J Biomol NMR (2014) 60:209–218
DOI 10.1007/s10858-014-9867-6

ARTICLE

"CON-CON" assignment strategy for highly flexible intrinsically disordered proteins

Alessandro Piai · Tomáš Hošek · Leonardo Gonnelli · Anna Zawadzka-Kazimierczuk · Wiktor Koźmiński · Bernhard Brutscher · Wolfgang Bermel · Roberta Pierattelli · Isabella C. Felli

Received: 27 August 2014 / Accepted: 10 October 2014 / Published online: 19 October 2014

Abstract Intrinsically disordered proteins (IDPs) are a class of highly flexible proteins whose characterization by NMR spectroscopy is complicated by severe spectral overlaps. The development of experiments designed to facilitate the sequence-specific assignment procedure is thus very important to improve the tools for the characterization of IDPs and thus to be able to focus on IDPs of increasing size and complexity. Here, we present and describe the implementation of a set of novel ^{1}H-detected 5D experiments, (HACA)CON(CACO)NCO(CA)HA, BT-(H)NCO(CAN)CONNH and BT-HN(COCAN)CONNH, optimized for the study of highly flexible IDPs that exploit the best resolved correlations, those involving the carbonyl and nitrogen nuclei of neighboring amino acids, to achieve sequence-specific resonance assignment. Together with the analogous recently proposed pulse schemes based on ^{13}C detection, they form a complete set of experiments for sequence-specific assignment of highly flexible IDPs. Depending on the particular sample conditions (concentration, lifetime, pH, temperature, etc.), these experiments present certain advantages and disadvantages that will be discussed. Needless to say, that the availability of a variety of complementary experiments will be important for accurate determination of resonance frequencies in complex IDPs.

Keywords Intrinsically disordered proteins · ^{13}C detection · $^{1}H^{N}$ detection · $^{1}H^{\alpha}$ detection · NUS · Multidimensional NMR experiment · BEST-TROSY · Backbone assignment

Introduction

The role of intrinsic disorder is essential in many cellular mechanisms (Tompa 2009, 2012; Uversky 2013a, b), which range from gene regulation to signaling processes. Intrinsically disordered proteins (IDPs) or intrinsically disordered protein regions (IDPRs) play key roles in many of these cellular events (Uversky and Dunker 2013), thanks to their local flexibility and conformational freedom, which allow them to modulate such processes. These evidences are pushing the scientific community towards a reconsideration of the structure/function paradigm, taking into account that many disordered proteins are able to carry out specific functions despite the absence of secondary and tertiary structures (Wright and Dyson 1999; Uversky et al. 2000; Dunker et al. 2001).

With X-ray crystallography being incapable to describe conformational disorder, nuclear magnetic resonance (NMR) spectroscopy becomes the most qualified tool of

Electronic supplementary material The online version of this article (doi:10.1007/s10858-014-9867-6) contains supplementary material, which is available to authorized users.

A. Piai · T. Hošek · L. Gonnelli · R. Pierattelli (✉) · I. C. Felli (✉)
CERM and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy
e-mail: felli@cerm.unifi.it

A. Zawadzka-Kazimierczuk · W. Koźmiński
Faculty of Chemistry, Biological and Chemical Research Centre, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw, Poland

B. Brutscher
Institut de Biologie Structurale, Université Grenoble 1, CNRS, CEA, Rue Jules Horowitz 41, 38027 Grenoble Cedex 1, France

W. Bermel
Bruker BioSpin GmbH, Silberstreifen, 76287 Rheinstetten, Germany

investigation, thanks also to its unique ability to provide information at atomic resolution (Konrat 2014; Felli and Pierattelli 2014; Nováček et al. 2014). However, in respect to what is found for well-folded proteins, the peculiar properties of IDPs introduce additional challenges that need to be overcome to prevent drastic reduction in the quality of the spectra. First, the conformational freedom experienced by IDPs leads to severe collapse of chemical shifts, since their values are mostly determined by the amino acid chemical structure and by the primary sequence while contributions deriving from the three-dimensional structure are missing. Secondly, the high rates of solvent exchange processes endured by amide protons, that in absence of a stable fold are largely solvent exposed, can dramatically compromise the information content of multidimensional spectra based on $^{1}H^{N}$ detection, since at physiological temperature and pH many signals may be broadened even beyond detection. Finally, IDPs are often characterized by repetitive amino acid sequences in which prolines are usually abundant, thus complicating the resonance assignment procedure. To overcome these limitations as much as possible, several IDP-dedicated NMR experiments have been recently proposed (Panchal et al. 2001; Hiller et al. 2007; Mäntylahti et al. 2010; Motáčkova et al. 2010; Narayanan et al. 2010; Mäntylahti et al. 2011; Nováček et al. 2011; Favier and Brutscher 2011; Bermel et al. 2012; Zawadzka-Kazimierczuk et al. 2012b; Pantoja-Uceda and Santoro 2013a, b; Nováček et al. 2013; Bermel et al. 2013b; Kazimierczuk et al. 2013; Solyom et al. 2013; Pantoja-Uceda and Santoro 2014); however, NMR experiments able to deal with all these complications still need to be fully developed.

To cope with the low chemical shift dispersion of IDPs/IDPRs, high spectral resolution is mandatory. Therefore, increasing the number of dimensions of NMR experiments provides a unique tool to enhance the resolution, provided non-uniform sampling (NUS) (Kazimierczuk et al. 2010a, 2012) and, when possible, longitudinal relaxation enhancement (LRE) (Pervushin et al. 2002; Schanda and Brutscher 2005; Schanda et al. 2005; Deschamps and Campbell 2006; Felli and Brutscher 2009; Solyom et al. 2013; Gil et al. 2013) are used to achieve high resolution also in indirect dimensions without exponentially increasing the duration of NMR experiments. In fact, since the number of hypercomplex points correlates with the number of cross-peaks, by adding another dimension the required experimental time is usually doubled. The exploitation of the heteronuclei and in particular the chemical shift labeling of ^{15}N and ^{13}C frequencies provide great benefits to reduce spectral crowding (Ikura et al. 1991; Sattler et al. 1999; Dyson and Wright 2004; Mittag and Forman-Kay 2007). In our previous paper (Bermel et al. 2013b), we proposed three novel 5D ^{13}C direct-detected NMR experiments in which two pairs of ^{15}N and ^{13}C frequencies are correlated in order to provide robust information for fast and reliable backbone assignment. ^{13}C detection was chosen to exploit at best the greater chemical shift dispersion and the reduced sensitivity to exchange broadening of carbons in respect to protons (Csizmok et al. 2008; Pérez et al. 2009; Hsu et al. 2009; Knoblich et al. 2009; O'Hare et al. 2009; Skora et al. 2010; Felli et al. 2013). Here, we extend the same experimental strategy also to proton detection, combining the advantages given by heteronuclei in terms of spectral resolution with the direct acquisition of protons, which are characterized by about a 4-times larger gyromagnetic ratio resulting in higher intrinsic sensitivity. The new experiments should not be considered a replacement of the ^{13}C detected ones, but valid alternatives that can be adopted when appropriate, as discussed later. Indeed, we strongly believe that ^{1}H and ^{13}C detection provide complementary tools that allow to extend the size and complexity of IDPs/IDPRs that can be investigated by NMR. In fact, the chemical shift assignment is a mandatory step towards the structural and dynamic characterization of proteins, as well as of their interactions. The usefulness of the new experiments is demonstrated on a paradigmatic IDP, human α-synuclein (14,460 Da).

Materials and methods

NMR samples

A sample of 0.9 mM uniformly ^{13}C, ^{15}N labeled human α-synuclein in 20 mM phosphate buffer at pH 6.5 was prepared as previously described (Huang et al. 2005). EDTA and NaCl were added to reach the final concentration of 0.5 and 200 mM, respectively, and 10 % D_2O was added for the lock (Sample *A*). An identical sample, but in fully deuterated solvent (99.99 % D_2O), was also prepared (Sample *B*). The H^{N} and H^{α}-based experiments were acquired on sample *A* and *B*, respectively, both in 3 mm NMR sample tubes to reduce the detrimental effects of high salt concentration. The temperature was set to 285.5 K for Sample *A*, to reduce the rate of chemical exchange of amide protons, and to 310.0 K for Sample *B*, to work as close as possible to physiological conditions.

NMR data acquisition

All the NMR experiments were performed at 22.3 T on a Bruker Avance III spectrometer operating at 950.20 MHz ^{1}H, 238.93 MHz ^{13}C and 96.28 MHz ^{15}N frequencies, equipped with a cryogenically cooled probehead. Parameters specific to each experiment are reported in the captions of the figures describing the pulse sequences, all reported in

the Supplementary Material. E-BURP2 (or time reversed E-BURP2) (Geen and Freeman 1991) and REBURP (Geen and Freeman 1991) shapes of durations of 1,200 and 1,180 μs, respectively, were employed for ^{1}H band-selective π/2 and π flip angle pulses; BIP-750-50-20 pulse shapes (Smith et al. 2001) of duration of 200 μs were used for broadband ^{1}H inversion. For ^{13}C band-selective π/2 and π flip angle pulses G4 (or time reversed G4) (Emsley and Bodenhausen 1990) and Q3 shapes (Emsley and Bodenhausen 1992) of durations of 260 and 161 μs, respectively, were used, except for the π pulses that should be band-selective on the C^{α} region (Q3, 667 μs) and for the adiabatic π pulse to invert both C' and C^{α} (smoothed Chirp 500 μs, 20 % smoothing, 80 kHz sweep width, 11.3 kHz RF field strength) (Böhlen and Bodenhausen 1993). The ^{13}C band selective pulses on C^{α} and C' were applied at the center of each region, respectively. For the $^{1}H^{\alpha}$ experiments, decoupling of ^{1}H and ^{13}C was achieved with DIPSI-2 (Shaka et al. 1988; Cavanagh and Rance 1992) (3.1 kHz) and GARP-4 (Shaka et al. 1985) (4.5 kHz) sequences, respectively. All gradients employed had a smoothed square shape. The parameters used for the acquisition of the 3D and 5D experiments are reported, respectively, in Table 1 and 2. Each experiment was acquired with the States or Echo-Antiecho method applied in the indirect dimensions to achieve quadrature detection. All the experiments were performed using on-grid non-uniform sampling (NUS). The on-grid "Poisson disk" sampling scheme (Kazimierczuk et al. 2008) was chosen to generate the time schedules with the *RSPack* program. The distribution was relaxation-optimized, i.e. the density of points was decaying according to the Gaussian distribution exp $(-t^2/\sigma^2)$, with $\sigma = 0.5$. All the spectra were acquired using *Bruker TopSpin 3.1* software.

NMR data processing and analysis

The experimental data were converted with *nmrPipe* (Delaglio et al. 1995). 3D data were processed using the multidimensional Fourier transform (MFT) algorithm implemented in the *ToASTD* program (Kazimierczuk et al. 2006). 5D data were processed using the Sparse MFT (SMFT) algorithm implemented in the *reduced* program (Kazimierczuk et al. 2009, 2010b). Both programs are available at http://nmr.cent3.uw.edu.pl/software. Finally, *Sparky* (Goddard and Kneller 2000) was used to analyze the spectra.

Handling and inspection of 5D NMR spectra

The SMFT algorithm (Kazimierczuk et al. 2009) was employed to process the 5D data sets. This method enables to process data with very high resolution in all dimensions by storing only the informative parts of multidimensional

Table 1 Experimental parameters used for the 3D experiments (providing the basis spectra)

	Spectral widths and maximal evolution times			No. of scans	Inter-scan delays (s)	No. of complex points (aq)	No. of hypercomplex points	Duration of the experiment	Relative data points density (%)
3D (HACACO)NCO (CA)HA	3.5 kHz (^{15}N) 35.7 ms	2.0 kHz ($^{13}C'$) 35.0 ms	13.3 kHz ($^{1}H^{\alpha}$)	4	1.0	1,024	1,450	8 h	20.3
5D BEST TROSY HNCO	1.8 kHz ($^{13}C'$) 41.7 ms	2.6 kHz (^{15}N) 40.4 ms	13.3 kHz ($^{1}H^{N}$)	4	0.2	1,024	1,575	3 h	20.0

Table 2 Experimental parameters used for the 5D experiments

	Spectral widths and maximal evolution times					No. of scans	Inter-scan delays (s)	No. of complex points (aq)	No. of hyper-complex points	Duration of the experiment	Relative data points density (%)
	Cross-sections		Dimensions shared with the basis spectrum								
5D (HACA) CON(CACO) NCO(CA)HA	2.0 kHz ($^{13}C'$) 35.0 ms	3.5 kHz (^{15}N) 35.7 ms	3.5 kHz (^{15}N) 35.7 ms	2.0 kHz ($^{13}C'$) 35.0 ms	13.3 kHz ($^{1}H^{\alpha}$)	8	1.0	1,024	1,150	2 days, 7 h	0.0015
5D BEST TROSY (H)NCO (CAN)CONNH	2.6 kHz (^{15}N) 40.4 ms	1.8 kHz ($^{13}C'$) 41.7 ms	1.8 kHz ($^{13}C'$) 41.7 ms	2.6 kHz (^{15}N) 40.4 ms	13.3 kHz ($^{1}H^{N}$)	4	0.2	1,024	1,200	11 h	0.0019
5D BEST TROSY HN (COCAN) CONNH	2.0 kHz ($^{1}H^{N}$) 40.0 ms	2.6 kHz (^{15}N) 40.4 ms	1.8 kHz ($^{13}C'$) 41.7 ms	2.6 kHz (^{15}N) 40.4 ms	13.3 kHz ($^{1}H^{N}$)	4	0.2	1,024	1,200	11 h	0.0018

spectra. Indeed processing NMR spectra with high resolution in all dimensions would provide prohibitively large data files for spectra with more than three dimensions. However, in most cases the number of expected cross-peaks in the spectra does not increase passing from a 3D to a higher dimensional experiment. Therefore, these high dimensional spaces are essentially "empty" with interesting information confined in very narrow regions, those where cross-peaks are observed in the lower-dimensionality 3D spectra. One interesting feature of SMFT is its unique ability to simplify the analysis of multidimensional spectra, by considering that a 5D spectrum can be decomposed into a 3D spectrum in which each detected correlation peak is associated to an additional 2D spectrum, contributing the remaining two chemical shifts. Actually, the 3D spectrum (to which we refer as "basis spectrum") can be conveniently obtained from a simpler (and thus also more sensitive) 3D experiment correlating three (out of the five) resonances. The two remaining frequencies, which are the new information-content of the 5D spectrum, are retrieved in a series of 2D spectra (to which we refer as "cross-sections"), which are the only spectra computed from the 5D data sets. Therefore, the analysis of a 5D spectrum is simplified to the inspection of a series of 2D cross-sections, to which the three "basic" frequencies, retrieved in the corresponding 3D spectrum, are joined.

Results and discussion

To acquire high dimensional NMR experiments (4D, 5D, etc.), the use of sparse sampling strategies is very convenient. In fact, data sampling at non-uniform intervals is not only an ingenious way to speed up the acquisition of multidimensional NMR experiments, but mostly a powerful method to significantly increase the achievable spectral resolution of the indirect dimensions, as demanded by IDPs. Therefore, all the spectra were acquired using NUS. An inconvenient feature of NUS data processing is the appearance of sampling artifacts in the spectrum. Sampling artifacts can be efficiently suppressed, for example by employing the signal separation algorithm (SSA) (Stanek et al. 2012). However, in the case of protein backbone assignment, with spectra characterized by a small number of correlation peaks, artifacts are of low intensity and can be safely ignored.

Using the experimental strategy outlined in the "Materials and methods" section, we recently built a set of 4D and 5D NMR experiments based on ^{13}C-direct detection, specifically tailored to facilitate the chemical shift assignment of IDPs (Bermel et al. 2012, 2013b). The high dimensional experiments were designed as expansion of the well-known 2D CON-IPAP and 3D CACON-IPAP experiments (used as basis spectra to process the 4D and 5D data sets, respectively). From our experience, the experiments in which two pairs of ^{15}N and ^{13}C′ frequencies belonging to two neighboring residues are correlated, in order to sequentially link the peaks retrieved in the 2D CON-IPAP spectrum, provide the best result in terms of rapidity and reliability of the sequential backbone assignment (Bermel et al. 2013b; Pantoja-Uceda and Santoro 2014). This assignment approach, to which we like to refer as "CON-CON strategy", takes advantage of the relatively good chemical shift dispersion of ^{15}N and ^{13}C′ frequencies (Bermel et al. 2013a) of IDPs, which significantly improve the resolution of the spectra together with the use of three or four indirect dimensions. The employment of ^{13}C-direct detection provides a series of additional benefits. First, it reduces the minimum number of dimensions needed to apply the CON-CON strategy to four, because one of the two ^{13}C′ dimensions is directly acquired. Secondly, it allows to detect correlation peaks for all residues, included prolines that are missing in amide ^{1}H-detected experiments. Finally, it allows to perform experiments at near physiological conditions (high temperature and pH), since carbon nuclei are not directly affected by chemical exchange with the solvent. The major drawback of ^{13}C′ direct-detected experiments is the ^{13}C intrinsic lower sensitivity, largely compensated by recent technological improvements, which however can still become a limiting factor for complete assignment of less-concentrated or short-lived protein samples. Herein we now propose a set of 5D NMR experiments that still exploit the CON-CON strategy, but make use of protons instead of carbons for detection.

In order to take advantage of the particular properties of aliphatic and amide ^{1}H, we developed two different pulse sequences in which either H^{α} or H^{N} nuclei are exploited for detection as well as for the starting source of magnetization. The new 5D experiments, named (HACA)CON(CA-CO)NCO(CA)HA, BT-(H)NCO(CAN)CONNH (and its variant BT-HN(COCAN)CONNH) will be discussed in the following in more detail. In all experiments, pairs of ^{15}N and ^{13}C′ resonances are collected in the indirect dimensions. Semi-constant time or real time chemical shift evolution is used to achieve highest possible spectral resolution in all dimensions, making transverse relaxation the only limitation for choosing the appropriate acquisition time.

For its similarity with carbon detection, alpha proton detection (Mäntylahti et al. 2011) was chosen in the design of the 5D (HACA)CON(CACO)NCO(CA)HA experiment. In fact, H^{α} are not influenced by chemical exchange with the solvent even at high temperature and pH, thus being very appropriate to investigate IDPs at near physiological conditions. Moreover, H^{α}-based experiments provide direct information on proline residues, giving the possibility to

obtain virtually complete assignment. The use of deuterated water as solvent (99.99 % D_2O) ensures to completely prevent missing signals under the residual water peak. It also provides further advantages that can enhance the spectrometer performance, like the possibility to work at receiver gain values where digitizer noise is not an issue and to avoid all the phenomena related to radiation damping which become an issue when working with cryoprobes and high salt sample conditions, often necessary when dealing with IDPs. A slight drawback is the increased viscosity of D_2O with respect to H_2O that results in increased transverse relaxation rates. This effect is, however, counterbalanced by the possibility to work at higher temperature, as there is no need to slow down solvent exchange processes.

The coherence transfer pathway and the information provided by the 5D (HACA)CON(CACO)NCO(CA)HA experiment are reported in Fig. 1, top panel. The flow of magnetization, which starts and ends on alpha protons, is established through ^{1}J scalar coupling transfer steps; in particular, the back-transfer step mediated by $^2J(N–C^{\alpha})$ scalar coupling is suppressed. Therefore, the experiment correlates C'_{i-1}–N_i–N_{i+1}–C'_i–H_i^{α} nuclei, making the magnetization transfer pathway unidirectional. In this way, the sequence-specific backbone assignment procedure becomes a "walk" through a series of CON cross-peaks. The sequence was designed to detect only one peak in each cross-section, reducing at maximum the risk of peak overlaps. For SMFT processing, a 3D (HACACO)NCO(CA)HA experiment (Mäntylahti et al. 2010) has to be performed, as it provides N_{i+1}–C'_i–H_i^{α} frequency correlations (basis spectrum), that are used to extract from the 5D spectrum the additional C'_{i-1}–N_i frequencies. Hence, the "CON-CON strategy" is carried out sequence-specifically linking the pairs of C'_i and N_{i+1} resonances, retrieved in the cross-sections, to the ones present in the basis spectrum, as shown in Fig. 2. To facilitate the mapping of the linked resonances on the amino acidic sequence, the 5D (HACA)CON(CACO)NCO(CA)HA experiment has been designed in a way that cross peaks for glycine residues (at the *i* position) are of opposite sign with respect to all others, as illustrated in Fig. 3 where six cross-sections of the 5D spectrum are shown. It is interesting to note that, with the resolution obtained in the direct ^{1}H dimension, doublets due to $^3J(H^{\alpha}–H^{\beta})$ scalar couplings are detected for alanine, valine, isoleucine and threonine residues. However, these line splittings do not significantly complicate the analysis of the spectra.

As an alternative to H^{α} detection, the 5D BT-(H)NCO(CAN)CONNH experiment was developed to exploit the same "CON-CON strategy", but using amide proton detection. The coherence transfer pathway and the information provided by this experiment are reported in Fig. 1, bottom panel. The experiment is a variant of the previously published 5D (H)NCO(NCA)CONH experiment (Zawadzka-Kazimierczuk et al. 2012b). However, in our new pulse sequence, the magnetization is transferred from and to neighboring amide protons by again only exploiting ^{1}J scalar coupling transfer steps, resulting in unambiguous N_{i+1}–C'_i–C'_{i-1}–N_i–H_i^N correlations. The experiment is implemented as a BEST-TROSY (BT) version, in order to benefit from the short selective longitudinal relaxation times of amide protons in IDPs (Solyom et al. 2013; Gil et al. 2013). The combination of BEST (Pervushin et al. 2002; Schanda 2009; Schanda et al. 2006) and TROSY (Pervushin et al. 1997), as described by Favier and Brutscher (Favier and Brutscher 2011), ensures highest

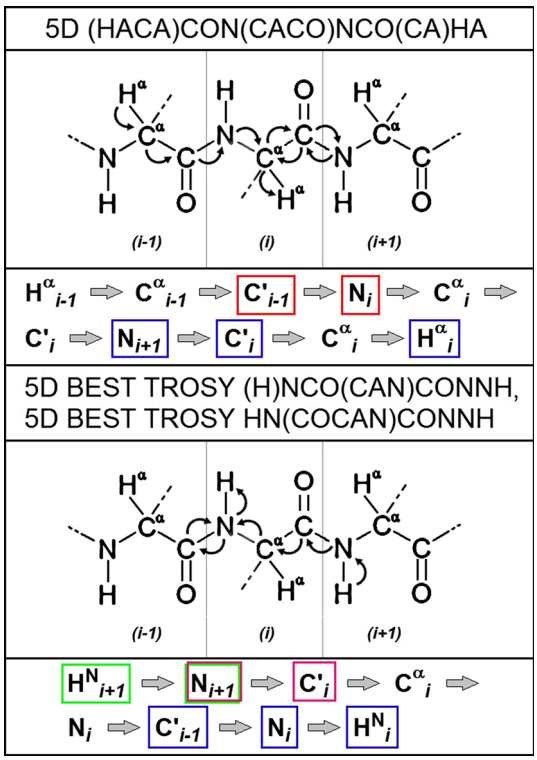

Fig. 1 The flow of the magnetization in the proposed experiments is schematically illustrated. The magnetization transfer pathway of the 5D (HACA)CON(CACO)NCO(CA)HA experiment is reported in the *top panel*, whereas that of the 5D BT-(H)NCO(CAN)CONNH and BT-HN(COCAN)CONNH experiments at the *bottom*. All *arrows* represent coherence transfer steps through ^{1}J scalar couplings. For each experiment, the correlations retrieved are also listed explicitly. Those embedded in colored boxes are the ones frequency-labeled within the experiments. *Blue rectangles* show the dimensions shared with the basis spectrum, whereas *red*, *magenta* and *green rectangles* indicate the dimensions of the 2D cross-sections of the 5D (HACA)CON(CACO)NCO(CA)HA, 5D BT-(H)NCO(CAN)CONNH and 5D BT-HN(COCAN)CONNH experiments, respectively. The two variants of the BEST TROSY experiments have in common four dimensions

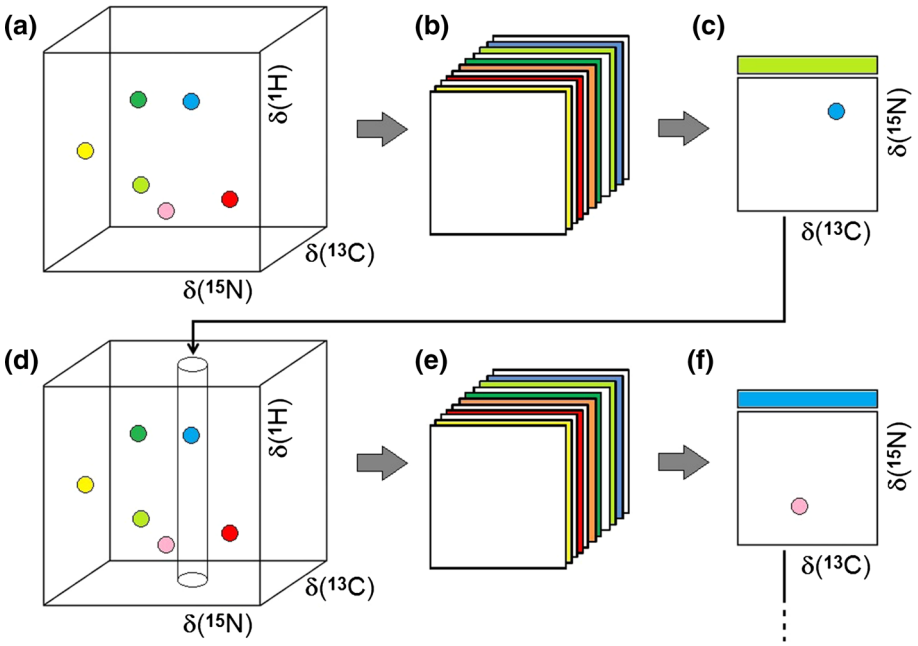

Fig. 2 Schematic illustration of the "CON-CON strategy". The sequence-specific assignment is performed through simultaneous exploitation of C′ and N resonances, linked between the 2D cross-sections of the 5D spectrum and the 3D basis spectrum. The same procedure can be used both for $^{1}H^{\alpha}$ and $^{1}H^{N}$ detected spectra, as described here schematically. First, a cross-peak (*light green* in the example) of the 3D basis spectrum is chosen (**a**). Thereafter, the related 2D cross-section of the 5D spectrum is inspected (**b**) to retrieve the frequencies of the CON peak immediately preceding or succeeding (for the $^{1}H^{\alpha}$ and $^{1}H^{N}$ spectrum, respectively) the one of the basis spectrum (**c**). Once the cross-peak (*blue* in the example) of the 3D basis spectrum with the same C′ and N frequencies is found (**d**), it becomes sequence-specifically linked to the one found in (**a**). Then, the procedure continues in (**e**) and (**f**) as explained in (**b**) and (**c**), forming chains of adjacent CON peaks in the primary sequence of the protein

experimental sensitivity and spectral resolution in very short overall acquisition times. As amide protons are affected by chemical exchange with the solvent, it is opportune to perform the experiment at lower temperature and pH in order to prevent extensive line broadening. In order to obtain the basis spectrum, correlating C'_{i-1}–N_i–H_i^N nuclei, an additional BT-HNCO (Solyom et al. 2013) spectrum is recorded. The frequency information, obtained from BT-HNCO, then allows to extract a series of 2D cross-sections from the 5D BT-(H)NCO(CAN)CONNH data set that adds the additional N_{i+1}–C'_i information (Fig. 4, panel a). Then, as described for the H^{α}-detected experiment, it is possible to establish sequential correlations linking the pairs of C'_i and N_{i+1} frequencies present in the 2D cross-sections of the 5D spectrum to those embedded in the 3D basis spectrum (Fig. 2). The only difference with the H^{α}-detected experiment is that the sequential walk now proceeds from the N- to the C-terminus, instead of the inverse. Because of the higher intrinsic sensitivity and shorter overall experimental time requirements, the 5D BT-(H)NCO(CAN)CONNH experiments is especially appealing for fast-degrading or low-concentrated IDP samples.

Finally, the above described "CON-CON strategy" can be extended to include the amide ^{1}H, thus becoming a "HNCO-HNCO strategy". For this purpose, the H_{i+1}^N needs to be added to the frequency information contained in the 5D BT-(H)NCO(CAN)CONNH spectrum. In principle, this can be achieved by recording a 6D BT-HNCO(CAN)CONNH correlation spectrum, adding an amide ^{1}H chemical shift evolution period at the beginning of the pulse sequence (Zawadzka-Kazimierczuk et al. 2012b). Here we have chosen an alternative approach that consists in recording a second 5D experiment, BT-HN(COCAN)CONNH, where chemical shift evolution of one of the carbonyls is replaced by that of the remaining amide proton, thus correlating H_{i+1}^N–N_{i+1}–C'_{i-1}–N_i–H_i^N nuclei (Fig. 1, bottom panel). The same 3D BT-HNCO spectrum can be used to compute the 2D cross-sections for this additional 5D experiment. Recording of the two 5D data sets with a different set of frequency-labeled nuclei is equivalent to recording a complete 6D data set as long as there is no degenerate N_{i+1}–C'_{i-1}–N_i–H_i^N frequency correlation. This strategy is well-known as projection spectroscopy (Kupce and Freeman 2003). The "HNCO-HNCO strategy" is likely to provide a high level of unambiguous sequential

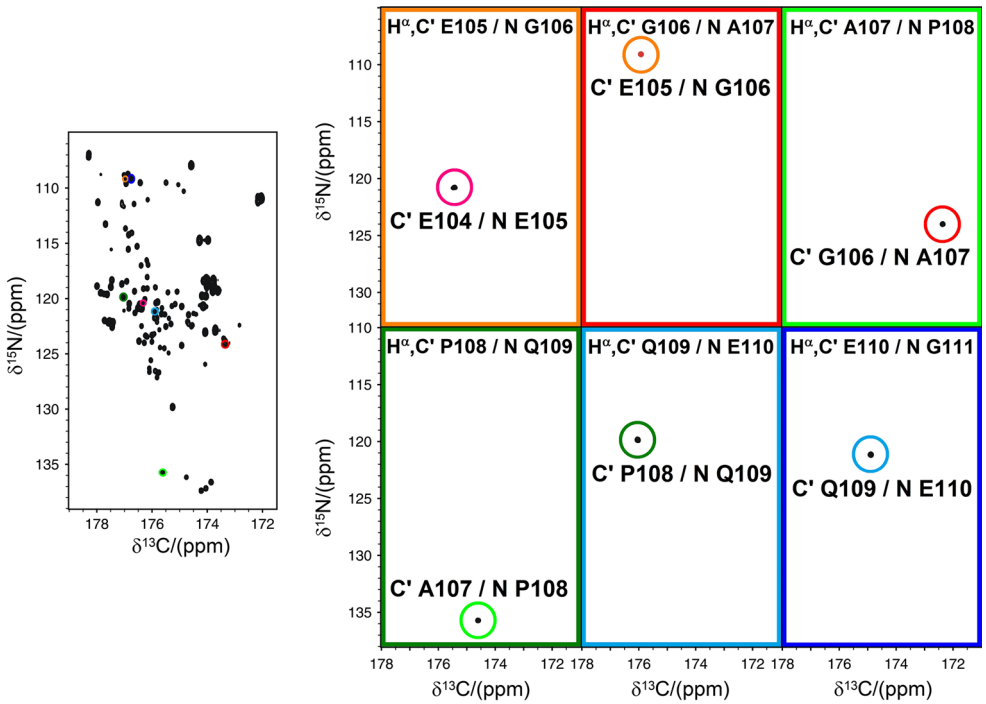

Fig. 3 Six 2D cross-sections of the 5D (HACA)CON(CACO)NCO (CA)HA spectrum are shown (*right side*), extracted at the N_{i+1}–C'_i–H^{α}_i frequencies reported on the top of each panel and schematically indicated by circles on the 2D C′–N projection of the 3D basis spectrum. In each cross-section, the observed cross-peak allows to identify the frequencies of the previous CON cross-peak (N_i–C'_{i-1}). *Black* peaks are positive, *red* negative. Inversion of peak sign is expected when a glycine is frequency-labeled. Therefore, negative peaks allow a straightforward identification of glycines. The sequential specific assignment is performed connecting the CON peak of a given cross-sections to the one of the basis spectrum with identical C′ and N frequencies, as explained in Fig. 2 and illustrated here with the use of colors, so that each color identifies a specific pair of N_{i+1}–C'_i frequencies. The simultaneous exploitation of carbonyl carbon and nitrogen makes the sequential assignment extremely reliable. As clearly visible, the use of five dimensions provide extremely-resolved cross-sections, minimizing the chance of possible overlaps

assignment, even for large IDPs with repetitive primary sequences, since it is very rare even for IDPs that two or more peaks retrieved in the 3D HNCO spectrum are completely overlapping. Analogously, a similar strategy could also be considered for the H^{α}-detected experiment although the added value is less pronounced due to the lower chemical shift dispersion of H^{α}.

Finally, it's worth mentioning that sequence-specific assignment of resonances based on the proposed experiments can be performed automatically, for example by using the *TSAR* program (Zawadzka-Kazimierczuk et al. 2012a), available at http://nmr.cent3.uw.edu.pl/software. However, the programs for automatic assignment usually require at least one experiment providing sequential connectivities and at least one experiment providing information on amino-acid type. Therefore automatic assignment procedures were not attempted in the present work, since all proposed experiments belong to the first category.

Conclusions

We have presented new high-dimensional (5D) experiments exploiting C'_i–N_{i+1} pairs for sequence-specific assignment, based on either H^{α} or H^{N} detection. These experiments provide new complementary tools to the previously described ^{13}C-detected experiments for highly efficient backbone resonance assignment of flexible IDPs. The "CON-CON strategy", including different experimental variants based on ^{13}C, $^{1}H^{\alpha}$ and $^{1}H^{N}$ detection, and its extension to the "HNCO-HNCO strategy", in case of $^{1}H^{N}$ detection, allow to establish almost unambiguous sequential correlations along the protein backbone. While the H^{α}–detected experiment is especially useful for IDP samples studied under near physiological conditions and/or characterized by proline-rich repetitive amino acid sequences, the H^{N}-detected experiments are expected to be favorable for low concentrated or short-lived IDP samples that can be studied under conditions where solvent exchange is sufficiently slowed down.

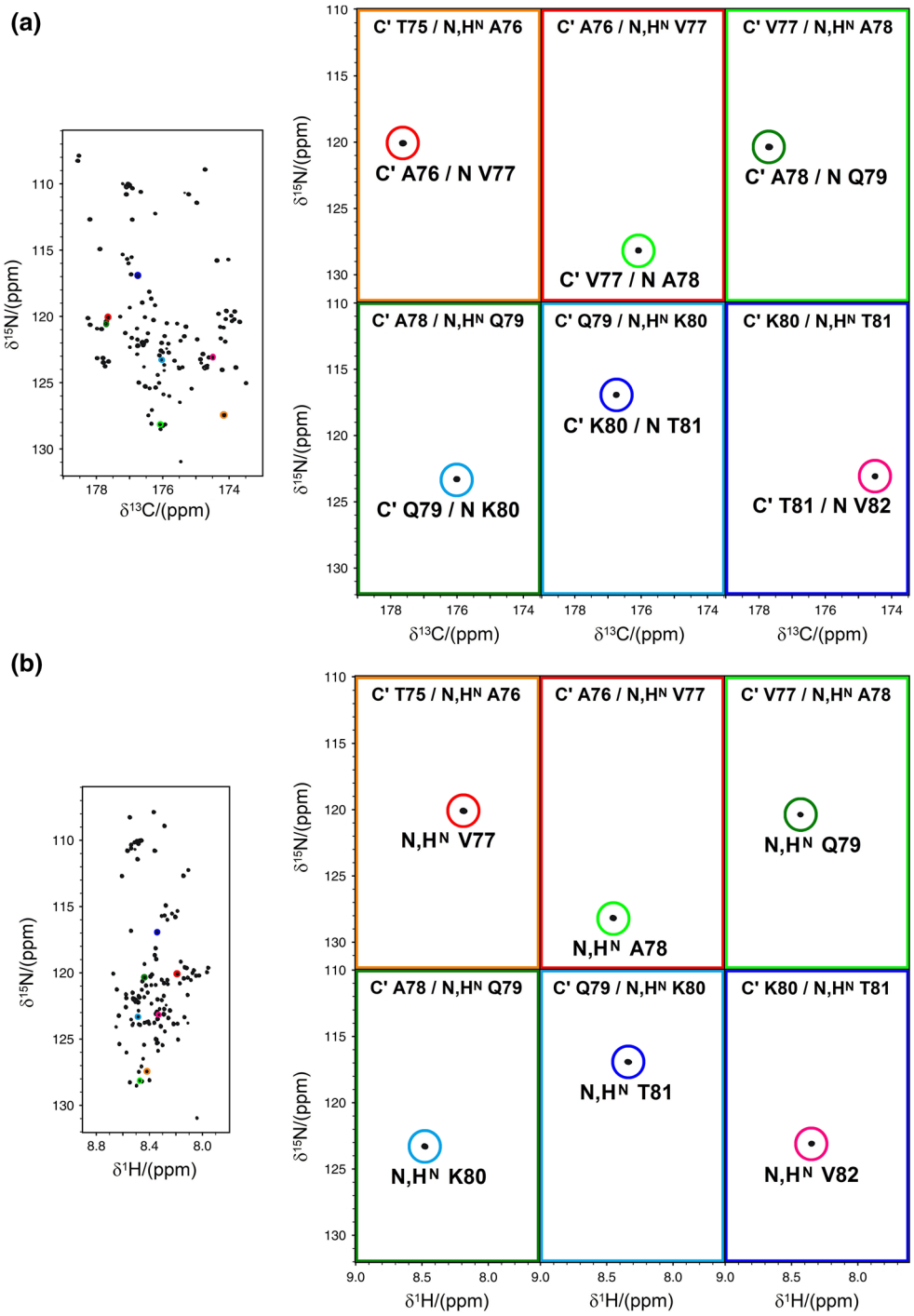

Fig. 4 Six 2D cross-sections of the 5D BT-(H)NCO(CAN)CONNH and 5D BT-HN(COCAN)CONNH spectra, extracted at C'_{i-1}–N_i–H^N_i frequencies, are shown respectively in (**a**) and (**b**). The color code and all the labels have the same meaning as described in the caption of Fig. 3. In each cross-section, the observed cross-peak allows to identify (**a**) the succeeding CON (C'_i–N_{i+1}) and (**b**) the succeeding HN (N_{i+1}–H^N_{i+1}). For this reason, on the left side the related C′–N (**a**) and H^N–N (**b**) projection of the common 3D HNCO basis experiment are reported. Since the cross-section of the two spectra share one nitrogen dimension, they can be exploited together to provide almost unambiguous assignment in the so-called "HNCO-HNCO strategy"

Acknowledgments This work has been supported in part by the European Commission Projects IDPbyNMR (Contract No. 264257), BioNMR (Contract No. 261863) and INSTRUCT (Contract No. 211252).

References

Bermel W, Bertini I, Gonnelli L, Felli IC, Koźmiński W, Piai A, Pierattelli R, Stanek J (2012) Speeding up sequence specific assignment of IDPs. J Biomol NMR 53:293–301

Bermel W, Bruix M, Felli IC, Kumar VMV, Pierattelli R, Serrano S (2013a) Improving the chemical shift dispersion of multidimensional NMR spectra of intrinsically disordered proteins. J Biomol NMR 55:231–237

Bermel W, Felli IC, Gonnelli L, Koźmiński W, Piai A, Pierattelli R, Zawadzka-Kazimierczuk A (2013b) High-dimensionality ^{13}C direct-detected NMR experiments for the automatic assignment of intrinsically disordered proteins. J Biomol NMR 57:353–361

Böhlen J-M, Bodenhausen G (1993) Experimental aspects of chirp NMR spectroscopy. J Magn Reson Ser A 102:293–301

Cavanagh J, Rance M (1992) Suppression of cross-relaxation effects in TOCSY spectra via a modified DIPSI-2 mixing sequence. J Magn Reson 96:670–678

Csizmok V, Felli IC, Tompa P, Banci L, Bertini I (2008) Structural and dynamic characterization of intrinsically disordered human securin by NMR. J Am Chem Soc 130:16873–16879

Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277–293

Deschamps M, Campbell ID (2006) Cooling overall spin temperature: protein NMR experiments optimized for longitudinal relaxation effects. J Magn Reson 178:206–211

Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC (2001) Intrinsically disordered protein. J Mol Graph Model 19:26–59

Dyson HJ, Wright PE (2004) Unfolded proteins and protein folding studied by NMR. Chem Rev 104:3607–3622

Emsley L, Bodenhausen G (1990) Gaussian pulse cascades: new analytical functions for rectangular selective inversion and in-phase excitation in NMR. Chem Phys Lett 165:469–476

Emsley L, Bodenhausen G (1992) Optimization of shaped selective pulses for NMR using a quaternion description of their overall propagators. J Magn Reson 97:135–148

Favier A, Brutscher B (2011) Recovering lost magnetization: polarization enhancement in biomolecular NMR. J Biomol NMR 49:9–15

Felli IC, Brutscher B (2009) Recent advancements in solution NMR: fast methods and heteronuclear direct detection. ChemPhysChem 10:1356–1368

Felli IC, Pierattelli R (2014) Novel methods based on ^{13}C detection to study intrinsically disordered proteins. J Magn Reson 241:115–125

Felli IC, Piai A, Pierattelli R (2013) Recent advances in solution NMR studies: ^{13}C direct detection for biomolecular NMR applications. Ann Rep NMR Spectroscop 80:359–418

Geen H, Freeman R (1991) Band-selective radiofrequency pulses. J Magn Reson 93:93–141

Gil S, Hošek T, Solyom Z, Kümmerle R, Brutscher B, Pierattelli R, Felli IC (2013) NMR studies of intrinsically disordered proteins near physiological conditions. Angew Chem Int Ed 52:11808–11812

Goddard TD, Kneller DG (2000) SPARKY 3. University of California, San Francisco

Hiller S, Wasmer C, Wider G, Wüthrich K (2007) Sequence-specific resonance assignment of soluble nonglobular proteins by 7D APSY-NMR spectroscopy. J Am Chem Soc 129:10823–10828

Hsu ST, Bertoncini CW, Dobson CM (2009) Use of protonless NMR spectroscopy to alleviate the loss of information resulting from exchange-broadening. J Am Chem Soc 131:7222–7223

Huang C, Ren G, Zhou H, Wang C (2005) A new method for purification of recombinant human alpha-synuclein in *Escherichia coli*. Protein Expr Purif 42:173–177

Ikura M, Spera S, Barbato G, Kay LE, Krinks M, Bax A (1991) Secondary structure and side-chain ^{1}H and ^{13}C resonance assignments of calmodulin in solution by heteronuclear multidimensional NMR spectroscopy. Biochemistry 30:9216–9228

Kazimierczuk K, Zawadzka A, Koźmiński W, Zhukov I (2006) Random sampling of evolution time space and Fourier transform processing. J Biomol NMR 36:157–168

Kazimierczuk K, Zawadzka A, Koźmiński W (2008) Optimization of random time domain sampling in multidimensional NMR. J Magn Reson 192:123–130

Kazimierczuk K, Zawadzka A, Koźmiński W (2009) Narrow peaks and high dimensionalities: exploiting the advantages of random sampling. J Magn Reson 197:219–228

Kazimierczuk K, Stanek J, Zawadzka-Kazimierczuk A, Koźmiński W (2010a) Random sampling in multidimensional NMR spectroscopy. Prog NMR Spectrosc 57:420–434

Kazimierczuk K, Zawadzka-Kazimierczuk A, Koźmiński W (2010b) Non-uniform frequency domain for optimal exploitation of non-uniform sampling. J Magn Reson 205:286–292

Kazimierczuk K, Misiak M, Stanek J, Zawadzka-Kazimierczuk A, Koźmiński W (2012) Generalized Fourier transform for non-uniform sampled data. Top Curr Chem 316:79–124

Kazimierczuk K, Stanek J, Zawadzka-Kazimierczuk A, Koźmiński W (2013) High-dimensional NMR spectra for structural studies of biomolecules. ChemPhysChem 14:3015–3025

Knoblich K, Whittaker S, Ludwig C, Michiels P, Jiang T, Schaffhausen B, Günther U (2009) Backbone assignment of the N-terminal polyomavirus large T antigen. Biomol NMR Assign 3:119–123

Konrat R (2014) NMR contributions to structural dynamics studies of intrinsically disordered proteins. J Magn Reson 241:74–85

Kupce E, Freeman R (2003) Projection-reconstruction of three-dimensional NMR spectra. J Am Chem Soc 125:13958–13959

Mäntylahti S, Aitio O, Hellman M, Permi P (2010) HA-detected experiments for the backbone assignment of intrinsically disordered proteins. J Biomol NMR 47:171–181

Mäntylahti S, Hellman M, Permi P (2011) Extension of the HA-detection based approach: (HCA)CON(CA)H and (HCA)NCO(CA)H experiments for the main-chain assignment of intrinsically disordered proteins. J Biomol NMR 49:99–109

Mittag T, Forman-Kay J (2007) Atomic-level characterization of disordered protein ensembles. Curr Opin Struct Biol 17:3–14

Motáčkova V, Nováček J, Zawadzka-Kazimierczuk A, Kazimierczuk K, Žídek L, Sanderová H, Krásný L, Koźmiński W, Sklenář V (2010) Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments. J Biomol NMR 48:169–177

Narayanan RL, Dürr HN, Bilbow S, Biernat J, Mendelkow E, Zweckstetter M (2010) Automatic assignment of the intrinsically disordered protein Tau with 441-residues. J Am Chem Soc 132:11906–11907

Nováček J, Zawadzka-Kazimierczuk A, Papoušková V, Žídek L, Sanderová H, Krásný L, Koźmiński W, Sklenář V (2011) 5D 13C-detected experiments for backbone assignment of unstructured proteins with a very low signal dispersion. J Biomol NMR 50:1–11

Nováček J, Janda L, Dopitová R, Žídek L, Sklenář V (2013) Efficient protocol for backbone and side-chain assignments of large, intrinsically disordered proteins: transient secondary structure analysis of 49.2 kDa microtubule associated protein 2c. J Biomol NMR 56:291–301

Nováček J, Žídek L, Sklenář V (2014) Toward optimal-resolution NMR of intrinsically disordered proteins. J Magn Reson 241:41–52

O'Hare B, Benesi AJ, Showalter SA (2009) Incorporating 1H chemical shift determination into 13C-direct detected spectroscopy of intrinsically disordered proteins in solution. J Magn Reson 200:354–358

Panchal SC, Bhavesh NS, Hosur RV (2001) Improved 3D triple resonance experiments, HNN and HN(C)N, for H^N and ^{15}N sequential correlations (^{13}C, ^{15}N) labeled proteins: application to unfolded proteins. J Biomol NMR 20:135–147

Pantoja-Uceda D, Santoro J (2013a) A suite of amino acid residue type classification pulse sequences for ^{13}C-detected NMR of proteins. J Magn Reson 234:190–196

Pantoja-Uceda D, Santoro J (2013b) Direct correlation of consecutive C′–N groups in proteins: a method for the assignment of intrinsically disordered proteins. J Biomol NMR 57:57–63

Pantoja-Uceda D, Santoro J (2014) New 13C-detected experiments for the assignment of intrinsically disordered proteins. J Biomol NMR 59:43–50

Pérez Y, Gairí M, Pons M, Bernadò P (2009) Structural characterization of the natively unfolded N-terminal domain of human c-Src kinase: insights into the role of phosphorylation of the unique domain. J Mol Biol 391:136–148

Pervushin K, Riek R, Wider G, Wüthrich K (1997) Attenuated T-2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci USA 94:12366–12371

Pervushin K, Vogeli B, Eletsky A (2002) Longitudinal (1)H relaxation optimization in TROSY NMR spectroscopy. J Am Chem Soc 124:12898–12902

Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progr NMR Spectrosc 34:93–158

Schanda P (2009) Fast-pulsing longitudinal relaxation optimized techniques: enriching the toolbox. Prog NMR Spectrosc 55:238–265

Schanda P, Brutscher B (2005) Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. J Am Chem Soc 127:8014–8015

Schanda P, Kupce E, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. J Biomol NMR 33:199–211

Schanda P, Van Melckebeke H, Brutscher B (2006) Speeding up three-dimensional protein NMR experiments to a few minutes. J Am Chem Soc 128:9042–9043

Shaka AJ, Barker PB, Freeman R (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. J Magn Reson 64:547–552

Shaka AJ, Lee CJ, Pines A (1988) Iterative schemes for bilinear operators; application to spin decoupling. J Magn Reson 77:274–293

Skora L, Becker S, Zweckstetter M (2010) Molten globule precursor states are conformationally correlated to amyloid fibrils of human beta-2-microglobulin. J Am Chem Soc 132:9223–9225

Smith MA, Hu H, Shaka AJ (2001) Improved broadband inversion performance for NMR in liquids. J Magn Reson 151:269–283

Solyom Z, Schwarten M, Geist L, Konrat R, Willbold D, Brutscher B (2013) BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins. J Biomol NMR 55:311–321

Stanek J, Augustyniak R, Koźmiński W (2012) Suppression of sampling artefacts in high-resolution four-dimensional NMR spectra using signal separation algorithm. J Magn Reson 214: 91–102

Tompa P (2009) Structure and function of intrinsically disordered proteins. CRC Press, Boca Raton

Tompa P (2012) Intrinsically disordered proteins: a 10-year recap. Trends Biochem Sci 37:509–516

Uversky VN (2013a) A decade and a half of protein intrinsic disorder: biology still waits for physics. Protein Sci 22:693–724

Uversky VN (2013b) Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. Chem Soc Rev 40:1623–1634

Uversky V, Dunker AK (2013) The case for intrinsically disordered proteins playing contributory roles in molecular recognition without a stable 3D structure. F1000 Biol Rep 5:1

Uversky VN, Gillespie JR, Fink AL (2000) Why are "natively unfolded" proteins unstructured under physiologic conditions? Proteins Struct Funct Genet 41:415–427

Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol 293:321–331

Zawadzka-Kazimierczuk A, Koźmiński W, Billeter M (2012a) TSAR: a program for automatic resonance assignment using 2D cross-sections of high dimensionality, high-resolution spectra. J Biomol NMR 54:81–95

Zawadzka-Kazimierczuk A, Koźmiński W, Sanderová H, Krásný L (2012b) High dimensional and high resolution pulse sequences for backbone resonance assignment of intrinsically disordered proteins. J Biomol NMR 52:329–337