J|A|C|S
ARTICLES
Published on Web 05/31/2007

Solution NMR Studies of the Maturation Intermediates of a 13 MDa Viral Capsid

Blair R. Szymczyna, Lu Gan, John E. Johnson, and James R. Williamson*

Contribution from the Departments of Molecular Biology and Chemistry, The Skaggs Institute for Chemical Biology and The Center for Integrative Molecular Biosciences, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received February 15, 2007; E-mail: jrwill@scripps.edu

Abstract: Dynamic regions of proteins often play essential roles in function, assembly, or maturation of macromolecular complexes. When X-ray crystallography and cryo-electron microscopy are used to investigate macromolecular structures, information about these dynamic regions is lost because of conformational disorder or flexibility. Structural studies on the precursor capsid conformations of the lambdoid bacteriophage HK97, a model system for macromolecular maturation, reveal that essential regions tend to be disordered at early maturation stages. In the Prohead II intermediate, the regions that have the greatest disorder are the N-terminal residues and the E-loop, a region involved in the formation of inter-subunit cross-links. The N-terminus of the subunits in the 13 MDa procapsid is sufficiently dynamic to be studied by solution nuclear magnetic resonance (NMR) spectroscopy. NMR measurements enabled the identification and assignment of resonances to specific residues, assessment of the region's behavior within the context of the capsid, and monitoring of changes in these residues during the maturation process. In the precursor Prohead II and immature EI-III states, the N-termini are found to make transient interactions with the interior capsid surface, while at least a subset of N-termini in EI-III becomes more flexible with time. No resonances are observed for the fully mature Head II capsid, which is consistent with its completely ordered structure. NMR spectroscopy complements the current X-ray crystallography and cryo-electron microscopy data of HK97 by providing key information about the behavior of essential dynamic regions only inferred by other techniques.

Introduction

Flexible segments of proteins or nucleic acids within large molecules or macromolecular assemblies are often essential for proper folding, assembly, and function. Many enzymes have flexible regions at or near the active site that become folded upon substrate interaction and the assembly of complex structures, such as the ribosome, involve a series of sequential RNA folding and protein binding events. The associated flexibility often results in the inability to detect these regions in electron density maps and assess their behavior within the structure. Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool that can observe and provide information about the structure and dynamics of such regions.

NMR spectroscopy is routinely used to determine structures and assess the dynamics of biomolecules and biomolecular complexes smaller than 30 kDa. Above this molecular weight, several factors limit the ability of NMR to yield useful information. An increase in the width of resonance signals, due to the rapid transverse relaxation rates associated with large correlation times, and an increase in the number of signals lead to complicated spectra with poor resolution. Much effort has been devoted to expanding the range of molecular sizes that could be investigated by NMR. Isotope-labeling strategies and multidimensional spectroscopy in combination with novel NMR techniques, such as transverse relaxation-optimized spectroscopy (TROSY),[1] cross-correlated relaxation-enhanced polarization transfer (CRINEPT)[2] and cross-correlated relaxation-induced polarization transfer (CRIPT),[2−4] enable studies of larger species.[5] Even when employing these techniques, however, structure determination by NMR is still limited to species smaller than 1 MDa.[6]

Despite these limitations, NMR can still be used to spectroscopically gather important information about species larger than 1 MDa. Flexible regions or domains tethered to the structure via a flexible linker may have favorable relaxation properties such that resonances can be observed. Two studies have probed the structure of the C-terminal domains of the L7/L12 stalk protein within the context of the 2.3 MDa *E. coli* 70S ribosome.[7,8] NMR was also used to assess the structure of the N-terminal 25 residues of the Cowpea Chlorotic Mottle Virus

(1) Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94* (23), 12366−71.
(2) Riek, R.; Wider, G.; Pervushin, K.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96* (9), 4918−23.
(3) Brüschweiler, R.; Ernst, R. R. *J. Chem. Phys.* **1992**, *96* (3), 1758−66.
(4) Dalvit, C. *J. Magn. Reson.* **1992**, *97* (3), 645−50.
(5) Riek, R.; Fiaux, J.; Bertelsen, E. B.; Horwich, A. L.; Wüthrich, K. *J. Am. Chem. Soc.* **2002**, *124* (41), 12144−53.
(6) Wider, G. *Methods Enzymol.* **2005**, *394*, 382−98.
(7) Mulder, F. A.; Bouakaz, L.; Lundell, A.; Venkataramana, M.; Liljas, A.; Akke, M.; Sanyal, S. *Biochemistry* **2004**, *43* (20), 5930−6.
(8) Christodoulou, J.; Larsson, G.; Fucini, P.; Connell, S. R.; Pertinhez, T. A.; Hanson, C. L.; Redfield, C.; Nierhaus, K. H.; Robinson, C. V.; Schleucher, J.; Dobson, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (30), 10949−54.

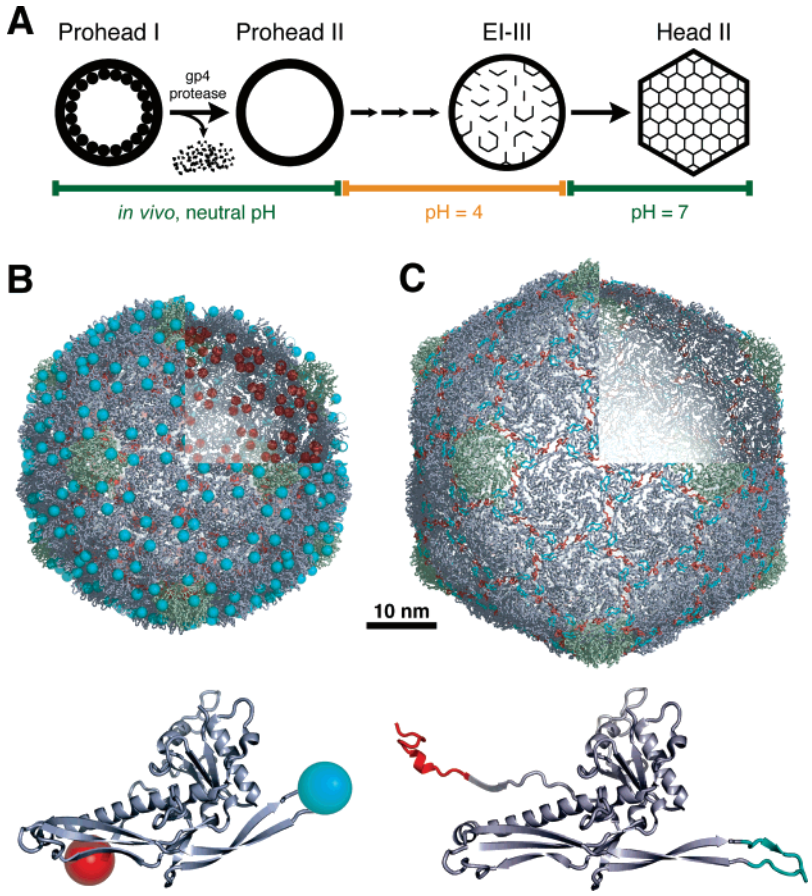

Figure 1. (A) Schematic demonstrating the progression of capsid maturation from Prohead II through to the fully mature Head II state. The N-terminal 102 amino acids in Prohead I are removed in vivo and cleaved into peptides by a viral protease, which subsequently degrades itself.[14,15] Further maturation of the resultant Prohead II particle proceeds through structurally distinct early intermediate stages to the EI-III state when the pH is lowered to 4.[17] The fully mature Head II state is obtained when the pH of the system is returned to neutrality. (B, C) X-ray crystal structures of the Prohead II[13] (B) and Head II[32] (C) capsid states reveal that the N-terminus (red) and E-loop (cyan) become ordered with maturation. Since the N-terminus and the apex of the E-loop have no observable electron density in Prohead II, spheres are used to represent these elements in the capsid (top) and subunit (bottom) structures. In both states, the N-terminus is in the cavity of the capsid, while the E-loop is surface exposed.

within the framework of the 3.6 MDa capsid.[9] In both systems, the dynamic regions are covalently tethered to the complex, but do not appear to significantly interact with the main body. The spectra of the independent regions are virtually identical to their spectra within the context of the system.

The capsid of the Hong Kong 97 (HK97) bacteriophage adopts a 13 MDa structure composed of 420 homogeneous subunits and is a well-studied model system for investigating the maturation of capsids and other macromolecular assemblies. The dodecameric portal complex, which is essential for the packaging and ejection of the genomic double-stranded DNA, is not necessary for correct assembly when the capsid is constructed from ectopically expressed proteins. The capsid is composed of 12 pentamers and 60 hexamers of a single capsid protein, gp5, that are packed in a $T = 7$*l* icosahedral arrangement. The fully mature Head II state of HK97 is a stable particle owing to 420 lysine-asparagine cross-links that catenate closed rings of five or six capsid subunits into a single network resembling chainmail.[10] Head II is the end product of an irreversible maturation pathway that involves many intermediate species (Figure 1A). All the subunit amino acids of Head II are ordered and resolved to the side-chain level in electron density maps, but some regions are not observed in the precursor particles. Extensive X-ray crystallography and cryo-electron microscopy studies have investigated the capsid structure at several stages of the maturation pathway. Dynamic or structurally disordered regions, however, are only implied by these methods and the timing of their ordering is unknown.[11−13]

The autonomous and irreversible HK97 maturation pathway is initiated from the metastable Prohead II capsid state, which is derived from Prohead I upon removal of the N-terminal 102 residues by the gp4 viral protease (Figure 1A).[14,15] In Prohead II, both biochemical and structural studies confirm that several amino acids in the residual N-terminus and E-loop are disordered and susceptible to proteolysis.[11,13,16] The crystal structure of Prohead II at a resolution of 5.2 Å suggests that the N-terminal 15 residues (104−118) and up to 11 residues within the E-loop (160−170) may be dynamic since no electron density is observed for these segments (red and blue spheres, respectively,

(9) Van der Graaf, M.; Kroon, G. J.; Hemminga, M. A. *J. Mol. Biol.* **1991**, *220* (3), 701−9.
(10) Duda, R. L. *Cell* **1998**, *94* (1), 55−60.
(11) Conway, J. F.; Wikoff, W. R.; Cheng, N.; Duda, R. L.; Hendrix, R. W.; Johnson, J. E.; Steven, A. C. *Science* **2001**, *292* (5517), 744−8.
(12) Helgstrand, C.; Wikoff, W. R.; Duda, R. L.; Hendrix, R. W.; Johnson, J. E.; Liljas, L. *J. Mol. Biol.* **2003**, *334* (5), 885−99.
(13) Gertsman, I.; Gan, L.; Wikoff, E. R.; Guttman, M.; Duda, R. L.; Hendrix, R. W.; Komives, E. A.; Johnson, J. E. To be submitted for publication.
(14) Duda, R. L.; Martincic, K.; Hendrix, R. W. *J. Mol. Biol.* **1995**, *247* (4), 636−47.
(15) Conway, J. F.; Duda, R. L.; Cheng, N.; Hendrix, R. W.; Steven, A. C. *J. Mol. Biol.* **1995**, *253* (1), 86−99.
(16) Hendrix, R. W.; Duda, R. L. *Adv. Virus Res.* **1998**, *50*, 235−88.

in Figure 1B).[13] Structural studies on the early intermediate set of precursors (EI-I, EI-II, EI-III, and EI-IV), which are obtained by lowering the pH of the system to 4, also reveal that the disordered regions may persist during maturation for at least a subset of the subunits.[11,17,18] Only in the catenated, fully mature Head II particle are all the residues ordered (Figure 1C).

In this study, we describe the properties and dynamics of the N-terminal residues of the HK97 capsid subunits at different stages of maturation. In the Prohead II and EI-III capsid states, NMR provides valuable information about the flexible N-terminus that could not be observed by cryo-electron microscopy or X-ray crystallography. A subset of N-termini behaves like a flexible peptide attached to a large complex in EI-III. Transient interactions, however, are observed between the N-terminal residues and the inner surface of the capsid in the Prohead II and 5-day old EI-III states. This phenomenon has not been observed in the NMR studies of other large macromolecular complexes reported to date.[7−9]

Materials and Methods

Sample Preparation. ^{15}N and ^{2}H, ^{15}N-labeled samples of wild-type and K169Y mutant Prohead II particles were generated by recombinantly expressing gene products 4 and 5 (gp4 and gp5) of the HK97 genome in M9 minimal media made up in H_2O or 99% D_2O, respectively. M9 media was prepared with $^{15}NH_4SO_4$ as the sole nitrogen source, unlabeled glucose, and supplemental trace metals and vitamins. Expression of the proteins from the pT7-Hd2.9 vector[14] in the BL21(DE3) pLysS strain of *E. coli* results in the in vivo assembly of a Prohead I intermediate that lacks a portal assembly, followed by proteolytic processing into the metastable Prohead II state (Figure 1A). Prohead II was purified from the cell lysate by PEG precipitation, sedimentation, and anion exchange chromatography, as described previously.[19] NMR samples of Prohead II were prepared in 50 mM or 300 mM KCl, 50 mM KH_2PO_4/K_2HPO_4 (pH 7.0), 0.02% NaN_3, and 10% D_2O/90% H_2O or 100% D_2O (pD 7.0).

The EI-III state was generated by slow titration of a solution containing 300 mM KCl and 2 M CD_3COOH (pH 3.6) into a Prohead II sample containing 300 mM KCl, lowering the pH to 4, and incubation at room temperature. The Head II capsid sample was prepared directly from the ^{15}N-labeled EI-III particles, after 2 days of incubation at pH 4, by increasing the pH to 7 with 1.5 M Tris (pH 9.5) buffer. Each of the capsid NMR samples has a final capsid subunit concentration between 1.0 and 1.5 mM (30−45 mg/mL).

The purity and the cross-linked state of the capsid were assessed by polyacrylamide gel electrophoresis. Samples were precipitated in 15% (w/v) TCA at 4 °C to prevent additional crosslinking.[19] The pellets were washed in acetone, dried, and resuspended in SDS sample buffer. Samples were run on a 10% SDS-polyacrylamide gel containing an acrylamide to methylene bisacrylamide ratio of 112:1[20] and stained with Brilliant Blue R dye (Sigma-Aldrich).

A peptide corresponding to the N-terminal 11 amino acids (104−114: SLGSDADSAGS) was synthesized by Sigma-Genosis. Natural abundance NMR experiments were conducted on samples containing 10 mM peptide in the same NMR buffers (pH 4.0 and 7.0) used for the capsid samples.

NMR Spectroscopy. Spectra of the HK97 capsid states were recorded using standard ^{1}H, ^{15}N-HSQC experiments employing water flip-back pulses[21] on a 900 MHz Bruker Avance spectrometer equipped with a triple-axis 5 mm TXI probe. All NMR experiments were acquired at 25 °C, except when the effect of temperature on the resonances was assessed between 20 and 40 °C. Doubling the number of points in the ^{15}N dimension by linear prediction increased the resolution of the spectra. All proton chemical shifts are relative to 2,2-dimethyl-2-silapentane-5-sulfonate.

Assignment of the resonances for the Prohead II and EI-III states required the collection of two and three-dimensional ^{15}N-edited NOESY experiments (T_{mix} = 100−200 ms) on ^{15}N and ^{2}H, ^{15}N-labeled samples. Homonuclear TOCSY experiments (T_{mix} = 51.8 ms), which aided in spin system identification for EI-III and synthetic peptide samples, were collected on a Bruker 500 MHz Avance spectrometer equipped with a TXI cryoprobe. Homonuclear ROESY experiments (T_{mix} = 300 ms), collected on a Bruker 600 MHz Avance spectrometer equipped with a TXI cryoprobe, enabled assignment of peptide resonances. Water suppression for all these experiments was achieved using the WATERGATE scheme.[22] The ^{15}N-edited NOESY experiment also included water flip-back pulses.[21] Unambiguous sequential assignments for both the capsid and synthetic peptide were made based upon NOE correlations between adjacent backbone amide groups and aliphatic protons of residues "i" and "i-1". Published random coil chemical shift values were used to help assign aliphatic resonances.[23−25] For data processing and analysis, the NMRPipe program package[26] and the XEASY[27] program were used.

Diffusion of the Prohead II and EI-III capsid states relative to a standard 2 mM lysozyme sample were measured on 800 or 900 MHz spectrometers equipped with triple-axis 5 mm TXI probes. A series of one-dimensional water sLED experiments were collected with varying x-gradient strengths, ranging from 70 to 330 mT/m.[28] Three individual experiments were collected in this range for Prohead II samples in 100% D_2O. For samples in H_2O, data from three to four spectra were collected with gradient strengths greater than 200 mT/m. The gradient strength range used for the peptide and lysozyme was limited to lower values because of relatively rapid signal attenuation and low signal-to-noise at higher strengths. Natural logarithms of the individual intensities, I_t, were plotted as a function of the square of the gradient strength, and the data from regression lines derived from different resonances were pooled, provided they had statistically similar slopes (F-test, $\alpha = 0.05$) [lysozyme and Prohead II in D_2O]. If the slopes were not significantly similar, the slopes of the three most intense resonances at the highest gradient strength were averaged (peptide, Prohead II, and EI-III in H_2O). The self-diffusion coefficient, D_s, of each species was calculated from the slope using

$$\ln(I_{(t)}) = \ln(I_o) - \gamma^2 G^2 D_s \delta^2 (\Delta - \delta/3)$$

where I_o is the initial intensity, γ is the gyromagnetic constant, G and δ are the magnitude and duration of the field-gradient pulses, and Δ is the delay between gradient pulses. The hydrodynamic radius was calculated from the Stokes−Einstein equation assuming that the HK97 capsid states have a spherical shape.

Nitrogen T_1, T_2 and steady-state heteronuclear NOE relaxation experiments were collected for ^{15}N-labeled Prohead II and EI-III

(17) Lata, R.; Conway, J. F.; Cheng, N.; Duda, R. L.; Hendrix, R. W.; Wikoff, W. R.; Johnson, J. E.; Tsuruta, H.; Steven, A. C. *Cell* **2000**, *100* (2), 253−63.
(18) Gan, L.; Speir, J. A.; Conway, J. F.; Lander, G.; Cheng, N.; Firek, B. A.; Hendrix, R. W.; Duda, R. L.; Liljas, L.; Johnson, J. E. *Structure* **2006**, *14* (11), 1655−65.
(19) Duda, R. L.; Hempel, J.; Michel, H.; Shabanowitz, J.; Hunt, D.; Hendrix, R. W. *J. Mol. Biol.* **1995**, *247* (4), 618−35.
(20) Dreyfuss, G.; Adam, S. A.; Choi, Y. D. *Mol. Cell. Biol.* **1984**, *4* (3), 415−23.
(21) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115* (26), 12593−4.
(22) Piotto, M.; Saudek, V.; Sklenár, V. *J. Biomol. NMR* **1992**, *2* (6), 661−5.
(23) Bundi, A.; Wüthrich, K. *Biopolymers* **1979**, *18* (2), 285−97.
(24) Merutka, G.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **1995**, *5* (1), 14−24.
(25) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5* (1), 67−81.
(26) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6* (3), 277−93.
(27) Bartels, C.; Xia, T.-H.; Billeter, M.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1995**, *5* (1), 1−10.
(28) Altieri, A. S.; Hinton, D. P.; Byrd, R. A. *J. Am. Chem. Soc.* **1995**, *117* (28), 7566−7.

samples and the unlabeled N-terminal synthetic peptide (pH 7) sample using sensitivity enhanced experiments.[29] The EI-III sample was incubated at pH 4 for more than 2 months to ensure a homogeneous population of N-terminal residues. The data were collected on 800 MHz Avance or DRX spectrometers, since both instruments yielded statistically identical results for the synthetic peptide (*T*-test, $\alpha = 0.05$). All experiments were collected at 24.6 °C, which was calibrated using methanol. T_1 relaxation delay times ranged from 10 to 640 ms for Prohead II, and 10 to 1280 ms for ^{15}N-labeled EI-III and the N-terminal synthetic peptide (pH 7). T_2 relaxation delay times ranged from 6 to 82 ms for Prohead II, 6 to 180 ms for ^{15}N-labeled EI-III, and 6 to 842 ms for the N-terminal synthetic peptide (pH 7). Delay time points were collected in a random order to avoid any systematic errors. The T_1 and T_2 values were calculated using a 2-parameter fit in the CurveFit program.[30] The saturated and unsaturated heteronuclear NOE spectra were collected in an interleaved manner for the capsid samples, and the ratio for each resonance was calculated using the signal intensities. Because of quantitative problems associated with slow relaxation, one-dimensional heteronuclear NOE spectra for the synthetic peptide were collected without nitrogen decoupling during acquisition. The average heteronuclear NOE ratio was calculated from the unassignable peak intensities. All spectra were processed using NMRpipe[26] and analyzed using NMRDRAW[26] or NMRVIEW.[31]

Results

Spectra of Viral Capsids. Amino acids at the N-terminus and in the E-loop of Prohead II have no associated electron density in the X-ray crystal structure, implicating static or dynamic disorder of these regions within the crystal.[13] Up to 26 residues are missing in the 5.2 Å resolution structure: 15 N-terminal residues and 11 residues within the E-loop. If there is sufficient flexibility, NMR spectroscopy of the procapsid states should provide information about the dynamics of these regions. ^{1}H, ^{15}N-HSQC spectra were collected for three ^{15}N-labeled viral capsid states of the HK97 bacteriophage (Figure 2). In each spectrum, only a small fraction (0−3%) of the backbone amide groups have an associated resonance. The distribution of the eight prominent resonances in the Prohead II spectrum is limited to the "random coil region" between 8 and 8.7 ppm in the proton dimension (Figure 2A). To assess if these resonances arise from the capsid itself or a proteolytic product within the sample, the NMR spectra of other capsid states were assessed. Transformation of the capsid into the EI-III state upon lowering the pH also reveals eight prominent resonances (Figure 2B). Although the resonances also appear in the unfolded region of the spectrum, their distribution is significantly different from that of Prohead II. The maturation of Prohead II into EI-III requires a change in both the salt concentration and the pH of the system, but an increase of the KCl concentration from 50 mM to 300 mM yields no change in the Prohead II spectrum. The spectral changes are only dependent upon the pH decrease.

In addition to pH, the spectra of the capsid states are dependent upon the molecular structure. Both Prohead II and Head II conformations exist at pH 7, yet maturation of Prohead II into Head II results in the disappearance of all the backbone resonances (Figure 2C). Only weak side-chain amide signals are detected in the Head II spectrum. The X-ray crystal structure of the Head II state supports these observations.[12,32] Continuous electron density is observed from residue 104 to 383, implying all the amino acids in the structure are ordered. Only the two C-terminal amino acids (384 and 385) are not visible in the density.

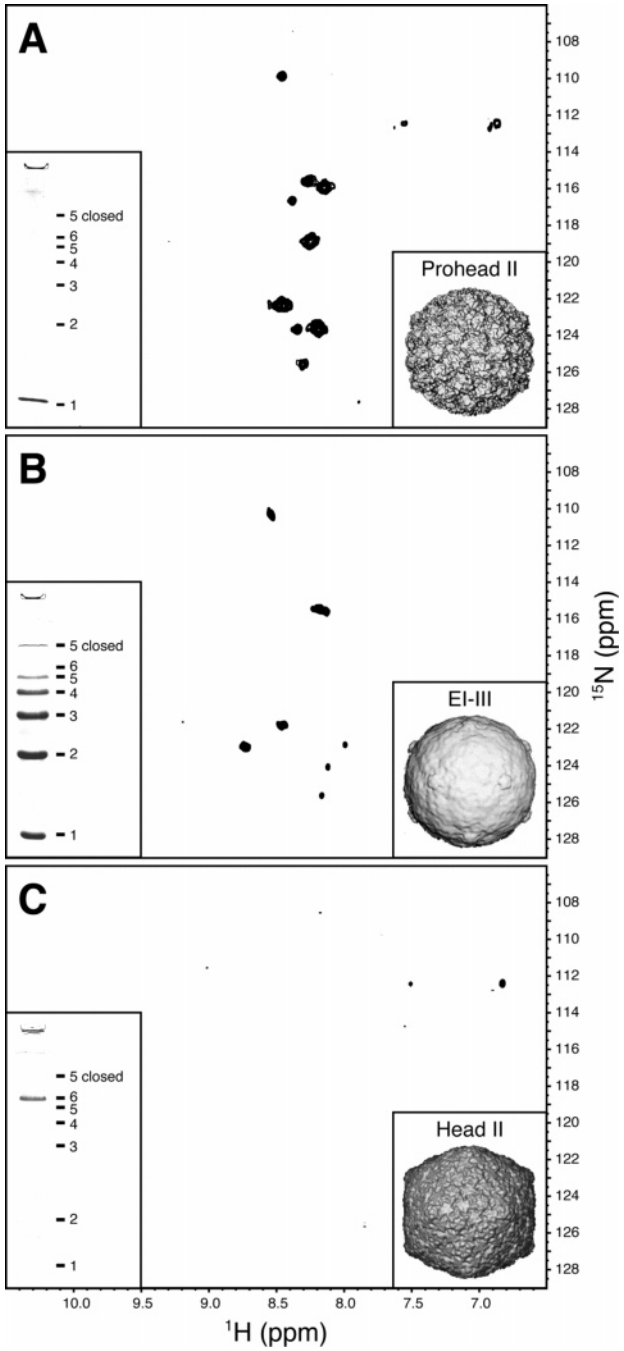

Figure 2. ^{1}H, ^{15}N-HSQC spectra of the ^{15}N-labeled HK97 capsid protein in the Prohead II (A, 40 mg/mL), EI-III (B, 39 mg/mL) and Head II (C, 30 mg/mL) capsid states. All spectra were collected on samples from the same capsid preparation with no additional purification steps. This demonstrates that the resonances do not arise from a contaminant in the system. The same parameters were used in the acquisition and processing of each spectrum, except that four times the number of scans were collected for the Head II experiment. SDS-PAGE gel electrophoresis was used to assess the degree of crosslinking in each capsid sample and is shown within the respective experiment. Also inset is the cryo-electron microscopy (A, B)[11,17] or crystal (C)[32] structures of each state.

(29) 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.; Kay, L. E. *Biochemistry* **1994**, *33* (19), 5984−6003.
(30) Mandel, A. M.; Akke, M.; Palmer, A. G., III. *J. Mol. Biol.* **1995**, *246* (1), 144−63.
(31) Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4* (5), 603−14.
(32) Wikoff, W. R.; Liljas, L.; Duda, R. L.; Tsuruta, H.; Hendrix, R. W.; Johnson, J. E. *Science* **2000**, *289* (5487), 2129−33.

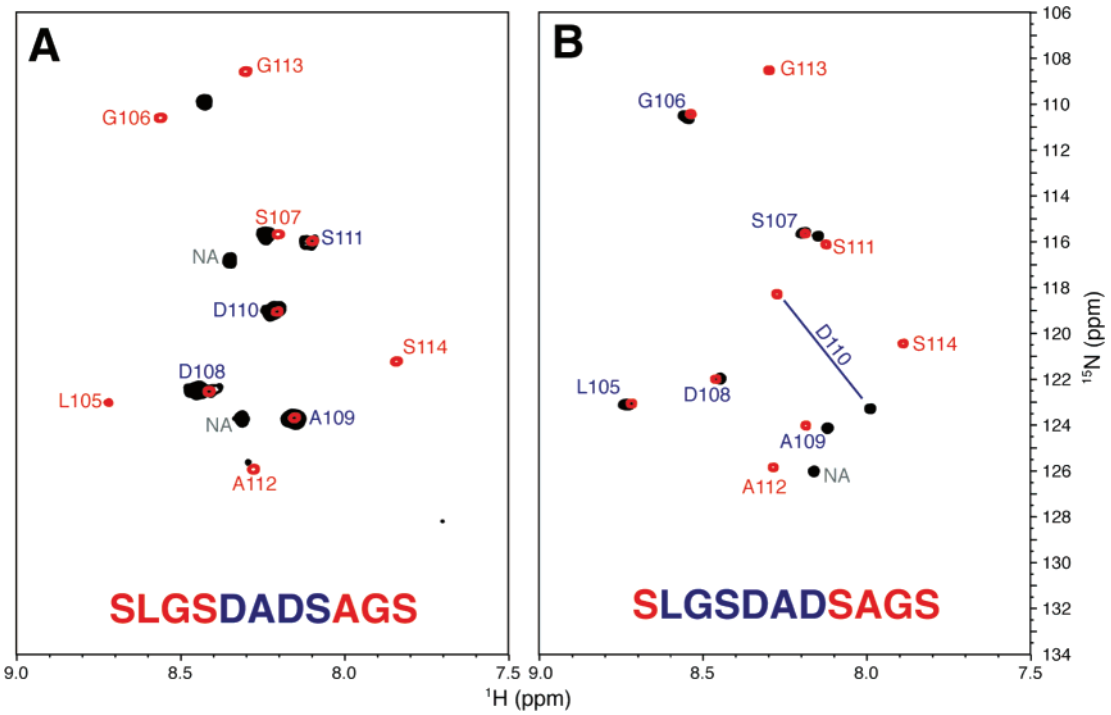

Figure 3. ^{1}H, ^{15}N-HSQC spectra (black) of ^{15}N-labeled Prohead II (A) and EI-III (B) compared to the natural abundance ^{1}H, ^{15}N-HSQC spectra (red) of the N-terminal synthetic peptide in the same buffer. Peptide resonance assignments are labeled in red type, while resonances that can be assigned for both the capsid and peptide are labeled in blue. Capsid resonances that could not be assigned are labeled with "NA".

Assignment of Capsid Resonances. Amino acids in the N-terminus and the E-loop regions are candidates for the observed Prohead II resonances. To discern which element is represented in the spectra, modified Prohead II particles were spectroscopically characterized. Since the unstructured E-loop of Prohead II is susceptible to proteolysis by trypsin between Lys166 and Ala167,[16] the ^{1}H, ^{15}N-HSQC spectrum of a trypsin-digested sample was collected. The resulting spectrum exhibited five new intense peaks in addition to the resonances previously observed for Prohead II, which remain unchanged (Supporting Information). Furthermore, ^{1}H, ^{15}N-HSQC spectra of Prohead II and EI-III that contain the K169Y mutation, which exists near the apex of the E-loop, are identical to those collected for the wild-type capsid states (Supporting Information). The spectra of the Prohead II particles containing the E-loop modifications imply that the resonances of the wild-type procapsid are not associated with the E-loop. In addition, the spectrum of wild type Prohead II is similar to the spectrum of a synthetic peptide composed of the N-terminal 11 amino acids (Figure 3). All these data suggest that the observed resonances are associated with flexible N-terminal residues.

To confirm this hypothesis, resonance assignments were determined for the Prohead II and EI-III capsid states. The small number of spin systems under consideration was amenable to standard homonuclear NMR techniques with ^{15}N-spectral editing. Two and three-dimensional ^{15}N-edited NOESY spectra of protonated and partially deuterated capsid samples were sufficient to provide many assignments. H_N-H_N NOE correlations and intraresidue and sequential H_N correlations to aliphatic resonances led to the assignment of Asp108-Ser111 and Leu105-Asp110 in the Prohead II and EI-III states, respectively (Figure 3). A homonuclear TOCSY experiment provided limited side-chain information for the EI-III capsid state, but was unhelpful for Prohead II.

The N-Terminus Is Sensitive to the Global Capsid Conformation. Homonuclear TOCSY and ROESY experiments enabled the assignment of all the N-terminal synthetic peptide resonances (red and blue labels, Figure 3). ROE correlations for amide protons included those to the previous and subsequent amide protons and aliphatic protons from the corresponding and previous residues. Only intraresidue correlations were observed for the aliphatic protons. No medium or long-range correlations were found, suggesting that the synthetic peptide is unstructured under the conditions in which HK97 particles are studied.

The peptide assignments reveal the unique properties of the N-terminal residues in the Prohead II and EI-III capsid states. First, there is little effect of pH on the ^{1}H and ^{15}N chemical shifts of the synthetic peptide. The most significant chemical shift changes (Δ_H < 0.1 ppm; Δ_N < 1 ppm) are associated with the amide protons of the aspartate residues, the residues C-terminal to aspartate, and the residue at the C-terminus. These assignments also validate the assignments of the capsid resonances, since many of the peaks that coincide in the ^{1}H, ^{15}N-HSQC spectra are assignable to the same amino acid in the sequence (blue labels, Figure 3). Amide resonances of amino acids Asp108 to Ser111 in Prohead II and their analogous peptide resonances have similar chemical shifts, but the spectra are not superimposable. In the spectrum of EI-III, the same correlation is observed for Leu105 to Asp108. Although Ala109 and Asp110 can also be assigned for the capsid state, the positions of their amide resonances are significantly different in the peptide spectrum.

Unlike the EI-III state, the N-terminal most amino acids of Prohead II (Leu105-Ser107) could not be assigned to a particular amide resonance. Also, no capsid resonance correlates in chemical shift with the Leu105 and Gly106 resonances for the synthetic peptide. The capsid resonance that does coincide with that of Ser107 from the peptide could not be unambiguously

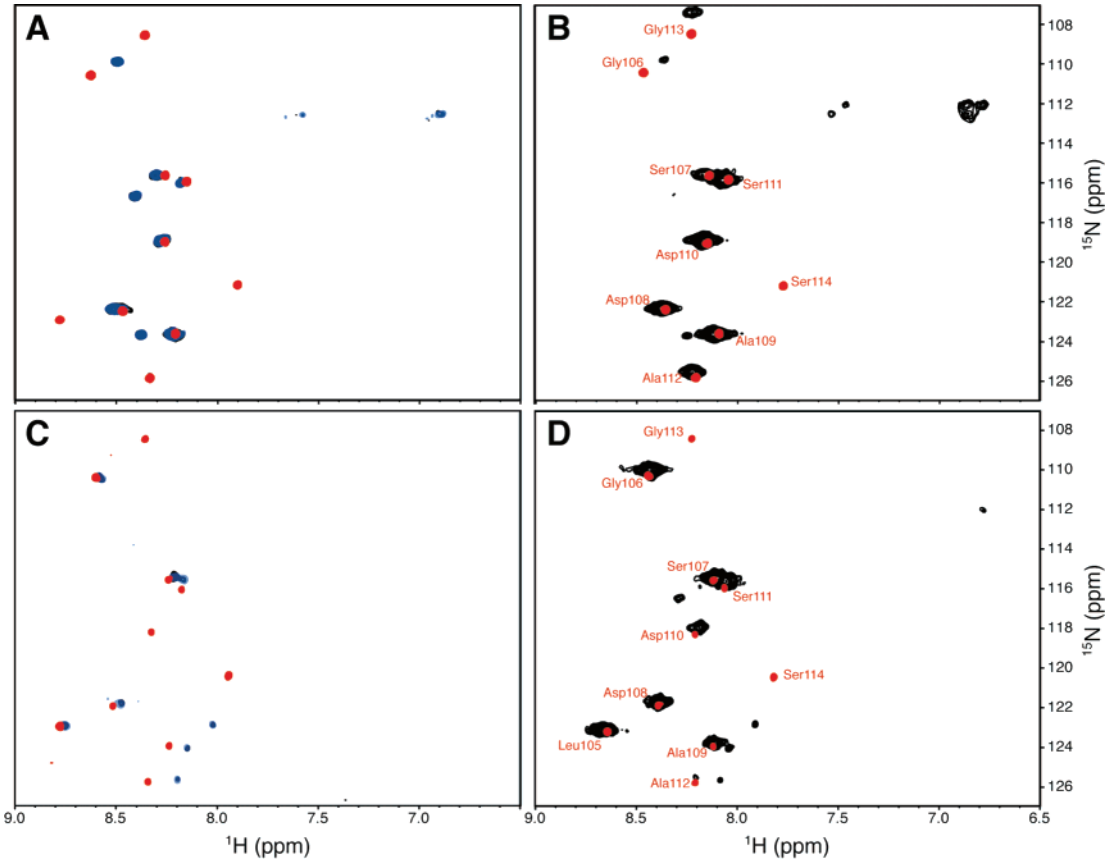

Figure 4. The similarities between the ^{1}H, ^{15}N-HSQC spectra of the Prohead II capsid state (black) and the N-terminal synthetic peptide (red) increase when the temperature is changed from 20 (A) to 40 °C (B). This temperature-dependent change is reversible since decreasing the temperature back to 20 °C yields a spectrum (blue) that is virtually identical to the original spectrum. A similar temperature dependence is observed for the 5-day old EI-III capsid state at 20 °C (C) and 40 °C (D).

attributed to this residue. The assignment of a few resonances could not be made because of ambiguity or a lack of observable NOE correlations (NA, Figure 3). In the EI-III spectrum, only the most downfield resonance in the nitrogen dimension could not be assigned, although a homonuclear TOCSY experiment suggests this resonance is an aspartic acid or asparagine residue.

Effect of Temperature on Capsid Spectra. A dramatic change in the ^{1}H, ^{15}N-HSQC spectrum of the Prohead II state is observed when the temperature of the system is altered (Figure 4). Small changes in the spectrum are observed when the temperature is dropped from 25 to 20 °C, since only the intensities of the assignable peaks (Asp108-Ser111) are reduced (Figure 4A). Increasing the temperature to 40 °C, however, leads to a dramatic increase in their intensities, while the unassignable resonances significantly decrease in intensity (Figure 4B). Resonance volumes of Asp108 to Ser111 increase 2−8 fold, and new signals appear near the positions of the synthetic peptide's Ala112 and Gly113 resonances. Two of the unassignable resonances disappear into the noise of the spectrum, while the intensities of resonances near the position of the peptide's Gly106 and Ser107 signals decrease to less than half the value measured at 20 °C. An additional effect of increasing the temperature of the system is the shifting of capsid resonances corresponding to Ser107 and Asp108 toward peptide chemical shift values. The temperature-dependent spectral changes are reversible since spectra of the capsid at 20 °C are identical both before and after raising the temperature to 40 °C (black and blue resonances, respectively, in Figure 4A).

A similar temperature-dependent phenomenon is observed for the EI-III capsid state, but two distinct sets of resonances can be identified. After a 5-day incubation at room temperature, the characteristic spectrum of the EI-III capsid state is observed at temperatures lower than 25 °C (Figures 2B and 4C). Like Prohead II, increasing the temperature to 40 °C results in spectral changes. The resonances of Leu105 to Asp108 increase in intensity by 3−6 fold relative to the spectrum collected at 20 °C, and resonances appear near the positions of the synthetic peptide's Ala109 and Asp110 signals. In contrast, the EI-III resonances that are assigned to Ala109 and Asp110 are relatively unresponsive to the temperature change, with less than a 40% intensity increase.

The population of N-termini in the EI-III capsid state that is not temperature-sensitive predominates with time. After 70 days, the intensities of all the resonances are nearly invariant with the temperature increase (Supporting Information). The assignable resonances increase in intensity by less than 20% with increasing temperature, and only Leu105 decreases in intensity by 25%. Comparatively, temperature-dependent spectral changes of the synthetic peptide at pH 4 or 7 involve a consistent decrease in signal intensity as the temperature is increased. The intensities of peptide resonances at 40 °C are decreased by less than 50% relative to the 20 °C spectrum, and, at pH 7, the N-terminal Leu105 resonance disappears altogether.

The strength of the hydrogen bonds between amides and water, ligands or other amino acids change with temperature and result in a characteristic shifting of the resonances, which

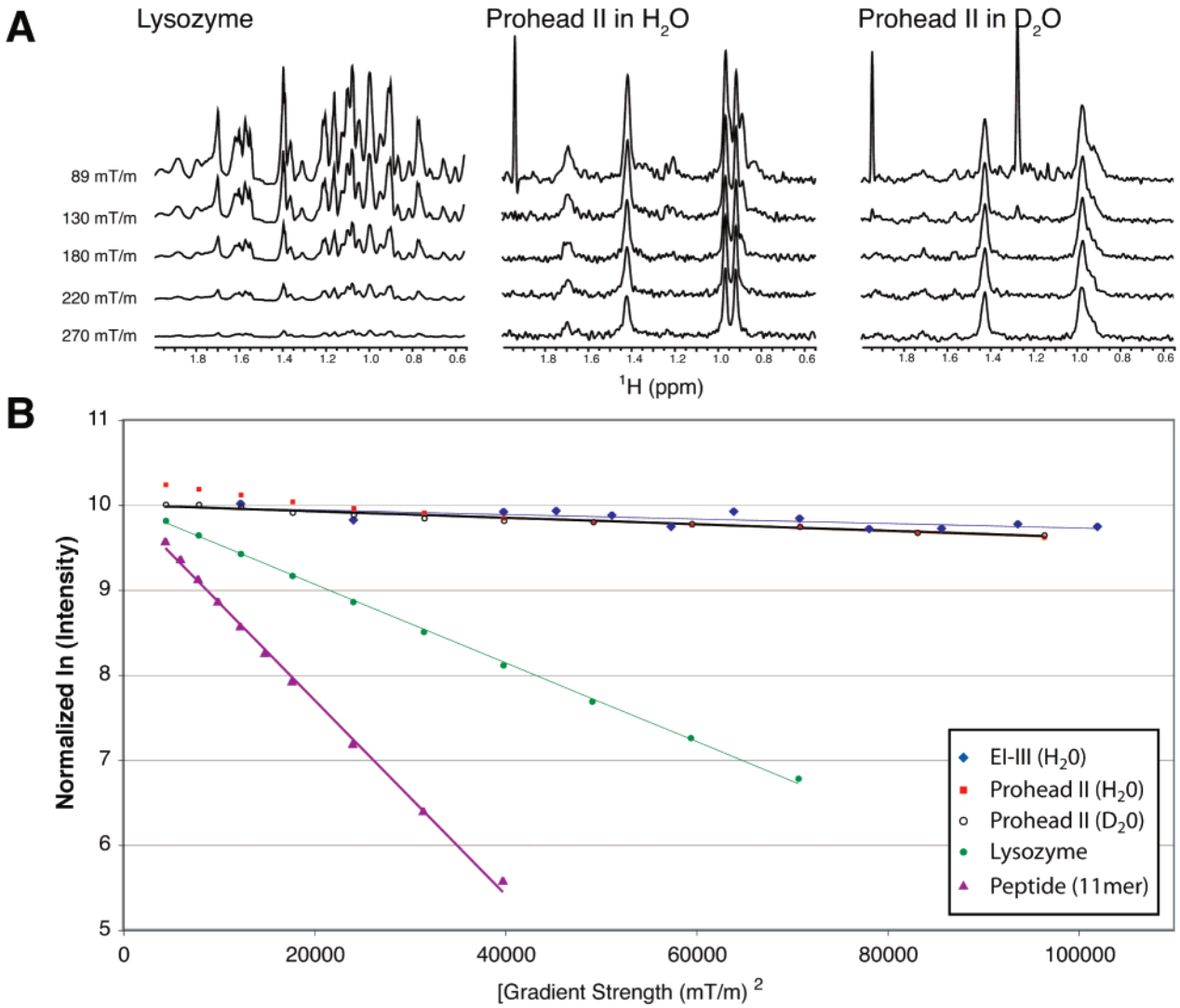

Figure 5. Diffusion results for the Prohead II and EI-III capsid states relative to a standard lysozyme sample and the N-terminal synthetic peptide at pH 7. (A) Aliphatic spectra of a standard 2 mM lysozyme sample and Prohead II samples in H_2O and D_2O (right to left) at different gradient strengths. The lysozyme resonances are attenuated at a much greater rate than those associated with Prohead II. (B) The translational diffusion constant of a molecule is proportional to the slope obtained when plotting the natural log of the intensity as a function of the square of the gradient strength. The small eleven-residue synthetic peptide translates more rapidly than the 14 kDa lysozyme protein. Relative to these species, shallow slopes are obtained for Prohead II in H_2O and D_2O and EI-III. At low gradient strengths, the intensities of some capsid associated resonances deviate from linearity when the species are solvated in 90% H_20/10% D_2O. This phenomenon appears to be associated with the presence of H_2O since it is not observed when the experiment is repeated in D_2O.

is represented by the temperature coefficient.[33,34] The temperature coefficients of the signals observed for the HK97 capsid precursors studied and the N-terminal synthetic peptide range between −5 and −8 ppb/K, values that are typically observed for flexible regions of proteins. Temperature coefficients for Ser107 and Asp108 in Prohead II could not be determined since their chemical shifts do not change linearly with temperature.

NMR Spectroscopy Detects Protons in Capsid Intermediates. It is remarkable that resonances associated with a 13 MDa particle can be observed, and it is important to rule out the presence of artifactual resonances due to peptides arising from proteolysis or degradation of the sample. To demonstrate that resonances of protons associated with the capsid are directly observable by NMR, pulse field gradient experiments were collected that quantitatively assess the rate of translational diffusion.[28] By varying the strength of the gradients applied before and after the fixed diffusion delay, the rate at which the signals disappear is proportional to the rate of diffusion (Figure 5A). Owing to increased spatial resolution, the application of steeper gradients lowers the likelihood that the magnetization of a molecule will refocus after the counter gradient is applied. Lysozyme, which has a known diffusion coefficient, was used as a standard to calculate the diffusion coefficients of the other molecules tested. Fast diffusion of the 11-residue N-terminal synthetic peptide, for instance, results in a rapid attenuation of associated signals, which yield a measured diffusion coefficient of $2.67 \times 10^{-6} \pm 4 \times 10^{-8}$ cm^2/s. For Prohead II and EI-III in H_2O, several signal intensities appear to plateau at large gradient strengths, while other signals (i.e., 1.95 ppm, Figure 5A) disappear at much lower gradient strengths. These rapidly attenuated resonances are likely associated with trace buffer components and may contribute to the deviation of procapsid signals from linearity at low gradient strengths. Representative spectral data for Prohead II in H_2O and D_2O relative to the 2 mM lysozyme standard sample is shown in Figure 5A.

(33) Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1988**, *201* (1), 161−200.
(34) Cierpicki, T.; Otlewski, J. *J. Biomol. NMR* **2001**, *21* (3), 249−61.

Table 1. Summary of Nitrogen Relaxation Data for the N-Terminal Amides in Different Precursor States of the HK97 Capsid

	Prohead II state	EI-III state	synthetic peptide (pH 7)[a]
T_1	640 ± 90 ms	800 ±100 ms	710 ± 70 ms
T_2	80 ± 40 ms	140 ± 40 ms	880 ± 90 ms
hNOE[b]	0.1 ± 0.2	−1.2 ± 0.2	−1.3 ± 0.3

[a] The C-terminal amide is excluded since its T_1 and T_2 values are significantly larger than that of the others, and it does not occur naturally in the capsid state. [b] Heteronuclear NOE relaxation results.

Because of low signal-to-noise, large errors are associated with the diffusion coefficients for the Prohead II and EI-III particles. For Prohead II in H_2O, the slopes associated with the three most intense resonances at the largest gradient strength yield an average diffusion coefficient of $1.1 \times 10^{-7} \pm 0.4 \times 10^{-7}$ cm^2/s. Application of the Stokes−Einstein equation, which relates molecular size with translational diffusion, reveals that this diffusion rate is associated with a molecule that has a hydrodynamic radius of 230 ± 80 Å. In D_2O, the diffusion coefficient and predicted radius of Prohead II are $8.9 \times 10^{-8} \pm 0.5 \times 10^{-8}$ cm^2/s and 280 ± 30 Å, respectively. These values corroborate the small-angle X-ray scattering (SAXS) results that reveal the average outer-radius of Prohead II is 247 Å.[35] Similarly, the hydrodynamic radius obtained for the EI-III state in H_2O is 210 ± 90 Å, which is consistent with the SAXS-determined average radius of 280 Å. These data strongly argue that the observed resonances for Prohead II and EI-III states arise from flexible residues associated with the capsid.

The comparison of Prohead II in H_2O and D_2O demonstrates the importance of focusing on the intensities at high gradient strengths (>200 mT/m) when measuring diffusion data for large particles in H_2O. The nonlinear phenomenon is not observed for the synthetic peptide or lysozyme and may partially arise from smaller, quickly diffusing components that complicate the spectrum at lower gradient strengths (Figure 5A). Nevertheless, water also appears to contribute additional intensity to the observed signals, likely via a nuclear Overhauser effect. Diffusion of Prohead II in a D_2O buffer results in a linear trend over the entire gradient strength window. Translational diffusion data for EI-III in D_2O could not be collected because the particle is unstable in the necessary buffer conditions.

Comparison of Nitrogen Relaxation Results. To better understand the properties of the N-terminal amino acids within the context of the Prohead II and EI-III states, nitrogen T_1 and T_2 relaxation measurements were compared to those acquired for the N-terminal synthetic peptide at pH 7. In accordance with its small size and expectedly rapid correlation time, the N-terminal peptide has both large T_1 and T_2 values (Table 1). The properties observed for the Prohead II and EI-III capsid states, however, are considerably different. The T_1 relaxation terms are similar to that of the synthetic peptide, but the T_2 relaxation values are considerably smaller. Despite the expectedly similar context of the N-terminal amino acids in the Prohead II and EI-III states, it is important to note that the T_2 relaxation times for Prohead II are significantly smaller than those associated with the EI-III state. Heteronuclear NOE values for the amide residues reveal that this difference is predominantly associated with the flexibility of the N-terminus. As predicted for flexible regions, the amides in the synthetic peptide and EI-III capsid state have a large negative heteronuclear NOE ratio, but the ratios for Prohead II are small and positive. This suggests that the N-terminus in the Prohead II state is more restrained than the equivalent residues in the EI-III state.

Discussion

Nuclear Magnetic Resonance Studies of Viral Capsids. The HK97 bacteriophage capsid is an excellent model system for studying macromolecular assembly and the intermolecular communication necessary for complex maturation. A total of 420 protein subunits irreversibly assemble into a homo-multimeric procapsid shell in vivo, and discrete maturation intermediates can be obtained by simple manipulation of the buffer conditions.[19,36] X-ray crystallography and cryo-electron microscopy have made much progress in determining the structures of several HK97 maturation intermediates.[11−13,15,17,19,32,36] This information provides an understanding of which molecular motions, adaptations, and folding events are necessary to progress along the in vitro maturation pathway. In the fully mature Head II state, the N-terminal amino acids and the E-loop residues make key inter-subunit contacts.[12,32] The E-loop contains lysine 169, one of the residues involved in the inter-subunit covalent bonds that catenate the structure. The N-terminal residues are involved in a complex network of hydrogen bonds with one E-loop and bridge it to an E-loop from a third subunit via polar contacts. The role and behavior of these critical regions in Prohead II and other early intermediates could not be addressed by these methods.

Nuclear magnetic resonance spectroscopy enabled the study of the N-terminus of the HK97 subunit at early stages of maturation. Since the viral complex has a mass of 13 MDa and an estimated correlation time of 14 μs at 25 °C, it is not surprising that most of the 282 expected resonances are too broad to be observed in a ^{1}H, ^{15}N-HSQC spectrum. In both the Prohead II and EI-III states, only resonances associated with the flexible N-termini from at least a fraction of the subunits are found. A combination of resonance properties and nitrogen relaxation, NOESY, and translational diffusion experiments demonstrate that resonances are associated with the capsid state and not a peptide contaminant within the system. In addition, maturation of the system to Head II results in the disappearance of all the backbone amide signals observed for Prohead II (Figure 2C). While a peptide contaminant may undergo spectral changes upon a pH change, the spectrum at a certain pH is expected to be independent of the capsid state.

Behavior of Dynamic N-Terminus Varies in Prohead II and EI-III Capsid States. NMR spectroscopy of Prohead II and EI-III reveals that the flexible N-termini in the two states behave quite differently (Figure 6). The degrees to which the N-termini interact with the inner surface of the capsid result in spectra that vary in temperature-dependence and similarity to the spectrum of the analogous synthetic peptide. These interactions with the capsid shell predominate in Prohead II (Figure 6A). The results for EI-III, on the other hand, are consistent with the behavior of a flexible, "peptide-like" free-end of a protein that is attached to a large particle (Figure 6B). A comparison of the EI-III and synthetic peptide (pH 4) ^{1}H, ^{15}N-HSQC spectra reveals that the amide resonances of the N-

(35) Lee, K. K.; Gan, L.; Tsuruta, H.; Hendrix, R. W.; Duda, R. L.; Johnson, J. E. *J. Mol. Biol.* **2004**, *340* (3), 419−33.

(36) Gan, L.; Conway, J. F.; Firek, B. A.; Cheng, N.; Hendrix, R. W.; Steven, A. C.; Johnson, J. E.; Duda, R. L. *Mol. Cell* **2004**, *14* (5), 559−69.

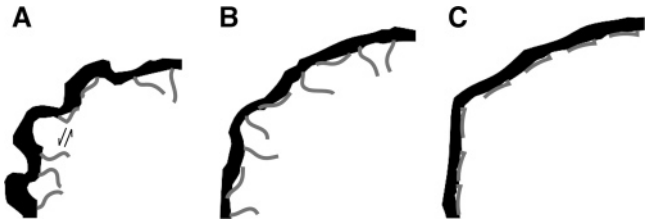

Figure 6. Model of N-terminus dynamics in Prohead II (A), EI-III (B), and Head II (C). Capsid outlines are derived from the structures presented in Figure 2. In Prohead II, the N-terminal residues are in equilibrium between ordered and flexible states because of transient interactions with the inner surface of the capsid (A). The equilibrium is shifted toward the flexible state in EI-III particles, but this may apply to only a fraction of the N-terminal residues (B). The N-terminal residues of all subunits are ordered in the fully mature Head II state (C).

terminal most residues superimpose, while the resonances of amino acids closer to the capsid surface have a larger chemical shift difference (Figure 3). The most C-terminal observable resonance is assigned to Asp110, which shifts 0.3 ppm upfield and 5 ppm downfield in the proton and nitrogen dimensions, respectively. The resonances of matured EI-III samples experience small changes in intensity with changing temperature. Heteronuclear NOE results corroborate this model since the large and negative ratios are typical of flexible termini in proteins.

The spectral properties of Prohead II are consistent with the N-terminus of the subunits interacting with the capsid surface (Figure 6A). Only residues Asp108 to Ser111 are assignable and correlate well with the peptide amide resonances. Residues both N- and C-terminal of this region could not be assigned and deviate from the peptide resonance positions in the ^{1}H, ^{15}N-HSQC spectrum. Since residues 108−111 maintain "peptide-like" chemical shifts, one or both of the N-terminal serine residue and the adjacent leucine likely mediates the interaction. An increase in temperature to 40 °C is enough to partially disrupt the contact, since resonances of many residues increase dramatically in intensity and shift to the positions of the corresponding synthetic peptide signals. This suggests that the N-terminus becomes more "peptide-like" at higher temperatures. The dramatic decrease in the intensity of the resonance near Gly106 may result from an increased rate of chemical exchange, which results in broader resonances. Chemical exchange may also impede the assignment of resonances to the N-terminal most amino acids.

Nitrogen relaxation measurements also support the transient interaction model for Prohead II. The T_2 relaxation values for resonances associated with the Prohead II state are almost half the values measured for the EI-III state (Table 1). The decrease in flexibility is also highlighted by the small and positive heteronuclear NOE ratios calculated for the observed resonances. Since the N-terminal amino acids have a similar context in both the Prohead II and EI-III states, nearly identical relaxation results would be expected if the interactions did not occur.

The transient interactions of the N-terminus with the capsid shell are also observed in the transition of Prohead II into EI-III upon acidification of the sample. Resonance positions for the EI-III capsid state are constant over time, and consistent with a population of N-termini that are unrestrained. Nevertheless, heating a 5-day old EI-III sample to 40 °C reveals that a significant fraction of N-termini are initially making such interactions. The resonances of residues 105−108 increase in intensity and new resonances appear at the positions of Ala109, Asp110, and Ser111 for the synthetic peptide. The resonances specifically assigned to Ala109 and Asp110 within the EI-III capsid state, however, differ in chemical shift from the peptide values, are insensitive to temperature changes, and typify a distinct population of N-termini that increases with time. These resonances are still associated with the capsid since their chemical shifts and nitrogen T_2 relaxation values differ considerably from those of the analogous peptide (Figure 3, Table 1).

Heterogeneity within Viral Capsids Structures. The model of N-terminal dynamics in the immature states of HK97 implies that there is more disorder in the N-terminus at a later stage of maturation. Typically, as complexes of molecules assemble and mature, there is a trend of increasing order and decreasing dynamics. This is in part the overall trend in HK97 maturation from Prohead II to Head II, since all the residues are ordered in the fully matured state. The apparent increased flexibility in the transition from Prohead II to EI-III, though, likely only occurs for a subset of the N-termini (Figure 6B). Temperature-sensitive and insensitive populations of N-termini are present in EI-III samples, revealing that a significant fraction of N-termini are interacting with the capsid surface after short incubation times at pH 4. Both populations are capable of assuming the fully mature Head II state after reneutralization (Figure 2C). Although the population of flexible, temperature insensitive N-termini increases with time, other N-termini may adopt ordered conformations that are unobservable by NMR. Far-ultraviolet circular dichroism studies reveal an increase in β-strand and type I β-turn content in the transition from Prohead II to EI-III, and the N-terminal residues are the only candidates to adopt these structures.[35] Additional order is observed in the number of N-terminal residues that are detected by NMR. The ^{1}H, ^{15}N-HSQC spectrum of Prohead II at 40 °C reveals amide resonances that correspond to residues 107 to 113, while amino acids C-terminal of Asp110 cannot be assigned to resonances in the spectrum of the EI-III state.

Biochemical studies demonstrate a large degree of nonicosahedrally arranged crosslinking of capsid proteins in EI-III, which likely results in the inhomogeneity of N-terminal residues. The ladder of bands in the SDS-PAGE gels indicates EI-III is composed of a mixture of all possible linear and covalently closed oligomers (Figure 2B). Additional evidence of inhomogeneity arises from the 7.5 Å resolution crystal structure of the EI-IV capsid state.[18] Two of the seven N-termini in the icosahedral asymmetric unit are considerably more disordered than the others.

The icosahedral asymmetric unit is an additional source of inhomogeneity that likely affects the NMR data of both the Prohead II and EI-III states. The asymmetric unit of HK97 consists of seven proteins, each in a unique microenvironment. One protein is associated with the penton capsomere (subunit G), while the other six (subunits A−F) comprise an asymmetric hexon capsomere in Prohead II and EI-III.[15,17] In addition to positional differences, X-ray crystallography reveals that the N-termini of different subunits within the Prohead II asymmetric unit have varying degrees of disorder.[13] The A subunit has continuous density between residues 119 and 130, and additional density extending to residue 114. Based upon the degree of electron density discontinuity between residues 119 and 130, the disorder of the N-termini in other subunits is greater, but dependent upon their position within the asymmetric unit. Since

the N-terminal amino acids make transient interactions with the capsid surface in different environments, it is not surprising that the Prohead II N-terminal amino acids may have slightly different chemical shifts. This could account for the nonuniform resonance shapes that we observe in the spectra of Prohead II (Figure 2A). Resonances of the EI-III state, on the other hand, are uniform in shape. This may result from the increased flexibility of the N-terminus, but may also arise if only a fraction of the N-termini in the asymmetric unit are observable by NMR.

Different environments within the asymmetric unit may also account for the appearance of the resonances that cannot be assigned. The one unassignable residue in the ^{1}H, ^{15}N-HSQC spectrum of the EI-III state appears to be an aspartic acid or asparagine residue, based upon homonuclear TROSY data. Although Asp108 and Asp110 are assignable to other resonances, this resonance may correspond to one of these residues at a different position within the asymmetric unit. Asp161 in the E-loop is the only other candidate in a flexible region that has not been assigned to a resonance. In this case, though, resonances arising from more apical amino acids in the E-loop would also be expected. The ability to unambiguously assign this resonance is hampered by the similar proton chemical shift of Ala109 ($\Delta < 0.5$ ppm).

Conclusion

In this work, we have shown that NMR spectroscopy can play a significant role in understanding the behavior of essential flexible regions in large macromolecular complexes. Although structure determination of large complexes by cryo-electron microscopy and X-ray crystallography is common, both methods are unable to observe dynamic regions within the individual subunits. The structures of several maturation intermediates of the HK97 bacteriophage capsid have been determined by these methods, but neither provided information about the residues in the disordered N-terminus and the E-loop regions. Both regions are ordered and make essential stabilizing contacts in the fully mature viral state, but their role and behavior in earlier maturation stages is unknown. Through the use of NMR, several resonances from the N-terminus of the capsid subunits in HK97 maturation intermediates were identified. In the Prohead II capsid state, the combination of nitrogen relaxation and temperature dependence data revealed that the N-terminus transiently interacts with the capsid surface. The relaxation values are small and the heteronuclear NOE values are near zero, consistent with relaxation properties of a sizable protein. An increase in temperature favors the unbound state of the N-terminus since the resonance signals become more intense and correlate better with peptide resonances. N-terminal residues of EI-III state, on the other hand, are considerably more flexible in at least a fraction of the subunits. The long nitrogen relaxation times and negative heteronuclear NOE values are consistent with flexible regions in most proteins. This study demonstrates that the utility of NMR can be extended to provide important information about the behavior of flexible regions in particles that are as large as 13 MDa in weight.

Acknowledgment. We thank John Chung and Gerard Kroon for helpful discussions, suggestions, and NMR support. We also thank Mirko Hennig, Gabriela Perez-Alvarado, Ilya Gertsman, Kelly Lee, and Rick Huang for their insightful contributions to the project. We further thank Gerard Kroon for helpful comments on the manuscript. This research is funded by a CIHR fellowship to B.R.S. and grants from the NIH, Grant AI-40101 to J.E.J. and Grant GM-53320 to J.R.W.

Supporting Information Available: Comparisons of ^{1}H, ^{15}N-HSQC spectra that demonstrate the observed resonances are associated with the N-terminal amino acids and temperature change has no effect on the spectral properties of fully matured EI-III particles. This material is available free of charge via the Internet at http://pubs.acs.org.

JA071118J