J Biomol NMR (2012) 53:25–41
DOI 10.1007/s10858-012-9618-5

ARTICLE

Conformational distributions of denatured and unstructured proteins are similar to those of 20 × 20 blocked dipeptides

Kwang-Im Oh · Young-Sang Jung · Geum-Sook Hwang · Minhaeng Cho

Received: 2 December 2011 / Accepted: 24 February 2012 / Published online: 18 March 2012

Abstract Understanding intrinsic conformational preferences of amino-acids in unfolded proteins is important for elucidating the underlying principles of their stability and re-folding on biological timescales. Here, to investigate the neighbor interaction effects on the conformational propensities of amino-acids, we carried out ^{1}H NMR experiments for a comprehensive set of blocked dipeptides and measured the scalar coupling constants between alpha protons and amide protons as well as their chemical shifts. Detailed inspection of these NMR properties shows that, irrespective of amino-acid side-chain properties, the distributions of the measured coupling constants and chemical shifts of the dipeptides are comparatively narrow, indicating small variances of their conformation distributions. They are further compared with those of blocked amino-acids (Ac–X–NHMe), oligopeptides (Ac–GGXGG–NH_2), and native (lysozyme), denatured (lysozyme and outer membrane protein X from *Escherichia coli*), unstructured (Domain 2 of the protein 5A of Hepatitis C virus), and intrinsically disordered (hNlg3cyt: intracellular domain of human NL3) proteins. These comparative investigations suggest that the conformational preferences and local solvation environments of the blocked dipeptides are quite similar to not only those of other short oligopeptides but also those of denatured and natively unfolded proteins.

Keywords Unfolded protein · Blocked dipeptide · NMR scalar coupling · Peptide backbone conformation

Electronic supplementary material The online version of this article (doi:10.1007/s10858-012-9618-5) contains supplementary material, which is available to authorized users.

K.-I. Oh · M. Cho (✉)
Department of Chemistry, Korea University, Seoul 136-701, Korea
e-mail: mcho@korea.ac.kr

Y.-S. Jung · G.-S. Hwang · M. Cho
Korea Basic Science Institute, Seoul 136-713, Korea
e-mail: gshwang@kbsi.re.kr

Introduction

The definition of the unfolded state of proteins is critical for understanding their stability and folding mechanism. Despite prolonged and extensive efforts over the decades, the underlying principle and hidden nature of denatured and intrinsically disordered proteins remains unclear. Particularly, even though the chemical diversity of natural amino-acid side-chains has been a key to the understanding of tertiary structures, dynamics, and functions of native proteins, it is not clear whether the same chemical diversity of side-chains plays any role in determining the structures and free energy landscapes of unfolded or intrinsically disordered proteins. An important and compelling hypothesis regarding to the backbone conformational distribution of unfolded proteins was suggested by Krimm and coworkers, which has been known as the PP_{II} hypothesis, i.e., 'the backbone conformations of unfolded proteins are not random coils but include short stretches of polyproline II (PP_{II}) structural motifs interspersed with turns and bends' (Avbelj et al. 2006; Eker et al. 2002; Shi et al. 2002a, b; Tiffany and Krimm 1968).

This PP_{II} hypothesis has been examined by numerous workers for a variety of oligopeptides and remains controversial (Avbelj et al. 2006; Dukor and Keiderling 1991; Eker et al. 2002; Hahn et al. 2004; Lee et al. 2006, 2007; Makowska et al. 2007; Oh et al. 2006, 2010; Schweitzer-Stenner and Measey 2007; Shi et al. 2002a; Shortle and

Ackerman 2001; Wright et al. 1988; Zagrovic et al. 2005; Zanni et al. 2001). There is experimental evidence suggesting that unfolded proteins are fairly compact and often contain residual secondary structure elements. However, for short oligopeptides, they exhibited strong propensities to adopt extended conformations such as PP_{II} and β-strand. It has been shown that the PP_{II} conformer is enthalpically more stable than the β-strand conformer while the β-strand conformer is entropically more stable than the PP_{II}. Thus, at a low temperature, the PP_{II} left-handed helix population is large with a little or no α-helical and β-strand structures. On the origin of the energy difference between PP_{II} and β-strand structures, a variety of hypotheses have been suggested and quite a large number of theoretical studies were reported before. For instance, Brant and Flory (Brant and Flory 1965) emphasized the importance of steric clash, dipole–dipole interactions, and torsional potentials of the dihedral internal rotation about the backbone peptide bonds. Sreerama and Woody (1999) found that water molecules actively participate in preferentially stabilizing the PP_{II} structure. Pappu and Rose (2002) emphasized the effects of steric clash and excluded volume on preferred backbone conformations of blocked alanine peptides, where the repulsive part of van der Waals potential was taken into consideration to explain such phenomenon. More recently, Raines and coworkers have shown that an electronic effect due to the $n \rightarrow \pi^*$ interaction between neighboring peptides on the stability of PP_{II} structure in unfolded proteins is quantitatively important and they pointed out that such electronic interactions have been completely ignored in the molecular mechanical studies of proteins (Woolfson et al. 2010).

Despite that conformational preferences of a few series of blocked amino-acids and oligopeptides have been studied before, we have carried out NMR, CD, and MD simulation studies for a complete set of 400 blocked dipeptides (Ac–X_jX_k–NH_2) recently (Oh et al. 2012). There, to avoid any possible complications due to the zwitterionic terminal groups and their interactions with amides and amino-acid side-chains, we specifically considered blocked dipeptides whose N- and C-terminal groups are acetylated and amidated, respectively. Such dipeptides constitute a comprehensive library of ideal model systems and possibly they can be considered to be minimal size building units of unfolded proteins as long as their conformational preferences are similar to those of the amino-acids in unfolded proteins. However, the latter possibility has not been thoroughly examined before. Recently, performing extensive spectroscopic studies of the dipeptides, we were able to extract critical information on the neighboring peptide–peptide interaction-induced effects on the intrinsic backbone conformational preferences of amino-acids and found that both PP_{II} and β-strand conformers are preponderant. The experimental results indicate that the dipeptide conformations are different from coils and that their aqueous solution structures are weakly dependent on the amino-acid side-chain properties. This suggests that the backbone peptide solvation and electronic effect are important in determining conformational distributions of unfolded proteins instead of side-chain properties and chemical diversity of amino-acids.

In the present paper, we specifically focus on the H^N–H^α scalar coupling constants and their chemical shifts, denoted as δ_{HN} and $\delta_{H\alpha}$, respectively, of the dipeptides and shall present comparative investigation results. The coupling constant, denoted as $^3J_{HN\alpha}$, of the two protons in a given H^N–N–C^α–H^α group is determined by its dihedral angle ϕ. In addition to such investigations utilizing experimentally measured $^3J_{HN\alpha}$ values, chemical shift-based method was also shown to be of use to identify secondary structural elements (Pastore and Saudek 1990; Wishart et al. 1991a, b, 1992; Wishart and Sykes 1994a, b; Williamson et al. 1995; Osapay and Case 1994), to estimate the secondary structural propensities of short peptides (Rizo et al. 1993; Shin et al. 1993; Merutka et al. 1995), and to understand and quantify main chain flexibility (Wishart et al. 1991a). A critical set of ingredients for the success of chemical shift-based techniques is the data on the chemical shifts of amino-acids in coil structures. Often, they were specifically defined as those in the context of a polypeptide that is free to access all sterically allowed regions of conformational space. In this regard, an old notion is that short and simple peptides, such as blocked or unblocked GGXA (Bundi and Wuthrich 1979), GXG, GGXGG, and GGXAGG, are excellent models of random coils in a variety of solvent conditions. For instance, Plaxco et al. (1997) studied the effects of chemical denaturant, guanidine hydrochloride, on the amino-acid conformational preferences in a series of blocked pentapeptides GGXGG, examining the ϕ angle-sensitive coupling constants and ROESY cross peak intensities. They found that the denaturant does not significantly perturb intrinsic backbone conformational preferences, whereas Avbelj et al. (2006) showed that the denaturant affects backbone conformational preferences of blocked amino acids whose side chains are polar. Nonetheless, the notion that short peptides have no secondary structure elements has proven to be invalid. A series of NMR and CD studies on alanine-based oligopeptides over the last decade have clearly shown that short oligopeptides have strong propensities to form PP_{II} and β-strand conformers (Graf et al. 2007; Shi et al. 2002a, 2006; Makowska et al. 2006; Schweitzer-Stenner and Measey 2007), though there could exist small but non-negligible populations of various turn conformations in some cases (Schweitzer-Stenner et al. 2011). Here, we shall show that the conformational preferences exhibited by selected sets of dipeptides among the complete 400 dipeptides are in

excellent correlations with those of amino-acids in unblocked and blocked GGXGG oligopeptides studied by Merutka et al. (1995) and Plaxco et al. (1997), respectively. Furthermore and more importantly, the H^N and H^α chemical shifts of the present dipeptides are directly compared with those of denatured, unstructured, and intrinsically disordered proteins to establish a possible relationship between the backbone conformational preferences of amino-acids in dipeptides and those in such unfolded proteins. Here, we show that the two chemical shifts δ_{HN} and $\delta_{H\alpha}$ are particularly good indicators on the extent of protein unfolding. In this regard, it is highly instructive to examine the H^N and H^α chemical shifts of both native and denatured hen egg white lysozymes, which were reported by Redfield and Dobson (1988) and Schlob et al. (2005), respectively (see Fig. 1). The distributions of δ_{HN} and $\delta_{H\alpha}$ chemical shifts of the *native* lysozyme are significantly broad, which reflect the conformational heterogeneity of the constituent peptides in the native form. On the other hand, the narrow distributions of δ_{HN} and $\delta_{H\alpha}$ values of the *denatured* lysozyme indicate that the local solvation environments and possibly backbone conformations of the peptides are quite similar to one another and homogeneous. In fact, such narrowing or collapsing of the (δ_{HN}, $\delta_{H\alpha}$)-distribution upon increasing denaturant concentration or temperature has been considered to be one of the most notable and quick diagnostic NMR signatures of protein unfolding. Nevertheless, no attempt to directly compare the (δ_{HN}, $\delta_{H\alpha}$)-distribution of denatured proteins with that of blocked amino-acids or oligopeptides has been made before. Based on the present comparative investigations along this line, we shall show that the blocked dipeptides that are minimal size model systems taking into account the neighbor interaction effects can serve as a full collection of structural motifs that can be used to describe conformational distributions of denatured and unstructured proteins.

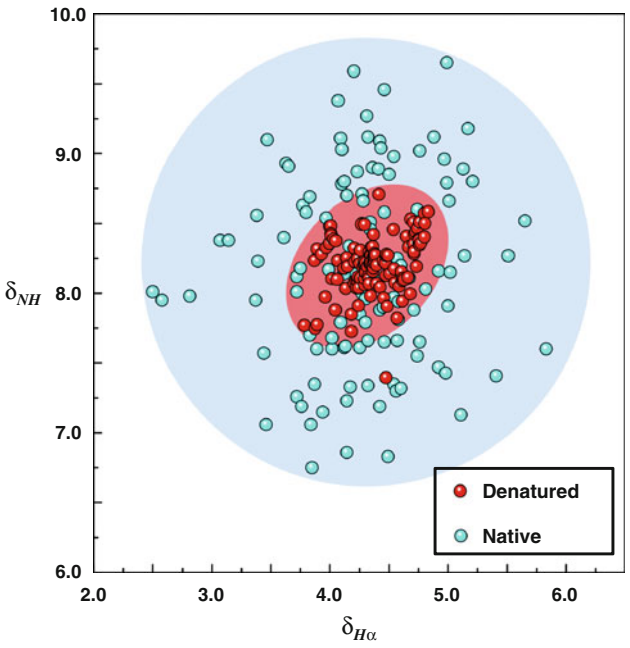

Fig. 1 NMR chemical shifts $\delta_{H\alpha}$ and δ_{HN} of both native and denatured hen egg white lysozymes. These data are taken from Redfield and Dobson (1988) and Schlorb et al. (2005)

Materials and methods

Peptide synthesis

The syntheses and characterizations of all the dipeptides were performed by Bead Tech (Seoul, Korea) with >95% purity. The blocked dipeptides were synthesized on the X-CTRa (Bead Tech, Seoul) with a LibrakitTM (Tokyo Rikakikai) solid-state synthesizer. Fmoc-amino acids were purchased from Bead Tech (Seoul, Korea). Fmoc-XX-OH was taken HBTU coupling for 1.5–2 h. Peptides were acetylated at the N terminus and amidated at the C terminus. All couplings were monitored using a ninhydrin test. Solvent washes with DMF, MeOH, CH_2Cl_2, and then DMF were performed right after deprotection, coupling, and capping. Peptide resin cleavage was performed with trifluoroacetic acid (TFA). Purity of the peptide lyophilized was determined using HPLC T1260 Infinity Micro-scale Purification/Spotting System (Agilent).

NMR spectroscopy and sample preparation

All the ^{1}H NMR spectra of the dipeptide aqueous solutions were measured on a Varian VnmrS 600 MHz NMR spectrometer equipped with a 5-mm $^1H\{^{13}C/^{15}N\}$ salt tolerant triple resonance cold probe. The Noesypresat pulse sequence was employed for water suppression to collect 1D ^{1}H NMR experiments, respectively. Temperature was controlled with a L99 temperature controller (Varian) and a TC-84 nitrogen air cooler (FTS Systems). Dipeptides were dissolved in D_2O/H_2O (1:9, pH 2) solution with 10–15 mM concentration at 25°C for 1D ^{1}H NMR spectra. A set of 38,462 complex data points was collected and 64 scans were averaged for aqueous solution. Before the Fourier transformation, the original free induction decay (FID) data were zero-filled to 524,288 points. The experimental NMR results were analyzed with ACD/SpecManager (ACD/Labs). The peak-to-peak frequency corresponding to $^3J_{HN\alpha}$ was estimated by carrying out spectral lineshape analysis with a Gaussian+Lorentzian function.

In total, we studied 361 (=19 × 19) blocked dipeptides ($Ac–X_1aa–X_2aa–NH_2$), where Xaa represents one of the nineteen amino acids (A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, Y) except for proline (P). In most cases, the

aqueous solution pH was adjusted to be 2 except for those containing acidic residues (D, E). In the latter cases, the solution pH was adjusted to be in the range from 4.1 to 5.0 (see Table S2 in Supplementary Material). In the cases of FF, FW, FV, FL, II, and MQ dipeptides, due to their low solubility in water, their NMR spectra couldn't be obtained. In some other cases, a small amount of DMSO (less than 10%) was used to prepare the solution samples (see Table S1 in Supplementary Material). For those dipeptides containing His residue, such as HH, KH, and TH, the two H^N NMR peaks of these dipeptides couldn't be distinguished from the H^N peak of the His imidazole ring, which appears at 8.61 ppm, so that they were not taken into consideration in the present data analysis. Furthermore, there are other cases (see the blanks in Table S5) that the H^α peaks of the dipeptides overlap with water NMR peak so that their corresponding chemical shift values couldn't be measured accurately. Finally, the GG dipeptide was excluded.

Results and discussion

Among the 400 dipeptides, those having very low solubility or containing Pro residue were excluded in the NMR experiments so that the NMR scalar coupling constants and chemical shifts were measured for 361 dipeptides (see "Materials and methods" section and Table S1 in Supplementary Material). In this paper, we present detailed NMR results on the chemical shifts of H^N and H^α as well as on the corresponding scalar coupling constants. We shall also examine possible correlations of these NMR properties with H–D exchange rates of amide protons, the δ_{HN} and $\delta_{H\alpha}$ values of other host–guest peptides such as GGXGG, and those of denatured and unstructured proteins.

Coupling constant $^3J_{HN\alpha}$

Protein backbone structure is mainly determined by two dihedral angles ϕ and ψ. They can be determined by using NMR method via measuring and analyzing the corresponding spin–spin coupling constants that are related to the dihedral angles by the appropriate Karplus equations. In particular, the $^3J_{HN\alpha}$ coupling constant of H^N–N –C^α–H^α directly provides us quantitative information on the backbone dihedral angle $\phi_i\left(C'_{i-1}-N_i-C^\alpha_i-C'_i\right)$. Thus, we carried out ^{1}H NMR measurements for the blocked dipeptides in D_2O/H_2O (=1:9) at 25°C and measured the $^3J_{HN\alpha}$ values of 361 dipeptides. For those dipeptides containing acidic residues, the pH of the aqueous solution was specifically controlled to be in the range from 4.1 to 5.0 to convert the side-chain COOH group into COO^- anion (see Table S2 in Supplementary Material). For varying pH at around the pK_a of the side chain COOH group, we measured the $^3J_{HN\alpha}$ coupling constants (see Table S3 in Supplementary Material) and found that they only change less than 0.3 Hz. In the present paper, we shall specifically consider those dipeptide solutions in the pH range from 4.1 to 5.0.

For each dipeptide Ac–X_jX_k–NH_2, there are two H^N–N–C^α–H^α groups so that two $^3J_{HN\alpha}$ values, denoted as $^3J^j_{HN\alpha}$ and $^3J^k_{HN\alpha}$, were extracted from the ^{1}H NMR spectrum. To examine the peptide backbone conformational distribution, we first plot the larger (orange circles) and smaller (grey circles) $^3J_{HN\alpha}$ of the peptides in Fig. 2. It is found that they are in the range from 5.3 to 8.5 Hz and the standard deviation (SD) is as small as about 0.59 Hz. For the sake of comparison, we also plot the reference $^3J_{HN\alpha}$ values, denoted as $^3J^{P_{II}}_{HN\alpha}$ and $^3J^{\beta}_{HN\alpha}$, respectively, of the PP_{II} and β-strand conformers of Gly-blocked amino-acids (Shi et al. 2005). In that paper by Shi et al., the corresponding ϕ-angles of the two conformers were determined by using the protein structures in coil library (Shi et al. 2005). Regardless of the side-chains of the amino-acids, the reference scalar coupling constants of PP_{II} and β-strand conformers are about 5.5 and 9.6 Hz, respectively. This indicates that the dipeptide backbone conformation is mainly a mixture of PP_{II} and β-strand structures, which is

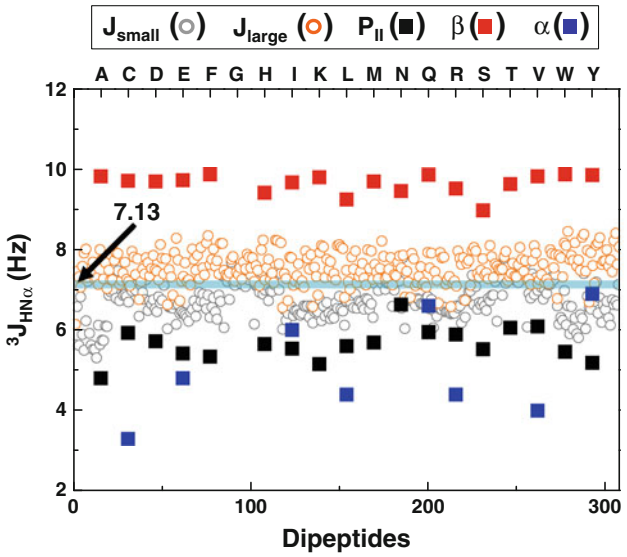

Fig. 2 $^3J_{HN\alpha}$ Coupling constants of 308 dipeptides. There are two amide groups for each dipeptide so that two different $^3J_{HN\alpha}$ coupling constants were obtained. J_{small} (*grey circles*) and J_{large} (*orange circles*) refer to the smaller and larger coupling constants, respectively. The average $\langle {}^3J_{HN\alpha}\rangle$ over all the $^3J_{HN\alpha}$ coupling constants is 7.13 Hz and the corresponding standard deviation is 0.59 Hz. In this figure, we also plot the reference $^3J_{HN\alpha}$ coupling constants of amino-acids in PP_{II} and β-strand conformations, which are denoted as $^3J^{P_{II}}_{HN\alpha}$ and $^3J^{\beta}_{HN\alpha}$ in the main text. The average $^3J^{P_{II}}_{HN\alpha}$ and $^3J^{\beta}_{HN\alpha}$ values for all the amino-acids are about 5.5 and 9.6 Hz. Thus, the fact that the $^3J_{HN\alpha}$ coupling constants of the dipeptides are distributed in the range from 5.5 to 9.6 Hz indicates that the backbone conformations of dipeptides are mixtures of PP_{II} and β-strand structures

consistent with the experimental finding by Avbelj et al. (2006) who showed that the fractional populations of PP_{II} and β-strand conformations are dominant in 19 blocked amino-acids (*N*-acetyl-X-*N*′-methylamide) except for Gly. In our previous work, we also showed that the CD spectra of the blocked dipeptides can be described as a linear combination of the CD eigenspectra of the PP_{II} and β-strand conformations (Oh et al. 2010, 2012) and further examined the validity of a two-state approximation by comparing our results with previous works done by other groups (Hagarman et al. 2010, 2011). However, it should also be emphasized that the fitting analysis method using reference values of two representative conformations such as PP_{II} and β-strand is intrinsically approximate in nature. The more general and probably ideal way to calculate an ensemble average value is to fully utilize the corresponding distribution function, which contains information on the conformational population distribution. Unfortunately, it is not possible to obtain the whole distribution function by any means. Despite the fact that a number of molecular dynamics simulation methods have been used to properly sample conformations of proteins and peptides, the simulated conformational distributions of short oligopeptides are not always in quantitative agreement with experimental results. Furthermore, none of the experimental methods including NMR, vibrational spectroscopy, CD, etc. provides direct and quantitative information on the distribution. Thus, it is perhaps inevitable to use a certain fitting analysis method using reference values corresponding to discrete representative structures.

Since the $^3J_{HN\alpha}$ values of various turn structures as well as α-helices are quantitatively similar to those of typical PP_{II} conformers, much detailed studies for determining conformational distributions of the dipeptides by using other spectroscopic means including IR (Grdadolnik et al. 2011), Raman (Hagarman et al. 2010), and two-dimensional IR spectroscopy (Cho 2008) need to be performed in the future. In the present paper, we shall compare our measured $^3J_{HN\alpha}$ values of the dipeptides with those of other short peptides to show that the conformational preferences of a variety of blocked amino-acids previously studied by other groups are in good correlation with those of even simpler systems like blocked GX, XG, and XX dipeptides.

Chemical shifts of peptide H^N and H^α

The NMR chemical shift is one of the most fundamental NMR parameters, and it provides structural information because local conformation and environment contributes to its magnitude. In fact, a number of experimental and computational investigations revealed that the NMR chemical shifts correlate with protein three-dimensional structures (Cornilescu et al. 1999; Han et al. 2011; Oldfield 1995; Osapay and Case 1994; Pastore and Saudek 1990; Shen et al. 2009; Szilagyi 1995; Williamson et al. 1995; Wishart and Nip 1998; Wishart and Sykes 1994a; Wishart et al. 1991a, b; Zhang et al. 2003). In particular, Wishart and coworkers reported the chemical shift indices (CSI) for amino-acids, which have been found to be useful to identify secondary structure contents of proteins (Wishart et al. 1992). They further established the RefDB (Re-referenced Protein Chemical Shift Database) to help NMR spectroscopists to derive chemical shift trends in proteins, where a few computational tools such as SHIFTX, SHIFTCOR, and SHIFTX2 were used (Han et al. 2011; Wishart and Nip 1998; Zhang et al. 2003). The NMR chemical shifts were also found to be useful to predict protein backbone torsion angles. Initially, Cornilescu et al. (1999) developed the TALOS program utilizing empirical relations of the ^{13}C, ^{15}N, and ^1H chemical shifts with the backbone torsion angles ϕ and ψ. More recently, Shen et al. (2009) developed an improved version TALOS+ program to predict protein backbone torsion angles from NMR chemical shifts. Also, the relationship between the H^N chemical shift and the secondary structure propensity (SSP) for amino-acids in solved proteins has been extensively discussed before. Often, the difference H^N chemical shift defined as the difference between the observed chemical shift and that of the random coil H^N chemical shift had been considered for theoretical and experimental studies (Oldfield 1995, 2002; Pastore and Saudek 1990; Wishart and Sykes 1994a; Wishart et al. 1992; Zhang et al. 2003; Wishart and Nip 1998; Lacroix et al. 1998; Asakura et al. 1995).

Here, it should be emphasized that, despite the fact that the term ‘random coil’ has been widely used before, its meaning was found to be different in various literatures. In old literatures, it was believed that the backbone conformations of short peptides such as GGXA and GGXGG should be random coils covering the entire Ramachandran space except for sterically disfavored regions. Consequently, the H^N and H^α chemical shifts and the corresponding scalar coupling constants obtained from their ^1H NMR spectra were used as a set of reference values for establishing relationships between such NMR properties and polypeptide secondary structures. However, over the last decade, it has been shown that the backbone conformations of such oligopeptides are not random but adopt PP_{II}, β-strand, and some turn structures that are in dynamic equilibria.(Avbelj et al. 2004, 2006; Cho et al. 2010; Pappu and Rose 2002; Schweitzer-Stenner et al. 2002, 2010, 2011; Schweitzer-Stenner and Measey 2007; Shi et al. 2002a, b) Thus, the notion that the conformation of small peptide in aqueous solution is a random coil has been proven to be incorrect. Although the scalar coupling constants of short peptides have been widely used to quantitatively determine the backbone conformational

distribution, the NMR chemical shifts of short peptide H^N and H^α have not been fully explored to elucidate the relationship between the backbone conformational distribution of unfolded proteins and that of short peptides. To achieve this goal, after analyzing experimentally measured chemical shifts of the H^N and H^α of the dipeptides, here we consider the average chemical shifts, defined as $\delta_{HN} = (\delta^j_{HN} + \delta^k_{HN})/2$ and $\delta_{H\alpha} = (\delta^j_{H\alpha} + \delta^k_{H\alpha})/2$, of the two H^N's and H^α's of each dipeptide Ac–X_jX_k–NH_2—note that, for series of GX and XG dipeptides, only X's H^N and H^α NMR peaks are taken into consideration.

The chemical shifts, δ_{HN} and $\delta_{H\alpha}$, of the dipeptides are plotted in Fig. 3 (see Tables S4 and S5 in Supplementary Material for the measured chemical shift data). From the H^N chemical shifts of the dipeptides, we found that the average chemical shift is 8.30 ppm and the SD is as small as 0.160 ppm. For the H^α chemical shifts, the average and the SD are found to be 4.39 and 0.181 ppm, respectively. In Fig. 3, for the sake of comparison, we also plot the δ_{HN} and $\delta_{H\alpha}$ values of amino-acids in the β-sheet, random coil, and α-helical conformations in proteins whose structures were previously determined (Zhang et al. 2003). Those ensemble-averaged chemical shifts for amino-acids are given in Table 1 for the sake of completeness. Here, the random coil refers to a non-α-helical and non-β-sheet conformation. For the chemical shifts for amino-acids in the PP_{II} conformation, we use the previously reported data obtained from p15 Fusion-Associated Small Transmembrane (FAST) protein (Top et al. 2011). As can be seen in Fig. 3a, b, the δ_{HN} and $\delta_{H\alpha}$ values of the dipeptides are generally smaller than those of amino-acids in the β-sheet conformation, whereas they are larger than those in the α-helical conformation. However, the δ_{HN} and $\delta_{H\alpha}$ values of amino-acids in the PP_{II} and random coil conformations of solved proteins are broadly distributed and their distributions significantly overlap with those of the dipeptides.

Since the ensemble-averaged chemical shifts for amino-acids of solved proteins (Zhang et al. 2003) are available (Table 1), we could calculate $\delta_{HN,\beta}$, $\delta_{HN,rc}$, $\delta_{HN,\alpha}$, $\delta_{H\alpha,\beta}$, $\delta_{H\alpha,rc}$, and $\delta_{H\alpha,\alpha}$ values for a given dipeptide X_1X_2, where $\delta_{HN,\beta} = \{\delta_{HN,\beta}(X_1) + \delta_{HN,\beta}(X_2)\}/2$ for example. Here, $\delta_{HN,\beta}(X_1)$ and $\delta_{HN,\beta}(X_2)$ are the chemical shifts of the H^N protons of the X_1 and X_2 amino-acids, which are given in Table 1. In Fig. 4, we compare these six chemical shifts with those of our dipeptides. In the cases of the H^N proton chemical shifts, our data are in excellent correlation with the $\delta_{HN,rc}$ values, where the correlation coefficient is 0.81. On the other hand, our H^N chemical shifts of the dipeptides do not exhibit any correlation with those obtained from the chemical shifts for amino-acids in the β-sheet or α-helical conformation. This observation suggests that the H^N chemical shift is sensitive to backbone conformation and the backbone conformation distribution of the dipeptides is likely to be similar to that of the so-called random coils. In contrast to the H^N chemical shifts, the H^α chemical shifts of the dipeptides are in good correlation with all three $\delta_{H\alpha,\beta}$, $\delta_{H\alpha,rc}$, and $\delta_{H\alpha,\alpha}$ values. This insensitivity of the H^α chemical shift on the backbone conformation suggests that the H^α chemical shift is mainly determined by its amino-acid type instead of the backbone conformation. This is

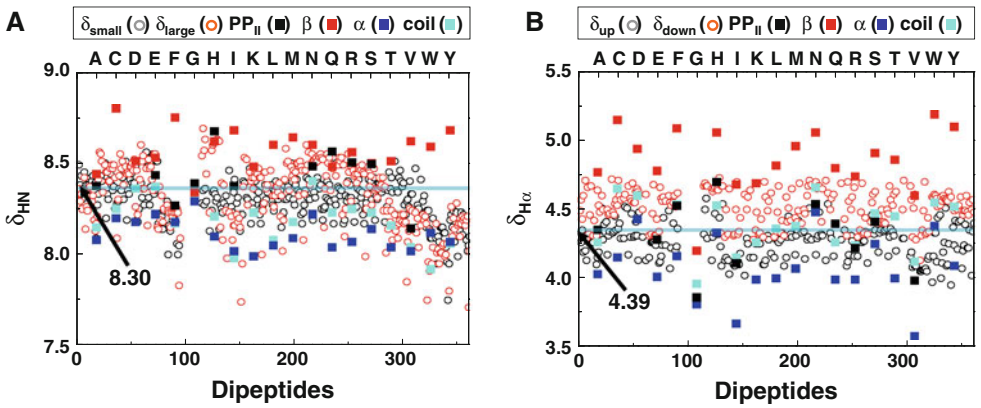

Fig. 3 The experimentally measured δ_{HN} and $\delta_{H\alpha}$ values of the dipeptides. There are two H^N and H^α protons for each dipeptide so that we have two different δ_{HN} and $\delta_{H\alpha}$ values experimentally measured, respectively. Similarly, we could measure two different $^3J_{HN\alpha}$ values for each dipeptide. In **a** δ_{small} (*black circles*) is the H^N chemical shift of the amino-acid whose $^3J_{HN\alpha}$ value is the smaller one among the two, whereas δ_{large} (*red circles*) is the H^N chemical shift of the amino-acid whose $^3J_{HN\alpha}$ value is the larger one among the two. We also plot the reference δ_{HN} values of amino-acids in the PP_{II}, β-sheet, α-helical, and coil conformations of solved proteins (see the main text). In **b**, δ_{up} (*grey circles*) and δ_{down} (*orange circles*) are the upfield and downfield chemical shifts of the two H^α protons of each dipeptide, respectively. The reference $\delta_{H\alpha}$ values of amino-acids in the PP_{II}, β-sheet, α-helical, and coil conformations of solved proteins are also plotted for the sake of comparisons. Here, the reference chemical shifts are taken from Top et al. (2011) (for PP_{II} conformation) and Zhang et al. (2003) (for β-sheet, α-helical, and coil conformations)

Table 1 Ensemble-averaged chemical shifts (in ppm) that were reported in Zhang et al. (2003) are given here for the sake of completeness and comparisons

	$^1H^N$			$^1H^\alpha$		
	Random coil	α-Helix	β-Strand	Random coil	α-Helix	β-Strand
A	8.15 (0.72)	8.08 (0.52)	8.44 (0.76)	4.26 (0.33)	4.03 (0.33)	4.77 (0.55)
C	8.25 (0.71)	8.20 (0.69)	8.80 (0.64)	4.65 (0.39)	4.15 (0.67)	5.15 (0.51)
D	8.36 (0.62)	8.18 (0.56)	8.51 (0.61)	4.60 (0.28)	4.43 (0.22)	4.94 (0.40)
E	8.37 (0.68)	8.22 (0.62)	8.53 (0.62)	4.28 (0.33)	4.01 (0.24)	4.78 (0.49)
F	8.17 (0.83)	8.18 (0.62)	8.75 (0.72)	4.54 (0.47)	4.16 (0.46)	5.09 (0.46)
G	8.33 (0.78)	8.29 (0.67)	8.34 (0.86)	3.96 (0.35)	3.81 (0.38)	4.20 (0.60)
H	8.21 (0.79)	8.10 (0.56)	8.62 (0.74)	4.53 (0.50)	4.33 (0.34)	5.06 (0.48)
I	7.98 (0.84)	8.02 (0.52)	8.68 (0.70)	4.15 (0.38)	3.67 (0.33)	4.68 (0.48)
K	8.23 (0.72)	7.99 (0.56)	8.48 (0.68)	4.26 (0.41)	3.99 (0.30)	4.69 (0.51)
L	8.08 (0.76)	8.05 (0.54)	8.60 (0.71)	4.36 (0.37)	4.00 (0.34)	4.82 (0.46)
M	8.18 (0.65)	8.09 (0.58)	8.64 (0.67)	4.38 (0.41)	4.07 (0.34)	4.96 (0.47)
N	8.40 (0.78)	8.22 (0.58)	8.60 (0.64)	4.66 (0.36)	4.48 (0.22)	5.06 (0.49)
Q	8.23 (0.65)	8.04 (0.55)	8.48 (0.66)	4.26 (0.34)	3.99 (0.28)	4.80 (0.49)
R	8.25 (0.67)	8.07 (0.55)	8.56 (0.64)	4.24 (0.43)	3.99 (0.32)	4.74 (0.50)
S	8.23 (0.65)	8.14 (0.56)	8.50 (0.67)	4.47 (0.35)	4.25 (0.25)	4.91 (0.48)
T	8.16 (0.69)	8.04 (0.51)	8.51 (0.61)	4.45 (0.36)	4.00 (0.34)	4.86 (0.46)
V	8.04 (0.65)	8.02 (0.65)	8.62 (0.69)	4.12 (0.41)	3.58 (0.36)	4.60 (0.48)
W	7.92 (0.89)	8.12 (0.74)	8.59 (0.83)	4.55 (0.48)	4.38 (0.37)	5.19 (0.50)
Y	8.06 (0.77)	8.07 (0.62)	8.68 (0.76)	4.52 (0.44)	4.09 (0.39)	5.10 (0.54)

The values in parenthesis are the standard deviations

quite consistent with the previous works showing that the H^α chemical shift is highly dependent on ring currents, torsion angles, H-bonds, nearest-neighbor effects, and solvation (Han et al. 2011).

Next, in Fig. 5a we compare the H^N chemical shifts with the H^α chemical shifts of the dipeptides. There appears no meaningful correlation between the two data. This is not surprising because the chemical shifts of the peptide H^N and H^α are dependent not only on backbone conformation but more strongly on solvent and other local environmental effects such as ring currents, torsion angles, nearest-neighboring effects, and so on. Therefore, we should rather consider the chemical shift difference between the observed chemical shift and the reference value. In the above, we have calculated the $\delta_{HN,\beta}$, $\delta_{HN,rc}$, $\delta_{HN,\alpha}$, $\delta_{H\alpha,\beta}$, $\delta_{H\alpha,rc}$, and $\delta_{H\alpha,\alpha}$ values for the dipeptides using the ensemble-averaged chemical shifts for amino-acids in solved proteins. Thus, the net chemical shift values after correcting these reference values are to be compared with each other to find out which reference data set serves as the best reference chemical shifts for structural analyses.

In Fig. 5b, the difference chemical shift defined as $\Delta\delta_{H\alpha,\beta} = \delta_{H\alpha} - \delta_{H\alpha,\beta}$ is compared with $\Delta\delta_{HN,\beta}$ ($\equiv \delta_{HN} - \delta_{HN,\beta}$). Hereafter, $\Delta\delta_{H,\beta}$ ($\Delta\delta_{H,rc}$) approximately represents the *net* chemical shift of the H-proton without any contribution from the chemical shift of β-sheet (random coil) conformation (see Tables S4 and S5 in Supplementary Material for all the measured chemical shifts for the dipeptides). As can be seen in Fig. 5b, the $\Delta\delta_{H\alpha,\beta}$ values barely correlate with the $\Delta\delta_{HN,\beta}$ values, where the Pearson correlation coefficient is about 0.64. However, as can be seen in Fig. 5c, d, there appear no notable correlations between $\Delta\delta_{H\alpha,rc}(\equiv\delta_{H\alpha} - \delta_{H\alpha,rc})$ and $\Delta\delta_{HN,rc}(\equiv\delta_{HN} - \delta_{HN,rc})$ as well as between $\Delta\delta_{H\alpha,\alpha}$ and $\Delta\delta_{HN,\alpha}$. These correlation analysis results indicate that a single set of reference chemical shifts is not sufficiently useful enough to establish the relationship between the peptide H^N and H^α chemical shifts and the corresponding backbone conformations of the dipeptides and furthermore the dipeptides do not adopt a single dominant conformation but a mixture of two or more conformations.

Peptide backbone conformations of Ac–X–NHMe, GX, XG, XX, GGXGG, and other coil libraries

Recently, we compared the NMR properties with those extracted from some coil libraries (Oh et al. 2012). In this paper, we additionally compare the $^3J_{HN\alpha}$ values of the dipeptides with those of the residues in the other coil libraries of proteins (Serrano 1995; Swindells et al. 1995) as well as those of short oligopeptides studied previousely (Avbelj et al. 2006; Bundi and Wuthrich 1979; Plaxco et al. 1997; Shi et al. 2005). Serrano (1995) reported the average $^3J_{HN\alpha}$-coupling

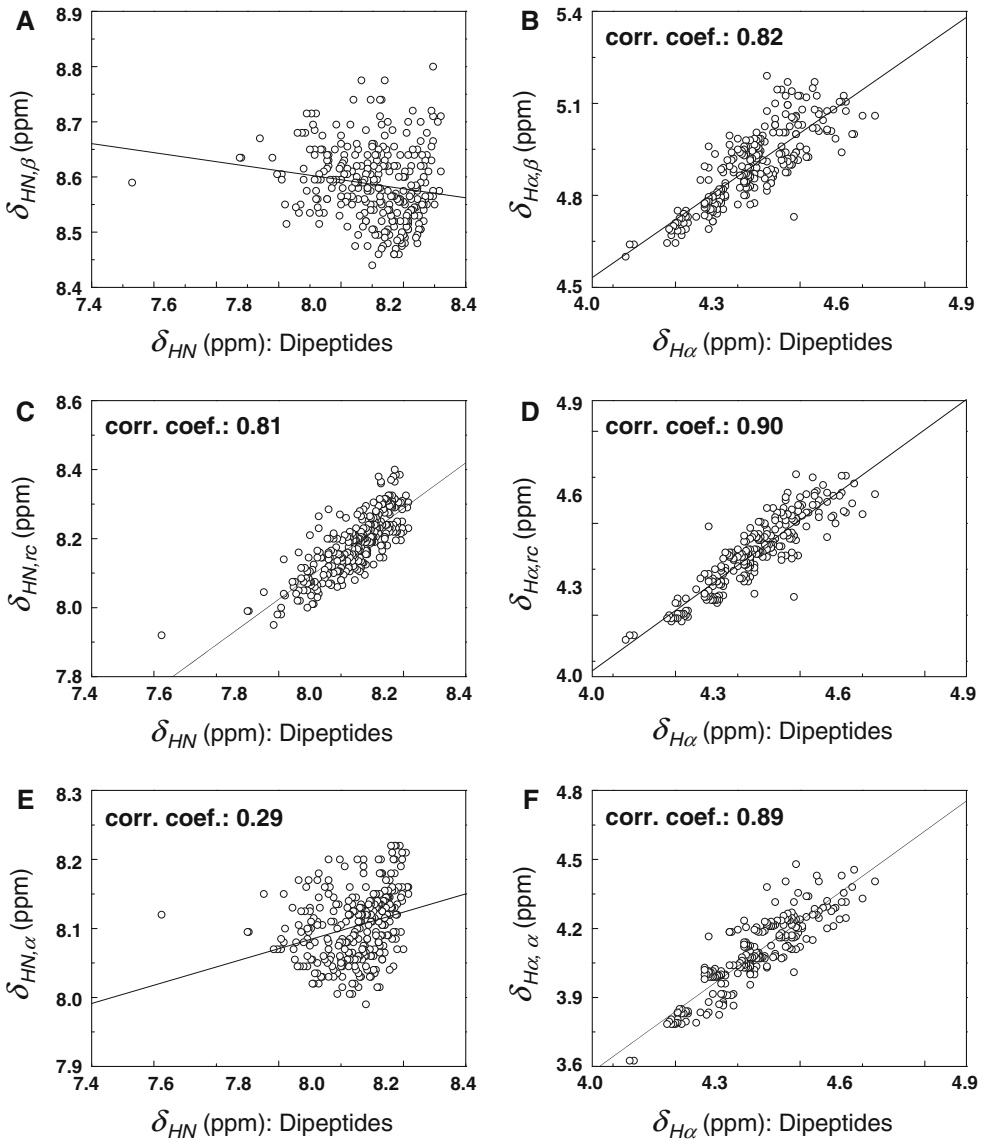

Fig. 4 Comparisons between the experimentally measured chemical shifts ($\delta_{H\alpha}$ and δ_{HN}) of the dipeptides and those of the dipeptides calculated by using the reference chemical shifts in Zhang et al. (2003). **a** δ_{HN} versus $\delta_{HN,\beta}$, **b** $\delta_{H\alpha}$versus $\delta_{H\alpha,\beta}$, **c** δ_{HN} versus $\delta_{HN,rc}$, **d** $\delta_{H\alpha}$ versus $\delta_{H\alpha,rc}$, **e** δ_{HN} versus $\delta_{HN,\alpha}$ and **f** $\delta_{H\alpha}$ versus $\delta_{H\alpha,\alpha}$. The correlation coefficients are given in the *upper-left* region of each plot

constants for amino-acids in the random coil conformation. We cannot directly compare their results with the experimentally measured $^3J_{HN\alpha}$ values of the present dipeptides, because the $^3J_{HN\alpha}$ value of a given X_1X_2 dipeptide is the average over the $^3J_{HN\alpha}$ values of the two H^N protons of the dipeptide and it does not represent the average coupling constant of a given amino-acid. Therefore, we need to calculate the doubly averaged $^3J_{HN\alpha}$-coupling constant for each amino-acid. First, let us denote $^3J_{HN\alpha}(X_1, X_2)$ to be the mean coupling constant of the X_1X_2 dipeptide, which is the experimentally measured value. Then, we could take an average over X_2 (C-terminal amino-acids) for a fixed N-terminal amino-acid Xaa as

$$^3J^{N}_{HN\alpha}(\mathrm{Xaa}) = \frac{1}{n}\sum_{X_2=A}^{Y} {}^3J_{HN\alpha}(\mathrm{Xaa}, X_2). \quad (1)$$

Similarly, we calculated an average over X_1 (N-terminal amino-acids) for a fixed C-terminal amino-acid Xaa as

$$^3J^{C}_{HN\alpha}(\mathrm{Xaa}) = \frac{1}{n}\sum_{X_1=A}^{Y} {}^3J_{HN\alpha}(X_1, \mathrm{Xaa}). \quad (2)$$

Thus, $^3J^{N}_{HN\alpha}$(Xaa) and $^3J^{C}_{HN\alpha}$(Xaa) are the average $^3J_{HN\alpha}$ values of the Xaa amino-acid when it is the N- and C-terminal residue in the dipeptides, respectively. Then, the mean value of $^3J^{N}_{HN\alpha}$(Xaa) and $^3J^{C}_{HN\alpha}$(Xaa), denoted as $^3\bar{J}_{HN\alpha}$, does not depend on whether it is the N- or C-terminal residue, and it is calculated by

$$^3\bar{J}_{HN\alpha}(\mathrm{Xaa}) = \{^3J^{N}_{HN\alpha}(\mathrm{Xaa}) + {}^3J^{C}_{HN\alpha}(\mathrm{Xaa})\}/2. \quad (3)$$

The quantity $^3\bar{J}_{HN\alpha}$(Xaa) is in fact a doubly averaged $^3J_{HN\alpha}$ coupling constant of Xaa.

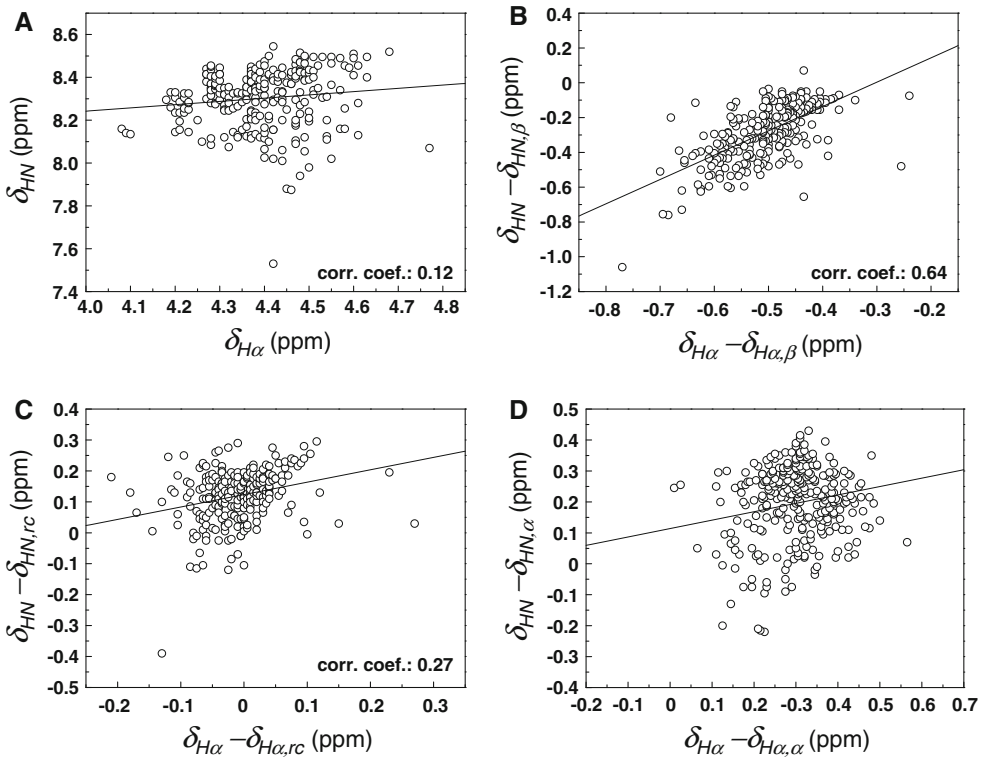

Fig. 5 The experimentally measured $\delta_{H\alpha}$ values of the dipeptides and the reference-corrected values are compared with the experimentally measured δ_{HN} values of the dipeptides and the reference-corrected ones. **a** δ_{HN} versus $\delta_{H\alpha}$, **b** $\delta_{HN} - \delta_{HN,\beta}$ versus $\delta_{H\alpha} - \delta_{H\alpha,\beta}$, **c** $\delta_{HN} - \delta_{HN,rc}$ versus $\delta_{H\alpha} - \delta_{H\alpha,rc}$, **d** $\delta_{HN} - \delta_{HN,\alpha}$ versus $\delta_{H\alpha} - \delta_{H\alpha,\alpha}$. Here, the reference chemical shifts, $\delta_{HN,\beta}$, $\delta_{HN,rc}$, $\delta_{HN,\alpha}$, $\delta_{H\alpha,\beta}$, $\delta_{H\alpha,rc}$, and $\delta_{H\alpha,\alpha}$, for the dipeptides are calculated by using the ensemble-averaged chemical shifts for amino-acids in solved proteins (Table 1)

Now, we can make direct comparisons of our ${}^3\bar{J}_{HN\alpha}(Xaa)$ values with Serrano's average ${}^3J_{HN\alpha}$ values for amino-acids found in the random coil regions of solved proteins (see the black circles in Fig. 6a). The mean ${}^3\bar{J}_{HN\alpha}(Xaa)$ values extracted from the coupling constants of our dipeptides are found to be in excellent correlation with the random coil data and the correlation coefficient is as large as 0.97. The red dots in Fig. 6a are those obtained by using Karplus equation with average ϕ angles of the amino-acids in the random coils of proteins. Again, the correlation between the two data sets is quantitative. In Fig. 6b–d, we compare the b-coil, B-coil, and β-strand propensities of amino-acids, which were reported by Swindells et al. (1995), with the ${}^3\bar{J}_{HN\alpha}$ values of amino-acids obtained from the ${}^3J_{HN\alpha}$ values of the dipeptides. Here, the b-coil refers to the conformations in the upper left region of the Ramachandran plot and it corresponds to typical β-sheet structures. The conformations around the PP_{II} structure were referred to as p-coils. Then, the B-coil region represents the sum of b-coil and p-coil regions. From Fig. 6b, one can find that the doubly averaged ${}^3\bar{J}_{HN\alpha}$ values of amino-acids correlate with their b-coil propensities (see Fig. 6b), whereas we couldn't find meaningful correlation between the ${}^3\bar{J}_{HN\alpha}$ values and B-coil propensities or between the ${}^3\bar{J}_{HN\alpha}$ values and β-strand propensities. These observations indicate that the conformational distribution of the dipeptides, which is deduced by the doubly averaged ${}^3\bar{J}_{HN\alpha}$ values for amino-acids in the dipeptides, is similar to those of random coils and b-coils (or extended structures).

One of the most thoroughly investigated model systems for elucidating backbone conformational propensities of amino-acids in short peptides or denatured proteins is a series of Gly-blocked amino-acids, e.g., GGXGG. Using NMR methods, Wuthrich and coworkers examined amino-acid conformational propensities by considering a series of tetrapeptides, GGXA, because it was believed that such short peptides could be good model systems for random coil peptides at that time (Bundi and Wuthrich 1979). Later, Merutka et al. (1995) studied GGXGG peptides with free amino and carboxyl groups, whereas Plaxco et al. (1997) and Shi et al. (2005) considered blocked GGXGG pentapeptides. In the latter studies on the GGXGG pentapeptides, the N-terminal amino group and the C-terminal carboxyl group were acetylated and amidated, respectively, to avoid possible complications originating from electrostatic interactions between charged terminal groups (ammonium and carboxylate) and backbone peptide bonds or polar side groups.

For the sake of direct comparisons, the GX, XG, and XX blocked dipeptides are specifically considered here to address the question about whether the longer peptides studied before have similar conformational preferences of even shorter dipeptides. First, the experimentally measured δ_{HN}'s of the blocked GX, XG, and XX dipeptides are compared with those of unblocked GGXGG pentapeptides

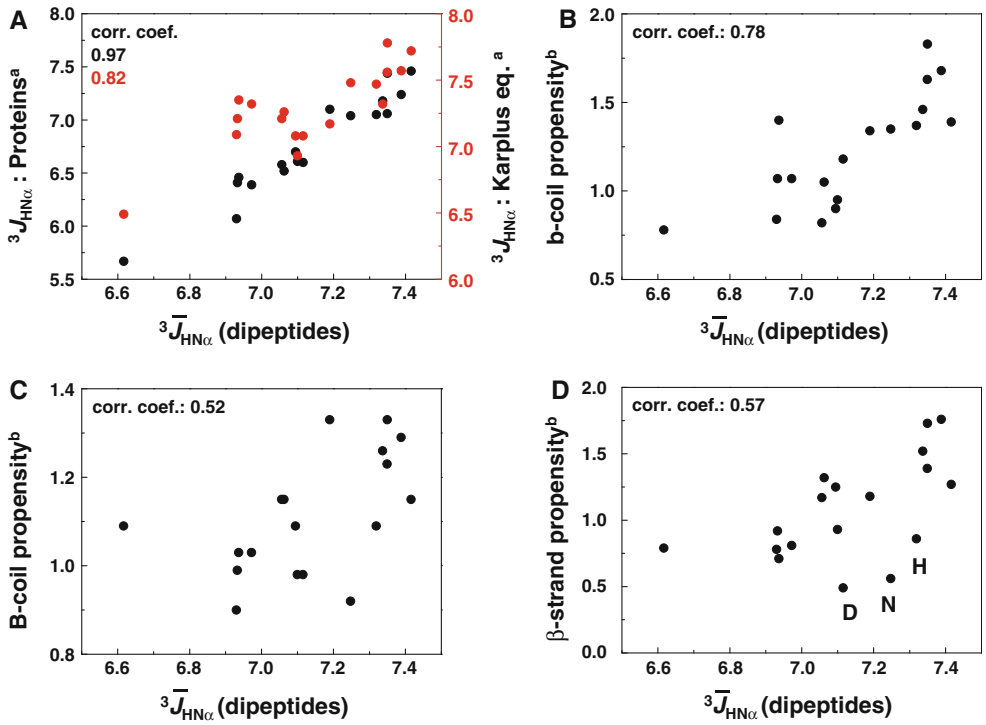

Fig. 6 Doubly averaged coupling constants, $^3\bar{J}_{HN\alpha}$, for amino-acids that are obtained from the $^3J_{HN\alpha}$-coupling constants of the dipeptides are compared with (**a**) the ensemble-averaged $^3J_{HN\alpha}$ values for amino-acids found in the random coil regions of solved proteins (*black circles*) and with the calculated $^3J_{HN\alpha}$ coupling constants for amino-acids (*red circles*) obtained by using the Karplus equation and average ϕ angles of amino-acids in random coils (Serrano 1995). The $^3\bar{J}_{HN\alpha}$ values are also compared with the b-coil (**b**), B-coil (**c**), and β-strand propensities of all the amino-acids that are taken from Swindells et al. (1995)

studied by Merutka et al. as well as with those of blocked GGXGG peptides studied by Plaxco et al. (see Fig. 7a, b, respectively). It is found that the H^N chemical shifts of the XG and GX series are in excellent agreement with those of the unblocked GGXGG series studied by Merutka et al.—note that the Pearson correlation coefficients are about 0.9. Second, the δ_{HN} values of the GX and XG dipeptides also correlate with those of the blocked GGXGG pentapeptides studied by Plaxco et al. (Fig. 7b). These observations indicate (1) that the N- and C-terminal amino and carboxyl groups do not make notable differences in the backbone conformational preferences of the middle amino-acids in these series of Gly-blocked pentapeptides at all and (2) that the local environmental and electronic effects on the H^N chemical shifts do not strongly depend on the blocking groups for each amino-acid. Furthermore, it is believed that the backbone conformational preferences of these short GX and XG dipeptides are similar to those of the GGXGG pentapeptides, irrespective of the additional blocking groups at the amino- and carboxyl-terminals.

We next compare the average H^N chemical shifts of all the homo-dipeptides (Ac–XX–NH_2) with those of unblocked and blocked GGXGG pentapeptides in Fig. 7c. Once the two amino-acids, Trp and Tyr, containing an aromatic ring are excluded in the correlation analyses, the relative correlation coefficients appear to be fairly large ($\sim$0.9). The deviations found in the two cases of the Trp and Tyr homo-dipeptides can be understood by noting that the aromatic rings in the Trp and Tyr residues can significantly affect on the peptide backbone H^N chemical shift due to its ring current effect. More specifically, a given aromatic side chain might influence the chemical shift of the H^N chemical shift of its *neighbor* residue, which is absent in the cases of the GGXGG pentapeptides as well as the GX and XG dipeptides.

Next, we compare the scalar coupling constants of the X amino-acids in the GX, XG, and XX blocked dipeptides with those of blocked GGXGG pentapeptides studied by Shi et al. and Plaxco et al. in Figs. 8a, b, respectively. In Fig. 8c, d, the coupling constants of the X amino-acids in the present XX dipeptides and in the Ac–X–NHMe (Avbelj et al. 2006) are compared with those in the GGXGG pentapeptides and with our XG and GX dipeptides, respectively. Again, they all show excellent correlations with one another, which clearly indicates that the amino-acid backbone conformations and the local environments around the peptides in the series of GX, XG, and XX blocked dipeptides are quite similar to those of the GGXGG pentapeptides.

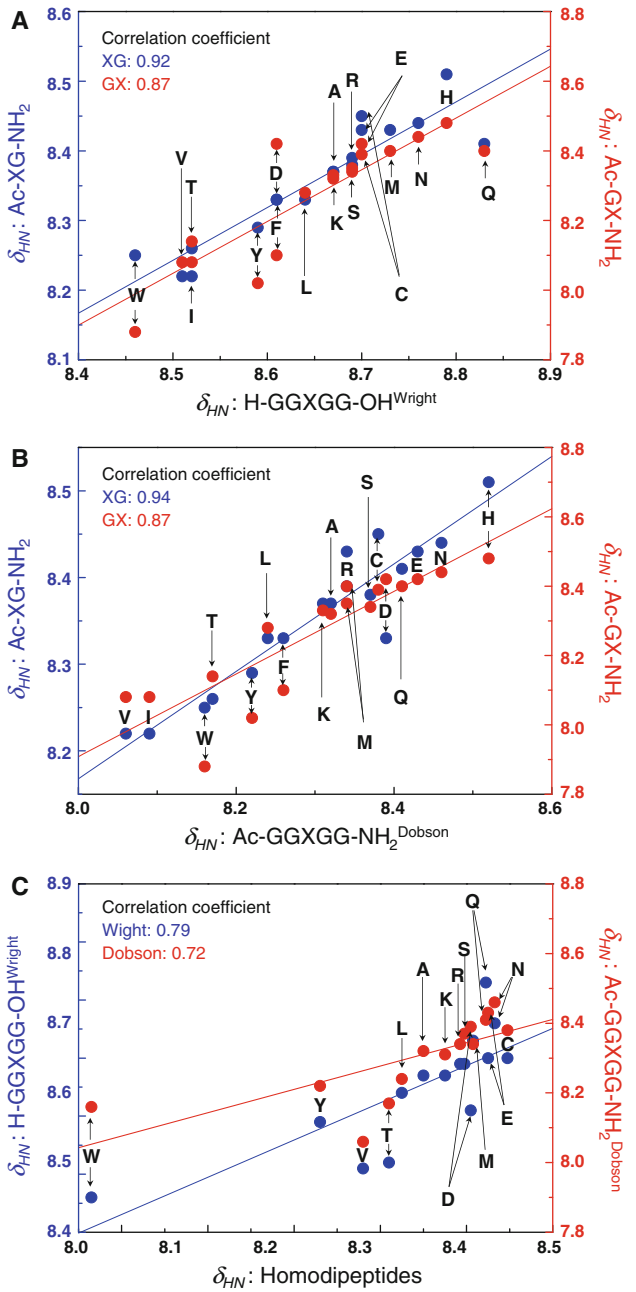

Fig. 7 The H^N chemical shifts of our blocked GX, XG, and XX dipeptides versus those of unblocked and blocked GGXGG pentapeptides. **a** The H^N chemical shifts of amino-acids in a series of unblocked GGXGG peptides, which were reported by Merutka et al. (1995) are compared with those of the Ac–GX–NH_2 and Ac–XG–NH_2 dipeptides studied in this work. Here, the cases that X = Gly and Pro are not considered. The correlation coefficients are given in each plot. **b** The H^N chemical shifts for amino-acids in a series of blocked Ac–GGXGG–NH_2 peptides, which were reported by Plaxco et al. (1997) are compared with those of Ac–GX–NH_2 and Ac–XG–NH_2 dipeptides. **c** The average H^N chemical shifts for blocked dipeptides (Ac–XX–NH_2) are compared with those of unblocked and blocked GGXGG peptides

Comparisons between chemical shifts and hydrogen exchange rates

Another useful information about the structure and conformational dynamics of proteins can be extracted by measuring the hydrogen exchange (HX) rates (Bai et al. 1993; Molday and Kallen 1972; Huyghues-Despointes et al. 1999). The HX rate depends on how strongly the corresponding peptide NH group is exposed to surrounding water molecules or catalytic hydronium or hydroxide ion. In particular, to extract structural information from the measured HX rates, one should compare them with reference values obtained from coils or other appropriately chosen reference peptide systems. Here, we shall compare the experimentally measured chemical shifts, δ_{HN} and $\delta_{H\alpha}$, of our blocked dipeptides with the HX rates reported by Bai et al. (1993).

Bai et al. carried out kinetic studies of the peptide NH → ND exchange processes with blocked amino-acids (Ac–X–NHMe). Since there are two N–H groups in each blocked amino-acid, they measured the HX rates of the left (L) and right (R) NH groups separately for varying pD. At a low (high) pD, the acid (base) catalysis mechanism dominates the HX process. Since the two (left and right NH groups) HX rates were measured for varying pD, they were able to obtain the acid-catalyzed, base-catalyzed, and water-catalyzed HX rates separately. The base-catalyzed process involves a direct abstraction of the H^N by OD^- or OH^- ion in the aqueous solution, whereas the acid-catalyzed HX reaction involves either deuteration or protonation of the peptide NH group by D^+ or H^+ ion. Therefore, these HX reactions regardless of acid- or base-catalysis are a first-order kinetic process with respect to the concentration of OD^- (or OH^-) or D^+ (or H^+) ion in the aqueous solution. Since such HX reactions require a chemical encounter between peptide NH group and these reactive ionic species, the HX rate reveals quantitative information on the extent of exposure of the target peptide NH group to solvent. In Bai et al. (1993), they particularly considered the HX rate increments for varying amino acids in reference to the HX rates of blocked alanine in a logarithmic form, i.e., log k_{ex}(Xaa) – log k_{ex}(Ala). Consequently, four different sets of data were given in their paper, which are (1) acid-catalyzed left (L) H^N HX rates, (2) acid-catalyzed right (R) H^N HX rates, (3) base-catalyzed left (L) H^N HX rates, and (4) base-catalyzed right (R) H^N HX rates.

In order to examine how strongly those difference HX rates correlate with the two chemical shifts δ_{HN} and $\delta_{H\alpha}$ of our blocked dipeptides, we plot them in Fig. 9. Figure 9a, b depict the acid-catalyzed HX rates versus the doubly averaged δ_{HN} and $\delta_{H\alpha}$ values of the dipeptides, respectively. They show no notable correlations at all. In Figs. 9c, d, the doubly averaged chemical shifts for amino-acids

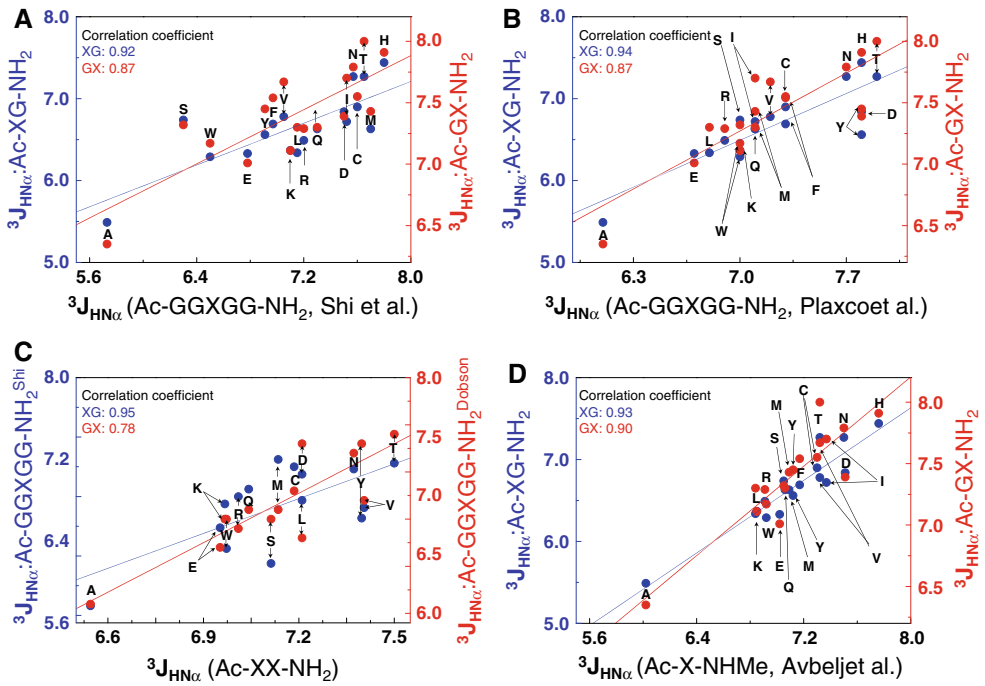

Fig. 8 The $^3J_{HN\alpha}$ coupling constants of our GX, XG, and XX dipeptides versus those of blocked GGXGG pentapeptides. **a** The $^3J_{HN\alpha}$ coupling constants for amino-acids in a series of blocked GGXGG peptides, which were reported by Shi et al. (2005), are compared with those of the Ac–GX–NH$_2$ and Ac–XG–NH$_2$ dipeptides studied in the present work. **b** The $^3J_{HN\alpha}$ coupling constants for amino-acids in a series of blocked Ac–GGXGG–NH$_2$ peptides, which were reported by Plaxco et al. (1997), are compared with those of the Ac–GX–NH$_2$ and Ac–XG–NH$_2$ dipeptides. **c** The average $^3J_{HN\alpha}$ coupling constants of our blocked homo-dipeptides (Ac–XX–NH$_2$) are compared with those of blocked GGXGG peptides reported by two different groups. Note that Plaxco et al. considered guanidine hydrochloride solutions of the pentapeptides. **d** The $^3J_{HN\alpha}$ coupling constants for blocked amino-acids reported in Avbelj et al. (2006) are compared with those of the Ac–GX–NH$_2$ and Ac–XG–NH$_2$ dipeptides

obtained from those of the dipeptides are compared with the base-catalyzed HX rate differences. There appears meaningful correlation between the two data sets. Noting that the base-catalyzed HX process occurs through a direct attack of the OH^- or OD^- anion to the peptide NH group, we believe that the present correlation analyses suggest that the magnitudes of the H^N and H^α chemical shifts are partly determined by how strongly each peptide is exposed to solvent. This result is in fact consistent with the observation made by Avbelj et al. (2004), where they showed that the H^α chemical shift is related to the peptide backbone solvation.

H^N and H^α chemical shifts of dipeptides and denatured proteins

Marsh et al. introduced the SSP score, which is obtained by considering the difference between the observed chemical shift and the reference chemical shift (α-helix and β-sheet). The sign and amplitude of the SSP score thus provides information on the α-helical and β-sheet propensities. The SSP score calculation method was used to explain the different fibrilization tendencies of α- and γ-synuclein, which are intrinsically disordered proteins (Marsh et al. 2006). In the present work, we show that the secondary structure propensities of residues in peptides and proteins can be compared with one another by plotting the distributions of the experimentally measured $\delta_{H\alpha}$ and δ_{HN} data for various peptides and proteins. More specifically, comparing the distribution of the experimentally measured $\delta_{H\alpha}$ and δ_{HN} values of the dipeptides with those of denatured or intrinsically disordered proteins, we shall show that the ensemble of the dipeptide conformations can be a representative basis set for describing the local backbone conformations of various denatured or unstructured proteins.

Particularly, we make direct comparisons of the chemical shifts δ_{HN} and $\delta_{H\alpha}$ of the dipeptides with those of denatured lysozyme, denatured OmpX (outer membrane protein X) from *Escherichia coli*, unstructured Domain 2 of the protein 5A(NS5A) of HCV (Hepatitis C virus), and intrinsically disordered hNlg3cyt (intracellular domain of human NL3). In Fig. 10, the distributions of the dipeptides' δ_{HN} and $\delta_{H\alpha}$ data points (red circles) are directly compared with those of the four proteins. In the same figure, we also plot the δ_{HN} and $\delta_{H\alpha}$ values of those peptides in α-helices (yellow) and β-sheets (cyan). Typically, the chemical shifts δ_{HN} and $\delta_{H\alpha}$ of the peptides in native proteins are found to be widely distributed, in comparison to those of denatured

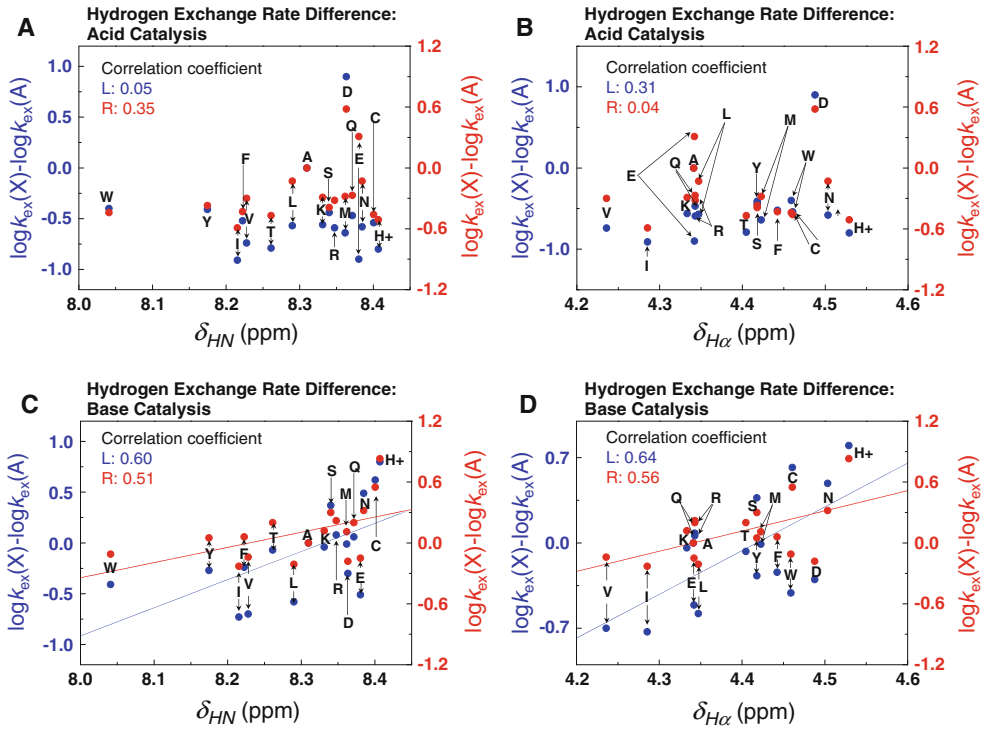

Fig. 9 The H–D exchange rate differences are compared with the NMR chemical shifts of the H^N and H^α protons of the dipeptides considered in this work. **a** The acid-catalyzed H–D exchange rate differences (Bai et al. 1993) defined as log k_{ex}(Xaa) – log k_{ex}(Ala) are compared with the doubly averaged H^N chemical shifts for amino-acids that are obtained from the chemical shifts of the dipeptides. **b** The acid-catalyzed H–D exchange rate differences are compared with the doubly averaged H^α chemical shifts. **c** The base-catalyzed H–D exchange rate differences are compared with the doubly averaged H^N chemical shifts. **d** The base-catalyzed H–D exchange rate differences are compared with the doubly averaged H^α chemical shifts

proteins (see Fig. 1 for example). Thus, such narrowing of the distribution of the chemical shifts δ_{HN} and $\delta_{H\alpha}$ has been considered to be a direct signature of protein denaturation. However, the hidden origin on such denaturation-induced narrowing of the chemical shift distributions has not been completely understood yet.

Now, for the sake of comparisons of the δ_{HN} and $\delta_{H\alpha}$ values of our dipeptides with those of the four proteins, we have used the following procedure. For a given protein with N amino-acids, let us denote its primary sequence as $X_1X_2\cdots X_N$. Then, from the known primary structure of the protein, one can consider $N-1$ dipeptide fragments such as X_1X_2, X_2X_3, and so on. For the jth dipeptide fragment, X_jX_{j+1}, the corresponding δ_{HN} and $\delta_{H\alpha}$ values of the protein are already known so that one can obtain the average chemical shifts as $\delta_{HN}^{j,j+1} = (\delta_{HN}^{j} + \delta_{HN}^{j+1})/2$ and $\delta_{H\alpha}^{j,j+1} = (\delta_{H\alpha}^{j} + \delta_{H\alpha}^{j+1})/2$. These values for the entire dipeptide fragments of a chosen protein are then compared with the δ_{HN} and $\delta_{H\alpha}$ values of our dipeptides, Ac–X_jX_{j+1}–NH_2 (see Fig. 10). Similarly, the corresponding δ_{HN} and $\delta_{H\alpha}$ values of the peptides in α-helices and β-sheets shown in Fig. 10 were obtained by using the δ_{NH}^{j} and $\delta_{H\alpha}^{j}$ values of the amino-acids in Zhang et al. (2003), which are presented in Table 1. As can be seen in Fig. 10, one can immediately find that the (δ_{HN}, $\delta_{H\alpha}$)-distribution of the present dipeptides significantly overlaps with those of the denatured and unstructured proteins—note that the x- and y-axis scales in Fig. 10 are far much narrower than those in Fig. 1.

To calculate the overlap between a pair of distributions in Fig. 10, we first obtain the distribution function for each set of data points, where the two-dimensional frequency distribution $P(\delta_{HN}, \delta_{H\alpha})$ is obtained by counting the number of data points in each two-dimensional grid with δ_{HN} and $\delta_{H\alpha}$ intervals of 0.1 ppm. Here, each individual 2D distribution is normalized by dividing the un-normalized distribution function with the total number of data points. Then, the overlap, denoted as S, between two normalized distributions is defined as $\sum_{\delta_{HN}}\sum_{\delta_{H\alpha}} P_1(\delta_{HN},\delta_{H\alpha})P_2(\delta_{HN},\delta_{H\alpha})$. Thus, the overlap S value is in between 0 and 1. The calculated overlap S values are summarized in Table 2. Although the four proteins are not homologous to one another at all, meaning that their primary, secondary, and tertiary structures are completely different from one another, the distributions of the H^N and H^α chemical shifts appear to be very similar ($S > 0.5$). Furthermore, the S values between those of proteins and that of our blocked

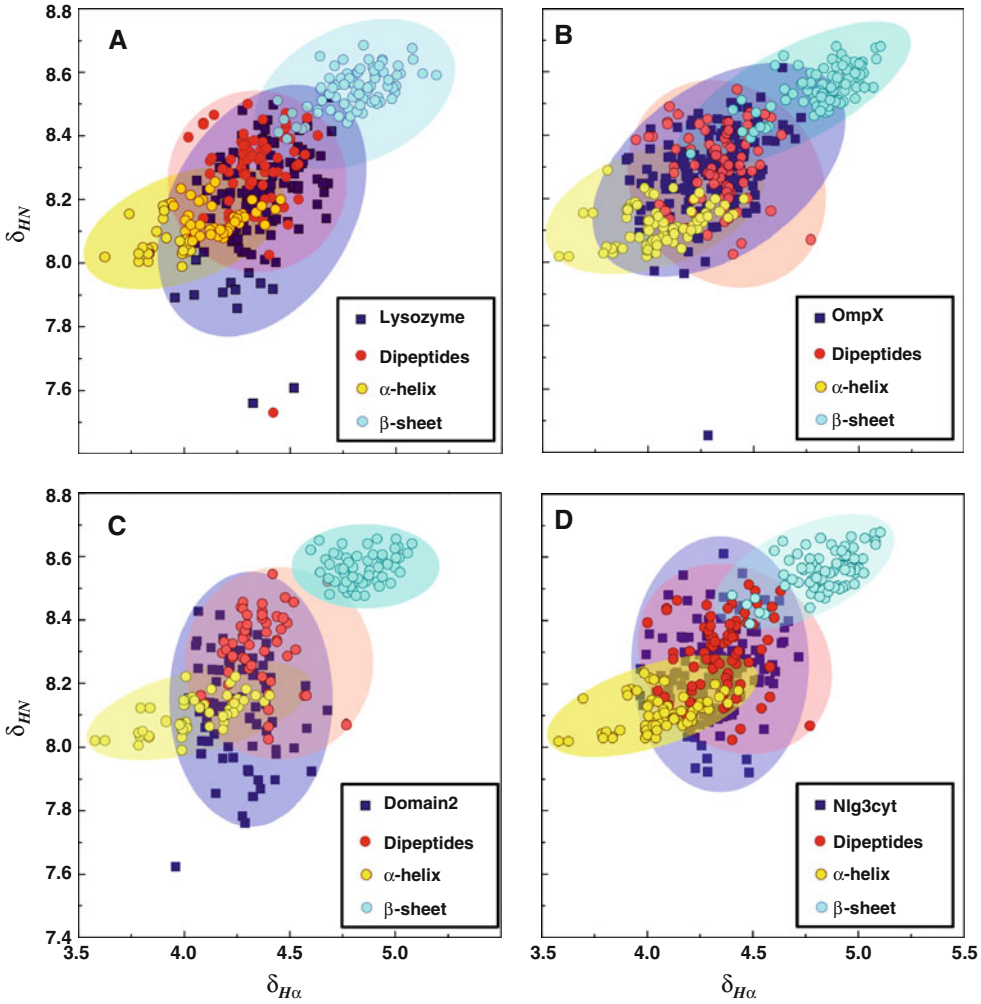

Fig. 10 Distributions of the NMR chemical shifts of the H^N and H^{α} protons of the present dipeptides, α-helix, β-sheets, denatured lysozyme (Schlorb et al. 2005), denatured OmpX (Tafer et al. 2004), unstructured Domain 2 of the protein 5A of Hepatitis C virus (Liang et al. 2007), and intrinsically disordered hNlg3cyt (Wood et al. 2011). The amino-acid sequence of each protein considered here is already known so that the primary structure information is used to obtain the chemical shifts of the jth dipeptide fragment (X_jX_{j+1}) (see the main text) for the sake of comparisons in this figure

Table 2 Calculated overlap (S) values, where S is defined as the product of a given pair of normalized $P(\delta_{HN}, \delta_{H\alpha})$ distributions

	Dipeptides	α-Helix	β-Sheet	OmpX	Domain2	hNlg3cyt
Lysozyme	0.60	0.30	0.04	0.73	0.59	0.85
OmpX	0.85	0.18	0.10	1	0.63	0.75
Domain2	0.52	0.40	0.01	0.63	1	0.64
hNlg3cyt	0.59	0.27	0.04	0.75	0.64	1

OmpX is the denatured outer membrane protein X from *E. coli*, Domain 2 protein is the unstructured domain 2 of the protein 5A of Hepatitis C virus, and hNlg3cyt is the intrinsically disordered intracellular domain of human NL3 protein

dipeptides are also fairly large, indicating conformational similarities. In contrast, the protein $P(\delta_{HN}, \delta_{H\alpha})$ distributions do not significantly overlap with those of amino-acids in α-helical and β-sheet structures. In addition, the S value between the $P(\delta_{HN}, \delta_{H\alpha})$ distribution of the dipeptides and that of *native* lysozyme is found to be as small as 0.20, which is in stark contrast with the S value of 0.60 between that of the dipeptides and that of the denatured lysozyme.

Instead of the above overlap calculation analyses, one can extract quantitative information about the degree of conformational disorder or unfolding by examining the mean values and SD's of the measured δ_{HN} and $\delta_{H\alpha}$ values. In Table 3, the corresponding mean values, denoted as $\bar{\delta}_{H\alpha}$ and $\bar{\delta}_{HN}$, and the SD's are summarized. Furthermore, those of the four proteins that are *predicted* by using our dipeptide library are given in the same table. Interestingly, the mean values, $\bar{\delta}_{H\alpha}$ and $\bar{\delta}_{HN}$, of the native lysozyme as well as the native OmpX are quantitatively similar to those of denatured and intrinsically disordered proteins. However, the SD's of the native lysozyme and OmpX are significantly larger than those of denatured and intrinsically

Table 3 Mean values and standard deviations of the H^{α} (upper half) and H^{N} (lower half) chemical shifts (in ppm)

	Proteins		Dipeptides	
	Mean value	SD	Mean value	SD
Lysozyme (N)[a]	4.27	0.459	–	–
Lysozyme (U)[b]	4.36	0.159	4.34	0.120
OmpX	4.30	0.174	4.31	0.165
Domain2	4.26	0.146	4.33	0.121
hNlg3cyt	4.35	0.154	4.33	0.129
Dipeptides (present)	–	–	4.37	0.128
Lysozyme (N)	8.17	0.513	–	–
Lysozyme (U)	8.20	0.167	8.30	0.135
OmpX (N)[c]	8.44	0.485	–	–
OmpX (U)	8.29	0.132	8.29	0.119
Domain2	8.11	0.188	8.31	0.106
hNlg3cyt	8.24	0.141	8.33	0.093
Dipeptides (present)	–	–	8.31	0.134

[a] Native lysozyme (Redfield and Dobson 1988)

[b] Denatured lysozyme (Schlorb et al. 2005)

[c] Native OmpX (Fernandez et al. 2001)

disordered proteins. More specifically, the corresponding SD's of denatured lysozyme, unfolded OmpX, Domain 2, and hNlg3cyt are significantly smaller than 0.2 ppm, whereas those of the native lysozyme and OmpX are larger than 0.45 ppm. Furthermore, the mean values and SD's of the δ_{HN} and $\delta_{H\alpha}$ distributions of unfolded (denatured and intrinsically disordered) proteins, which were calculated (not measured) by using the chemical shift values obtained from our dipeptides, are in excellent agreement with those of the four proteins considered here.

The present statistical analysis results therefore led us to reach a conclusion that the peptide backbone conformational distributions and local environments of these four denatured or unstructured proteins are quite similar to one another as well as to those of the blocked dipeptides. For the dipeptides, we have shown that the predominant conformations of dipeptides in aqueous solutions are enthalpically stable PP_{II} and entropically favored β-strand. At room temperature, the population of PP_{II} conformer is on average larger than that of β-strand structure, though certain turn structures might be present in some cases. Thus, the local peptide backbone conformations of denatured proteins could have significant PP_{II} content with a truly unordered form in which the polypeptide chain samples all of accessible Ramachandran space. In fact, the latter form including turns and bends along a polypeptide chain has been shown to play a role in making the unfolded or denatured proteins relatively compact in shape.

In summary, the present comparative investigation shows that the chemical shifts δ_{HN} and $\delta_{H\alpha}$ that are essentially determined by backbone structure and local electronic and solvation effects become fairly uniform for most of the residues in denatured and unstructured proteins and that their distributions are similar to those of dipeptides studied here. This indicates that, upon denaturation of a given protein, (1) their backbone peptide conformations become comparatively homogeneous with a significant population of PP_{II} conformation and (2) the ensemble-average solvation environment around the peptides in denatured and even natively unfolded proteins is similar to that around blocked dipeptides. Thus, we anticipate that detailed structural analyses of short peptides by using a variety of spectroscopic methods will still provide invaluable information on conformational preferences of amino-acids even in unfolded and intrinsically disordered proteins.

Acknowledgments This work was supported by the National Research Foundation of Korea (NRF) grants funded by the Korea government (MEST) (No. 20090078897 and 20110020033) to MC. Also, we thank the financial supports from the Korea Basic Science Institute (T31401) grant to GH and MC. KIO thanks the financial support by Korea University.

References

Asakura T, Taoka K, Demura M, Williamson MP (1995) The relationship between amide proton chemical-shifts and secondary structure in proteins. J Biomol NMR 6(3):227–236

Avbelj F, Kocjan D, Baldwin RL (2004) Protein chemical shifts arising from alpha-helices and beta-sheets depend on solvent exposure. Proc Natl Acad Sci USA 101(50):17394–17397. doi:10.1073/pnas.0407969101

Avbelj F, Grdadolnik SG, Grdadolnik J, Baldwin RL (2006) Intrinsic backbone preferences are fully present in blocked amino acids. Proc Natl Acad Sci USA 103(5):1272–1277. doi:10.1073/pnas.0510420103

Bai YW, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen-exchange. Proteins Struct Funct Genet 17(1):75–86

Brant DA, Flory PJ (1965) Configuration of random polypeptide chains. 2. Theory. J Am Chem Soc 87(13):2791–2800

Bundi A, Wuthrich K (1979) H-1-NMR parameters of the common amino-acid residues measured in aqueous-solutions of the linear tetrapeptides H-Gly-Gly-X-L-Ala-Oh. Biopolymers 18(2):285–297

Cho M (2008) Coherent two-dimensional optical spectroscopy. Chem Rev 108(4):1331–1418

Cho M, Oh KI, Lee KK, Park EK, Yoo DG, Hwang GS (2010) Circular dichroism eigenspectra of polyproline II and beta-strand conformers of trialanine in water: singular value decomposition analysis. Chirality 22(1E):E186–E201. doi:10.1002/chir.20870

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13(3):289–302

Dukor RK, Keiderling TA (1991) Reassessment of the random coil conformation—vibrational CD study of proline oligopeptides and related polypeptides. Biopolymers 31(14):1747–1761

Eker F, Cao XL, Nafie L, Schweitzer-Stenner R (2002) Tripeptides adopt stable structures in water. A combined polarized visible

Raman, FTIR, and VCD spectroscopy study. J Am Chem Soc 124(48):14330–14341. doi:10.1021/Ja027381w

Fernandez C, Adeishvili K, Wuthrich K (2001) Transverse relaxation-optimized NMR spectroscopy with the outer membrane protein OmpX in dihexanoyl phosphatidylcholine micelles. Proc Natl Acad Sci USA 98(5):2358–2363

Graf J, Nguyen PH, Stock G, Schwalbe H (2007) Structure and dynamics of the homologous series of alanine peptides: a joint molecular dynamics/NMR study. J Am Chem Soc 129(5): 1179–1189. doi:10.1021/Ja0660406

Grdadolnik J, Mohacek-Grosev V, Baldwin RL, Avbelj F (2011) Populations of the three major backbone conformations in 19 amino acid dipeptides. Proc Natl Acad Sci USA 108(5): 1794–1798. doi:10.1073/pnas.1017317108

Hagarman A, Measey TJ, Mathieu D, Schwalbe H, Schweitzer-Stenner R (2010) Intrinsic propensities of amino acid residues in GxG peptides inferred from amide I′ band profiles and NMR scalar coupling constants. J Am Chem Soc 132:540–551

Hagarman A, Mathieu D, Toal S, Measey TJ, Schwalbe H, Schweitzer-Stenner R (2011) Amino acids with hydrogen-bonding side chains have an intrinsic tendency to sample various turn conformations in aqueous solution. Chem Eur J 17(24): 6789–6797. doi:10.1002/chem.201100016

Hahn S, Lee H, Cho M (2004) Theoretical calculations of infrared absorption, vibrational circular dichroism, and two-dimensional vibrational spectra of acetylproline in liquids water and chloroform. J Chem Phys 121(4):1849–1865. doi:10.1063/1.1763889

Han B, Liu YF, Ginzinger SW, Wishart DS (2011) SHIFTX2: significantly improved protein chemical shift prediction. J Biomol NMR 50(1):43–57. doi:10.1007/s10858-011-9478-4

Huyghues-Despointes BMP, Scholtz JM, Pace CN (1999) Protein conformational stabilities can be determined from hydrogen exchange rates. Nat Struct Biol 6(10):910–912

Lacroix E, Viguera AR, Serrano L (1998) Elucidating the folding problem of alpha-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. J Mol Biol 284(1):173–191

Lee KK, Hahn S, Oh K-I, Choi JS, Joo C, Lee H, Han HY, Cho M (2006) Structure of N-acetylproline amide in liquid water: experimentally measured and numerically simulated infrared and vibrational circular dichroism spectra. J Phys Chem B 110(38): 18834–18843. doi:10.1021/Jp055846+

Lee KK, Joo C, Yang S, Han H, Cho M (2007) Phosphorylation effect on the GSSS peptide conformation in water: infrared, vibrational circular dichroism, and circular dichroism experiments and comparisons with molecular dynamics simulations. J Chem Phys 126(23):235102-1–235102-14. doi:10.1063/1.2738472

Liang Y, Ye H, Kang CB, Yoon HS (2007) Domain 2 of nonstructural protein 5A (NS5A) of hepatitis C virus is natively unfolded. Biochemistry 46(41):11550–11558. doi:10.1021/Bi700776e

Makowska J, Rodziewicz-Motowidlo S, Baginska K, Vila JA, Liwo A, Chmurzynski L, Scheraga HA (2006) Polyproline II conformation is one of many local conformational states and is not an overall conformation of unfolded peptides and proteins. Proc Natl Acad Sci USA 103(6):1744–1749. doi:10.1073/pnas.0510549103

Makowska J, Rodziewicz-Motowidlo S, Baginska K, Makowski M, Vila JA, Liwo A, Chmurzynski L, Scheraga HA (2007) Further evidence for the absence of polyproline II stretch in the XAO peptide. Biophys J 92(8):2904–2917. doi:10.1529/biophysj.106.097550

Marsh JA, Singh VK, Jia ZC, Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation. Protein Sci 15(12):2795–2804. doi:10.1110/Ps.062465306

Merutka G, Dyson HJ, Wright PE (1995) Random coil H-1 chemical-shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. J Biomol NMR 5(1):14–24

Molday RS, Kallen RG (1972) Substituent effects on amide hydrogen-exchange rates in aqueous-solution. J Am Chem Soc 94(19):6739–6745

Oh K-I, Han J, Lee KK, Hahn S, Han H, Cho M (2006) Site-specific hydrogen-bonding interaction between N-acetylproline amide and protic solvent molecules: comparisons of IR and VCD measurements with MD simulations. J Phys Chem A 110(50):13355–13365. doi:10.1021/jp065475c

Oh KI, Lee KK, Park EK, Jung Y, Hwang GS, Cho M (2012) A comprehensive library of blocked dipeptides reveals intrinsic backbone conformational propensities of unfolded proteins. Proteins Struct Funct Genet. doi:10.1002/prot.24000

Oldfield E (1995) Chemical-shifts and 3-dimensional protein structures. J Biomol NMR 5(3):217–225

Oldfield E (2002) Chemical shifts in amino acids, peptides, and proteins: from quantum chemistry to drug design. Annu Rev Phys Chem 53:349–378. doi:10.1146/annurev.physchem.53.082201.124235

Osapay K, Case DA (1994) Analysis of proton chemical-shifts in regular secondary structure of proteins. J Biomol NMR 4(2):215–230

Pappu RV, Rose GD (2002) A simple model for polyproline II structure in unfolded states of alanine-based peptides. Protein Sci 11(10):2437–2455. doi:10.1110/ps.0217402

Pastore A, Saudek V (1990) The relationship between chemical-shift and secondary structure in proteins. J Magn Reson 90(1): 165–176

Plaxco KW, Morton CJ, Grimshaw SB, Jones JA, Pitkeathly M, Campbell ID, Dobson CM (1997) The effects of guanidine hydrochloride on the 'random coil' conformations and NMR chemical shifts of the peptide series GGXGG. J Biomol NMR 10(3):221–230

Redfield C, Dobson CM (1988) Sequential H-1-NMR assignments and secondary structure of hen egg-white lysozyme in solution. Biochemistry 27(1):122–136

Rizo J, Blanco FJ, Kobe B, Bruch MD, Gierasch LM (1993) Conformational behavior of *Escherichia coli* OmpA signal peptides in membrane mimetic environments. Biochemistry 32(18):4881–4894

Schlorb C, Ackermann K, Richter C, Wirmer J, Schwalbe H (2005) Heterologous expression of hen egg white lysozyme and resonance assignment of tryptophan side chains in its non-native states. J Biomol NMR 33(2):95–104. doi:10.1007/s10858-005-2063-y

Schweitzer-Stenner R, Measey TJ (2007) The alanine-rich XAO peptide adopts a heterogeneous population, including turn-like and polyproline II conformations. Proc Natl Acad Sci USA 104(16):6649–6654. doi:10.1073/pnas.0700006104

Schweitzer-Stenner R, Hagarman A, Measey TJ, Mathieu D, Schwalbe H (2010) Intrinsic propensities of amino acid residues in GxG peptides inferred from amide I′ band profiles and NMR scalar coupling constants. J Am Chem Soc 132(2):540–551. doi:10.1021/ja9058052

Serrano L (1995) Comparison between the phi distribution of the amino-acids in the protein database and NMR data indicates that amino-acids have various phi propensities in the random coil conformation. J Mol Biol 254(2):322–333

Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 44(4):213–223. doi:10.1007/s10858-009-9333-z

Shi ZS, Olson CA, Rose GD, Baldwin RL, Kallenbach NR (2002a) Polyproline II structure in a sequence of seven alanine residues.

Proc Natl Acad Sci USA 99(14):9190–9195. doi:10.1073/pnas.112193999

Shi ZS, Woody RW, Kallenbach NR (2002b) Is polyproline II a major backbone conformation in unfolded proteins? Adv Protein Chem 62:163–240

Shi ZS, Chen K, Liu ZG, Ng A, Bracken WC, Kallenbach NR (2005) Polyproline II propensities from GGXGG peptides reveal an anticorrelation with beta-sheet scales. Proc Natl Acad Sci USA 102(50):17964–17968. doi:10.1073/pnas.0507124102

Shi ZS, Chen K, Liu ZG, Kallenbach NR (2006) Conformation of the backbone in unfolded proteins. Chem Rev 106(5):1877–1897. doi:10.1021/Cr040433a

Shin HC, Merutka G, Waltho JP, Wright PE, Dyson HJ (1993) Peptide models of protein-folding initiation sites. 2. The G-H turn region of myoglobin acts as a helix stop signal. Biochemistry 32(25):6348–6355

Shortle D, Ackerman MS (2001) Persistence of native-like topology in a denatured protein in 8 M urea. Science 293(5529):487–489

Sreerama N, Woody RW (1999) Molecular dynamics simulations of polypeptide conformations in water: a comparison of alpha, beta, and poly(pro)II conformations. Proteins Struct Funct Genet 36(4):400–406

Swindells MB, Macarthur MW, Thornton JM (1995) Intrinsic phi, psi propensities of amino-acids, derived from the coil regions of known structures. Nat Struct Biol 2(7):596–603

Szilagyi L (1995) Chemical-shifts in proteins come of age. Prog Nucl Magn Reson Spectrosc 27:325–443

Tafer H, Hiller S, Hilty C, Fernandez C, Wuthrich K (2004) Nonrandom structure in the urea-unfolded *Escherichia coli* outer membrane protein X (OmpX). Biochemistry 43(4):860–869. doi:10.1021/Bi0356606

Tiffany ML, Krimm S (1968) Circular dichroism of poly-L-proline in an unordered conformation. Biopolymers 6(12):1767–1770

Top D, Read JA, Dawe SJ, Syvitski RT, Duncan R (2011) Cell–cell membrane fusion induced by the p15 fusion-associated small transmembrane (FAST) protein requires a novel fusion peptide motif containing a myristoylated polyproline type II helix. J Biol Chem. doi:10.1074/jbc.M111.305268

Williamson MP, Kikuchi J, Asakura T (1995) Application of H-1-NMR chemical-shifts to measure the quality of protein structures. J Mol Biol 247(4):541–546

Wishart DS, Nip AM (1998) Protein chemical shift analysis: a practical guide. Biochem Cell Biol-Biochimie Et Biologie Cellulaire 76(2–3):153–163

Wishart DS, Sykes BD (1994a) The C-13 chemical-shift index—a simple method for the identification of protein secondary structure using C-13 chemical-shift data. J Biomol NMR 4(2):171–180

Wishart DS, Sykes BD (1994b) Chemical-shifts as a tool for structure determination. Methods Enzymol 239:363–392

Wishart DS, Sykes BD, Richards FM (1991a) Relationship between nuclear-magnetic-resonance chemical-shift and protein secondary structure. J Mol Biol 222(2):311–333

Wishart DS, Sykes BD, Richards FM (1991b) Simple techniques for the quantification of protein secondary structure by H-1-NMR spectroscopy. FEBS Lett 293(1–2):72–80

Wishart DS, Sykes BD, Richards FM (1992) The chemical-shift index—a fast and simple method for the assignment of protein secondary structure through NMR-spectroscopy. Biochemistry 31(6):1647–1651

Wood K, Paz A, Dijkstra K, Scheek RM, Otten R, Silman I, Sussman JL, Mulder FA (2011) Backbone and side chain NMR assignments for the intrinsically disordered cytoplasmic domain of human neuroligin-3. Biomol NMR Assign. doi:10.1007/s12104-011-9315-4

Woolfson DN, Bartlett GJ, Choudhary A, Raines RT (2010) n → pi* interactions in proteins. Nat Chem Biol 6(8):615–620. doi:10.1038/NCHEMBIO.406

Wright PE, Dyson HJ, Lerner RA (1988) Conformation of peptide-fragments of proteins in aqueous-solution—implications for initiation of protein folding. Biochemistry 27(19):7167–7175

Zagrovic B, Lipfert J, Sorin EJ, Millettt IS, van Gunsteren WF, Doniach S, Pande VS (2005) Unusual compactness of a polyproline type II structure. Proc Natl Acad Sci USA 102(33):11698–11703. doi:10.1073/pnas.0409693102

Zanni MT, Stenger J, Asplund MC, Hochstrasser RM (2001) Solvent dependent conformational dynamics of dipeptides studied with two-dimensional infrared spectroscopy. Biophys J 80(1):8A–9A

Zhang HY, Neal S, Wishart DS (2003) RefDB: a database of uniformly referenced protein chemical shifts. J Biomol NMR 25(3):173–195