

Article

Not peer-reviewed version

The Temperature Dependence of Hydrogen Bonds Is More Uniform in Stable Proteins: Analysis of NMR $^h3J_{NC}$ Couplings in Four Protein States

[Andrei T Alexandrescu](#) * and Aurelio J Dregni

Posted Date: 21 May 2024

doi: 10.20944/preprints202405.1360.v1

Keywords: protein folding; m-value; chi-1 angle temperature dependence; structure conservation; NMR structure



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

The Temperature Dependence of Hydrogen Bonds Is More Uniform in Stable Proteins: Analysis of NMR $^3\text{J}_{\text{NC}}$ Couplings in Four Protein States

Andrei T. Alexandrescu ^{1,*} and Aurelio J. Dregni ²¹ Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269, USA² Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

* Correspondence: Andrei T. Alexandrescu; Tel: (860) 486-4414, Fax: (860) 486-4331, E-mail: andrei.alexandrescu@uconn.edu

Abstract: Long-range HNCO NMR spectra for proteins show crosspeaks due to $^1\text{J}_{\text{NC}}$, $^2\text{J}_{\text{NC}}$, $^3\text{J}_{\text{NCg}}$ and $^3\text{J}_{\text{NC}}$ couplings. The $^3\text{J}_{\text{NC}}$ couplings are transmitted through hydrogen bonds and are correlated to hydrogen bond lengths. We collected long-range HNCO data at a series of temperatures for four protein states. P22i and CUS-3i are six-stranded beta-barrel I-domains from phages P22 and CUS-3 that share less than 40% sequence identity. The *cis* and *trans* states of the C-terminal domain from pore-forming toxin hemolysin II (HlyIIC) arise from isomerization of a single G404-P405 peptide bond. For P22i and CUS-3i the hydrogen bonds detected by NMR agree with those observed in the corresponding domains from cryoEM structures of the two phages. Hydrogen bond lengths derived from the $^3\text{J}_{\text{NC}}$ couplings, however, are poorly conserved between the distantly related CUS-3i and P22i domains and show differences even between the *cis* and *trans* states of HlyIIC. This is consistent with hydrogen bond lengths being determined by local differences in structure rather than the overall folding topology. With increasing temperature hydrogen bonds typically show an apparent increase in length, which has been attributed to protein thermal expansion. Some hydrogen bonds are invariant with temperature, however, while others show apparent decreases in length suggesting they become stabilized with increasing temperature. Considering the data for the three proteins in this study and previously published data for ubiquitin and GB3, lowered protein folding stability and cooperativity correlates with a larger range of temperature responses for hydrogen bonds. This suggests a partial uncoupling of hydrogen bond energetics from global unfolding cooperativity as protein stability decreases.

Keywords: protein folding; *m*-value; chi-1 angle temperature dependence; structure conservation; NMR structure

1. Introduction

Hydrogen bonds (H-bonds) are critical components of protein secondary and tertiary structure. A main driver of protein secondary structure is that backbone H-bonds compensate for the partially polar character of the protein mainchain when it crosses through the hydrophobic core of a protein [1]. H-bonds involving sidechains and/or solvent play additional roles in protein structure specificity, stability, and function. Individually, a H-bond confers only about 1-2 kcal/mol to the stability of the protein, the weakness of the interaction allowing it to be easily formed and broken under physiological conditions [2–4]. However, the tens to hundreds of H-bonds in a typical protein, together with other non-covalent interactions, provide the enthalpic stabilization of the structure that counteracts the loss of conformational entropy accompanying protein folding. The degree to which individual H-bonds contribute additively to protein stability or as part of a cooperative network of noncovalent interactions remains an open question [5–7].

Initially, H-bonds were viewed as an electrostatic interaction between a hydrogen bonded to an electronegative donor atom and a second electronegative acceptor atom [3,8]. In the 1930's Pauling suggested that H-bonds have partial covalent character (first edition of [2]). This was finally confirmed in the late 1990's by NMR detection of J-couplings transmitted through H-bonds [9–14]

and Compton X-ray scattering anisotropy in ice [3,15]. In proteins, direct NMR investigations of H-bond couplings are usually carried out with the long-range HNCO (lrHNCO) experiment that measures a three-bond ${}^{\text{H}}\text{J}_{\text{NC}'}$ through-hydrogen-bond scalar coupling between amide ${}^{15}\text{N}$ and carbonyl ${}^{13}\text{C}'$ nuclei [9]. The ${}^{\text{H}}\text{J}_{\text{NC}'}$ couplings are small with values typically >1 Hz, requiring high-sensitivity and usually necessitating perdeuteration of proteins larger than ~ 10 kDa [13]. However, the ${}^{\text{H}}\text{J}_{\text{NC}'}$ couplings provide unambiguous identification of H-bond donors and acceptors, without requiring knowledge about the rest of the molecular structure. This contrasts with hydrogen exchange protection studies [16] that identify the H-bond donor but not the acceptor. Thus, our lab and others have found through-H-bond couplings extremely useful restraints in protein NMR structure calculations [17–23].

Another interesting but lesser explored application of ${}^{\text{H}}\text{J}_{\text{NC}'}$ couplings is to investigate how structure varies with conditions – a unique strength of NMR in structural biology. Arguably ${}^{\text{H}}\text{J}_{\text{NC}'}$ couplings because of their short distance range are less complicated by averaging and dynamics than other types of NMR restraints such as RDCs, dihedral restraints, and NOEs (that can also be confounded by spin diffusion). The detection of a ${}^{\text{H}}\text{J}_{\text{NC}'}$ coupling unambiguously establishes that an H-bond between 2.5 and 3.5 Å is present, with a relatively sensitive and straightforward dependence on backbone H-bond length [24,25]. By contrast, side-chain H-bonds are seen infrequently in lrHNCO experiments, possibly due to their increased flexibility, and have a more complex dependence on H-bond length [24]. Studies on the variability of H-bonds have investigated how ${}^{\text{H}}\text{J}_{\text{NC}'}$ couplings respond to pressure [26], kosmotropic solvents like trifluoroethanol (TFE) [27], and ligand binding [28].

To date there have been only two studies of the temperature dependence of ${}^{\text{H}}\text{J}_{\text{NC}'}$ couplings, for the proteins ubiquitin [29] and GB3 [30]. Both proteins showed a small weakening of ${}^{\text{H}}\text{J}_{\text{NC}'}$ couplings with increasing temperatures that was interpreted in terms of lengthening of $\text{N}\cdots\text{O}$ H-bond distances due to the thermal volume expansion of the proteins [29]. In ubiquitin, a few individual residues showed different behavior than the average H-bond expansion, characterized by H-bonds that became stronger or were invariant with temperature. These exceptions were attributed to localized partially autonomous folding regions at the ends of regular secondary structure that became more stable with increasing temperature [29]. Both the ubiquitin and G3 proteins are small (< 8.5 kDa), have the same protein folding motif, and are very stable proteins ($T_{\text{m}} \geq 80$ °C). We therefore wanted to investigate how generalizable the temperature dependence of H-bonds is to proteins with other structural motifs and stabilities. Since we were interested in the conservation of H-bonds between proteins with similar structures [25], we selected two pairs of structures for these studies (Figure 1). P22i and CUS-3i are autonomously folding “insertion domains” from the coat proteins of the related bacteriophages P22 and CUS-3 [21]. The two domains have 40% sequence identity and the same overall 6-strand b-barrel fold, although there are structural differences in the lengths of b-strand elements and intervening loops [21]. The *cis* and *trans* states of the C-terminal domain of hemolysin II (HlyIIC) are two slightly different structural forms of the same protein sequence brought about by isomerization of the G404-P405 peptide bond. The HlyIIC domain has a pseudo-barrel fold comprised of five beta strands and two α -helices [19]. Since the *cis* and *trans* forms are in slow exchange on the NMR timescale, the two closely related structures which differ mainly in the position of the proline-bearing loop, give separate NMR signals for about half of the residues in the protein [31]. The four protein states considered in this work were used to probe the conservation of H-bonds and their dependence on temperature in distantly and closely related protein structures.

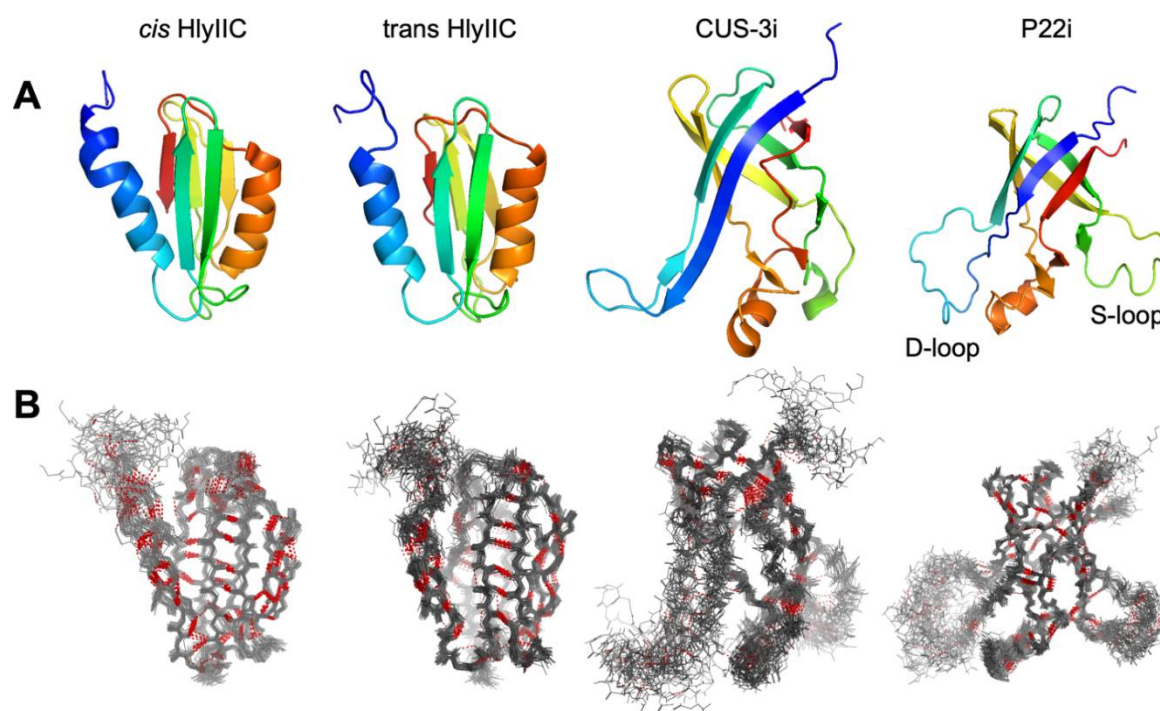


Figure 1. Solution structures of the four protein states studied in this work. (A) Cartoons of the NMR conformer closest to the ensemble average, shown with a color ramp from the N (blue) to the C-terminus (red). For HlyIIC, the largest difference between the *cis* and *trans* structures [31] is in the loop between helix a2 (orange) and strand b5 (red). The D- and S-loops in the P22i structure are dynamically disordered [43]. In Cus3i, the segment corresponding to the P22i loop becomes structured to form an extension of the b1-b2 hairpin [21]. (B) NMR backbone ensembles of the four protein states. Backbone H-bonds identified with the PyMol program are shown with dotted red lines.

2. Results and Discussion

2.1. IrHNCO Experiments Detect H-bond $^3J_{\text{NC}'}^{\text{H}}$ and Sidechain $^3J_{\text{NCg}}$ Scalar Couplings

Figure 2 shows representative data from IrHNCO experiments obtained at 5 different temperatures for the protein HlyIIC. In Figure 2A, ^{15}N planes near 116.5 ppm are superposed for the different temperatures. In this portion of the spectrum, none of the six residues have resolvable signals from the *cis* and *trans* states of the protein since they are far away in the structure from the site of isomerization, P405. However, separate *cis* and *trans* signals are observed for K355 in Figure 2B. Four types of correlations are observed in the IrHNCO spectra (Figure 2A).

The first type of crosspeak, is an ~ 11 Hz: $^1J_{\text{NC}}$ coupling connecting the amide nitrogen with the carbonyl of the preceding residue [32], exemplified for M349, F384, V390. These couplings are detected in standard HNCO experiments but are suppressed in the IrHNCO experiment by setting the N to C' dephasing time $2T$ to a value of 133 ms corresponding to $2/(^1J_{\text{NC}})$ [9]. For comparison a value of ~ 16 ms corresponding to $1/(2(^1J_{\text{NC}}))$ is used in the standard HNCO experiment [33]. Despite being suppressed, strong $^1J_{\text{NC}'}$ couplings can nevertheless persist in the IrHNCO experiment (Figure 2). The $^1J_{\text{NC}}$ couplings are correlated with the strengths of H-bonds [32,34,35].

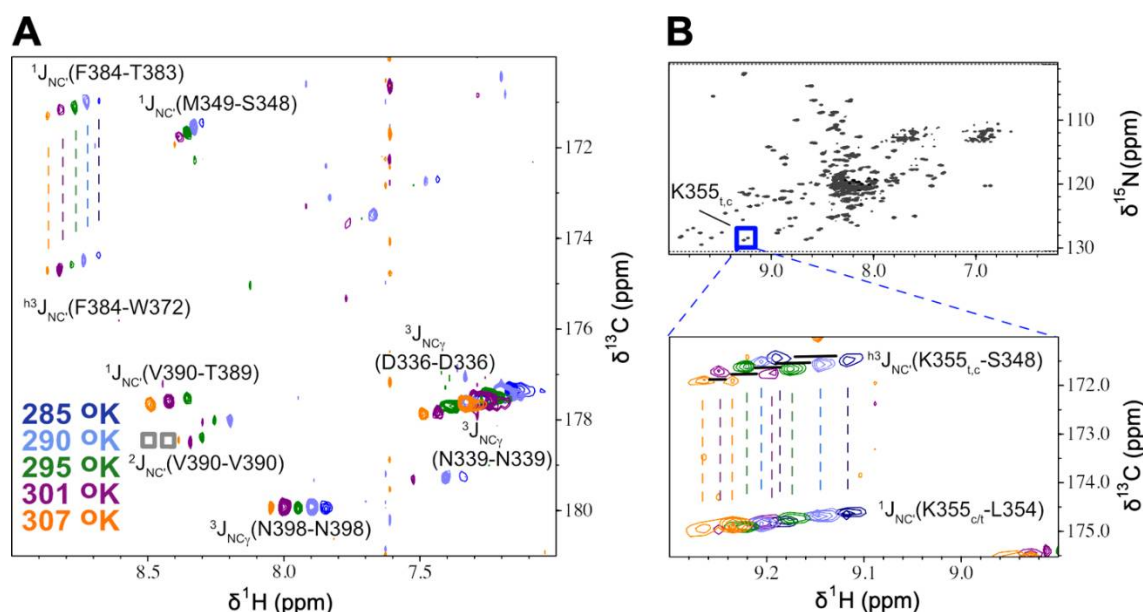


Figure 2. Superposed ^{15}N planes from IrHNCO spectra for HlyIIC at five temperatures. (A) ^{15}N plane at 116.6 ppm illustrating the various types of correlations observed: $^1\text{J}_{\text{NC}}$, $^{\text{h}3}\text{J}_{\text{NC}}$, and $^3\text{J}_{\text{NCg}}$. Two weak intrasidue $^2\text{J}_{\text{NC}}$ crosspeaks are also present in this plane but at contour levels lower than shown (gray boxes near 8.5 ppm). (B) ^1H - ^{15}N HSQC of HlyIIC highlighting crosspeaks due to *trans* (left) and *cis* (right) states of the protein. The expansion shows superpositions of 3D IrHNCO and reference HNCO planes at ^{15}N 128.4 ppm for five different temperatures, using the same coloring scheme for contour levels as in A. At each temperature a pair of $^1\text{H}_{\text{N}}$ resonances is observed making it possible to investigate the temperature dependence of $^{\text{h}3}\text{J}_{\text{NC}}$ couplings for K355 in both the *cis* and *trans* states. There were ten residues in HlyIIC for which separate mainchain $^{\text{h}3}\text{J}_{\text{NC}}$ couplings could be resolved from the *cis* and *trans* states and fifteen for which NMR signals from the two conformations were unresolved (Table S4, S6).

The second type of crosspeak $^{\text{h}3}\text{J}_{\text{NC}}$ is the three-bond through-Hbond coupling between the N atom of the H-bond donor and the C' of the acceptor (Figure 2), that the IrHNCO experiment is intended to detect [9]. The $^{\text{h}3}\text{J}_{\text{NC}}$ coupling constants in units of Hz were calculated from cross-peak volumes in H-bond (V_{HB}) and reference (V_{ref}) versions of long-range HNCO experiments according to the formula:

$$^{\text{h}3}\text{J}_{\text{NC}} = \left(\left(\frac{V_{\text{HB}}}{V_{\text{ref}}} \right) / \left(\frac{N_{\text{ref}}}{N_{\text{HB}}} \right) \right)^{1/2} / (2pT) \quad \text{Eq. 1}$$

where N_{ref} and N_{HB} were the number of scans per FID in the two experiments, and T is the 66 msec delay for the N to C' INEPT refocusing period used to detect H-bond couplings [9,25]. The $^{\text{h}3}\text{J}_{\text{NC}}$ couplings are typically less than 1 Hz, with smaller values for H-bonds in α -helices than β -sheets [13]. It has been empirically established that the sizes of the $^{\text{h}3}\text{J}_{\text{NC}}$ couplings are inversely correlated with N-to-O distances across H-bonds [24,25] as:

$$r_{\text{NO}} = 2.75 - 0.25 * (\ln |^{\text{h}3}\text{J}_{\text{NC}}|) \quad \text{Eq. 2}$$

assuming that the $^{\text{h}3}\text{J}_{\text{NC}}$ coupling constant depends only on the N-O distance [24].

The third type of crosspeak that occurs in IrHNCO spectra is the intrasidue two-bond $^2\text{J}_{\text{NC}}$ coupling between the N and the C' atoms within the same residue [9,25]. These couplings are not seen in standard HNCO experiments, but their small sizes of 1.5 to 0 Hz are comparable to the $^{\text{h}3}\text{J}_{\text{NC}}$ couplings, so that both can appear in IrHNCO experiments [9,25]. The size of the $^2\text{J}_{\text{NC}}$ coupling depends on the angle between sequentially adjacent peptide groups [36].

Finally, a set of relatively strong crosspeaks are seen in the IrHNCO spectra from the backbone N to the sidechain carbonyl Cg atoms of Asp and Asn residues, for example D336, N339, and N398 in Figure 2A. These correlations are due to $^3\text{J}_{\text{NCg}}$ couplings with sizes between 0 and 3 Hz. The $^3\text{J}_{\text{NCg}}$ couplings depend on the ϕ_1 dihedral angles of Asp/Asn residues defined by the atoms N-Ca-Cb-Cg [37,38]. The $^3\text{J}_{\text{NCg}}$ coupling constant has the largest values near 3 Hz for the *trans* conformation ($\phi_1 = 180^\circ$), decreasing to near zero for the *gauche* conformations ($\phi_1 \pm 180^\circ$). Motional averaging of the ϕ_1 angle is expected to give $^3\text{J}_{\text{NCg}}$ values near 1.4 Hz. To characterize the $^3\text{J}_{\text{NCg}}$ couplings we supplemented

complete mainchain assignments for the four protein states with side-chain assignments for carbonyl containing Asp, Asn, Glu, and Gln using 3D HNCO and the sidechain-HCACO experiment [39]. These side-chain assignments are given in Tables S1-S3. The $^3\text{J}_{\text{NC}'}$ and $^3\text{J}_{\text{NCg}}$ coupling constants determined from IrHNCO experiments at various temperatures are summarized in Tables S4-S5. The $d(\text{rNO})/dT$ slopes defining the temperature dependence of mainchain H-bond distances calculated from $^3\text{J}_{\text{NC}'}$ coupling constants are given in Tables S6-S7, and Tables S8-S9 compare H-bonds detected by NMR for CUS-3i and P22i with those from the cryoEM structures of the corresponding phages.

2.2. Agreement of H-Bonds Measured by NMR and cryoEM

None of the four protein states studied have X-ray structures, however, there are cryoEM structures for phages CUS-3 (PDB 8SKG, resolution of 2.8 Å) and P22 (PDB 5UU5 and 8I1V, resolutions of 3.3 and 2.6 Å, respectively). We compared the H-bonds detected in IrHNCO experiments in this work to the H-bonds in the I-domains of the cryoEM phage structures.

The NMR and cryoEM data are in good agreement for CUS-3i with 31 of the 48 mainchain H-bonds (65%) seen by both NMR and cryoEM (Table S8). There were an additional 15 main-chain H-bonds in the cryoEM structure for which $^3\text{J}_{\text{NC}'}$ couplings were not detected in the IrHNCO experiment. Two of these could not be identified due to NMR spectral overlap, one is at the dynamic N-terminus of CUS-3i that is free in the NMR fragment but covalently attached to the coat protein in the cryoEM phage structure. An additional 10 H-bonds in the cryoEM structure are in turns connecting residues separated by less than three sequence position. These are probably an artefact of overly close contacts in the cryoEM structure, since the H-bonds have marginal $\angle\text{DHA}$ (donor-H•••acceptor) angles near 120° that would probably not qualify them as true H-bonds based on energetic considerations [12,13,40]. Only three genuine H-bonds disagreed between the NMR and cryoEM data. Two H-bonds G261(N)-A259(O) and R291(N)-S303(O) were observed in the IrHNCO spectra but were not present in the cryoEM structure, with both at the end of regular b-sheet secondary structure. The Q288(N)-V273(O) H-bond in the cryoEM structure was not observed by NMR.

For P22i, the H-bonds detected in IrHNCO experiments are also consistent with those in the cryoEM structures determined at 2.6 Å-resolution [41] and an earlier 3.3 Å-resolution structure [42]. A complicating factor for comparison is that two long D- (a.a.239-254) and S-loops (a.a. 281-291) are disordered when the P22i domain is studied in isolation by solution NMR [43], but become involved in H-bonded b-sheet structure that stabilizes the icosahedral capsid when the P22i domain is studied in the context of the intact P22 phage structure determined by cryoEM [41,42]. For the 2.6 Å resolution cryoEM P22 structure 25 of 37 (68%) mainchain H-bonds are detected by NMR, and 25 of 32 (78%) for the lower 3.3 Å-resolution structure. Conversely, excluding the D- and S-loops, 27 of 30 (90 %) mainchain H-bonds detected by NMR for P22i are seen in one of the corresponding domains from the cryoEM structures of phage P22. As with Cus-3, the differences between the H-bonds detected by NMR and cryoEM for P22i are largely due to differences in dynamics, H-bonds that cannot be detected by NMR due to spectral overlap, and H-bonds with marginal $\angle\text{DHA}$ angles in the cryoEM structures that often involve turns shorter than 4 residues.

By comparison, H-bonds involving sidechains show much poorer agreement between the NMR and cryoEM data. In CUS-3i, only two N-H•••O=C H-bonds involving sidechains were detected in the IrHNCO experiments and only one of these is observed in the cryoEM structure. Conversely, six of the sidechain-involving H-bonds in the cryoEM structure do not give detectable $^3\text{J}_{\text{NC}'}$ couplings. For P22i, none of the four sidechain H-bonds observed by NMR are detected in the 2.6 Å-resolution 8I1V cryoEM structure, and only one is seen in the 3.3 Å-resolution 5UU5 structure. Most of the sidechain H-bonds in the cryoEM structures are not seen in the IrHNCO NMR experiments, and there is also poor internal agreement for the side-chain H-bonds between the two cryoEM structures of P22i (Table S9). H-bonds involving sidechains are typically more difficult to detect via $^3\text{J}_{\text{NC}'}$ couplings than their backbone counterparts, possibly to due to their more dynamic character [24].

We next investigated the agreement of mainchain H-bond distances calculated from $^3\text{J}_{\text{NC}'}$ couplings according to Eq. 2 [24] with the corresponding N-to-O distances across H-bonds in the cryoEM structures. We did not observe a correlation between distances calculated from NMR and cryoEM structures similar to those observed with X-ray structures [24,25], probably due to the more limited resolutions of the cryoEM structures. Nevertheless, when examining the RMS differences in

N-to-O distances across H-bonds calculated from $^3\text{J}_{\text{NC}}$ couplings and cryoEM, these were 0.13 Å for the 2.8 Å-resolution CUS-3 structure (31 H-bonds), 0.21 Å for the 2.6 Å-resolution P22 structure (21 H-bonds), and 0.16 Å for the 3.3 Å-resolution P22 structure (25 H-bonds). Thus, the H-bond distances from $^3\text{J}_{\text{NC}}$ couplings agree with those from the cryoEM structures on average within about 0.1 to 0.2 Å.

The temperature dependence of H-bonds $d(\text{NO})/dT$ obtained for the proteins in this work are as large as 0.01 to 0.02 Å /°K, although for the more stable ubiquitin the values are smaller, on average 0.0005 Å /°K (see below). To determine cryoEM structures, samples prepared at physiological temperatures are vitrified by plunging them into cryogens such as liquid ethane. Vitrification necessitates a rapid cooling rate between 10^5 to 10^8 °K/s⁻¹ to bring samples to a typical temperature of ~80 °K for cryoEM data collection [44]. The H-bond distances obtained from $^3\text{J}_{\text{NC}}$ couplings and cryoEM structures agree within 0.2 Å, even though the data collection temperatures for the two methods differ by more than 200 °K. This suggests the cryoEM samples are trapped by rapid vitrification in conformations similar to those present at physiological temperatures.

2.3. H-Bonds and Their Temperature Dependence Are Poorly Conserved between Related Proteins Structures

To obtain information on the temperature dependence of H-bonds we recorded lrHNCO and reference HNCO spectra for the three proteins studied in this work at five or six different temperatures. $^3\text{J}_{\text{NC}}$ coupling constants were calculated from peak volumes in the H-bond and reference spectra according to Eq. 1 [9,25]. The resulting coupling constants were used to calculate N-O distances across H-bonds according to Eq. 2 [24]. The derived individual H-bond distances are shown as a function of temperature in Figure 3. H-bond distances for ubiquitin, calculated the same way from published $^3\text{J}_{\text{NC}}$ coupling constants [29] are also included for reference. The H-bond distances show three types of behaviors. Most of the distances increase with temperature, indicated by the red symbols. Some of the H-bonds become shorter with increasing temperature, indicated by the green symbols. The last class of H-bonds with slopes smaller than the uncertainty of the slope and R-values < 0.6 indicated with "X" symbols, were assigned as temperature invariant within experimental uncertainty. The proportions of different kinds of H-bond temperature responses vary among the different proteins studied (Figure 3).

In addition to the main-chain H-bonds we also looked at the temperature dependence of the sidechain $^3\text{J}_{\text{NCg}}$ couplings that to our knowledge has not been described before. Of 18 Asp/Asn sidechains analyzed in four protein states, six (33%) showed decreases in $^3\text{J}_{\text{NCg}}$ couplings with increasing temperature and the rest were invariant within experimental uncertainty (Table S4, S5). Most of the six residues that experienced a decrease in the coupling constant with temperature had a large sidechain $^3\text{J}_{\text{NCg}}$ above 2 Hz at low temperature characteristic of a *trans* conformation ($\phi = 180^\circ$), that decreased towards the ~1.4 Hz limit expected for ϕ dihedrals undergoing conformational averaging [38].

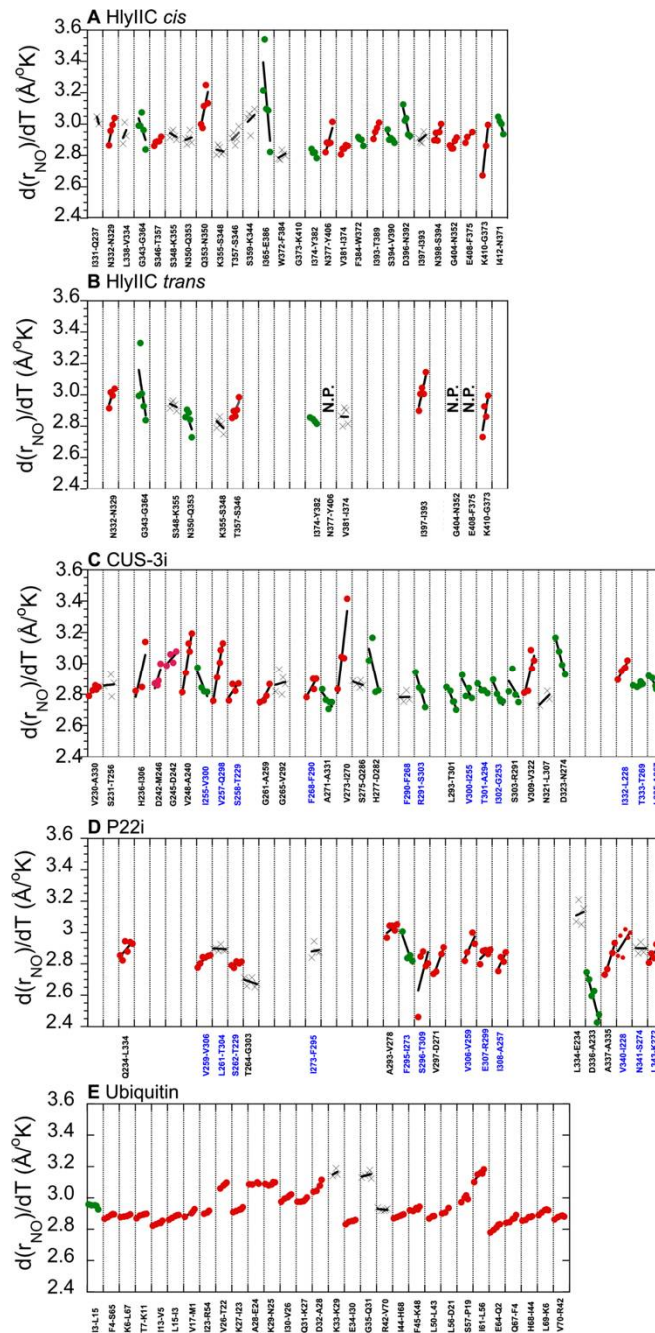


Figure 3. Temperature dependence of mainchain H-bond distances derived from $^{13}\text{J}_{\text{NC}'}$ couplings. (A) HlyIIC *cis* state (left to right 285, 290, 295, 301, 307 °K), (B) HlyIIC *trans* state (same temperatures as *cis*), (C) CUS-3i (274, 286, 295, 298, 305 °K), (D) P22i (274, 282, 290, 298, 307, 314 °K), (E) ubiquitin (278, 298, 318, 328, 338 °K). The data for ubiquitin is derived from a previously published paper [29]. Each H-bond is labeled as donor(N)-acceptor(O). H-bonds that show an increase with temperature, decrease, or no change within uncertainty are shown with red, green, and gray “X” symbols, respectively. Linear fits are shown for all H-bonds but in some cases the lines are obscured by the data points. For Cus3i and P22i, H-bonds labeled blue are structurally equivalent after superposition of the structures. N.P. indicates 3 H-bonds that are present in the *cis* but not *trans* state of HlyIIC.

For the analysis of main-chain H-bonds we first looked at conservation between related structures. The phage I-domains CUS-3i and P22i share 40% sequence identity and the same 6-strand b-barrel folding motif but have differences in secondary structure elements, differences in loop dynamics, as well as markedly different surface electrostatics [21]. The cryoEM structures of the two proteins (PDB codes 8SKG, 5UU5) align with an RMSD of 0.9 Å, allowing for the comparison of 12 structurally equivalent H-bonds after best-fit superposition. The r_{NO} distances (Eq. 2) for equivalent

H-bonds show only a moderate correlation between the two proteins (R-value = 0.76, $p = 0.0038$), that could be due to the proteins sharing similar secondary structures and the fact that H-bonds are shorter in β -sheets. The $d(r_{\text{NO}})/dT$ slopes describing the changes in H-bond distances with temperature are not significantly correlated between the two proteins (R-value = -0.50, $p = 0.093$).

We next examined the *cis* and *trans* states of HlyIIC related by isomerization about the G404-P405 peptide bond. The two states are in slow exchange on the NMR timescale, giving separate NMR signals for about half of the residues in the protein [45]. The main difference between the structures of the two states is in the orientation of the loop between helix $\alpha 2$ and strand $\beta 5$ that harbors P405 [31]. We were able to resolve and analyze ten structurally equivalent H-bonds in the *cis* and *trans* states. The H-bond distances at 307 °K are moderately correlated (R-value = 0.69, $p = 0.026$). However, in the same 3D ^1H NMR spectrum where the *cis* and *trans* state can be analyzed simultaneously, three of 28 H-bonds in the *cis* state are not seen in the *trans* state (N377-Y406, G404-N352, E408-F375). The three H-bonds involve residues near the P405 site of isomerization that are lost in the *trans* state due to the structural differences accompanying isomerization. The $d(r_{\text{NO}})/dT$ slopes for the ten structurally equivalent H-bonds are only moderately correlated between the *cis* and *trans* states (R-value = 0.80, $p = 0.0059$). In several cases N350-Q353, T357-S346, V381-I374, I397-I393 the H-bonds show markedly different $d(r_{\text{NO}})/dT$ slopes between the two states.

The relatively weak conservation of H-bond distances and temperature responses between closely similar protein structures such as the *cis* and *trans* states of HlyIIC, suggest that the length and temperature dependence of an H-bond is determined mostly by short-range interactions in its immediate vicinity and less by the overall protein fold. This is further supported by the observation that three H-bonds are lost near P405 in the *trans* compared to the *cis* state of HlyIIC due to the conformational differences localized to the loop bearing the proline.

2.4. The Variability of H-Bond Temperature Responses Is Inversely Correlated with Global Folding Stability

Data on the stabilities of the three proteins used to carry out temperature dependent H-bond studies in this work, together with ubiquitin from a previously published study [29] are given in Table 1. A plot summarizing the variability in $d(r_{\text{NO}})/dT$ slopes is shown in Figure 4A. The proteins in Figure 4A are arranged in order of increasing stability to unfolding from left to right (Table 1). The average $d(r_{\text{NO}})/dT$ slope near $(10.7 \pm 4.6) \times 10^{-4} \text{ \AA/K}$ is similar for all the proteins. This average probably reflects the thermal volume expansion coefficient, an intrinsic property that has a conserved value of $\sim 5.2 \times 10^{-4} \text{ 1/K}$ for a variety of proteins [29]. The spread in $d(r_{\text{NO}})/dT$ slopes, however, increases with decreasing stability to unfolding. For the moderately stable protein HlyIIC, the range of $d(r_{\text{NO}})/dT$ slopes between 0.019 and -0.022 \AA/K is more than 10-fold larger than for the stable protein ubiquitin where $d(r_{\text{NO}})/dT$ slopes vary between 0.0015 and -0.0004 \AA/K . This is illustrated in Figure 4B where the standard deviation of the $d(r_{\text{NO}})/dT$ slopes is correlated with the ΔG_{unf}^0 and the m -values obtained for the three proteins from equilibrium unfolding experiments (Table 1). Studies of the temperature dependence of H-bond $^3\text{J}_{\text{NC}}$ couplings were done for the additional protein GB3 [30] but are not included in Figure 4 because data for the couplings of individual H-bonds were not available. Nevertheless, the range of $d(^3\text{J}_{\text{NC}})/dT$ slopes between 0.001 and -0.0003 Hz/K for GB3 [30] are very similar to those for ubiquitin [29]. The GB3 protein has a similar fold to ubiquitin [46], and a high thermal stability estimated to be $\geq 90 \text{ }^\circ\text{C}$ [47,48]. Therefore, the data for GB3 also qualitatively supports that $d(r_{\text{NO}})/dT$ slopes are more uniform for stable proteins.

Table 1. Folding stability information on proteins used for IrHNCO studies^a.

Protein	ΔG^0_{unf}	m -value	T_{melt}	References
	(kcal/mol)	(kcal/mol•M)	(°C)	
HlyIIC	3.0 ± 0.2	1.3 ± 0.1	52	[19,31]
CUS-3i	3.3 ± 0.6	1.3 ± 0.2	48	[61]
P22	5.8 ± 0.3	1.6 ± 0.1	54	[53]
ubiquitin	9.9 ± 1.0	2.4 ± 0.2	> 90	[62,63]
GB3	~ 6 to 7	N.D.	≥ 90	[47,48]

^aThe ΔG^0_{unf} and m values are from equilibrium denaturation experiments using urea, except for ubiquitin where guanidine chloride was used as a denaturant, and GB3 where values were estimated from the highly homologous GB1 and GB2 domains [47,48]. The T_{melt} values are midpoints for thermal unfolding.

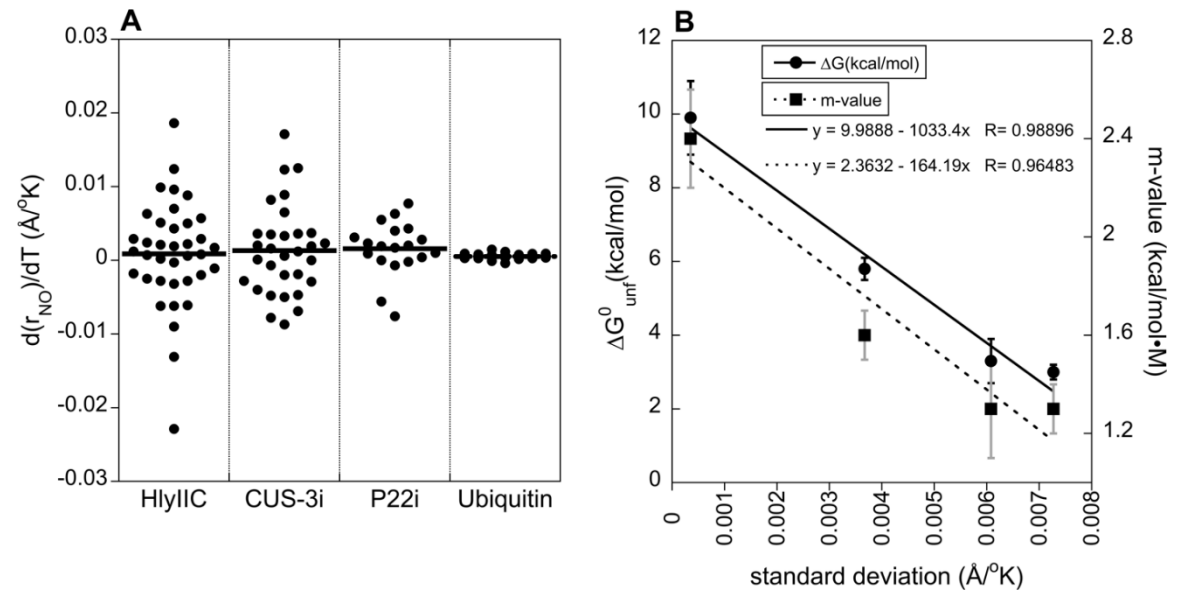


Figure 4. Variability in the temperature dependence of H-bonds as a function of folding stability. (A) Distribution of $d(r_{\text{NO}})/dT$ slopes in four proteins arranged in order of increasing stability to unfolding (least stable right, most stable left). (B) Correlation between the standard deviation in $d(r_{\text{NO}})/dT$ slopes and parameters related to protein stability: ΔG^0_{unf} and m -values (from Table 1).

The ΔG^0_{unf} parameter is the change in the Gibbs free energy for unfolding, measured in equilibrium denaturation experiments, and reflects the stability of the folded compared to the unfolded state. The m -value is a descriptor of the slope or steepness of the unfolding transition [49–51]. There are two interpretations of the m -value. The first, is that it describes the change in solvent accessible surface between the folded and unfolded state [49,50]. The second is that it describes the cooperativity of the unfolding transition [51,52]. The two interpretations are largely equivalent, in as much as cooperative all-or-none unfolding transitions lead to a large change in accessible surface area whereas non-cooperative partial unfolding leads to a smaller change in accessible surface area. It is well established that for most proteins ΔG^0_{unf} and the m -value for unfolding are highly correlated [50,53]. The m -values tend to be small for poorly structured proteins such as molten globules, and

large for stable proteins [51,54]. The ΔG_{unf}^0 and m -values are also highly correlated for the four proteins considered in Table 1 and Figure 4 (R -value = 0.991, p = 0.0088).

The correlation between increased variability in H-bond $d(r\text{NO})/dT$ slopes and lower m -values (Figure 4B) is consistent with the interpretation of the latter in terms of cooperativity. For the highly cooperative and stable protein ubiquitin all H-bonds show nearly the same small 5×10^{-4} Å/K expansion that coincides with the value for the thermal volume expansion of the protein. For the least stable proteins CUS-3 and HlyIIC with the smallest m -values indicative of diminished unfolding cooperativity, the range of $d(r\text{NO})/dT$ slopes is about 10-fold larger and includes an increasing proportion of H-bonds with negative slopes, shown in green in Figure 3. Most of the H-bonds with negative $d(r\text{NO})/dT$ slopes occur in irregular or α -helical secondary structure, or at the edges corresponding to the first or last residues of β -strands. Few occur in the middle of β -strands, except F295(N)-I273(O) in strand b4 of P22i and several H-bonds connecting strands b3-b4-b5 of CUS3-i. The segment b3-b4-b5 in CUS-3i behaves as an autonomously folding sub-domain under acid denaturing conditions, retaining β -sheet structure while the rest of the protein is unfolded (A.T.A, unpublished observations).

Residues with negative $d(r\text{NO})/dT$ slopes tend to correspond to the some of the shortest H-bonds in the proteins HlyIIC, CUS-3i and P22i (green in Figure 3). This suggests these H-bonds are energetically favorable but opposed by strain arising from the remainder of non-covalent interactions stabilizing the structure. In weakly cooperative proteins, as the structure becomes increasingly dynamic at higher temperature, the strain due to rigid side-chain packing could become abated allowing some of the H-bonds to move towards their energetic optimum. By contrast in a highly cooperative stable systems like ubiquitin nearly all H-bonds experience the same increase driven by thermal expansion, with the few exceptions of H-bonds in the β -strand near the N-terminus that are invariant or become more stable at higher temperature [29].

The H-bonds may not be experiencing changes in length with temperature but changes in the ratio of conformers that have the H-bonds formed or broken. These conformers are in fast exchange on the NMR timescale, so that the measured $^3J_{\text{NC}}$ coupling will be a population-weighted average. When α -helical structure of the RNaseA S-peptide is stabilized by increasing concentrations of the kosmotropic solvent TFE, the $^3J_{\text{NC}}$ coupling constants show apparent increases suggestive of decreasing H-bond distances [27]. However, circular dichroism ellipticity becomes more negative [55] and the NMR α secondary chemical shifts increase [27]. Both observations are more consistent with an increasing fraction of molecules adopting H-bonded α -helical conformers than with a shortening of pre-existing H-bonds. Similarly, in proteins with low m -values indicative of reduced unfolding cooperativity, the responses of H-bonds to changes in temperature will be more varied if the populations of individual sub-structures become uncoupled from the unfolding of the overall global structure.

3. Materials and Methods

3.1. Samples

Double-labeled $^{13}\text{C}/^{15}\text{N}$ and triple-labeled $^{13}\text{C}/^{15}\text{N}/^2\text{H}$ samples of wild type HlyIIC [19], P22i [56], and CUS-3i [57], were prepared as described previously. For long-range HNCO experiments to measure H-bonds, triple-labeled $^{13}\text{C}/^{15}\text{N}/^2\text{H}$ proteins in 90% $\text{H}_2\text{O}/10\%$ D_2O were used with the following sample conditions: 1.0 mM HlyIIC in 20 mM NaH_2PO_4 with 1mM EDTA and 1 mM PMSF at pH 6.6; 1.7 mM P22i in 20 mM NaH_2PO_4 at pH 6.1, 1.6 mM CUS-3i in 20 mM NaH_2PO_4 at pH 6.3. For sidechain-HCACO experiments to obtain Asp/Asn and Glu/Gln assignments, double-labeled $^{13}\text{C}/^{15}\text{N}$ samples were lyophilized and re-dissolved in 99.96% D_2O with the following sample conditions: 1 mM HlyIIC at pH* 6.0, 3.8 mM P22i at pH* 6.1, 0.7 mM CUS-3i at pH* 6.3, where pH* is the pH measured in D_2O with a glass pH electrode.

3.2. NMR Data Acquisition and Analysis

Experiments were done on 600 MHz Bruker Avance and 600 MHz Varian Inova spectrometers, both equipped with cryogenic probes. For the Bruker instrument we modified the BEST TROSY HNCO H-bond experiment [58] from the Bruker IBS library (pulse sequence *BT_HNCO_hbonds*) to work on Topsis 2.1. At each temperature we obtained long-range HNCO ($d23$ = 66 ms) and reference HNCO ($d24$ = 16.5 ms) versions of the experiments to quantify $^3J_{\text{NC}}$ couplings [9,25,27]. For

HlyIIC we recorded 3D spectra on the Bruker 600 MHz spectrometer at temperatures of 285, 290, 295, 301, and 307 °K, with $32(t_1, C') \times 16(t_2, N) \times 1024(t_3, H)$ complex points and acquisition times of 13, 9, and 114 ms in the t_1 , t_2 , and t_3 dimensions. Total experiment times were on the order of 46 and 4 h for the H-bond and reference spectra, respectively. For CUS-3i we recorded 3D spectra on the Bruker 600 MHz spectrometer at temperatures of 274, 286, 295, 298, and 305 °K, with $64(t_1, C) \times 16(t_2, N) \times 1024(t_3, H)$ complex points and acquisition times of 26, 10, and 107 ms in the t_1 , t_2 , and t_3 dimensions. Total experiment times were about 54 and 8 h for each of the H-bond and reference spectra, respectively. Since P22i had the best dispersion of the three proteins we recorded 2D versions of the TROSY-HNCO experiment (pulse sequence *best_trospy_hbonds*) on the Varian 600 MHz spectrometer at temperatures of 274, 282, 290, 298, 307, and 314 °K. For the reference experiment we modified the Varian *best_trospy_hbonds* pulse sequence to shift the $^{13}C'$ 180°-pulses by 16.5 ms with respect to the ^{15}N 180°-pulses in the INEPT steps as described in the literature [9]. The 2D data sets on P22i were recorded with $64(t_1, C) \times 512(t_2, H)$ complex points with acquisition times of 14 (t_1) and 107 ms (t_2). Total acquisition times were 11 h for the H-bond and 1 h for the reference experiments. NMR sample temperatures were calibrated using 100% methanol ($T < 300$ °K) and 100% ethylene glycol ($T \geq 300$ °K) standards, as described in the Bruker VT-calibration manual.

Sidechain carbonyl and amide resonances were assigned from 3D HNCO and HCACO experiments. The 3D HCACO experiment (Bruker pulse sequence *hcacogp3d*) was modified for sidechains as described in the literature [39], namely the aliphatic ^{13}C center was placed at ~39 ppm, the carbonyl ^{13}C center at ~177 ppm, and the delay $1/[4(J_{HC})]$ (called d4 in the *hcacogp3d* pulse sequence) was set to 1.8 ms for methylene protons rather than 3.3 ms in the standard experiment. The 3D HCACO spectrum for HlyIIC was recorded at a sample temperature of 37 °C on a 500 MHz instrument with $50(t_1, C) \times 32(t_2, C') \times 1024(t_3, H)$ complex points, and acquisition times of 20, 12, and 122 ms in the t_1 , t_2 , and t_3 dimensions, respectively. The total acquisition time was 34 h. The 3D HCACO spectrum for CUS-3i was recorded at a temperature of 25 °C on a 500 MHz magnet with $16(t_1, C) \times 11(t_2, C') \times 1024(t_3, H)$ complex points, and acquisition times of 4 (t_1), 3 (t_2), and 157 (t_3) ms; total 30 h. The 3D HCACO for P22i was at 37 °C on a 600 MHz spectrometer with $48(t_1, C) \times 32(t_2, C') \times 1024(t_3, H)$ complex points, and acquisition times of 29 (t_1), 13 (t_2), and 122 (t_3) ms; total 64 h.

NMR spectra were processed with iNMR 6.3 (Mestrelab Research, Santiago de Compostela, Spain) and analyzed with CcpNmr Analysis version 2.5.2 [59]. Published assignments were used to analyze the NMR data for *cis* HlyIIC (BMRB 19461), *trans* HlyIIC (BMRB 19462), CUS-3i (BMRB 25263) and P22i (BMRB 18566).

$^{13}J_{NC'}$ coupling constants were calculated according to Eq. 1. To estimate uncertainties in $^{13}J_{NC'}$ values experiments were replicated on two separate samples at one temperature for each protein: HlyIIC (307 °K), CUS-3i (298 °K), and P22i (307 °K). The RMS differences between the duplicate $^{13}J_{NC'}$ values were 0.10, 0.06 and 0.04 Hz for HlyIIC, CUS-3i, and P22i, respectively. H-bond distances were calculated using the empirical relationship (Eq. 2) established by Bax and co-workers [24].

3.3. Structure Analysis

To correlate NMR with structural parameters we used the segment corresponding to CUS-3i (Chain A of the asymmetric unit, residues 223-337) from the 2.8 Å-resolution cryoEM structure of phage CUS-3 (Richard Whitehead, Carolyn Teschke, and A.T.A, in preparation; PDB accession code 8SKG, coordinates are available from the corresponding author until release). For P22i we used the 3.3 Å-resolution cryoEM structure of phage P22 (PDB 5UU5, chain E of asymmetric unit, residues 223-345 [42]), and the 2.6 Å-resolution cryoEM structure of phage P22 (PDB 8I1V, chain A of asymmetric unit, residues 223-345 [41]). Interatomic distances were calculated with the "get_distance" command of PyMol (Schrödinger software, New York). H-bonds were calculated with the program HBPLUS v.3.06 [60].

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Author Contributions: Conceptualization, A.T.A.; methodology, A.T.A. and A.J.D.; formal analysis, A.T.A.; resources, A.T.A.; writing – original draft preparation, A.T.A.; writing – reviewing and editing, A.T.A. and A.J.D. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: We thank Richard Whitehead III and Carol Teschke for the coordinates of the cryoEM CUS-3 phage structure. We thank Anne Kaplan and Rich Olson for samples of HlyIIC, and Carol Teschke and Therese Tripler for samples of CUS-3i and P22i. A.T.A. thanks Prof. Mei Hong and the Hong lab for the use of their solution 600 MHz instrument during his sabbatical. The NMR experiments used equipment at the MIT-Harvard Center for Magnetic Resonance, which is supported by the P41 grant GM132079. M.H. is partially supported by NIH grant AG059661. A.J.D. was partially supported by an NIH fellowship F31AG069418.

Conflict of Interest: The authors declare no potential conflict of interest.

Abbreviations

2D, 3D, one-, two, three-dimensional; cryoEM, cryogenic electron microscopy; CUS-3i, insertion domain from the coat protein of phage CUS-3; FID, free induction decay, H-bond, hydrogen bond; HlyIIC, hemolysin II C-terminal domain; INEPT, insensitive nuclei enhancement by polarization transfer; lrHNCO, long-range HNCO experiment, NMR, nuclear magnetic resonance; P22i, insertion domain from the coat protein of phage P22; RMS, root-mean-square, RMSD, root-mean-square deviation; TFE, trifluoroethanol; TROSY, transverse relaxation optimized spectroscopy.

References

1. Branden, C.; Tooze, J., *Introduction to Protein Structure*. Second ed.; Garland Science: New York, 1998.
2. Pauling, L., The hydrogen bond. In *The nature of the chemical bond*, 3rd ed.; Pauling, L., Ed. Cornell University Press: New York, 1960; pp 449-504.
3. Martin, T. W.; Derewenda, Z. S., The name is bond--H bond. *Nat Struct Biol* **1999**, 6, (5), 403-6.
4. Pace, C. N.; Fu, H.; Lee Fryar, K.; Landua, J.; Trevino, S. R.; Schell, D.; Thurlkill, R. L.; Imura, S.; Scholtz, J. M.; Gajiwala, K.; Sevcik, J.; Urbanikova, L.; Myers, J. K.; Takano, K.; Hebert, E. J.; Shirley, B. A.; Grimsley, G. R., Contribution of hydrogen bonds to protein stability. *Protein Sci* **2014**, 23, (5), 652-61.
5. Boresch, S.; Archontis, G.; Karplus, M., Free energy simulations: the meaning of the individual contributions from a component analysis. *Proteins* **1994**, 20, (1), 25-33.
6. Brady, G. P.; Szabo, A.; Sharp, K. A., On the decomposition of free energies. *J Mol Biol* **1996**, 263, (2), 123-5.
7. Mark, A. E.; van Gunsteren, W. F., Decomposition of the free energy of a system in terms of specific interactions. Implications for theoretical and experimental studies. *J Mol Biol* **1994**, 240, (2), 167-76.
8. Pauling, L., The Shared-Electron Chemical Bond. *Proc Natl Acad Sci U S A* **1928**, 14, (4), 359-62.
9. Cordier, F.; Grzesiek, S., Direct observation of hydrogen bonds in proteins by interresidue ³HJNC' scalar couplings. *J Am Chem Soc* **1999**, 121, 1601-1602.
10. Cordier, F.; Rogowski, M.; Grzesiek, S.; Bax, A., Observation of through-hydrogen-bond 2hJHC' in a perdeuterated protein. *J Magn Reson* **1999**, 140, (2), 510-2.
11. Cornilescu, G.; Hu, J. S.; Bax, A., Identification of the hydrogen bonding network in a protein by scalar couplings. *J Am Chem Soc* **1999**, 121, 2949-2950.
12. Dingley, A. J.; Grzesiek, S., Direct Observation of Hydrogen Bonds in Nucleic Acid Base Pairs by Internucleotide 2JNN Couplings. *J Am Chem Soc* **1998**, 120, (33), 8293-8297.
13. Grzesiek, S.; Cordier, F.; Jaravine, V. A.; Barfield, M., Insights into biomolecular hydrogen bonds from hydrogen bond scalar couplings. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2004**, 45, 275-300.
14. Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wuthrich, K., NMR scalar couplings across Watson-Crick base pair hydrogen bonds in DNA observed by transverse relaxation-optimized spectroscopy. *Proc Natl Acad Sci U S A* **1998**, 95, (24), 14147-51.
15. Isaacs, E. D.; Shukla, A.; Platzman, P. M.; Hamann, D. R.; Barbiellini, B.; Tulk, C. A., Covalency of the Hydrogen Bond in Ice: A Direct X-Ray Measurement. *Physical Review Letters* **1999**, 82, (3), 600-603.
16. Englander, S. W.; Sosnick, T. R.; Englander, J. J.; Mayne, L., Mechanisms and uses of hydrogen exchange. *Curr Opin Struct Biol* **1996**, 6, (1), 18-23.
17. Alexandrescu, A. T., Strategy for supplementing structure calculations using limited data with hydrophobic distance restraints. *Proteins* **2004**, 56, (1), 117-29.
18. Sheftic, S. R.; Garcia, P. P.; White, E.; Robinson, V. L.; Gage, D. J.; Alexandrescu, A. T., Nuclear magnetic resonance structure and dynamics of the response regulator Sma0114 from *Sinorhizobium meliloti*. *Biochemistry* **2012**, 51, (35), 6932-41.
19. Kaplan, A. R.; Kaus, K.; De, S.; Olson, R.; Alexandrescu, A. T., NMR structure of the *Bacillus cereus* hemolysin II C-terminal domain reveals a novel fold. *Sci Rep* **2017**, 7, (1), 3277.
20. Newcomer, R. L.; Schrad, J. R.; Gilcrease, E. B.; Casjens, S. R.; Feig, M.; Teschke, C. M.; Alexandrescu, A. T.; Parent, K. N., The phage L capsid decoration protein has a novel OB-fold and an unusual capsid binding strategy. *Elife* **2019**, 8.

21. Tripler, T. N.; Kaplan, A. R.; Alexandrescu, A. T.; Teschke, C. M., Conservation and Divergence of the I-Domain Inserted into the Ubiquitous HK97 Coat Protein Fold in P22-Like Bacteriophages. *J Virol* **2019**, 93, (9).
22. Kanelis, V.; Rotin, D.; Forman-Kay, J. D., Solution structure of a Nedd4 WW domain-ENaC peptide complex. *Nat Struct Biol* **2001**, 8, (5), 407-12.
23. Jee, J. G.; Ikegami, T.; Hashimoto, M.; Kawabata, T.; Ikeguchi, M.; Watanabe, T.; Shirakawa, M., Solution structure of the fibronectin type III domain from *Bacillus circulans* WL-12 chitinase A1. *J Biol Chem* **2002**, 277, (2), 1388-97.
24. Cornilescu, G.; Ramirez, B. E.; Frank, M. K.; Clore, G. M.; Gronenborn, A. M.; Bax, A., Correlation between $^3\text{hJNC}'$ and hydrogen-bond length in proteins. *J Am Chem Soc* **1999**, 121, 6275-6279.
25. Alexandrescu, A. T.; Snyder, D. R.; Abildgaard, F., NMR of hydrogen bonding in cold-shock protein A and an analysis of the influence of crystallographic resolution on comparisons of hydrogen bond lengths. *Protein Sci* **2001**, 10, (9), 1856-68.
26. Li, H.; Yamada, H.; Akasaka, K.; Gronenborn, A. M., Pressure alters electronic orbital overlap in hydrogen bonds. *J Biomol NMR* **2000**, 18, (3), 207-16.
27. Jaravine, V. A.; Alexandrescu, A. T.; Grzesiek, S., Observation of the closing of individual hydrogen bonds during TFE-induced helix formation in a peptide. *Protein Sci* **2001**, 10, (5), 943-50.
28. Cordier, F.; Wang, C.; Grzesiek, S.; Nicholson, L. K., Ligand-induced strain in hydrogen bonds of the c-Src SH3 domain detected by NMR. *J Mol Biol* **2000**, 304, (4), 497-505.
29. Cordier, F.; Grzesiek, S., Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR. *J Mol Biol* **2002**, 317, (5), 739-52.
30. Hong, J.; Jing, Q.; Yao, L., The protein amide (1)H(N) chemical shift temperature coefficient reflects thermal expansion of the N-H...O=C hydrogen bond. *J Biomol NMR* **2013**, 55, (1), 71-8.
31. Kaplan, A. R.; Olson, R.; Alexandrescu, A. T., Protein yoga: Conformational versatility of the Hemolysin II C-terminal domain detailed by NMR structures for multiple states. *Protein Sci* **2021**, 30, (5), 990-1005.
32. Juranic, N.; Ilich, P. K.; Macura, S., Hydrogen Bonding Networks in Proteins As Revealed by the Amide $1\text{JNC}'$ Coupling Constant. *J Am Chem Soc* **1995**, 117, (1), 405-410.
33. Kay, L. E.; Ikura, M.; Tschudin, R.; Bax, A., Three-dimensional triple-resonance NMR Spectroscopy of isotopically enriched proteins. 1990. *J Magn Reson* **2011**, 213, (2), 423-41.
34. Juranic, N.; Moncrieffe, M. C.; Likic, V. A.; Prendergast, F. G.; Macura, S., Structural dependencies of $^3\text{hJNC}'$ scalar coupling in protein H-bond chains. *J Am Chem Soc* **2002**, 124, (47), 14221-6.
35. Juranic, N.; Macura, S., Correlations among $(1\text{J}(\text{NC})')$ and $(^3\text{hJ}(\text{NC})')$ coupling constants in the hydrogen-bonding network of human ubiquitin. *J Am Chem Soc* **2001**, 123, (17), 4099-100.
36. Juranic, N.; Dannenberg, J. J.; Cornilescu, G.; Salvador, P.; Atanasova, E.; Ahn, H. C.; Macura, S.; Markley, J. L.; Prendergast, F. G., Structural dependencies of protein backbone $2\text{JNC}'$ couplings. *Protein Sci* **2008**, 17, (4), 768-76.
37. Benirschke, R. C.; Thompson, J. R.; Nomine, Y.; Wasielewski, E.; Juranic, N.; Macura, S.; Hatakeyama, S.; Nakayama, K. I.; Botuyan, M. V.; Mer, G., Molecular basis for the association of human E4B U box ubiquitin ligase with E2-conjugating enzymes UbcH5c and Ubc4. *Structure* **2010**, 18, (8), 955-65.
38. Juranic, N.; Atanasova, E.; Moncrieffe, M. C.; Prendergast, F. G.; Macura, S., Calcium-binding proteins afford calibration of dihedral-angle dependence of $3\text{J}(\text{NC}(\gamma))$ coupling constant in aspartate and asparagine residues. *J Magn Reson* **2005**, 175, (2), 222-5.
39. Yamazaki, T.; Nicholson, L. K.; Torchia, D. A.; Wingfield, P.; Stahl, S. J.; Kaufman, J. D.; Eyermann, C. J.; Hodge, C. N.; Lam, P. Y. S.; Ru, Y.; Jadhav, P. K.; Chang, C.-H.; Weber, P. C., NMR and X-ray evidence that the HFV protease catalytic aspartyl groups are protonated in the complex formed by the protease and a non-peptide cyclic urea-based inhibitor. *J Am Chem Soc* **1994**, 116, 10791-10792.
40. Wood, P. A.; Allen, F. H.; Pidock, E., Hydrogen-bond directionality at the donor H atom—analysis of interaction energies and database statistics. *CrystEngComm* **2009**, 11, 1563-1571.
41. Xiao, H.; Zhou, J.; Yang, F.; Liu, Z.; Song, J.; Chen, W.; Liu, H.; Cheng, L., Assembly and Capsid Expansion Mechanism of Bacteriophage P22 Revealed by High-Resolution Cryo-EM Structures. *Viruses* **2023**, 15, (2).
42. Hryc, C. F.; Chen, D. H.; Afonine, P. V.; Jakana, J.; Wang, Z.; Haase-Pettingell, C.; Jiang, W.; Adams, P. D.; King, J. A.; Schmid, M. F.; Chiu, W., Accurate model annotation of a near-atomic resolution cryo-EM map. *Proc Natl Acad Sci U S A* **2017**, 114, (12), 3103-3108.
43. Rizzo, A. A.; Suhanovsky, M. M.; Baker, M. L.; Fraser, L. C.; Jones, L. M.; Rempel, D. L.; Gross, M. L.; Chiu, W.; Alexandrescu, A. T.; Teschke, C. M., Multiple functional roles of the accessory I-domain of bacteriophage P22 coat protein revealed by NMR structure and CryoEM modeling. *Structure* **2014**, 22, (6), 830-41.
44. Weissenberger, G.; Henderikx, R. J. M.; Peters, P. J., Understanding the invisible hands of sample preparation for cryo-EM. *Nat Methods* **2021**, 18, (5), 463-471.
45. Kaplan, A. R.; Maciejewski, M. W.; Olson, R.; Alexandrescu, A. T., NMR assignments for the cis and trans forms of the hemolysin II C-terminal domain. *Biomolecular NMR assignments* **2014**, 8, (2), 419-23.

46. Kraulis, P. J., Similarity of protein G and ubiquitin. *Science* **1991**, 254, (5031), 581-2.
47. Alexander, P.; Fahnestock, S.; Lee, T.; Orban, J.; Bryan, P., Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures. *Biochemistry* **1992**, 31, (14), 3597-603.
48. Granata, D.; Camilloni, C.; Vendruscolo, M.; Laio, A., Characterization of the free-energy landscapes of proteins by NMR-guided metadynamics. *Proc Natl Acad Sci U S A* **2013**, 110, (17), 6817-22.
49. Myers, J. K.; Pace, C. N.; Scholtz, J. M., Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci* **1995**, 4, (10), 2138-48.
50. Shortle, D., Staphylococcal nuclease: a showcase of m-value effects. *Adv Protein Chem* **1995**, 46, 217-47.
51. Alexandrescu, A. T.; Jaravine, V. A.; Dames, S. A.; Lamour, F. P., NMR hydrogen exchange of the OB-fold protein LysN as a function of denaturant: the most conserved elements of structure are the most stable to unfolding. *J Mol Biol* **1999**, 289, (4), 1041-54.
52. Carra, J. H.; Privalov, P. L., Thermodynamics of denaturation of staphylococcal nuclease mutants: an intermediate state in protein folding. *FASEB J* **1996**, 10, (1), 67-74.
53. Newcomer, R. L.; Fraser, L. C. R.; Teschke, C. M.; Alexandrescu, A. T., Mechanism of Protein Denaturation: Partial Unfolding of the P22 Coat Protein I-Domain by Urea Binding. *Biophysical journal* **2015**, 109, (12), 2666-2677.
54. Smith, L. J.; Alexandrescu, A. T.; Pitkeathly, M.; Dobson, C. M., Solution structure of a peptide fragment of human alpha-lactalbumin in trifluoroethanol: a model for local structure in the molten globule. *Structure* **1994**, 2, (8), 703-12.
55. Nelson, J. W.; Kallenbach, N. R., Stabilization of the ribonuclease S-peptide alpha-helix by trifluoroethanol. *Proteins* **1986**, 1, (3), 211-7.
56. Rizzo, A. A.; Fraser, L. C.; Sheftic, S. R.; Suhanovsky, M. M.; Teschke, C. M.; Alexandrescu, A. T., NMR assignments for the telokin-like domain of bacteriophage P22 coat protein. *Biomolecular NMR assignments* **2013**, 7, (2), 257-60.
57. Tripler, T. N.; Maciejewski, M. W.; Teschke, C. M.; Alexandrescu, A. T., NMR assignments for the insertion domain of bacteriophage CUS-3 coat protein. *Biomolecular NMR assignments* **2015**.
58. Favier, A.; Brutscher, B., Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J Biomol NMR* **2011**, 49, 9-15.
59. Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D., The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **2005**, 59, (4), 687-96.
60. McDonald, I. K.; Thornton, J. M., Satisfying hydrogen bonding potential in proteins. *J Mol Biol* **1994**, 238, (5), 777-93.
61. Tripler, T. N. Conservation and divergence between P22-like bacteriophages coat protein's I-domains and procapsid-like particles. Doctoral Dissertation, University of Connecticut, 2019.
62. Surana, P.; Das, R., Observing a late folding intermediate of Ubiquitin at atomic resolution by NMR. *Protein Sci* **2016**, 25, (8), 1438-50.
63. Wintrod, P. L.; Makhatadze, G. I.; Privalov, P. L., Thermodynamics of ubiquitin unfolding. *Proteins* **1994**, 18, (3), 246-53.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.