the dimerization of Gag through protein−protein contacts in the capsid domain (CA), followed by higher order assembly to form the immature virus particle. NMR relaxation studies were conducted to investigate the initial steps of Rous sarcoma virus (RSV) assembly by examining the association with nucleic acid of a fragment of Gag comprising the C-terminal domain of CA (CTD) postulated to mediate Gag dimerization, the spacer region between CA and NC (SP), and NC. This fragment, CTD-SP-NC (residues 394−577), spans the critical SP region and allows assessment of this key Gag−nucleic acid interaction in the context of the Gag polyprotein rather than the isolated domains. Main-chain amide relaxation of CTD-SP-NC was measured in the absence and presence of $(GT)_4$, an 8-mer DNA oligonucleotide that binds tightly to the polyprotein but is too short to promote Gag dimerization. The results show that the CTD and NC domains tumble independently. In contrast, the two zinc finger domains within NC are rotationally coupled in both the unbound and bound states, even though only the first zinc finger appears to make direct contact with $(GT)_4$. In addition, the NMR data indicate that SP and flanking residues undergo a conformational exchange process that is slowed in the presence of $(GT)_4$. This region around SP where relaxation is strongly affected by $(GT)_4$ binding is nearly identical to the assembly domain defined previously by mutagenesis studies. Other changes in relaxation induced by $(GT)_4$ implicate conformational perturbations of helices 1 and 4 in CTD. On the basis of the combined data, we propose a model for the promotion of Gag dimerization by RNA association in which NC-RNA binding disrupts an assembly inhibitory, intramolecular interaction involving SP and CTD. Disruption of this intramolecular interaction is proposed to enhance the accessibility of the Gag dimer contact surface and release the assembly domain to promote intermolecular oligomerization.

The Gag polyprotein, comprising matrix (MA),[1] capsid (CA), nucleocapsid (NC), and other smaller proteins depending on the viral genus, is the necessary and sufficient protein component for retroviral assembly. In immature virus particles, the Gag polyprotein is arranged radially with the N-terminus of Gag associated with the membrane (*1*, *2*). Assembly of the immature particle involves interaction of the Gag NC domain with viral RNA, accumulation of Gag at the plasma membrane through the association of the N-terminal MA domain, and oligomerization through Gag-Gag interactions in the CA domain and closely adjoining regions to form the protein core of an immature virion. NC-RNA association promotes by an unknown mechanism the oligomerization mediated by CA-CA domain contacts and higher order assembly to form the immature viral core. Late in the budding process Gag is cleaved by the viral protease into the three structural proteins, MA, CA, and NC. This proteolytic processing leads to a change in virus particle morphology and maturation of the virus into its infectious form. During maturation the liberated CA forms a shell that surrounds the NC-RNA complex and the viral enzymes reverse transcriptase and integrase.

The amino acid sequence similarity between the different genera of retroviruses is less than 25−30%; however, the tertiary structures of the three major structural proteins of Gag are highly conserved (*3*−*7*). The isolated domain structures of individual Gag proteins have been determined in both solid and solution states for diverse retroviruses. MA is predominantly α-helical (*8*−*11*). CA is a two-domain α-helical protein consisting of an N-terminal domain (NTD) and a C-terminal (CTD) domain that are connected by a flexible linker (*3*, *4*, *12*−*14*). Isolated NC is a highly flexible protein possessing one or two 14-residue CCHC zinc finger domains, which are the only structured regions of NC (*15*−*19*).

[†]This work was supported by the National Institutes of Health (AI45976 to C.B.P. and CA20081 to V.M.V.), the Markey Center for Structural Biology, and the Purdue University Center for Cancer Research (CA 23568).

*Corresponding author: 765-494-5980 (phone); 765-496-1189 (fax); cbp@purdue.edu (e-mail).

[1]Abbreviations: CA, capsid; CSA, chemical shift anisotropy; CTD, C-terminal domain of capsid; CTD-SP-NC, C-terminal domain of capsid−spacer−nucleocapsid; DTT, DL-dithiothreitol; EM, electron microscopy; HIV-1, human immunodeficiency virus 1; I/I_0, NOE-1; IPTG, isopropyl β-D-thiogalactoside; kDa, kilodalton; MA, matrix; MHz, megahertz; NC, nucleocapsid; NH_4Cl, ammonium chloride; NMR, nuclear magnetic resonance; NTD, N-terminal domain of capsid; NaCl, sodium chloride; PEI, polyethylenimine; RSV, Rous sarcoma virus; R_1, longitudinal relaxation rates; R_2, transverse relaxation rates; SP, RSV spacer; SP-1, HIV-1 spacer; S^2, generalized order parameter; τ_m, correlation time; VLP, virion-like particle; XNOE, $\{^1H\}-^{15}N$ steady-state heteronuclear NOE; ZF1, N-terminal zinc finger; ZF2, C-terminal zinc finger; $ZnCl_2$, zinc chloride.

 Published on Web 04/13/2010

Besides HIV-1 (genus lentiviruses), one of the most important model systems to study retrovirus assembly is Rous sarcoma virus (RSV, genus alpharetroviruses or avian sarcoma and leukemia viruses). *In vitro* assembly studies have established that association of Gag with nucleic acid creates an assembly "building block" to initiate RSV assembly. Gag interaction with either RNA or DNA for *in vitro* assembly promotes dimerization of two Gag molecules through protein−protein contacts postulated to occur in the CTD (*20*). Artificially induced dimerization of Gag molecules lacking an NC domain also leads to assembly and budding in cells (*21*). Oligomerization and assembly to regular higher order structure require as well the presence of the NTD, and it is suggested that NTD-NTD contacts lead to Gag assembly into virion-like particles (VLP) while CTD-CTD contacts stabilize the transient dimer building block (*22*−*24*). Due to structural similarities between retroviral CA and SCAN domains, it has been suggested that CA dimerization may involve domain swapping, similar to that occurring between two SCAN domains (*25*). Clearly, dimerization of CA is not a simple process. While RSV CTD and CA are monomeric in solution even at millimolar concentrations, HIV-1 CA dimerizes at modest concentrations. Further, RSV CTD dimers observed from crystals obtained at different conditions of pH form five unique interfaces (*26*), one of which is similar to that observed for HIV-1 CTD crystals (*6*) and two closely resemble the interface of HIV-1 CTD in crystals of the complex with a peptide that inhibits assembly (*27*).

In many retroviruses, including HIV-1 and RSV, a short "spacer" sequence separates the CA and NC domains of Gag and is proteolytically removed during virus maturation. This sequence, called SP in RSV (12 residues) and SP-1 or p2 in HIV-1 (14 residues), has been shown by mutagenesis to be absolutely required for proper immature assembly. RSV SP also contains an internal protease cleavage site (*28*). In HIV-1 (*29*, *30*) the short stretch of 11 or 12 residues between the last helix of the CTD and SP-1 is also critical for the correct formation of immature virus particles, while in the case of RSV, an "assembly domain" consisting of the last residues of CA, all of SP, and the first four residues of NC is critical (*31*). In structures determined for CA or CTD from either HIV-1 (*6*) or RSV (*3*, *4*), these residues are disordered and without stable secondary or tertiary structure. The mechanistic role of spacer in assembly is unknown, although part or all of this sequence in HIV-1 is predicted to be helical, which implicates a possible helix−coil transition (*32*−*34*).

In electron cryotomography studies of immature HIV-1 particles, density was observed below the hole made by the CA NTD hexamer corresponding to SP and appears to be holding the Gag hexamer together (*35*, *36*). The density was fit with a six-helix bundle, consistent with previous data suggesting that SP-1 forms an α-helix (*29*−*31*). Briggs et al. point out that this is a working model and at the current resolution other arrangements of SP1 can be fit into the density (*35*).

We report the results from NMR studies to investigate the initial steps of assembly by examining the association of RSV Gag with nucleic acid. ^{15}N relaxation studies were conducted to characterize the dynamics of an RSV Gag fragment in the absence and presence of $(GT)_4$, an 8-mer DNA oligonucleotide of nonspecific sequence that binds tightly to the polyprotein but is too short to promote Gag dimerization (*22*, *37*, *38*). The aim was to examine properties in the context of the Gag polyprotein, rather than the isolated domains, and thus a Gag fragment that spans the critical SP region and comprises CTD (394−476), SP (477−488), and NC (489−577) (CTD-SP-NC) (Figure 1) was used. An analogous Gag fragment from HIV-1 has been characterized in the free state by NMR relaxation (*39*). The overall trends observed for unbound RSV CTD-SP-NC relaxation are generally consistent with the conclusions for HIV-1, although certain differences exist in the details of the motional properties, particularly those of SP. New to the work reported here are the results that characterize the effect of binding $(GT)_4$ on CTD-SP-NC dynamics. These data, along with the identification of an assembly domain by mutagenesis (*21*, *31*), and the discovery of inhibition of HIV-1 assembly by a peptide that binds to the CTD (*27*, *40*) suggest a model for how nucleic acid binding to Gag initiates Gag dimerization and in turn leads to assembly.

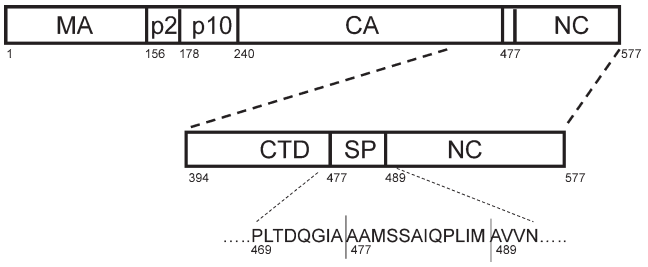


FIGURE 1: Schematic representation of the RSV Gag polyprotein and the CTD-SP-NC polyprotein construct used in this study. CTD is from residue 394 to 476, SP from 477 to 488, and NC from 489 to 577. NC has two zinc fingers, ZF1 (509−522) and ZF2 (535−548) that together makeup the NC core (509−548). The horizontal lines in SP represent the PR cleavage sites between CTD and NC at 476 and 488. A third cleavage site also exists within SP at 479. The amino acid sequence comprising the "assembly domain" from residue 469 at the C-terminus of CTD to residue 492 at the N-terminus of NC is shown.

MATERIALS AND METHODS

Expression and Purification of Recombinant RSV CTD-SP-NC Polyprotein. Protein expression and purification of RSV CTD-SP-NC were carried out as described by Ma and Vogt (*22*) with some modifications. For protein expression in *Escherichia coli*, the RSV pET-CTD-NC plasmid was transformed into BL21(DE3) pLysS cells. Uniformly ^{15}N- and ^{15}N/^{13}C-labeled proteins were prepared by growing the cells in M9 minimal media with $^{15}NH_4Cl$ (1 g/L) and [$^{12}C_6$]-D-glucose (2 g/L) or $^{15}NH_4Cl$ (1 g/L) and [$^{13}C_6$]-D-glucose (2 g/L), respectively, supplemented with 100 μg/L ampicillin and 34 μg/L chloramphenicol. A starter culture of 25 mL of M9 inoculated from a single colony was grown at 37 °C overnight. The starter culture was added to 1 L of M9 supplemented with 0.1 mM $CaCl_2$ and grown at 37 °C to an OD_{600} of 0.6 before induction with 1 mM IPTG (isopropyl β-D-thiogalactopyransoide). Induction was carried out for 4 h before the cells were harvested by centrifugation (6000*g*) at 4 °C and stored at −20 °C. The cells were thawed on ice, suspended in lysis buffer (20 mM Tris, pH 7.5, 100 mM NaCl, 10% glycerol, 0.1% NP-40, 50 mM DTT, 20 μM $ZnCl_2$, and protease inhibitor cocktail for His tag proteins (Sigma-Aldrich)), and sonicated on ice by three bursts of 2 min. Cell debris was pelleted by centrifugation at 45K for 3 h in a Type 60 Ti rotor (Beckman). The remaining nucleic acid was removed by polyethylenimine (PEI) precipitation using a final concentration of 0.3% (w/v) PEI. The protein was then precipitated with a final concentration of 40% (w/v) ammonium sulfate and dialyzed against low salt buffer (20 mM Tris, pH 7.5, 100 mM NaCl,

50 mM DTT, 20 μM $ZnCl_2$, and protease inhibitor cocktail) overnight. After dialysis, the protein was applied to a pre-equilibrated phosphocellulose column and eluted from the column with high salt buffer (20 mM Tris, pH 7.5, 500 mM NaCl, 50 mM DTT, 20 μM $ZnCl_2$). Purified protein was concentrated to approximately 1 mM and dialyzed against NMR buffer (10 mM sodium phosphate, 150 mM NaCl, 20 μM $ZnCl_2$, 1 mM DTT). Prior to NMR data collection, 10% (v/v) ^2H_2O was added to the samples.

NMR Spectroscopy. All NMR spectra were acquired at 25 °C on a Varian Inova 600 spectrometer equipped with 5 mm [^{1}H,^{15}N,^{13}C] triple-resonance *z* axis pulsed-field gradient probes. Backbone assignments were made by using standard triple-resonance experiments 3D HNCACB, CBCA(CO)HN, C(CO)-NH, and HC(CO)HN recorded on uniformly ^{15}N/^{13}C-labeled RSV CTD-SP-NC. NMR data were processed with NMRPIPE (*41*) and analyzed with SPARKY (T. D. Goddard and D. G. Kneller, University of California, San Francisco). Longitudinal relaxation rates (R_1), transverse relaxation rates (R_2), and $\{^1H\}-^{15}N$ steady-state heteronuclear NOE (XNOE) for backbone ^{15}N nuclei were measured with 2D inverse-detected water flip-back pulse sequences using conventional procedures (*42*, *43*) on uniformly ^{15}N-labeled RSV CTD-SP-NC. The R_1 data were collected with 10, 20, 40, 80, 160, 320, 640, 1200, and 2000 ms relaxation recovery delays, and the R_2 data were collected with 10, 30, 50, 70, 90, 110, 130, and 190 ms relaxation recovery delays. A recovery delay of 1.5 s was applied between scans. The $\{^1H\}-^{15}N$ NOEs were measured from ^{15}N-HSQC spectra recorded with and without ^{1}H presaturation. A window function with Gaussion line broadening of −15 Hz was applied in the ^{1}H dimension to resolve resonances in the crowded core region of the 2D $^1H-^{15}N$ spectra.

Heteronuclear Relaxation Rates. ^{15}N relaxation data were measured at 600 MHz field strength and in duplicate. The intensities of the signals of the cross-peaks in the 2D $^1H-^{15}N$ spectra were measured with SPARKY (T. D. Goddard and D. G. Kneller, University of California, San Francisco). The R_1 and R_2 relaxation rates of the ^{15}N nuclei were estimated by a least-squares fit of a two-parameter exponential function. Only well-resolved peaks with well-fit decay curves were selected for further analysis. Uncertainties of the R_1 and R_2 rates were estimated from the root-mean-square noise level in the spectra. The $\{^1H\}-^{15}N$ steady-state heteronuclear NOE interactions are reported using the relationship NOE-1 = I/I_0, where I and I_0 are signal intensities corresponding to spectra with and without ^{1}H saturation, respectively. Uncertainties in the $\{^1H\}-^{15}N$ steady-state heteronuclear NOE values were estimated from individual data sets collected in duplicate.

Heteronuclear Relaxation Analysis. R_1 and R_2 relaxation rates and the steady-state nuclear Overhauser enhancement of the ^{15}N nucleus from dipolar interaction with the bonded proton and from ^{15}N chemical shift anisotropy are given by (*44*)

$$\begin{aligned} R_1 &= \frac{d^2}{4}[J(\omega_H-\omega_N)+3J(\omega_N)+6J(\omega_H+\omega_N)]+c^2J(\omega_N) \\ R_2 &= \frac{d^2}{8}[4J(0)+J(\omega_H-\omega_N)+3J(\omega_N)+6J(\omega_H) \\ &\quad +6J(\omega_H+\omega_N)]+\frac{c^2}{6}[4J(0)+3J(\omega_N)] \\ \frac{I}{I_0} &= 1+\frac{d^2}{4R_1}\left(\frac{\gamma_H}{\gamma_N}\right)[6J(\omega_H+\omega_N)-J(\omega_H-\omega_N)] \end{aligned} \quad (1)$$

where the dipolar constant, d, and chemical shift anisotropy (CSA) interaction constant, c, are

$$\begin{aligned} d &= \left(\frac{\mu_0}{4\pi}\right)\hbar\gamma_H\gamma_N\langle r_{NH}{}^{-3}\rangle \\ c &= \frac{\omega_N}{\sqrt{3}}(\sigma_\parallel-\sigma_\perp) \end{aligned} \quad (2)$$

μ_0 is the permeability of free space, γ_H and γ_N are the gyromagnetic ratio of ^{1}H and ^{15}N, respectively, $\hbar$ is Planck's constant divided by 2π, γ_{NH} is the average N−H bond length, 1.02 Å, and $\sigma_\parallel - \sigma_\perp$ is the difference between parallel and perpendicular components of the axially symmetric chemical shift tensor for amide ^{15}N and equal to −160 ppm. In the event that a conformational exchange process occurs on a time scale longer than the overall rotational correlation time, an additional term, R_{ex} (not shown in eq 1), contributes to the measured R_2.

Motional Models. General relaxation behavior and the trends associated with $(GT)_4$ were considered using four models for molecular motion: a single correlation time where overall rotation is either (1) isotropic or (2) anisotropic with axial symmetry; (3) Lipari−Szabo model-free analysis of two independent motions, one for overall rotation and one for internal motion on a much faster time scale (> 50×) than overall rotation; and (4) was the same as (3) except that the internal motion is only somewhat faster than overall rotation (10−50×).

(*A*) *Motional Model 1*. In the simplest analysis, relaxation is accounted for by one rotational correlation function with correlation time τ_m. For a globular domain, this model corresponds to a spherical, rigid body that tumbles in solution with isotropic rotation and no significant internal motion. For nonglobular portions of the polyprotein, the model applies to flexible regions with dynamics that can be fit by a single time dependence on a relatively fast time scale, less than 1−2 ns. The spectral density for isotropic rotation

$$J(\omega) = \frac{2}{5}\left[\frac{\tau_m}{1+\omega^2\tau_m{}^2}\right] \quad (3)$$

The overall rotational correlation time, τ_m, is related to the isotropic rotational diffusion rate, D^{iso}, by $\tau_m = 1/6D^{iso}$.

(*B*) *Motional Model 2*. A rigid-body, axially symmetric, folded domain tumbles anisotropically, as described by two diffusion constants. For an axially symmetric diffusion tensor, $D_{zz} = D_\parallel$ and $D_{xx} = D_{yy} = D_\perp$, where D_{xx}, D_{yy}, and D_{zz} are the principal components of the molecular rotational diffusion tensor. The spectral density comprises three components (*45*, *46*):

$$J(\omega) = \frac{2}{5}\sum_{i=1}^{3}\frac{A_iD_i}{(D_i)^2+\omega^2} \quad (4)$$

where

$$\begin{aligned} D_1 &= (5D_\perp+D_\parallel) \\ D_2 &= (2D_\perp+4D_\parallel) \\ D_3 &= 6D_\perp \end{aligned} \quad (5)$$

and the coefficients

$$\begin{aligned} A_1 &= 3z_d{}^2(1-z_d{}^2) \\ A_2 &= \frac{3}{4}(1-z_d{}^2)^2 \\ A_3 &= \frac{1}{4}(3z_d{}^2-1)^2 \end{aligned} \quad (6)$$

with z_d being the z-component of the NH unit vector in the principal axis frame of the diffusion tensor. The relative correlation time is maximal or minimal for an N−H vector aligned along xy or z, respectively.

(*C*) *Motional Model 3*. Internal mobility is taken into account using the model-free approach (*47*). Internal motion alters the spectral density depending on the amplitude and the time scale of the motion. For very fast (relative to τ_m) internal motion, the spectral density is scaled by the generalized order parameter, S^2, with no time dependence. We examine the effect of internal motion for the case of isotropic overall rotation. (Fast internal motion with anisotropic overall rotation is also a reasonable model, but the combined effects of internal motion and rotational anisotropy are not required for the analysis here.) For fast internal motion and isotropic rotation, $J(\omega)$ becomes

$$J(\omega) = \frac{2}{5}S^2\left[\frac{\tau_m}{1+\omega^2{\tau_m}^2}\right] \quad (7)$$

(*D*) *Motional Model 4*. For the case of longer time scale internal motion, an additional time-dependent term is required (*47*):

$$J(\omega) = \frac{2}{5}S^2\left[\frac{\tau_m}{1+\omega^2{\tau_m}^2}\right] + \frac{2}{5}(1-S^2)\left(\frac{\tau_{eff}}{1+\omega^2{\tau_{eff}}^2}\right) \quad (8)$$

where ${\tau_{eff}}^{-1} = {\tau_m}^{-1} + {\tau_{int}}^{-1}$ and τ_{int} is the correlation time of the internal motion. The Lipari−Szabo formalism requires that the internal and overall tumbling motions are uncoupled. The condition that τ_{int} is much shorter than τ_m satisfies this requirement and is needed to accurately determine τ_{int} from experiment (*48*). Given $\tau_m \sim 1-10$ ns, eq 7 applies for $\tau_{int} < 10$ ps, while eq 8 holds for $\tau_{int} \sim 10-100$ ps.

(*E*) *Application*. Model 1 or 2 is applicable to rigid core regions of globular domains with spherical or axially symmetric shape, respectively, and correlation times typically longer than 4−5 ns. Models 3 and 4 correspond to regions within a globular domain that exhibit internal mobility on time scales faster than the overall rotation of the domain. Flexible linkers that behave as random-flight chains, without folded structure, are treated using model 1 with correlation times less than 1−2 ns. More elaborate analyses of domain motions are the "extended" model-free approach (*49*) and wobble-in-a-cone model (*50*, *51*), which in principle could be applied to CTD-SP-NC. Nevertheless, the differences in predicted relaxation rates compared to the extended model-free approach are small considering the descriptive analysis sought in this study. Analysis using the wobble-in-a-cone model demands additional relaxation data and is an area for future studies. An overall rotational correlation time (τ_m) was estimated from the ^{15}N relaxation data for structured domains in CTD-SP-NC by the numerical procedure in TENSOR (*52*). Motional parameters for linker regions were estimated graphically based on agreement of the ^{15}N relaxation data with theoretical plots.

Molecular Weight Dependence of τ_m. The expected value of τ_m for a folded domain of a given molecular mass polypeptide is taken from the data reported by Garcia de la Torre (*53*). Correlation times were determined from experimental measurements of rotational diffusion and were predicted by hydrodynamic calculations for proteins ranging in molecular mass from 3 to 27 kDa, and the values are plotted in Figure 2. The data were fit by a least-squares linear regression to give the dependence of τ_m on molecular weight: $\tau_m = 0.68$ ns/kDa.

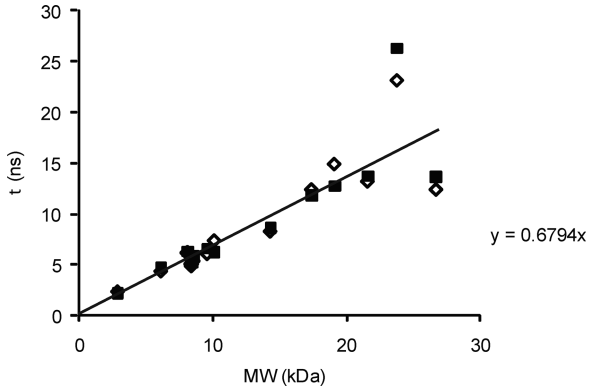


FIGURE 2: Dependence of molecular correlation time on molecular mass determined empirically. Values taken from ref *53* for estimates from analysis of experimental data (solid squares) and calculated from structure (open diamonds). Linear regression: τ (ns) = 0.68 MW (kDa).

RESULTS AND DISCUSSION

Heteronuclear Relaxation of Unbound CTD-SP-NC. ^{15}N longitudinal (R_1) and transverse (R_2) relaxation rates and steady-state nuclear Overhauser effects (NOE+1 = I/I_0) were measured in duplicate at a magnetic field strength of 600 MHz (11.74 T) for resolved peaks of CTD-SP-NC in the absence (Figure 3) and presence of $(GT)_4$ (see below). In both the free and $(GT)_4$ complex of the polyprotein, data points are absent for residues lacking backbone assignments, which include residues in the inter zinc finger linker (527, 528, 531, 532) and residues at the carboxy terminus of NC (552−559, and 564−574). Values are also missing because of peak overlap (residues 414, 462, 430, 490, 482, 509) or peak broadening upon addition of $(GT)_4$ (residues 478−496).

As expected for a polyprotein, CTD-SP-NC exhibits complex relaxation behavior (Figure 3) consistent with multiple domains of differing degrees of flexibility, rather than a single globular unit. The rates plotted by residue for unbound CTD-SP-NC in Figure 3 show large variations in R_1, R_2, and I/I_0 values. In particular, SP and the flanking regions of the C-terminus of CTD and N-terminus of NC (residues 468−507) have I/I_0 values near 0.0 while CTD and the NC core, comprising ZF1 and ZF2, have near limiting values of 0.8 for I/I_0. The overall trends in dynamics define three regions in RSV CTD-SP-NC: the CTD core region (residues 394−467), the NC core region (residues 509−548), and finally SP and its flanking regions (residues 468−507). Average values for these regions, grouped based on their dynamics, are shown in Table 1.

Relaxation Rates Predicted for Alternative Motional Models. The objective of the study reported here is to characterize the interaction of CTD-SP-NC with nucleic acid by examining hydrodynamic behavior of various segments of RSV Gag polyprotein. We therefore focus on dynamic properties averaged over these three polypeptide regions, rather than a detailed interpretation of residue-level mobility. Correlation times for a domain are estimated for the clearly globular CTD and NC core regions by numerically fitting a subset of R_1, R_2, and I/I_0 values from residues within these regions to dynamic models with anisotropic rotation and internal motion described in Materials and Methods. For SP and flanking regions, it cannot be assumed a priori that the dynamics of this linker region are similar to that of a flexible loop in a globular protein with two characteristic time constants for overall slow tumbling and a

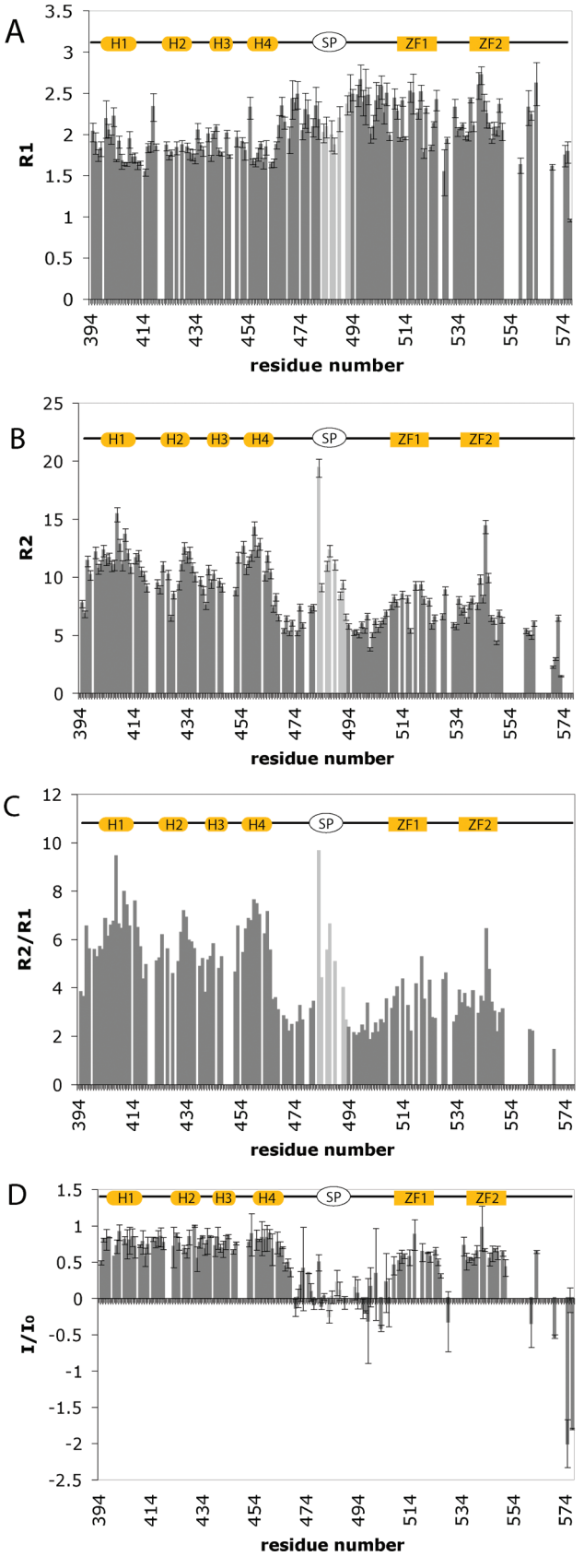


FIGURE 3: ^{15}N relaxation data at 600 MHz for RSV CTD-SP-NC in the unbound state. (A) Longitudinal relaxation rates, R_1 ($1/T_1$). (B) Transverse relaxation rates R_2 ($1/T_2$). (C) R_2/R_1 ratio. (D) I/I_0 values. α1 = 402−416; α2 = 420−436; α3 = 437−447; α4 = 453−465. Helices in CTD are indicated at the top with cylinders, SP by an oval, and ZF1 and ZF2 by rectangles. Residues 483−493 (light gray) at the boundary of SP and the N-terminus of NC have a peak-like R_2 profile with a maximum at I487.

faster internal motion. Alternatively, SP motion may be appropriately modeled as an independent linker by using a single, fast time scale correlation function. Dynamics of this apparently unstructured, highly flexible Gag region is therefore interpreted with theoretical curves for R_1, R_2, and I/I_0 values calculated for the four motional models to obtain a region-level characterization. Overall, this assessment of CTD-SP-NC yields conclusions in terms of an effective hydrodynamic molecular weight and general mobility of regions composing CTD-SP-NC. Of particular interest are the changes in dynamics upon $(GT)_4$ binding assessed in this manner for giving insights into the assembly mechanism. Spectral density mapping (*54*), a structure-independent approach, was also conducted, and those results are provided in Supporting Information.

We recall the theoretical relaxation behavior for four motional models (see Materials and Methods for details) to guide the interpretation of the relaxation exhibited by each of the three regions of CTD-SP-NC. Theoretical curves to illustrate the dependence of R_2/R_1 and I/I_0 on anisotropic rotation for nonspherical shapes and picosecond to nanosecond internal motion are compared in Figure 4 to curves for the reference rigid-body spherical rotation. The curves were calculated from spectral densities corresponding to the four motional models described in Materials and Methods.

Relaxation rates for isotropic (model 1) and anisotropic, axially symmetric (model 2) rigid rotation are compared in Figure 4A,B, where R_2/R_1 and I/I_0, respectively, are plotted as a function of overall correlation time $\tau_m = 1/6D^{iso}$ for isotropic (solid) and $\tau_m = 1/[2(D_{xx} + D_{yy} = D_{zz})]$ for axially symmetric (dashed and dotted) diffusion. The two curves for axially symmetric rotation are for an NH bond vector aligned parallel (dashed) or perpendicular (dotted) to the z-axial component of the diffusion tensor and delimit the range of possible R_2/R_1 values. The anisotropy for model 2 is $D_\parallel/D_\perp = 1.75$, which is an upper limit based on previously reported values for retroviral CA and NC proteins. Studies report $D_\parallel/D_\perp$ values of 1.4 for CTD in the context of CA (*4*) and 1.2−1.6 for individual zinc finger domains in the context of NC (*16*, *55*). With anisotropic (axially symmetric) rotation, R_2/R_1 is observed in Figure 4A to depend strongly on the NH vector orientation relative to the diffusion axes when $\tau_m > 3$ ns, while the effect on I/I_0 due to anisotropic rotation (Figure 4B) is insignificant in this slow tumbling regime. For small values of τ_m where the slope is steep, I/I_0 does depend on the NH vector orientation so that somewhat larger I/I_0 values result for vectors oriented parallel to the z diffusion-tensor component while perpendicular vectors exhibit smaller I/I_0 values.

Internal motion within a globular domain (models 3 and 4) is evaluated using the model-free formalism of Lipari and Szabo (*47*). Theoretical curves calculated for very fast internal motion, with no dependence on the time scale of the internal motion (model 3, eq 7, squares), and for a time-dependent fast internal motion (model 4, eq 8, dotted or dashed) are compared to those for rigid-body rotation (model 1, solid curves) in Figure 4C,D. The parameters for model 4 are $\tau_{int} = 0.1$ ns and $S^2 = 0.85$ (dashed), 0.6 (dot-dash), or 0.5 (dotted). The value $\tau_{int} = 0.1$ ns is the slowest value reasonably consistent with the assumption in Lipari−Szabo theory of fast internal motion with respect to overall rotation on a time scale of approximately 1−10 ns, and $S^2 = 0.5$ indicates considerable reorientation of the NH bond vector within the molecular framework, and thus together, these parameters reasonably illustrate limiting effects that arise from reorientation of the amide NH bond due to internal motion.

Table 1: Average ^{15}N Backbone Relaxation Rates and I/I_0 (=NOE-1) for Regions of RSV CTD-SP-NC in the Unbound State[a]

	residues	$\bar{R}_1$ (s^{-1})	$\bar{R}_2$ (s^{-1})	$\bar{R}_2/\bar{R}_1$	$\bar{I}/\bar{I}_0$	τ_m (ns)[b]
CTD	397−467	1.8 ± 0.2	10.6 ± 2.1	5.9 ± 1.4	0.8 ± 0.1	6.7
SP + flanking	468−507					
C-term CTD	468−477[c]	2.2 ± 0.3	6.4 ± 0.6	2.8 ± 1.5	0.2 ± 0.2	ND
SP[b]	479−493[c,d]	2.2 ± 0.1	10.2 ± 4.2	4.9 ± 2.2	0.06 ± 0.2	ND
N-term NC	495−507[c]	2.4 ± 0.2	5.5 ± 1.6	2.3 ± 0.8	−0.02 ± 0.2	ND
NC core	509−548	2.2 ± 0.3	7.4 ± 1.7	3.4 ± 0.9	0.6 ± 0.01	5.1
ZF1	509−522	2.3 ± 0.3	8.0 ± 1.1	3.7 ± 0.9	0.6 ± 0.1	4.9
ZF2	535−548	2.2 ± 0.2	7.7 ± 2.2	3.5 ± 0.9	0.6 ± 0.1	5.1

[a]The standard deviation of the average value is listed. [b]Calculated using TENSOR by an overall fit of relaxation data from residues listed in column 2. [c]Regions defined by common relaxation behavior. [d]Region colored in gray in Figure 3. SP* residue range is based on relaxation behavior and differs somewhat from SP residues 477−488.

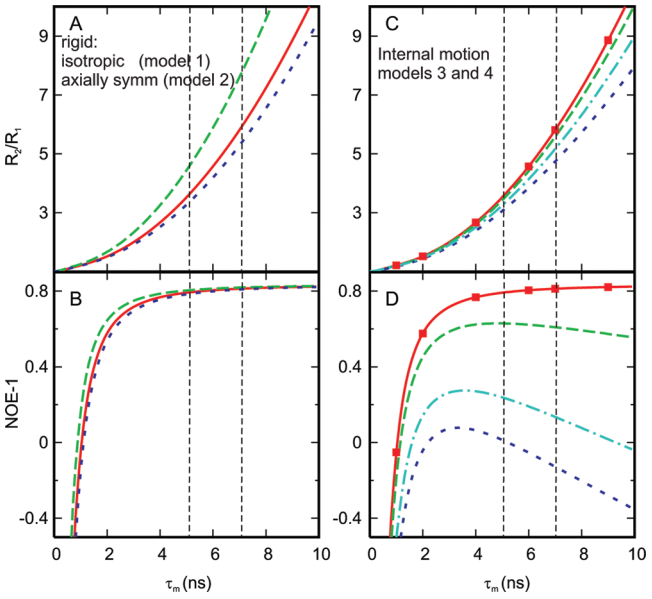


FIGURE 4: Theoretical R_2/R_1 and I/I_0 at 600 MHz ^{1}H frequency as a function of overall correlation time, τ_m, for four motional models. Model 1: Isotropic rotation with a single overall rotational correlation time (solid, red curve) is shown in all panels for reference. (A, B) Axially symmetric rotation with $D_{zz} = D_{\parallel} = 1.3$ and $D_{xx} = D_{yy} = D_{\perp} = 0.85$, and $\tau_m = 1/[2(D_{xx} = D_{yy} = D_{zz})]$. Two curves delimit the relaxation for N−H bond vectors aligned parallel (dashed, green) or perpendicular (dotted, blue) to the z-axial component of the diffusion tensor. (C, D) Overall isotropic rotation with a fast internal motion. Model 3: Very fast internal motion squares, $S^2 = 0.5$ (filled squares, red). Model 4: Internal motion with $\tau_{int} = 0.1$ ns and $S^2 = 0.85$ (dashed, green), 0.6 (dash-dot, cyan), and 0.5 (dotted, blue). The vertical lines denote τ_m values estimated for CTD and NC core regions in unbound CTD-SP-NC.

As previously recognized (*56*), R_2/R_1 (Figure 4C) is insensitive to fast internal motion (model 3, squares), but when the internal motion time scale approaches overall tumbling, R_2/R_1 is somewhat reduced at longer τ_m (model 4, dashed, dot-dash, or dotted). (The reader is reminded that the theoretical plots for R_2/R_1 in Figure 4 do not include R_{ex}, which accounts for slower exchange processes and that microsecond to millisecond motions would contribute by addition of an R_{ex} component to an observed R_2.) As with R_2/R_1, I/I_0 (Figure 4D) is quite insensitive to very fast internal motion ($\tau_{int} < 0.1\tau_m$) of model 3: no change is observed even though $S^2 \ll 1$. On the other hand, I/I_0 is diminished significantly if internal motion occurs on a time scale such that τ_{int} approaches the overall correlation time. This time-dependent behavior depends on the amplitude S^2 and internal correlation time τ_{int} (Figure 4D, dashed, dash-dot, or dotted); the diminution in I/I_0 increases with slower, larger amplitude internal motion, i.e., as τ_{int} increases and S^2 decreases. It is noted that the curves in Figure 4D show that the internal motion must result in extensive reorientation ($S^2 < 0.6$, $\tau_{int} \sim 0.1$ ns) in order to reduce I/I_0 to values below 0.3 for τ_m between 5 and 10 ns.

Dynamics of the CTD Core in Unbound CTD-SP-NC. The heteronuclear relaxation data of CTD core residues in CTD-SP-NC were analyzed using the previously determined structure of RSV CTD (PDB code 1EOQ) (*3*) and axially symmetric diffusion, model 2. The chemical shifts of CTD in the Gag construct are similar to those in the isolated CTD domain, excluding the C-terminal residues, indicating that CTD remains close to the 3D structure determined for the isolated domain. Axially symmetric diffusion was also found previously to be the best fit to relaxation rates of CTD in HIV-1 (*4*). Values of R_2 and R_1 for residues 394−467 of CTD are well fit by the program TENSOR with $\tau_m = 6.7$ ns (Table 1). (Residues in the C-terminus of CTD exhibit different dynamics and are discussed below.) The large range of R_2/R_1 ~4−7 observed for CTD core residues 394−464 (Figure 3C) is consistent with the theoretical curves for axially symmetric rotation of a well-ordered domain given $\tau_m \approx 7$ ns (Figure 4A). The average I/I_0 equals 0.75 with small variation among these residues. This value is only slightly less than the limiting value of 0.83, indicating that CTD core residues have only limited internal mobility. Given the empirical plot for the dependence of τ_m on molecular weight in Figure 2 (*53*), the value $\tau_m \approx 7$ ns corresponds to a 90-residue protein, only slightly larger in size than CTD alone. A value for τ_m previously determined from RSV CTD relaxation was in the context of full-length CA and is a larger value, ~12 ns (*4*), which implies that rotational diffusion of the two domains NTD and CTD is coupled. By contrast, in the context of CTD-SP-NC, CTD rotates faster, more like an isolated domain, and is marginally affected by the presence of NC. Thus CTD is a well-ordered globular domain with axially symmetric rotational diffusion (model 2) nearly independent of SP and NC.

Dynamics of the NC Core in Unbound CTD-SP-NC. The NC core, comprising ZF1 and ZF2, of RSV is folded and likely nonspherical, but because the structure is unknown, the relaxation rates are analyzed based only on an isotropic rotation (model 1) to estimate an overall correlation time for each zinc finger separately and together with the connecting segment. The purpose of this analysis is to determine the degree of rotational coupling of the two zinc fingers by comparing the diffusional behavior of the two ZF domains analyzed independently and combined.

The τ_m value fitted using TENSOR from R_1 and R_2 for residues in ZF1 (509−522), ZF2 (535−548), or the longer NC

core region (509−548) is 4.9, 4.9, and 5.1 ns, respectively. The globally fit value for τ_m of 5.1 ns is confirmed with the theoretical plot (Figure 4A); the average R_2/R_1 values for ZF1, ZF2, and NC core equal to 3.7, 3.5, and 3.4, respectively, correspond to $\tau_m \approx$ 5 ns. The similarity in the correlation time for both ZF1 and ZF2, as well as the combined NC core, indicates that this central region of NC rotates as a unit and that the two ZFs are coupled. Further, the large τ_m of 4.9 ns for the NC core corresponds to a 65-residue protein based on the hydrodynamic data of de la Torre (*53*) (Figure 2). NC is 89 residues, and the NC core region spans approximately 40 amino acids, with each ZF being 14 residues. Thus the magnitude of the correlation time also supports the notion that ZF1 and ZF2 rotate more as a single unit than two independent units and are hydrodynamically coupled. In the case of HIV-1 NC, $\tau_m \sim 6$ and 5.7 ns are reported for ZF1 and ZF2, respectively, and it was concluded that the two ZF's are rotationally independent based on consistently lower values of R_1 for ZF2 with respect to ZF1 (*55*). We do not find this type of systematic variation in the data for RSV Gag.

The average for I/I_0 over the NC core residues is approximately 0.6, somewhat smaller than the limiting high molecular weight value of 0.83 expected for $\tau_m = 5$ ns. Neither anisotropic rotation nor very fast internal motion ($\tau_{int} \ll \tau_m$, model 3) reduces I/I_0; therefore, we conclude the NC core experiences slower time scale internal mobility with more extensive motional averaging than CTD.

Dynamics of SP and Flanking Regions in Unbound CTD-SP-NC. The relaxation rates of the SP linker and flanking regions from the C-terminus of CTD and the N-terminus of NC (residues 468−507) indicate local conformational flexibility for this region of CTD-SP-NC. High mobility can result from a fully flexible linker described by a single, fast correlation time or by a more complex motion with a longer overall domain-like rotation time plus an internal motion on a faster time scale. We assess the local dynamics of this region based on general relaxation behavior and consistent agreement for I/I_0, R_1, and R_2 between predicted and experimental values.

A notable feature of the SP and flanking region (Figure 3D) is the near zero I/I_0 values, whereas I/I_0 values exhibited by residues in CTD and the NC core are close to the maximal limit expected for globular proteins. Near zero values for I/I_0 are often observed for terminal residues of a folded protein but are less common for internal residues. This small value can be accounted for by either model 1, with a single rotational correlation time near $\tau_m \sim 1$ ns (Figure 4B, solid) representative of random and flexible chains, or model 4, with a longer overall τ_m and a time-dependent internal motion. To consider viable estimates of the motional parameters of model 4, we use for τ_m the range 5−7 ns for the NC core and CTD (Table 1) and find that $S^2 = 0.5-0.6$ and $\tau_{int} \sim 0.1$ ns (Figure 4D, dashed and dot-dash) are consistent with $I/I_0 \sim 0$. Other combinations of S^2 and τ_{int} are of course also possible. Anisotropic rotation and very fast internal motion have negligible effect on I/I_0 (Figure 4B,D, squares), and thus models 2 and 3 cannot account for the near zero value for I/I_0.

A second notable feature in the relaxation behavior of unbound CTD-SP-NC is the peak-like profile in R_2 with a maximum at I487 (Figure 3B, gray bars) for residues 483−493 at the boundary of SP and NC. This region also has larger R_2/R_1 than the immediate upstream or downstream residues, which suggests that exchange (R_{ex}) contributes to R_2. A maximum in the residue profile for R_2 has been interpreted to correspond to the site at which a conformational transition takes place (*57*). Here, the maximum in the R_2 profile occurs at residue I487. RSV has a

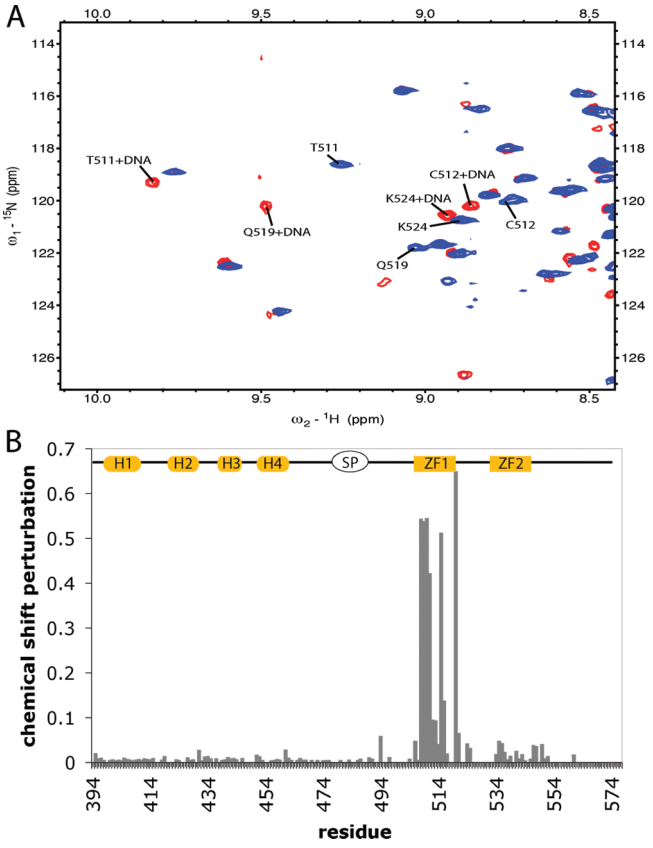


FIGURE 5: Chemical shift perturbation of RSV CTD-SP-NC upon $(GT)_4$ binding. (A) Representative region of ^{15}N-HSQC of CTD-SP-NC bound (red) and unbound (blue). (B) Plot of chemical shift perturbation of RSV CTD-SP-NC upon binding to $(GT)_4$ oligonucleotide. CSP = $(0.5[\Delta\delta_H^2 + (\Delta\delta_N/5)^2])^{1/2}$. Structural elements drawn at the top are as in Figure 3.

conserved proline residue at position 485 (Figure 1) near the maximum in the profile, so that exchange could involve a cis−trans isomerization of P485 in SP. This conformational exchange process is affected by $(GT)_4$ binding, as discussed below.

To decipher the motional properties in general terms of the SP and flanking regions of Gag polyprotein, we consider the R_1 and R_2 profiles (Figure 3) and the fact that near zero I/I_0 values are consistent with either a single, shorter correlation time (model 1, $\tau_m \sim 1$ ns) for a fully flexible linker or a longer overall τ_m plus a time-dependent internal motion (model 4). Of these, model 4 is most consistent based on the R_2/R_1 values observed for SP and the flanking regions; the R_2/R_1 values are too large to be consistent with fast tumbling $\tau_m \sim 1$ ns (Figure 4C, Table 1). The individual R_2 and R_1 values also rule out model 1 as detailed in Supporting Information. Further, the presence of a conformational exchange process cannot be readily reconciled with the behavior of a fully flexible linker associated with model 1.

We conclude that relaxation of SP and the flanking regions (residues 468−507) is most consistent with a motional model described by two correlation times (model 4). For a limiting case of the overall rotational diffusion similar to either CTD or NC, the slow internal motion is approximated by $S^2 = 0.5-0.6$ and $\tau_{int} \sim$ 0.1. Importantly, a conformational exchange process is needed to account for the residue variation and peak maximum in R_2.

CTD-SP-NC Association with $(GT)_4$. Perturbations in chemical shift (CSP) of CTD-SP-NC in the presence of $(GT)_4$ were used to identify the likely site of interaction with $(GT)_4$. Chemical shifts were measured from ^{15}N-HSQC (Figure 5A) at

a ratio of $(GT)_4$ to protein of 0.5, 1.0, and 2.0. At a ratio of $(GT)_4$ to protein of 0.5, two sets of frequencies are observed for residues involved in binding while a single set of frequencies is observed at ratios of 1.0 and 2.0, indicating that binding is in slow exchange. The observation of slow exchange and that no differences in resonance frequencies were observed for equimolar and 2-fold excess molar equivalents of $(GT)_4$ is consistent with the previous estimate for the $(GT)_4$ binding association constant $K_a = (1.5-1.7) \times 10^5\ M^{-1}$ (*20*) and that $(GT)_4$ binds to a single site.

Upon $(GT)_4$ binding, the resonances for SP and the N-terminus of NC are broadened so that no intensity is observed for most of the residues 478−496. $(GT)_4$ binding also affects chemical shifts for NC resonances; however, only the resonances from ZF1 residues 509−524 have large perturbations in chemical shift upon binding $(GT)_4$ (Figure 5B). Small perturbations are observed for resonances from ZF2 while those of CTD are relatively unaffected by the presence of $(GT)_4$. These data suggest that the $(GT)_4$ binding site is localized to ZF1 and are consistent with previous biochemical data showing that ZF1 of HIV-1 plays a more significant role in RNA encapsidation (*58*).

Effects of $(GT)_4$ Binding on Heteronuclear Relaxation. Nucleic acid association is thought to induce conformational changes in Gag that are necessary to promote assembly (*20*, *22*). Alterations in protein dynamics can reflect such conformational changes.

The ^{15}N relaxation rate constants R_1 and R_2 and steady-state enhancement I/I_0 were measured for CTD-SP-NC plus $(GT)_4$ at a 1:1 molar ratio and a 600 MHz Larmor frequency. The residue profiles for these values, along with the R_2/R_1 ratio, are shown in Figure 6, and the averages over regions of the protein are listed in Table 2. The differences in relaxation rates between free and bound states are indicated with black bars in Figure 6 to facilitate the comparison of these two states of CTD-SP-NC.

(A) Association Affects Apparent Mass of both ZF1 and ZF2 but Not CTD. Nucleic acid binding alters R_2/R_1 relaxation of not only ZF1, the region that exhibits large chemical shift changes, but also ZF2 (Figure 6 and Table 2). The value of τ_m estimated from R_2/R_1 for ZF1 and ZF2 is 7.9 and 7.8 ns, an increase of 3.0 and 2.7 ns, respectively, for the $(GT)_4$ complex relative to unbound CTD-SP-NC. For the NC core in the presence of $(GT)_4$, R_2/R_1 becomes 7.1 ns, an increase of 2.0 ns. This increase in the R_2/R_1 ratios is roughly consistent with the predicted increase in the overall correlation time of approximately 1.8 ns for the larger molecular mass of 2.7 kDa corresponding to $(GT)_4$ (Figure 2). Further, the concomitant increase in R_2 and decrease in R_1 values expected for this change in τ_m are observed (Table 2 and Supporting Information). That the presence of $(GT)_4$ similarly alters both ZF1 and ZF2 relaxation indicates a comparable increase in apparent mass for both zinc fingers. On the other hand, only ZF1 appears to make direct contact based on chemical shift perturbations. Together, these observations strongly support the postulate that the NC core is a rotationally linked domain.

In the case of the CTD domain, the changes in overall rotational time observed upon $(GT)_4$ binding are minimal. There is a small increase in the average R_2/R_1 from 5.9 to 6.0, which corresponds to an increase in τ_m from 6.7 to 7.2 ns (Table 2). We view this small change in the estimated value of τ_m to be attributed to an increase in the mass of CTD-SP-NC by association of $(GT)_4$ at a remote site on the polyprotein, but CTD and NC remain largely rotationally uncoupled in the complex.

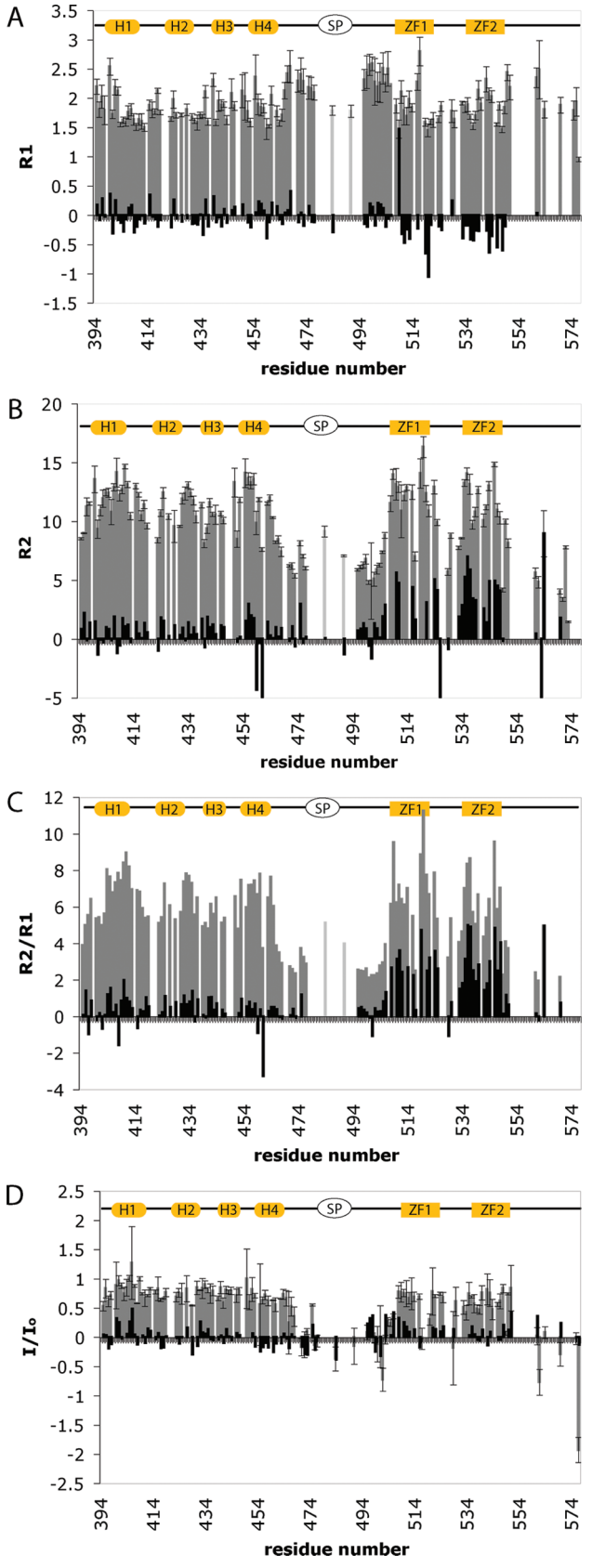


FIGURE 6: ^{15}N relaxation data at 600 MHz for RSV CTD-SP-NC in the bound form (gray bars) and the difference equal to bound minus unbound values (black bars). (A) Longitudinal relaxation rates, R_1. (B) Transverse relaxation rates R_2. (C) R_2/R_1 ratio. (D) $^1H-^{15}N$ I/I_0 values. Structural elements drawn at the top are as in Figure 3. Except for Q484 and V491, residues 479−493 in SP and the N-terminus of NC, colored light gray in Figure 3, are exchange broadened and not detected.

Table 2: Average ^{15}N Backbone Relaxation Rates and I/I_0 (=NOE-1) for Regions of RSV CTD-SP-NC in the $(GT)_4$ Bound State[a]

	residues	$\bar{R}_1$ (s^{-1})	$\bar{R}_2$ (s^{-1})	$\bar{R}_2/\bar{R}_1$	$\bar{I}/\bar{I}_0$	τ_m (ns)[b]
CTD	397−467	1.9 ± 0.3	11.2 ± 2.1	6.0 ± 1.5	0.7 ± 0.1	7.2
SP + flanking	468−507					
C-term CTD	468−477[c]					
SP*	479−493[c,d]					
N-term NC	495−507[c]	2.4 ± 0.2	6.2 ± 0.8	2.6 ± 0.5	0.02 ± 0.2	ND
NC core	509−548	1.8 ± 0.3	11.3 ± 2.4	6.3 ± 1.9	0.6 ± 0.2	7.1
ZF1	509−522	1.9 ± 0.4	12.1 ± 2.2	6.4 ± 2.2	0.6 ± 0.2	7.9
ZF2	535−548	1.8 ± 0.3	11.2 ± 1.7	6.1 ± 1.6	0.6 ± 0.1	7.8

[a]The standard deviation of the average value is listed. [b]Calculated using TENSOR by an overall fit of relaxation data from residues listed in column 2. [c]Regions defined by common relaxation behavior. [d]Region colored in gray in Figure 3. SP* residue range is based on relaxation behavior and differs somewhat from SP residues 477−488.

(*B*) *Slowed Exchange for SP*. A result of $(GT)_4$ association on the relaxation properties of CTD-SP-NC is that the peaks for residues in SP and the N-terminus of NC broaden and are no longer detected. This intermediate exchange behavior is observed for the amide resonances of residues 479−493, excluding residues Q484 and V491. The broadening of peaks for SP and the N-terminus of NC supports the proposal that the relatively large R_2 values measured in the free state for this region are due to exchange contributions to R_2. That is, these resonances in the free state are in a fast exchange regime, and the broadened line widths in the bound state indicate that $(GT)_4$ binding alters the underlying conformational exchange and slows the time scale to the intermediate regime. The reason resonances Q484 and V491 are not broadened is unknown at this time.

(*C*) *Effect of* $(GT)_4$ *Binding on Two CTD Helices*. Significant features due to $(GT)_4$ binding are the decreases in I/I_0 and R_2/R_1 observed for two CTD helices (Figure 6C,D). For I/I_0, helix 4, as well as residues in the C-terminus of CTD, exhibits a small but consistent trend to smaller I/I_0 values (Figure 6D). Given $\tau_m \sim 7$ ns, I/I_0 can be diminished only by a time-dependent internal motion indicating that the mobility of helix 4 and the C-terminus is enhanced upon binding. In the case of R_2/R_1, association of $(GT)_4$ leads to reduced R_2/R_1 values for two residues in helix 1 and two residues in helix 4 (Figure 6C), while all other residues in CTD and the NC core show an increase. As discussed above, a small increase in R_2/R_1 values is expected from the added mass of $(GT)_4$ at a distant site. For axially symmetric rotation, assuming the domain structure of CTD and principal axis system of the diffusion tensor are unchanged by $(GT)_4$ binding so that the orientation of an NH bond vector is constant within this framework, then R_2/R_1 should increase uniformly according to τ_m (Figure 4A). Two explanations are proposed for the observed decrease in R_2/R_1 of certain residues. First, if an NH vector reorients within the principal axis framework, then R_2/R_1 is not predicted by a change in τ_m and can either increase or decrease. As such, a decrease in R_2/R_1 upon ligand binding can occur by reorientation of the NH group, so that the smaller values of R_2/R_1 for residues in helices 1 and 4 infer that $(GT)_4$ binding induces a conformational rearrangement of these helices within the frame of CTD. Second, an exchange process could exist in the free state of CTD-SP-NC, leading to increased R_2 values due to R_{ex}. Binding of $(GT)_4$ could therefore diminish an exchange contribution and reduce R_2. Together, these data suggest that $(GT)_4$ binding has a subtle effect on the conformational equilibrium for an area that encompasses helix 1, helix 4, and the C-terminus of CTD and thereby implicates this area to have a role in assembly.

Model for Autoinhibition of Assembly. We propose a molecular model for autoinhibition of assembly and how nucleic acid binding releases this inhibition to initiate assembly. The model is based on the NMR results reported here and earlier mutagenesis studies (*20*, *22*) and is also motivated by a discovered inhibition of HIV-1 assembly conferred by a peptide that binds CTD (*27*, *40*).

Nearly an identical region near SP is distinguished by both NMR relaxation behavior and earlier mutagenesis studies of Gag polyprotein. As described here, the NMR relaxation behavior of SP and flanking regions (residues 468−507) differs distinctly from that of CTD and NC; collectively, the relaxation data are best fit by motion modeled with two correlation times corresponding to a slower overall rotation similar to the CTD and NC domains, plus an internal motion (model 4). In addition, the NMR data show that $(GT)_4$ binding slows a conformational exchange process involving residues 479−493 (SP*). Previously reported mutagenesis of Gag polyprotein identified the "assembly domain" (residues 469−492); the insertion or substitution of the flexible peptide sequences GSGSG or GG into SP, the eight C-terminal residues of CTD, or the first four residues of NC alter budding VLPs from an immature spherical to a mature tubular morphology, and the same mutations totally disrupt assembly *in vitro* (*21*, *31*). Thus the assembly domain identified by mutagenesis to be critical for assembly is nearly identical to the region of SP and the flanking residues where an exchange process is observed. The exchange process reported on by NMR relaxation is therefore likely a functional change in conformational equilibrium associated with assembly.

The residues found to have relaxation rates altered by $(GT)_4$ binding in a manner that suggests a conformation change by reorientation with respect to the overall diffusion tensor are located in a region that is key for protein interactions of assembly. These residues are in helices 1 and 4, which lie on the same surface of CA and form two edges of a hydrophobic patch (Figure 7, blue surface) conserved among retroviral CA proteins (*3*). CTD residues with smaller R_2/R_1 values in the presence of $(GT)_4$ are colored in green in Figure 7. Recent electron cryocrystallography studies of two-dimensional crystals of HIV-1 CA demonstrate that this conserved hydrophobic patch forms a heterodimerization interface between intermolecular NTD-CTD domains in the mature hexameric lattice (*24*). The conserved hydrophobic patch was also found by crystallography and NMR to bind a peptide (CAI) (*27*) that inhibits both immature and mature HIV-1 assembly *in vitro* (*40*). The inhibitory CAI peptide alters the homodimerization interface of CTD observed in the crystal (*27*, *40*). This CTD-CTD homodimeric interface is involved in connecting neighboring hexamers in the mature CA lattice of HIV (*24*)

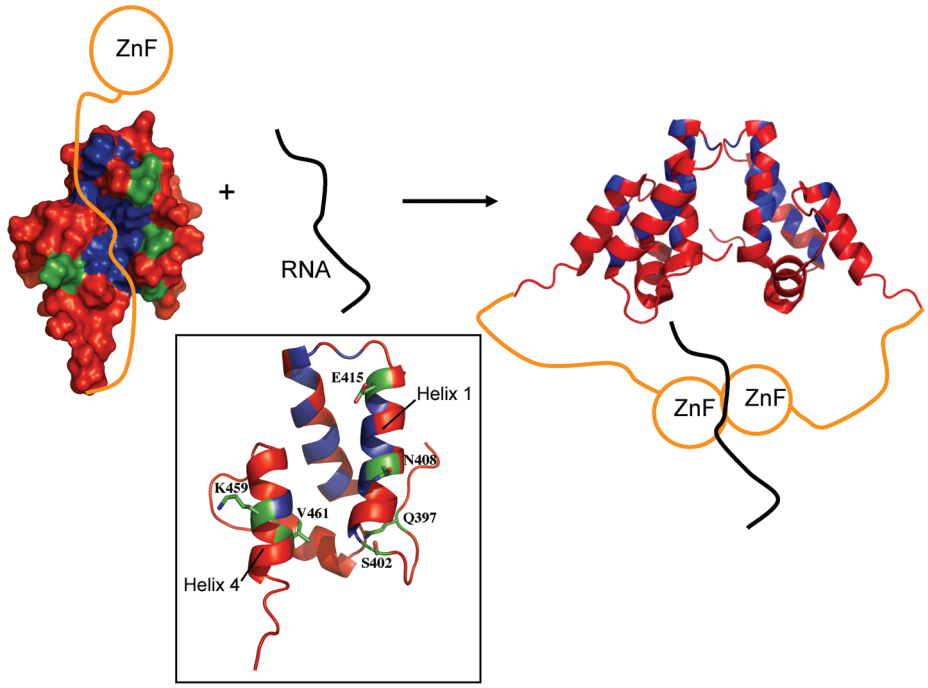


FIGURE 7: Model for initiation of assembly. In the free state of Gag, C-terminal residues of CTD associate with the hydrophobic area of CTD between helix 1 and helix 4 (colored blue, residues L410, L411, L419, P421, A423, A425, I428, I429, *F*432, G455, and I458) and localize SP at the top of the groove near helices 1 and 2. The intramolecular association disfavors a functional dimer contact (as does the inhibitory CAI peptide). Binding of NC (ZF1) to RNA "drags on" SP and its flanking region to dislodge the C-term and SP from CTD. This action allows a proper dimerization unit to form. Residues with reduced R_2/R_1 values in helices 1 and 4 are colored in green (residues Q397, S402, N408, E415, K459, and V461).

and may be involved in formation of the immature hexameric lattice (*23*). The crystallographic structure of RSV CTD at low pH (*26*) finds a dimeric interface that closely resembles the HIV-1 interface, and thus similar interfaces likely stabilize the RSV lattice. Mutagenesis of conserved residues within the CAI binding site of HIV-1 supports the idea that this hydrophobic patch is important in regulating the conformation necessary for virus assembly and maturation (*59*). It was concluded that the inhibitory activity from CAI binding may result from direct steric hindrance of NTD-CTD contacts or indirectly by stabilizing a CTD conformation that is not compatible with the CTD-CTD dimer formed in the mature particle.

Together, these results suggest a model for autoinhibition of assembly and release triggered by RNA association with Gag. We postulate that, in the absence of nucleic acid, formation of assembly-competent CTD dimers is autoinhibited by an interaction of residues at the C-terminus of CTD with the vicinity of the conserved hydrophobic patch between helices 1 and 4 (Figure 7, left-hand side) and localizes SP at the top of the groove near the loop between helix 1 and helix 2. This groove contact is speculated to disfavor Gag dimerization in a manner similar to that achieved by the inhibitory CAI peptide (*40*). NMR relaxation of SP and flanking regions is consistent with this region of Gag being associated with CTD in that the overall relaxation is best fit by model 4 with a longer (~5 ns) rotational correlation time similar to CTD, plus a faster time scale motion corresponding to a conformational exchange process. Although the nature of this conformational change cannot yet be defined, it is clear that nucleic acid binding alters the exchange given the loss of these peaks by exchange broadening in the complex of CTD-SP-NC with $(GT)_4$.

Intramolecular association of the C-terminus of CTD and the assembly domain with the hydrophobic patch between helix 1 and helix 4 of CTD would destabilize the Gag dimeric interface and sequester the assembly domain from intermolecular associations compatible with higher order oligomerization for assembly. The ZF1-mediated association of RNA with NC is postulated to dislodge the C-terminus of CTD and SP from the hydrophobic groove (Figure 7, right-hand side), thereby affecting helices 1 and 4 consistent with the observed changes in NMR relaxation, and would promote the proper dimer contact surface of Gag necessary for immature assembly. It is noted that the system studied here corresponds only to the initial steps in assembly in that oligonucleotides as short as $(GT)_4$ do not actually lead to RSV Gag dimerization shown in Figure 7 given that the formation of the dimer building block requires $(GT)_8$ (*20*) or oligomers long enough to bind two Gag molecules. During assembly, the close proximity of Gag upon binding one $(GT)_8$, or longer oligonucleotide, increases the effective local Gag concentration and is thought to stabilize a Gag-Gag association that is too low affinity to form otherwise. As such, the function of Gag association with nucleic acid may be 2-fold: first, to remove the inhibitory intramolecular interaction between the assembly domain and the hydrophobic groove of CTD, and second, to drive a low-affinity intermolecular association.

CONCLUSIONS

^{15}N heteronuclear relaxation and chemical shifts indicate that the CTD and NC regions of the Gag polyprotein are rotationally independent domains. The correlation time of CTD in CTD-SP-NC is only slightly longer than that estimated for isolated CTD and varies independent of that for NC. Chemical shifts of CTD in CTD-SP-NC and from the isolated CTD domain indicate the CTD structure is largely preserved in the context of the polyprotein. Small disparities in chemical shifts,

however, do suggest that the conformation of the C-terminal helix 4 of CTD is altered by the covalent attachment of SP and NC. Based on the characterization of rotational tumbling in both the free and $(GT)_4$-bound state, the NC core appears to be a single domain so that $(GT)_4$ binding slows the tumbling and increases the apparent mass of both ZF1 and ZF2 even though direct contact is suggested by chemical shift perturbation to occur only with ZF1.

Residues in SP and the flanking regions exhibit greater variation in relaxation rates, and thus more complex dynamics, than either CTD or the NC core. Moreover, the dynamics are sensitive to $(GT)_4$ binding, even though no direct interaction with $(GT)_4$ is suggested from chemical shift perturbations; the only region of CTD-SP-NC with large perturbations (>0.1 ppm) is ZF1. In particular, SP and the flanking regions, which correspond to the assembly domain defined by mutagenesis (*31*), undergo a relatively fast conformational exchange process in the free state and a slower one in the presence of $(GT)_4$. Comparison of the RSV relaxation profiles reported here with those of HIV-1 (*39*) finds that SP and the analogous SP-1 region of HIV are both conformationally flexible, but the R_2 relaxation profile for RSV SP shows a strong maximum at residue I487 while the SP-1 profile is notably uniform. Cis−trans isomerization of Pro485, a conserved residue in RSV located at the C-terminus of SP, may be the origin of the observed R_2 maximum. There are no conserved Pro residues in SP-1 of HIV-1. RSV SP also differs from HIV-1 SP-1 in that a trend in chemical shifts of residues in the N-terminal region of SP-1suggests transient formation of a helix (*39*). Such a trend is not observed in the chemical shift profile of RSV SP. Whether the observed differences in NMR relaxation behavior reflect significant distinctions between RSV and HIV-1 or are due to slight differences in equilibrium or time dependences of transitions, or dependences on other factors such as pH, is unknown at this time. It is important to recall that virus assembly is driven by the combined forces of many weak interactions, so that a structural feature important for stabilizing the assembled particle may not be realized under the conditions in which the isolated protein or protein complex is studied.

A central outcome of this work is the proposal for a model on the autoinhibition of assembly and how nucleic acid binding to Gag polyprotein can trigger dimerization to promote assembly. The fundamental features of the model (Figure 7), suggested by the NMR relaxation data and mutagenesis effects on assembly, are the intramolecular association of the SP assembly domain with CTD and the disruption of this association by intermolecular binding of RNA to NC. SP binding to the hydrophobic surface patch of CTD is postulated to inhibit dimerization, analogous to the inhibition of assembly by the CAI that binds in a similar location on CTD (*27*, *40*), and to sequester the assembly domain to prevent premature intermolecular association. The model provides a new framework to test with future experiments and should aid progress in defining the molecular mechanism of retroviral assembly.

SUPPORTING INFORMATION AVAILABLE

Values from spectral density mapping and generalized order paramters averaged over domains for CTD-SP-NC unligated and bound to $(GT)_4$ and theoretical curves of R_1 and R_2 relaxation rates shown as a function of correlation time. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Fuller, S. D., Wilk, T., Gowen, B. E., Kräusslich, H.-G., and Vogt, V. M. (1997) Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Curr. Biol. 7*, 729–738.
2. Yeager, M., Wilson-Kubalek, E. M., Weiner, S. G., Brown, P. O., and Rein, A. (1998) Supramolecular organization of immature and mature murine leukemia virus revealed by electron cryo-microscopy: Implications for retroviral assembly mechanisms. *Proc. Natl. Acad. Sci. U.S.A. 95*, 7299–7304.
3. Kingston, R. L., Fitzon-Ostendorp, T., Eisenmesser, E. Z., Schatz, G. W., Vogt, V. M., Post, C. B., and Rossmann, M. G. (2000) Structure and self-association of the Rous sarcoma virus capsid protein. *Structure 8*, 617–628.
4. Campos-Olivas, R., Newman, J. L., and Summers, M. F. (2000) Solution structure and dynamics of the Rous sarcoma virus capsid protein and comparison with capsid proteins of other retroviruses. *J. Mol. Biol. 296*, 633–649.
5. Murray, P. S., Li, Z., Wang, J., Tang, C. L., Honig, B., and Murray, D. (2005) Retroviral matrix domains share electrostatic homology: Models for membrane binding function throughout the viral life cycle. *Structure 13*, 1521–1531.
6. Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997) Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science 278*, 849–853.
7. Darlix, J.-L., Lapadat-Tapolsky, M., de Rocquigny, H., and Roques, B. P. (1995) First glimpses at structure-function relationships of the nucleocapsid protein of retroviruses. *J. Mol. Biol. 254*, 523–537.
8. McDonnell, J. M., Fushman, D., Cahill, S. M., Zhou, W., Wolven, A., Wilson, C. B., Nelle, T. D., Resh, M. D., Wills, J., and Cowburn, D. (1998) Solution structure and dynamics of the bioactive retroviral M domain from Rous sarcoma virus. *J. Mol. Biol. 279*, 921–928.
9. Massiah, M. A., Starich, M. R., Paschall, C., Summers, M. F., Christensen, A. M., and Sundquist, W. I. (1994) Three-dimensional structure of the human immunodeficiency virus type 1 matrix protein. *J. Mol. Biol. 244*, 198–223.
10. Christensen, A. M., Massiah, M. A., Turner, B. G., Sundquist, W. I., and Summers, M. F. (1996) Three-dimensional structure of the HTLV-II matrix protein and comparative analysis of matrix proteins from the different classes of pathogenic human retroviruses. *J. Mol. Biol. 264*, 1117–1131.
11. Rao, Z., Belyaev, A. S., Fry, E., Roy, P., Jones, I. M., and Stuart, D. I. (1995) Crystal structure of SIV matrix antigen and implications for virus assembly. *Nature 378*, 743–747.
12. Gamble, T. R., Vajdos, F. F., Yoo, S., Worthylake, D. K., Houseweart, M., Sundquist, W. I., and Hill, C. P. (1996) Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell 87*, 1285–1294.
13. Gitti, R. K., Lee, B. M., Walker, J., Summers, M. F., Yoo, S., and Sundquist, W. I. (1996) Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science 273*, 231–235.
14. Nandhagopal, N., Simpson, A. A., Johnson, M. C., Francisco, A. B., Schatz, G. W., Rossmann, M. G., and Vogt, V. M. (2004) Dimeric Rous sarcoma virus capsid protein structure relevant to immature Gag assembly. *J. Mol. Biol. 335*, 275–282.
15. Gao, Y., Kaluarachchi, K., and Giedroc, D. P. (1998) Solution structure and backbone dynamics of Mason-Pfizer monkey virus (MPMV) nucleocapsid protein. *Protein Sci. 7*, 2265–2280.
16. Klein, D. J., Johnson, P. E., Zollars, E. S., De Guzman, R. N., and Summers, M. F. (2000) The NMR structure of the nucleocapsid protein from the mouse mammary tumor virus reveals unusual folding of the C-terminal zinc knuckle. *Biochemistry 39*, 1604–1612.
17. Summers, M. F., Henderson, L. E., Chance, M. R., Bess, J. W., South, T. L., Blake, P. R., Sagi, I., Perez-Alvarado, G., Sowder, R. C., and et al. (1992) Nucleocapsid zinc fingers detected in retroviruses: EXAFS studies of intact viruses and the solution-state structure of the nucleocapsid protein from HIV-1. *Protein Sci. 1*, 563–574.
18. Summers, M. F., South, T. L., Kim, B., and Hare, D. R. (1990) High-resolution structure of an HIV zinc fingerlike domain via a new NMR-based distance geometry approach. *Biochemistry 29*, 329–340.
19. Morellet, N., Jullian, N., Derocquigny, H., Maigret, B., Darlix, J. L., and Roques, B. P. (1992) Determination of the structure of the nucleocapsid protein NCP7 from the human-immunodeficiency-virus type-1 by H-1-NMR. *EMBO J. 11*, 3059–3065.
20. Ma, Y. M., and Vogt, V. M. (2002) Rous sarcoma virus Gag protein-oligonucleotide interaction suggests a critical role for protein dimer formation in assembly. *J. Virol. 76*, 5452–5462.

21. Johnson, M. C., Scobie, H. M., Ma, Y. M., and Vogt, V. M. (2002) Nucleic acid-independent retrovirus assembly can be driven by dimerization. *J. Virol. 76*, 11177–11185.
22. Ma, Y. M., and Vogt, V. M. (2004) Nucleic acid binding-induced Gag dimerization in the assembly of Rous sarcoma virus particles in vitro. *J. Virol. 78*, 52–60.
23. Wright, E. R., Schooler, J. B., Ding, H. J., Kieffer, C., Fillmore, C., Sundquist, W. I., and Jensen, G. J. (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO J. 26*, 2218–2226.
24. Ganser-Pornillos, B. K., Cheng, A., and Yeager, M. (2007) Structure of full-length HIV-1 CA: A model for the mature capsid lattice. *Cell 131*, 70–79.
25. Ivanov, D., Stone, J. R., Maki, J. L., Collins, T., and Wagner, G. (2005) Mammalian SCAN domain dimer is a domain-swapped homolog of the HIV capsid C-terminal domain. *Mol. Cell 17*, 137–143.
26. Bailey, G. D., Hyun, J. K., Mitra, A. K., and Kingston, R. L. (2009) Proton-linked dimerization of a retroviral capsid protein initiates capsid assembly. *Structure 17*, 737–748.
27. Ternois, F., Sticht, J., Duquerroy, S., Krausslich, H.-G., and Rey, F. A. (2005) The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor. *Nat. Struct. Mol. Biol. 12*, 678–682.
28. Pepinsky, R. B., Papayannopoulos, I. A., Chow, E. P., Krishna, N. K., Craven, R. C., and Vogt, V. M. (1995) Differential proteolytic processing leads to multiple forms of the CA protein in avian sarcoma and leukemia viruses. *J. Virol. 69*, 6430–6438.
29. Liang, C., Hu, J., Whitney, J. B., Kleiman, L., and Wainberg, M. A. (2003) A structurally disordered region at the C terminus of capsid plays essential roles in multimerization and membrane binding of the Gag protein of human immunodeficiency virus type 1. *J. Virol. 77*, 1772–1783.
30. Melamed, D., Mark-Danieli, M., Kenan-Eichler, M., Kraus, O., Castiel, A., Laham, N., Pupko, T., Glaser, F., Ben-Tal, N., and Bacharach, E. (2004) The conserved carboxy terminus of the capsid domain of human immunodeficiency virus type 1 Gag protein is important for virion assembly and release. *J. Virol. 78*, 9675–9688.
31. Keller, P. W., Johnson, M. C., and Vogt, V. M. (2008) Mutations in the spacer peptide and adjoining sequences in Rous sarcoma virus Gag lead to tubular budding. *J. Virol. 82*, 6788–6797.
32. Cheslock, S. R., Poon, D. T. K., Fu, W., Rhodes, T. D., Henderson, L. E., Nagashima, K., McGrath, C. F., and Hu, W.-S. (2003) Charged assembly helix motif in murine leukemia virus capsid: An important region for virus assembly and particle size determination. *J. Virol. 77*, 7058–7066.
33. Liang, C., Hu, J., Russell, R. S., Roldan, A., Kleiman, L., and Wainberg, M. A. (2002) Characterization of a putative {alpha}-helix across the capsid-SP1 boundary that is critical for the multimerization of human immunodeficiency virus type 1 Gag. *J. Virol. 76*, 11729–11737.
34. Morellet, N., Druillennec, S., Lenoir, C., Bouaziz, S., and Roques, B. P. (2005) Helical structure determined by NMR of the HIV-1 (345−392)Gag sequence, surrounding p2: Implications for particle assembly and RNA packaging. *Protein Sci. 14*, 375–386.
35. Briggs, J. A. G., Riches, J. D., Glass, B., Bartonova, V., Zanetti, G., and Krausslich, H. G. (2009) Structure and assembly of immature HIV. *Proc. Natl. Acad. Sci. U.S.A. 106*, 11090–11095.
36. Wright, E. R., Schooler, J. B., Ding, H. J., Kieffer, C., Fillmore, C., Sundquist, W. I., and Jensen, G. J. (2007) Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. *EMBO J. 26*, 2218–2226.
37. Campbell, S., and Vogt, V. M. (1995) Self-assembly in vitro of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. *J. Virol. 69*, 6487–6497.
38. Fisher, R. J., Rein, A., Fivash, M., Urbaneja, M. A., Casas-Finet, J. R., Medaglia, M., and Henderson, L. E. (1998) Sequence-specific binding of human immunodeficiency virus type 1 nucleocapsid protein to short oligonucleotides. *J. Virol. 72*, 1902–1909.
39. Newman, J. L., Butcher, E. W., Patel, D. T., Mikhaylenko, Y., and Summers, M. F. (2004) Flexibility in the P2 domain of the HIV-1 Gag polyprotein. *Protein Sci. 13*, 2101–2107.
40. Sticht, J., Humbert, M., Findlow, S., Bodem, J., Muller, B., Dietrich, U., Werner, J., and Krausslich, H.-G. (2005) A peptide inhibitor of HIV-1 assembly in vitro. *Nat. Struct. Mol. Biol. 12*, 671–677.
41. Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277–293.
42. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry 33*, 5984–6003.
43. Ma, L., Hass, M. A. S., Vierick, N., Kristensen, S. M., Ulstrup, J., and Led, J. J. (2003) Backbone dynamics of reduced plastocyanin from the cyanobacterium *Anabaena variabilis*: Regions involved in electron transfer have enhanced mobility. *Biochemistry 42*, 320–330.
44. Abragam, A. (1961) The Principles of Nuclear Magnetism, Clarendon Press, Oxford.
45. Ghose, R., Fushman, D., and Cowburn, D. (2001) Determination of the rotational diffusion tensor of macromolecules in solution from NMR relaxation data with a combination of exact and approximate methods—Application to the determination of interdomain orientation in multidomain proteins. *J. Magn. Reson. 149*, 204–217.
46. Woessner, D. E. (1962) Nuclear spin relaxation in ellipsoids undergoing rotational Brownian motion. *J. Chem. Phys. 37*, 647–654.
47. Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc. 104*, 4546–4559.
48. Jin, D., Figueirido, F., Montelione, G. T., and Levy, R. M. (1997) Impact of the precision in NMR relaxation measurements on the interpretation of protein dynamics. *J. Am. Chem. Soc. 119*, 6923–6924.
49. Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. *J. Am. Chem. Soc. 112*, 4989–4991.
50. Baber, J. L., Szabo, A., and Tjandra, N. (2001) Analysis of slow interdomain motion of macromolecules using NMR relaxation data. *J. Am. Chem. Soc. 123*, 3953–3959.
51. Chang, S.-L., and Tjandra, N. (2001) Analysis of NMR relaxation data of biomolecules with slow domain motions using wobble-in-a-cone approximation. *J. Am. Chem. Soc. 123*, 11484–11485.
52. Dosset, P., Hus, J.-C., Blackledge, M., and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR 16*, 23–28.
53. García de la Torre, J., Huertas, M. L., and Carrasco, B. (2000) HYDRONMR: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson. 147*, 138–146.
54. Peng, J. W., and Wagner, G. (1992) Mapping of the spectral densities of nitrogen-hydrogen bond motions in Eglin c using heteronuclear relaxation experiments. *Biochemistry 31*, 8571–8586.
55. Lee, B. M., De Guzman, R. N., Turner, B. G., Tjandra, N., and Summers, M. F. (1998) Dynamical behavior of the HIV-1 nucleocapsid protein. *J. Mol. Biol. 279*, 633–649.
56. Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: Application to staphylococcal nuclease. *Biochemistry 28*, 8972–8979.
57. Schwalbe, H., Fiebig, K. M., Buck, M., Jones, J. A., Grimshaw, S. B., Spencer, A., Glaser, S. J., Smith, L. J., and Dobson, C. M. (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. *Biochemistry 36*, 8977–8991.
58. Gorelick, R. J., Chabot, D. J., Rein, A., Henderson, L. E., and Arthur, L. O. (1993) The two zinc fingers in the human immunodeficiency virus type 1 nucleocapsid protein are not functionally equivalent. *J. Virol. 67*, 4027–4036.
59. Bartonova, V., Igonet, S. B., Sticht, J., Glass, B. R., Habermann, A., Vaney, M.-C., Sehr, P., Lewis, J., Rey, F. A., and Krasslich, H.-G. (2008) Residues in the HIV-1 capsid assembly inhibitor binding site are essential for maintaining the assembly-competent quaternary structure of the capsid protein. *J. Biol. Chem. 283*, 32024–32033.

Supplementary Material.

NMR relaxation studies of an RNA-binding segment of the Rous sarcoma virus Gag polyprotein in free and bound states - a model for autoinhibition of assembly

Gwen M. Taylor[1], Lixin Ma[1], Volker M. Vogt[2] and Carol Beth Post[1]

[1]*Department of Medicinal Chemistry and Molecular Pharmacology, Markey Center for Structural Biology and Purdue Cancer Center, Purdue University, West Lafayette, Indiana 47907*

[2]*Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853*

keywords: retrovirus, virus assembly, gag polyprotein, multidomain protein dynamics

Spectral density mapping for unbound and $(GT)_4$ bound CTD-SP-NC

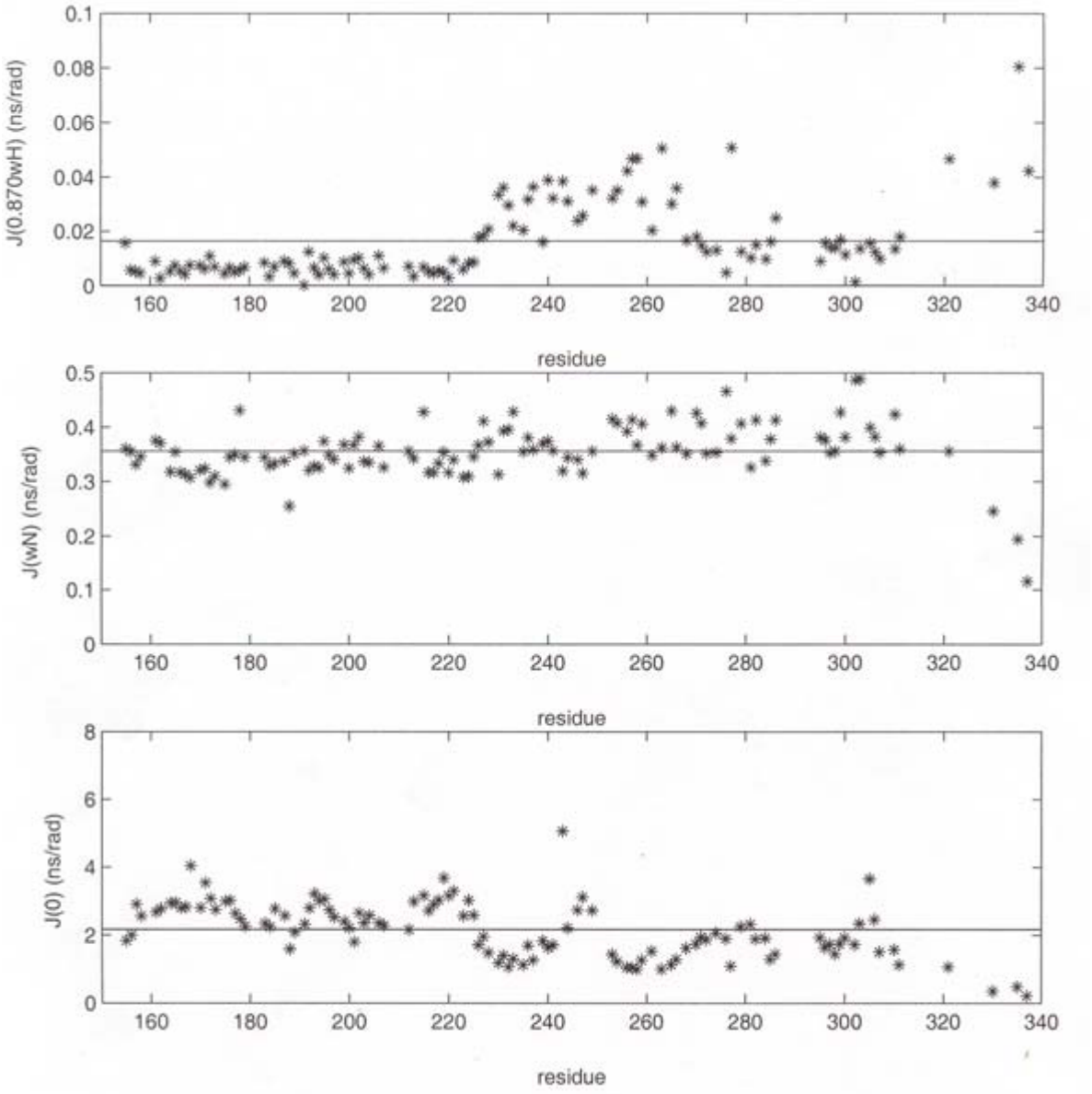
Main chain dynamics of CTD-SP-NC measured by ^{15}N heteronuclear relaxation were characterized by spectral density mapping(*1*), a structure-independent approach, and the results are shown in Supplementary Figures 1 and 2. Averages over regions grouped based on relaxation behavior are listed in Supplementary Tables 1 and 2.

Dynamics of SP and flanking region in unbound CTD-SP-NC

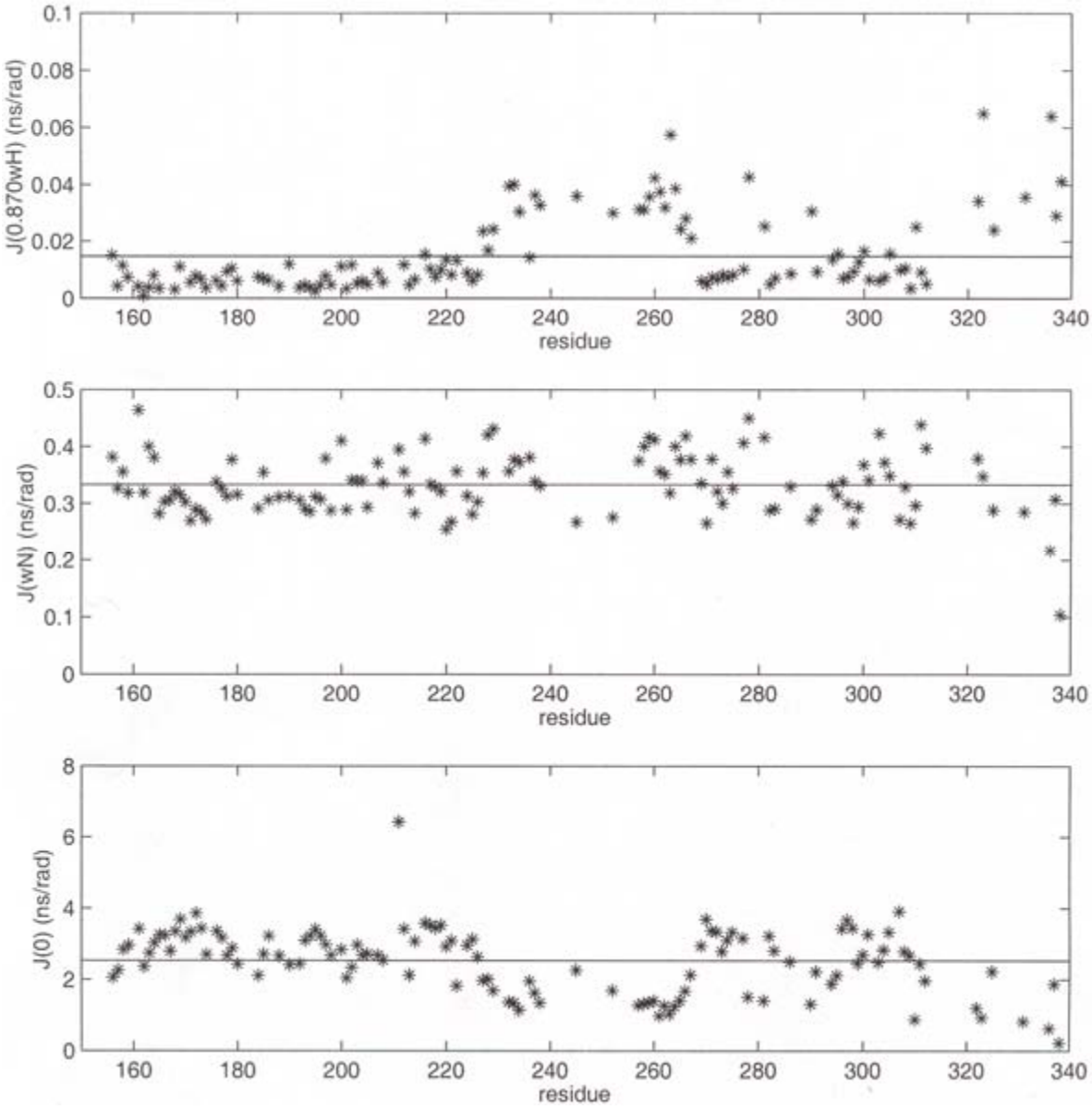
Theoretical curves for R_2 and R_1 as a function of overall correlation time are shown in Supplementary Figure 3 for the four motional models described in the article. Anisotropic rotation, shown in the left panel for axially symmetric rotation, affects the value of R_2 opposite in direction to that of R_1. Internal motion decreases both relaxation rates.

A general description of the dynamics of SP and flanking regions is concluded based on the combined average values measured for I/I_0, and R_1 and R_2. The observed average value is $I/I_0 \sim 0$. Neither model 2 (anisotropic rotation) nor model 3 (very fast internal motion, single overall correlation time) can account for the observation that $I/I_0 \sim 0$. For model 1, $I/I_0 \sim 0$ occurs at $\tau_m = 1$ ns, a relatively short correlation time corresponding to a fully flexible linker. This value $\tau_m \sim 1$ ns falls to the left of the R_1 maximum so that the predicted R_1 and R_2 would be similar in value (for $\tau_m = 1$ ns, $R_1 = 2.0$ and R_2 2.5 s^{-1}). For model 4, R_2 is greater than R_1 and we use the example S^2=0.5-0.6, $\tau_{int} \sim 0.4$ ns, $\tau_m = 5 - 7$ ns for the predicted R_1 and R_2 values listed in Supplementary Table 4. For SP and flanking regions, the observed R_1 ranges from 1.8 – 2.1 s^{-1} and R_2 ranges from 6 – 13 s^{-1}. While R_1 cannot readily discriminate models 1 and 4, the observed R_2 values are significantly greater and more variable than the values expected for model 1 and a fully flexible chain characterized by $\tau_m \sim 1$ ns. Model 4 more closely reproduces the overall relaxation behavior of the SP and flanking region; however, experimental R_2 values for several residues are underestimated by the theoretical relaxation curves. That R_2 values for this region are underestimated suggests that exchange (R_{ex}) contributes to R_2.

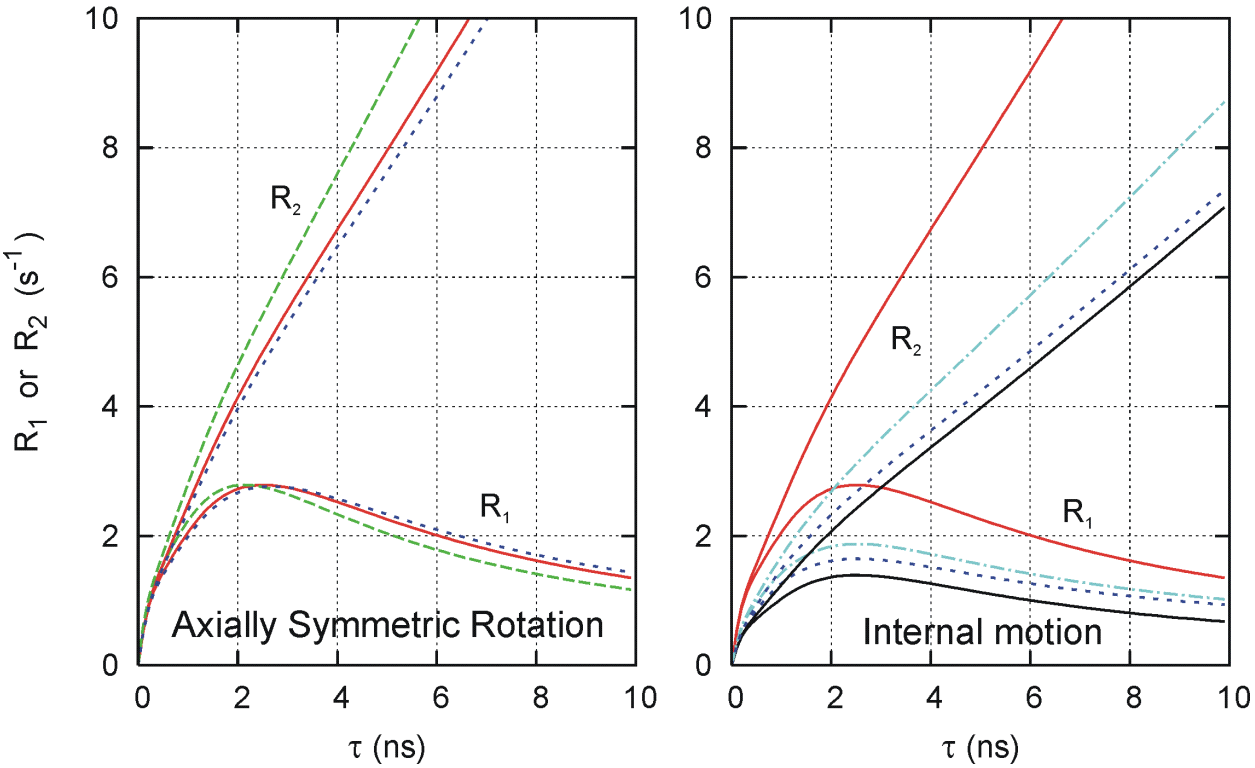
1. Peng, J. W., and Wagner, G. (1992) Mapping of the spectral densities of nitrogen-hydrogen bond motions in Eglin c using heteronuclear relaxation experiments, *Biochemistry 31*, 8571-8586.



Supplementary Figure 1. Spectral density mapping of RSV CTD-SP-NC in the apo form at 11.74 T. The average J(0.87ωH) is 1.64 x 10^{-11}, J(ωN) is 3.33 x 10^{-10}, and J(0) is 2.16 x 10^{-9}.



Supplementary Figure 2. Spectral density mapping of RSV CTD-SP-NC in the bound form at 11.74 T. The average J(0.87ωH) is 1.47 x 10^{-11}, J(ωN) is 3.33 x 10^{-10}, and J(0) is 2.52 x 10^{-9}.



Supplementary Figure 3. R_2 and R_1 as a function of overall correlation time,τ_m, for four motional models. Model 1, isotropic rotation with a single overall rotational correlation time (solid, red) is shown in both panels for reference. Left panel: Axially symmetric rotation with $D_{zz} = D_{\parallel}$=1.3 and $D_{xx} = D_{yy} = D_{\perp}$=0.85, and $\tau_m = 1/\left[2\left(D_{xx} + D_{yy} + D_{zz}\right)\right]$. Two curves delimit the relaxation for N-H bond vectors aligned parallel (dashed, green) or perpendicular (dotted, blue) to the z-axial component of the diffusion tensor. Right panel: Overall isotropic rotation with faster internal motion. Model 3, a single correlation function including very fast internal motion squares, S^2=0.5 (solid, black). Model 4, internal motion with τ_{int}=0.1 ns and S^2= 0.6 (dash-dot, cyan) or 0.5 (dotted, blue).

Supplementary Table 1. Average spectral density data for RSV CTD-SP-NC in the unbound state.

	residues	J(0) (ns/rad)	$J(\omega_N)$ (ns/rad 10^{-1})	$J(0.87\omega_H)$ (ns/rad 10^{-2})
CTD	394 - 474	2.5 ± 0.6	3.4 ± 0.3	0.9 ± 0.7
SP	475 - 489	2.4 ± 1.1	3.5 ± 0.2	3.1 ± 0.7
N-term NC	490 - 505	1.2 ± 0.2	3.9 ± 0.3	3.5 ± 1.1
NC core	509 - 548	1.9 ± 0.5	3.9 ± 0.4	1.5 ± 0.9
ZF 1	509 - 522	1.9 ± 0.4	3.9 ± 0.4	1.7 ± 0.1
ZF 2	535 - 548	2.0 ± 0.7	4.0 ± 0.5	1.2 ± 0.4

Supplementary Table 2. Average spectral density data for RSV CTD-SP-NC in the bound state.

	residues	J(0) (ns/rad)	$J(\omega_N)$ (ns/rad 10^{-1})	$J(0.87\omega_H)$ (ns/rad 10^{-2})
CTD	394 - 474	2.8 ± 0.8	3.3 ± 0.4	0.9 ± 0.8
SP	475 - 489	N.D.	N.D.	N.D.
N-term NC	490 - 505	1.7 ± 0.8	3.6 ± 0.5	3.0 ± 1.4
NC core	509 - 548	2.8 ± 0.7	3.3 ± 0.5	1.2 ± 0.9
ZF 1	509 - 522	2.8 ± 0.7	3.5 ± 0.6	1.3 ± 1.2
ZF 2	535 - 548	3.1 ± 0.5	3.2 ± 0.5	1.0 ± 0.4

Supplementary Table 3. S^2 values calculated by TENSOR

	residues	S^2
CTD	397 – 467	0.874
SP+flanking regions	468 - 507	
C-term CTD	468 - 477	
SP	479 - 493	
N-term NC	495 - 507	
NC core	509 - 548	0.642
ZF 1	509 - 522	0.707
ZF 2	535 - 548	0.671

Supplementary Table 4. Observed R_1 and R_2 values for SP* (residues 479 - 493) and the flanking regions (residues 468 - 477 and 495 - 507) compared to predicted values for model 1 or 4 at τ_m values with the condition I/I_0 is near zero.

	R_1 (s^{-1})	R_2 (s^{-1})
Observed		
Flanking C-term CTD	~2.2	~6.4
SP* [c]	1.8 to 2.2	6 to 13
Flanking N-term NC	~2.4	~5.5
Predicted		
Model 1, expected for τ_m = 1 ns[a]	2.0	2.5
Model 4, expected for τ_m = 5-7 ns[b]	1.9	4 to 7

[a] Predicted values (Figure S3) calculated for a single overall correlation time, no internal motion.
[b] Predicted values (Figure S3) calculated for slow internal motion parameters are S^2=0.5 to 0.6 and τ_{int}=0.4 ns.
[c] SP* residue range is based on relaxation behavior and differs somewhat from SP residues 477-488.