Probing Local Backbone Geometries in Intrinsically Disordered Proteins by Cross-Correlated NMR Relaxation**
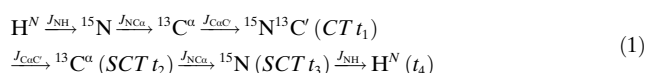
Jan Stanek, Saurabh Saxena, Leonhard Geist, Robert Konrat, and Wiktor Koźmiński**

Proteins that lack stable tertiary structures are attracting growing interest because of their involvement in fundamental biological processes.[1] The traditional conceptual view correlating structures with functions does not seem to grasp the essential properties of these protein families.[2] A fundamental problem in the structural characterization of intrinsically disordered proteins (IDPs) is the definition of the conformational ensemble sampled by the polypeptide chain in solution. There is growing evidence that the popular dichotomic partitioning into ordered and disordered proteins does not adequately grasp the dynamic conformational ensembles of biological polypeptides and their mutual similarities.[3] The recently developed *meta*-structure approach provided first insight into the heterogeneity of folding funnels of disordered proteins suggesting that IDPs populate a diverse conformational space of thermally accessible (sub-)states.[3] Often the interpretation relies on the concept of residual structure. The observation of a (residue-specific) conformational preference and deviation from an idealized random coil devoid of any structural propensity is interpreted as a prevalence of residual structures. Structural preferences predominantly occur on a secondary structure level. Hydrophobic collapse of preformed local structure elements leads to transiently formed conformations with varying degrees of compaction (for example, molten globules, compact conformations with distinct packing of side chains). NMR spectroscopy has not only been developed into a powerful structural biology technique complementing protein X-ray crystallography, but additionally, it offers unique opportunities for structural and dynamical studies of intrinsically disordered/unfolded proteins. Residue-specific information (with atomic resolution) about hydrogen exchange rates,[4] chemical shifts,[5] heteronuclear relaxation rates,[6] and scalar and dipolar couplings[7] have provided unprecedented insight into the conformational properties of unfolded proteins in solution. Most importantly, unique long-range structural information is obtained using paramagnetic relaxation enhancement (PRE) data.[8]

Cross-correlated NMR relaxation (CCR) has attracted considerable interest in the past as it offers unique possibilities to probe structural dynamics in proteins.[9] The method is based on correlated fluctuations of relaxation relevant interaction tensors (for example, dipole–dipole interactions, chemical shift anisotropy (CSA), and quadrupolar interactions). To date, a series of NMR pulse sequences exist that quantitatively measures dipole–dipole (DD), dipolar–CSA (D-CSA), and CSA–CSA cross-correlations. Using these methods, valuable information about structurally relevant dihedral angles of the protein backbone have been obtained. For example, intraresidue ^{1}H(i)–^{15}N(i)–^{13}C′(i) dipolar–CSA interference was proposed to determine the orientation of the ^{1}H(i)–^{15}N(i) dipole vector in the principal frame of the intraresidue ^{13}C′(i) CSA tensor and discriminate between type I and type II β turns in proteins.[10] As IDPs populate Ramachandran space in a rather unique way and substantially sample β-turn (I,II) and polyproline II helical conformations, novel experimental approaches are highly desirable that allow assess to these (non-α-helical, non-β-strand) conformations in IDPs. Applications of multidimensional (triple-resonance) NMR spectroscopy to IDPs are challenging because of the limited signal dispersion owing to the lack of stably formed tertiary structures. These experimental limitations have been largely overcome by the introduction of higher-dimensionality experiments combined with non-uniform sampling. Herein, we extend the cross-correlated relaxation method to highly crowded NMR spectra typically encountered for IDPs. The novel experiment allows for the quantification of intraresidue ^{1}H(i)–^{15}N(i)–^{13}C′(i) dipolar–CSA interference terms in IDPs exploiting the outstanding spectral resolution of higher-dimensionality NMR experiments.

An outline of the experiment is given in Equation (1). A timing diagram and experimental details of the pulse sequence can be found in the Supporting Information. Briefly, the coherence transfer pathway is as follows:

$$H^N \xrightarrow{J_{NH}} {}^{15}N \xrightarrow{J_{NC\alpha}} {}^{13}C^\alpha \xrightarrow{J_{C\alpha C'}} {}^{15}N^{13}C' \, (CT \, t_1)$$
$$\xrightarrow{J_{C\alpha C'}} {}^{13}C^\alpha \, (SCT \, t_2) \xrightarrow{J_{NC\alpha}} {}^{15}N \, (SCT \, t_3) \xrightarrow{J_{NH}} H^N \, (t_4) \tag{1}$$

where ^{15}N^{13}C′ denotes zero- and double-quantum ^{15}N–^{13}C′ coherences and CT and SCT are constant and semi-constant

[*] J. Stanek, S. Saxena, Prof. W. Koźmiński
Faculty of Chemistry, University of Warsaw
Pasteura 1, 02093 Warsaw (Poland)
E-mail: kozmin@chem.uw.edu.pl

L. Geist, Prof. R. Konrat
Department of Structural and Computational Biology
Max F. Perutz Laboratories, University of Vienna
Campus Vienna Biocenter 5, 1030 Vienna (Austria)
E-mail: robert.konrat@univie.ac.at

time evolution periods, respectively. In terms of the magnetization transfers, the pulse sequence follows the conventional 3D HN(CA)CO experiment[11] and its four-dimensional extension.[12] In the CT period, zero- and double-quantum ^{15}N–$^{13}C'$ coherences are evolved as suggested previously;[10] however, only in-phase J_{NH}-separated doublets are recorded. It is noteworthy that both $^{13}C^\alpha$ and ^{15}N spins are evolved in the semi-constant time manner for sensitivity enhancement. Interresidual peaks which provide less structural information are usually of smaller intensity due to weaker two-bond coupling $^2J_{C\alpha N} \approx 4$–7 Hz.

As a first application of the new pulse sequence, we provide CCR data for BASP1 (270 a.a. chicken brain acid soluble protein 1) at different pH values. Transformation of chicken embryonic fibroblasts by the protein product (c-Myc) of the protooncogene *c-myc* leads to the suppression of BASP1. Strikingly, ectopic expression of BASP1 renders fibroblasts resistant to subsequent cell transformation by c-Myc.[13] It has been described to be involved in neurite outgrowth and plasma membrane organization, and non-myristoylated BASP1 has been discovered as a co-suppressor of WT1 function (Wilms' tumor suppressor protein 1).[13] NMR data obtained for human BASP1 indicate that BASP1 is lacking a well-defined tertiary structure in solution and displays a biophysical behavior reminiscent of an IDP.[14]

Despite extensive spectral crowding in 2D ^{15}N-HSQC spectrum, presence of both intra- and interresidual peaks, and J_{NH} splitting of each resonance, it was possible to determine the CCR rate for about 82% of BASP1 residues with amide protons in slow exchange regime at both pH 2 and 6. Clearly, the superior resolution of the four-dimensional spectrum allows data to be extracted without the need to record and combine both in-phase and anti-phase J_{NH}-modulated data sets.[10] The closer inspection of the spectra revealed that most of spectral crowding results from overlap of intra- and interresidual resonances, that would not be eliminated by the separation of upfield and downfield doublet lines. Interestingly, error analysis of $\Gamma_{H(i)N(i),C'(i)}$ (given by Eq. (1) in Ref. [10]) leads to the conclusion that optimal precision of cross-correlated relaxation rate is obtained for CT duration comparable to mean T_2 of zero- and double-quantum ^{15}N–$^{13}C'$ coherences. Owing to the favorable relaxation properties of intrinsically disordered proteins, CT lengths can be as large as the $T_C = 90$ ms used herein. Clearly, the full utilization of this extraordinarily long constant-time period can only be accomplished using non-uniform sampling. It should be emphasized that the data sets were collected using as few as only 0.24% of sampling points required conventionally. Very-high-resolution spectra were obtained by employing the signal separation algorithm (SSA) that is known to faithfully reproduce relative signal intensities.[15] Recently, consistent findings were reported by Coggins and co-workers for a very similar algorithm called SCRUB.[16] Application of SSA has led to virtually complete removal of sampling artifacts and recovery of natural sensitivity of the experiments. In comparison to zero-augmented Fourier transformation, the average signal-to-noise ratio measured for all relevant intraresidue resonances increased by a factor of 2.66 and 1.31 for BASP1 at pH 6 and pH 2, respectively. Different gain in S/N reflects

intrinsic sensitivity associated with various sample concentrations.

As shown with representative C'–C^α planes for Thr108 on the Figure 1, inclusion of an extra $^{13}C^\alpha$ evolution period to the original 3D experiment greatly enhances the spectral resolution even if $^{13}C^\alpha$ chemical shifts in IDPs tend to cluster in the regions specific for a given amino acid type. Seemingly, this is due to the fact that $^{13}C^\alpha$ is barely *correlated* to $^{13}C'$, ^{15}N, and $^1H^N$ chemical shifts.

Of particular relevance is the fact that intraresidue $^1H(i)$–$^{15}N(i)$–$^{13}C'(i)$ dipolar–CSA cross-correlation rates have sensitive dependencies on local β-turn geometries.[10] While slightly positive values are found in the case of type I β turns, alternating signs of cross-correlation rates are observed for the central residues in a type II β turn (structural information available from this experiment is further discussed in the Supporting Information).[10] For BASP1 at pH 6, predominantly positive cross-correlation rates were found. In contrast, at low pH (pH 2) significant changes and alternating signs of the cross-correlation rates were found (Figure 2). As can be seen, the changes are distributed along the entire backbone. A closer inspection revealed significant variation of CCR rate for residues 70–80. In the surrounding region acidic amino acids (Glu, Asp) are especially abundant and showed most pronounced chemical shift changes upon lowering the pH (Supporting Information, Figure S2). We thus conclude that the observed cross-correlation rate changes
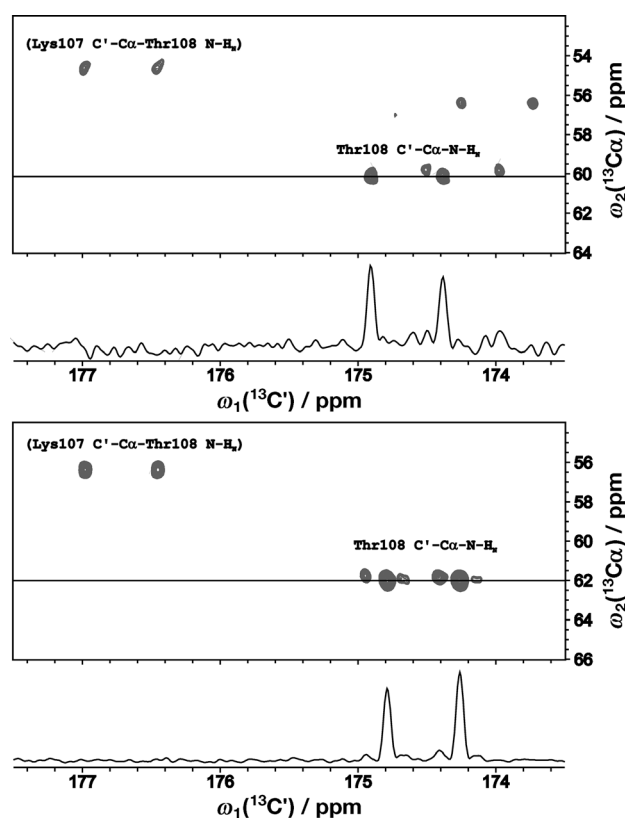

Figure 1. F_1 (C')–F_2 (C^α) planes from the 4D HNCACO-CCR(C'/NH$_N$) spectrum of cBASP1 at pH 2 (top) and pH 6 (bottom) showing J_{NH}-resolved doublets for threonine 108 residue. Labels for interresidual cross-peaks (irrelevant here) are given in parentheses. Noteworthy is the significant difference in relative intensity of the doublet lines.
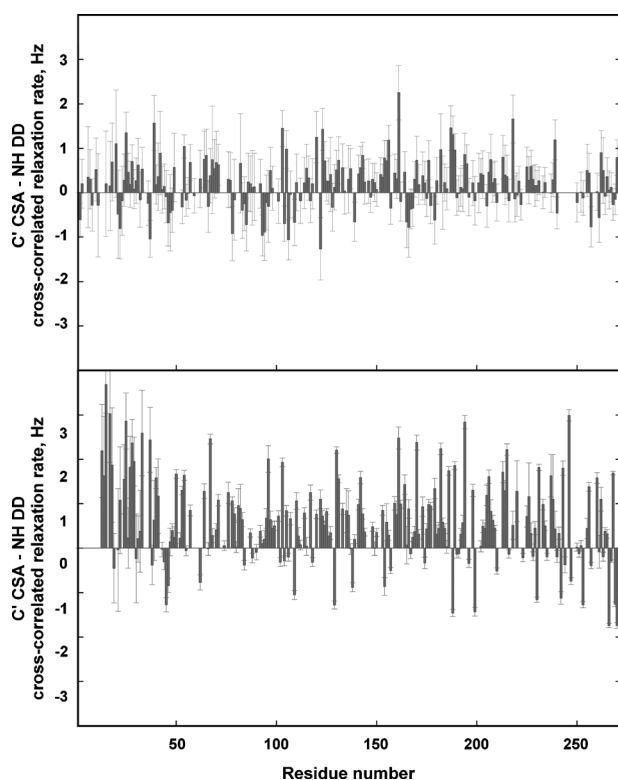
Figure 2. C′ CSA–NH DD cross-correlated relaxation rates for individual residues of BASP1 at pH 2 (top) and pH 6 (bottom).

indeed reflect local structural changes induced by protonation of side-chain carboxylic groups. Additionally, changes in the average overall structural ensemble (for example, spatial extension of the conformational ensemble) was probed by DOSY-based measurements of the hydrodynamic radius. Normalized to dioxan, the hydrodynamic radius dropped from 42 Å at pH 6 to 27 Å at pH 2. Furthermore, this compaction was independently verified by SOFAST-HSQC experiments under saturation of aliphatic side-chain protons. Details of this overall structural change is beyond the scope of this current work and will be published in due course elsewhere. Suffice to say here that the pH-induced changes in β-turn geometries can be efficiently detected by the novel NMR approach, suggesting that: 1) subtleties of local backbone structures in IDPs can be assessed using this technique; and that 2) in IDPs a subtle interplay exists between local β-turn geometries and the average extension of the conformational space accessible under changing environmental conditions.

In summary, we have presented a novel NMR pulse sequence for probing local backbone structure in IDPs. Employing high dimensionality and non-uniform sampling, the novel method allows for the quantification of intraresidue $^1H(i)$–$^{15}N(i)$–$^{13}C'(i)$ dipolar–CSA interference in highly over-lapped NMR spectra of IDPs. Data obtained on the IDP BASP1 at different pH values illustrated the applicability of the approach to elucidate subtle changes in local backbone geometry under different environmental conditions. Specifically, it was found that in the case of BASP1, lowering the pH induces a population shift of β-turn geometries; at neutral pH,

BASP1 predominantly samples type I β turns and polyproline II helices, lowering the pH leads to an increased population of type II β turns accompanied by an unexpected overall compaction of the BASP1 structural ensemble. Given the outstanding sensitivity and spectral resolution of the NMR approach, we envisage useful applications for the characterization of the structural dynamics of IDPs in solution. The detailed data obtained in this study has also provided compelling evidence that intrinsically disordered proteins are characterized by extremely heterogenous energy landscapes, allowing for subtle conformational transitions under changing environmental conditions. Thus, to fully grasp the richness of the structural dynamics of these proteins and their potential role in biological processes, more sophisticated experimental methods as well as theoretical concepts surmounting the conventional order–disorder partitioning are required.

[1] H. J. Dyson, P. E. Wright, *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208.
[2] P. Tompa, *FEBS Lett.* **2005**, *579*, 3346–3354.
[3] a) R. Konrat, *Cell. Mol. Life Sci.* **2009**, *66*, 3625–3639; b) R. Konrat, *Structure* **2010**, *18*, 416–419.
[4] P. A. Jennings, P. E. Wright, *Science* **1993**, *262*, 892–896.
[5] N. A. Farrow, O. W. Zhang, J. D. Forman-Kay, L. E. Kay, *Biochemistry* **1995**, *34*, 868–878.
[6] L. J. Smith, K. A. Bolin, H. Schwalbe, M. W. MacArthur, J. M. Thornton, C. M. Dobson, *J. Mol. Biol.* **1996**, *255*, 494–506.
[7] a) P. Bernado, L. Blanchard, P. Timmins, D. Marion, R. W. H. Ruigrok, M. Blackledge, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17002–17007; b) T. Mittag, J. D. Forman-Kay, *Curr. Opin. Struct. Biol.* **2007**, *17*, 3–14.
[8] J. R. Gillespie, D. Shortle, *J. Mol. Biol.* **1997**, *268*, 170–184.
[9] a) B. Reif, M. Hennig, C. Griesinger, *Science* **1997**, *276*, 1230–1233; b) D. W. Yang, R. Konrat, L. E. Kay, *J. Am. Chem. Soc.* **1997**, *119*, 11938–11940; c) K. Kloiber, W. Schüler, R. Konrat, *J. Biomol. NMR* **2002**, *22*, 349–363.
[10] K. Kloiber, R. Konrat, *J. Am. Chem. Soc.* **2000**, *122*, 12033–12034.
[11] R. T. Clubb, V. Thanabal, G. Wagner, *J. Magn. Reson.* **1992**, *97*, 213–217.
[12] A. Zawadzka-Kazimierczuk, K. Kazimierczuk, W. Koźmiński, *J. Magn. Reson.* **2010**, *202*, 109–116.
[13] a) M. Hartl, A. Nist, M. I. Khan, T. Valovka, K. Bister, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5604–5609; b) R. M. Epand, P. Vuong, C. M. Yip, S. Maekawa, R. F. Epand, *Biochem. J.* **2004**, *379*, 527–532; c) B. Carpenter, K. J. Hill, M. Charalambous, K. J. Wagner, D. Lahiri, D. I. James, J. S. Andersen, V. Schumacher, B. Royer-Pokora, M. Mann, A. Ward, S. G. E. Roberts, *Mol. Cell. Biol.* **2004**, *24*, 537–549.
[14] L. Geist, A. Zawadzka-Kazimierczuk, S. Saxena, S. Żerko, W. Koźmiński, R. Konrat, *Biomol. NMR Assigm.* **2013**, DOI: 10.1007s12104-012-9436-4.
[15] J. Stanek, W. Koźmiński, *J. Biomol. NMR* **2010**, *47*, 65–77.
[16] B. E. Coggins, J. W. Werner-Allen, A. Yan, P. Zhou, *J. Am. Chem. Soc.* **2012**, *134*, 18619–18630.