J Biomol NMR (2014) 58:49–60
DOI 10.1007/s10858-013-9804-0

ARTICLE

# Bridge over troubled proline: assignment of intrinsically disordered proteins using (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H experiments concomitantly with HNCO and i(HCA)CO(CA)NH

**Maarit Hellman · Henni Piirainen · Veli-Pekka Jaakola · Perttu Permi**

Received: 17 October 2013 / Accepted: 9 December 2013 / Published online: 18 December 2013


**Abstract** NMR spectroscopy is by far the most versatile and information rich technique to study intrinsically disordered proteins (IDPs). While NMR is able to offer residue level information on structure and dynamics, assignment of chemical shift resonances in IDPs is not a straightforward process. Consequently, numerous pulse sequences and assignment protocols have been developed during past several years, targeted especially for the assignment of IDPs, including experiments that employ $H^N$, $H^\alpha$ or $^{13}C$ detection combined with two to six indirectly detected dimensions. Here we propose two new HN-detection based pulse sequences, (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H, that provide correlations with $^1H^N(i-1)$, $^{13}C'(i-1)$ and $^{15}N(i)$, and $^1H^N(i+1)$, $^{13}C'(i)$ and $^{15}N(i)$ frequencies, respectively. Most importantly, they offer sequential links across the proline bridges and enable filling the single proline gaps during the assignment. We show that the novel experiments can efficiently complement the information available from existing HNCO and intraresidual i(HCA)CO(CA)NH pulse sequences and their concomitant usage enabled >95 % assignment of backbone resonances in cytoplasmic tail of adenosine receptor A2A in comparison to 73 % complete assignment using the HNCO/i(HCA)CO(CA)NH data alone.

M. Hellman · P. Permi (✉)
Program in Structural Biology and Biophysics, Institute of Biotechnology/NMR Laboratory, University of Helsinki, P.O. Box 65, 00014 Helsinki, Finland
e-mail: perttu.permi@helsinki.fi

H. Piirainen · V.-P. Jaakola
Department of Biochemistry, University of Oulu, P.O. Box 3000, 90014 Oulu, Finland

H. Piirainen · V.-P. Jaakola
Biocenter Oulu, University of Oulu, P.O. Box 5000, 90014 Oulu, Finland

**Keywords** Assignment · Adenosine receptor A2A · HN-detection · (HCA)CON(CAN)H · (HCA)N(CA)CO(N)H · Intrinsically disordered proteins · IDP

## Introduction

NMR spectroscopy is the most versatile biophysical method for characterization of intrinsically disordered proteins (IDPs), involved in numerous cellular processes that have growing biomedical relevance (Dyson and Wright 2001, 2005; Kosol et al. 2013; Uversky 2013). First obstacles in NMR-based structural characterization of IDPs are already met in the initial phase of NMR study, the resonance assignment conundrum, governed by severely clustered chemical shifts due to fast conformational averaging on NMR timescale, faster sample degradation due to susceptibility to proteolysis as well as exchange broadening of amide proton (HN) resonances at alkali pH (Bai et al. 1993; Grzesiek et al. 1997; Hu et al. 2007). Moreover, proline is the most abundant amino acid residue in IDPs and due to N-substitution it lacks amide proton. Owing to these premises, paradigm forming HN-detected experiments that correlate $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts of the residue $i$ to $^1H^N$ and $^{15}N$ frequencies of the residue $i$ and residue $i+1$ are less than optimal for the main-chain assignment of IDPs (Mäntylahti et al. 2010). Recently, Yao and co-workers showed that dispersion of $^{13}C'$ and $^{15}N$ chemical shifts sustains even in IDPs, thanks to neighboring residue effect on their resonance frequencies (Yao et al. 1997). Another hurdle stems from bidirectional coherence transfer in $^{15}N$–$^{13}C\alpha$ coupling mediated experiments such as HNCACB or HN(CA)CO (for review, see e.g., Sattler et al. 1999; Permi and Annila 2004). Owing to similar size

of $^{1}J_{NC\alpha}$ and $^{2}J_{NC\alpha}$ couplings (Delaglio et al. 1991), both intraresidual and sequential correlations can be observed in these experiments, which may hamper the assignment procedure especially for the overlapping cross peaks. As a remedy, we and others have proposed a unidirectional coherence transfer scheme, based on solely intraresidual and sequential coherence transfer schemes such as iHNCA and HN(CO)CA pair (Brutscher 2002; Nietlispach 2002; Permi 2002). This asset has been utilized in several assignment protocols developed for IDPs, including number of 4D, 5D and 6D experiments (Fiorito et al. 2006; Kazimierczuk et al. 2013). In addition to HN-detection based approaches, also methods relying on Hα and $^{13}C'$-detection have been proposed (Bottomley et al. 1999; Kanelis et al. 2000; Bermel et al. 2006, 2009; Mäntylahti et al. 2010, 2011; Nováček et al. 2012; Permi and Hellman 2012). While merits of these approaches for the assignment of proline-rich polypeptide sequences are undeniable, we strived for developing a pulse sequence that would enable efficient, yet simple assignment of IDPs, based on a set of familiar HN-detected 3D experiments, resemblance with assignment of scheme of globular proteins. To this end, an assignment protocol based on unidirectional coherence transfer to $^{13}C'$ spin was developed. A pair of HN-detected experiments, HNCO and i(HCA)CO(CA)NH, each providing correlations solely between $^{13}C'(i-1)$, N($i$), $^{1}$HN($i$) and $^{13}C'(i)$, N($i$), $^{1}$HN($i$) spins, respectively, was introduced (Mäntylahti et al. 2009). Except for the first 10 N-terminal residues, this approach was successful for the assignment of PAGE5 (Prostate Associated Gene 5 protein), a 110-residue IDP (Hellman et al. 2011). However, it did not allow us to assign a 134-residue construct of A2A-ct, a cytoplasmic tail of adenosine receptor A2A. As proposed for the assignment of proline-rich sequences using the Hα-detected experiments that utilize both $^{15}$N and $^{13}C'$ frequencies as an assignment pathways (Mäntylahti et al. 2011), we introduce here two new unidirectional NH-detected experiments, dubbed as (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H, that facilitate linking of residues $i$ $^{15}$N frequencies to $^{13}C'$ and $^{1}H^{N}$ frequencies of residue $i-1$, and residues $i$ $^{13}C'$ and $^{15}$N frequencies to residue $i+1$ $^{1}H^{N}$ frequency. Although the proposed (HCA)-CON(CAN)H experiment provides similar $H^{N}(i-1)$–$^{13}C'(i-1)$–$^{15}$N($i$) correlations as earlier proposed hNCOcanH experiment (Kumar and Hosur 2011), it minimizes the resonance overlap in IDPs and most importantly correlates also $^{15}$N frequencies of prolines at the C-terminal position. Together with the novel (HCA)N(CA)CO(N)H scheme that provides $H^{N}(i+1)$–$^{13}C'(i)$–$^{15}$N($i$) cross peaks and correlates $^{15}$N frequencies of prolines at the N-terminal position, these new pulse sequences establish links to across the proline residues and complements the HNCO and i(HCA)CO(CA)NH experiments, which together significantly facilitate the assignment of IDPs with numerous proline residues.

## Materials and methods

The novel (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H experiments, and earlier proposed hNCOcanH experiment (Kumar and Hosur 2011) were tested on 0.5 mM uniformly $^{15}$N/$^{13}$C labeled A2A-ct, dissolved in 50 mM HEPES, pH 7.2, 250 mM NaCl, 1 mM TCEP-HCl with 7 % $D_2O$ in a 250 µl Shigemi micro-cell. All data were recorded at 25 °C, on a Varian INOVA 800 NMR spectrometer, equipped with a cryogenically cooled $^{15}$N/$^{13}$C/{$^{1}$H} triple-resonance probehead and an actively shielded z-axis gradient system. Two-dimensional $^{15}$N-$^{1}$H correlation spectra shown in Fig. 1a–d were measured using the hNcocanH scheme with delay $2T_{CAN} = 56$ ms (Fig. 1a), the novel (HCACO)N(CAN)H experiment with two $2T_{CAN}$ delay values, 56 ms and 28 ms (Fig. 1b, c), and the new (HCA)N(CON)H scheme with $T_{CAN} = 56$ ms (Fig. 1d). All two-dimensional spectra were measured using 80 and 852 complex points in $t_1$ and $t_2$, respectively, corresponding to acquisition times of 30 and 85 ms in $t_1$ ($^{15}$N) and $t_2$ ($^{1}H^{N}$) dimensions. All spectra were accumulated with 32 transients per FID using relaxation time of 1 s. Total experimental time for each spectrum was 1 h and 50 min.

For the assignment of A2A-ct, a three-dimensional (HCA)CON(CAN)H ($2T_{CAN} = 56$ ms), (HCA)N(CA)-CO(N)H ($T_{CAN} = 56$ ms) and (HCA)N(CO)NH ($T_{CAN} = 56$ ms) spectra were measured and supplemented with additional i(HCA)CO(CA)NH (Mäntylahti et al. 2009), HNCO (Muhandiram and Kay 1994), iHNCACB (Tossavainen and Permi 2004) and CBCA(CO)NH (Rios et al. 1996) experiments. For the (HCA)CON(CAN)H experiment 43, 110 and 852 complex points in $t_1$ ($^{13}$C), $t_2$ ($^{15}$N), and $t_3$ ($^{1}H^{N}$) were collected, corresponding to acquisition times of 27, 28 and 85 ms, respectively. Spectrum was accumulated with 4 transients using relaxation time of 1 s, resulting in experimental time of 24.5 h. For (HCA)N(-CA)CO(N)H, 80, 32 and 852 complex points in $t_1$ ($^{15}$N), $t_2$ ($^{13}C'$), and $t_3$ ($^{1}H^{N}$) was collected, corresponding to acquisition times of 30, 20 and 85 ms, respectively. Experimental time, with four transients per FID using relaxation time of 1 s, was 14 h 50 min. The (HCA)N(-CO)NH implementation (Fig. 2d′) was collected using 80, 32 and 852 complex points in $t_1$ ($^{15}$N), $t_2$ ($^{15}$N), and $t_3$ ($^{1}H^{N}$), respectively. This translates into acquisition times of 30, 12 and 85 ms in $t_1$, $t_2$ and $t_3$, respectively. Four transients per FID and relaxation time of 1 s were used, resulting in total experimental time of 14 h and 50 min.

Spectra were processed using the standard VNMRJ 2.2 revision D software package (Varian Associates 2006) and

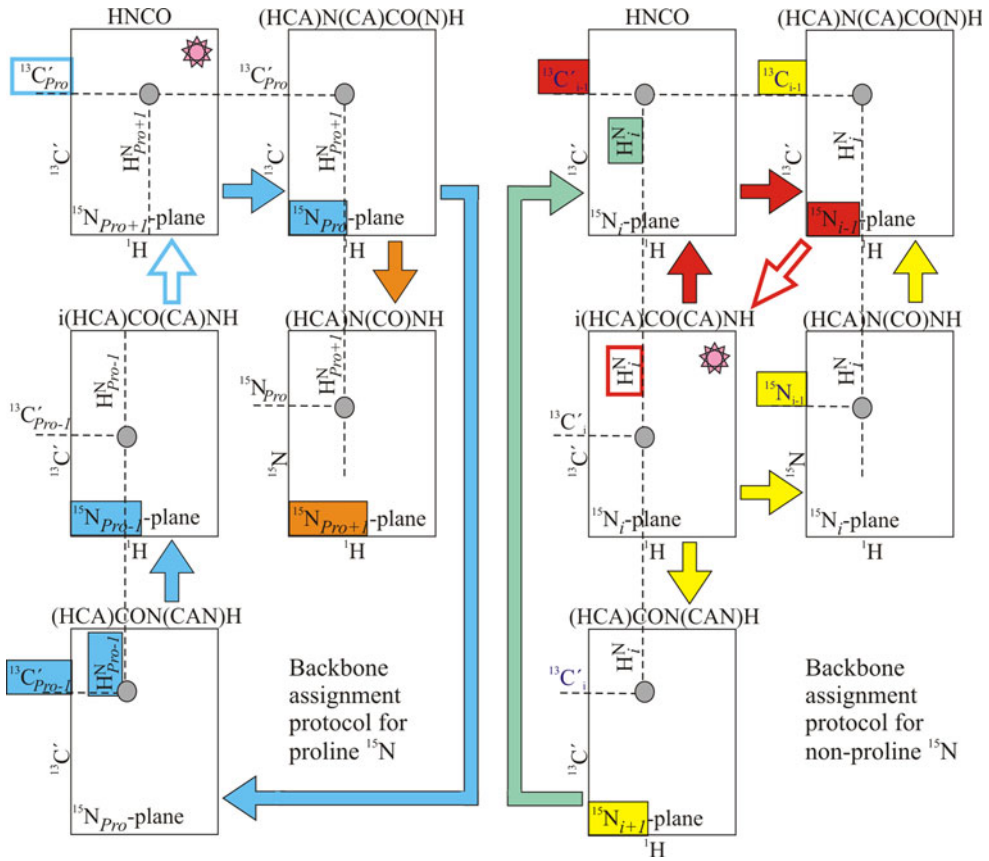



**Fig. 1** Schematic presentation of sequential assignment procedure based on i(HCA)CO(CA)NH, HNCO, and novel (HCA)N(CA)-CO(N)H, (HCA)CON(CAN)H and (HCA)N(CACO)NH spectra. The spectrum which can be used as a starting point for assignment is marked with a *pink star*. Different assignment routes are shown by *arrows*, which are marked with individual colors. By following the route i.e., moving to the next spectrum in particular order shown by *colored arrows*, yields one new resonance frequency, colored according to an associated *arrow*. Moving from one spectrum to another, two resonance frequencies are known and only one frequency is unknown except for when crossing the proline bridge, where only the proline $^{15}$N is known a priori. **a** Assignment procedure for crossing the proline bridge. *Light blue arrows* indicate the assignment route, which can be followed when assigning the proline $^{15}$N resonances and connecting them to their neighboring residues. Assignment route, marked with *orange arrow*, can be used to complete assignments in cases with resonance overlapping. **b** Assignment procedure for non-proline amino acids is indicated by *red arrows*. Complementary assignment routes, marked with *yellow* and *light green arrows*, can be used in cases with resonance overlap

analyzed with VNMRJ 2.2 revision D and Sparky 3.1.10 (Goddard and Kneller 2004). Prior to zero-filling to two-dimensional 1,024 × 4,096 data matrix (Fig. 3), or three-dimensional 512 × 512 × 2,048 data matrix (Fig. 4) followed by the Fourier transform, the data were weighted with a shifted squared sine-bell functions applied to all three dimensions.

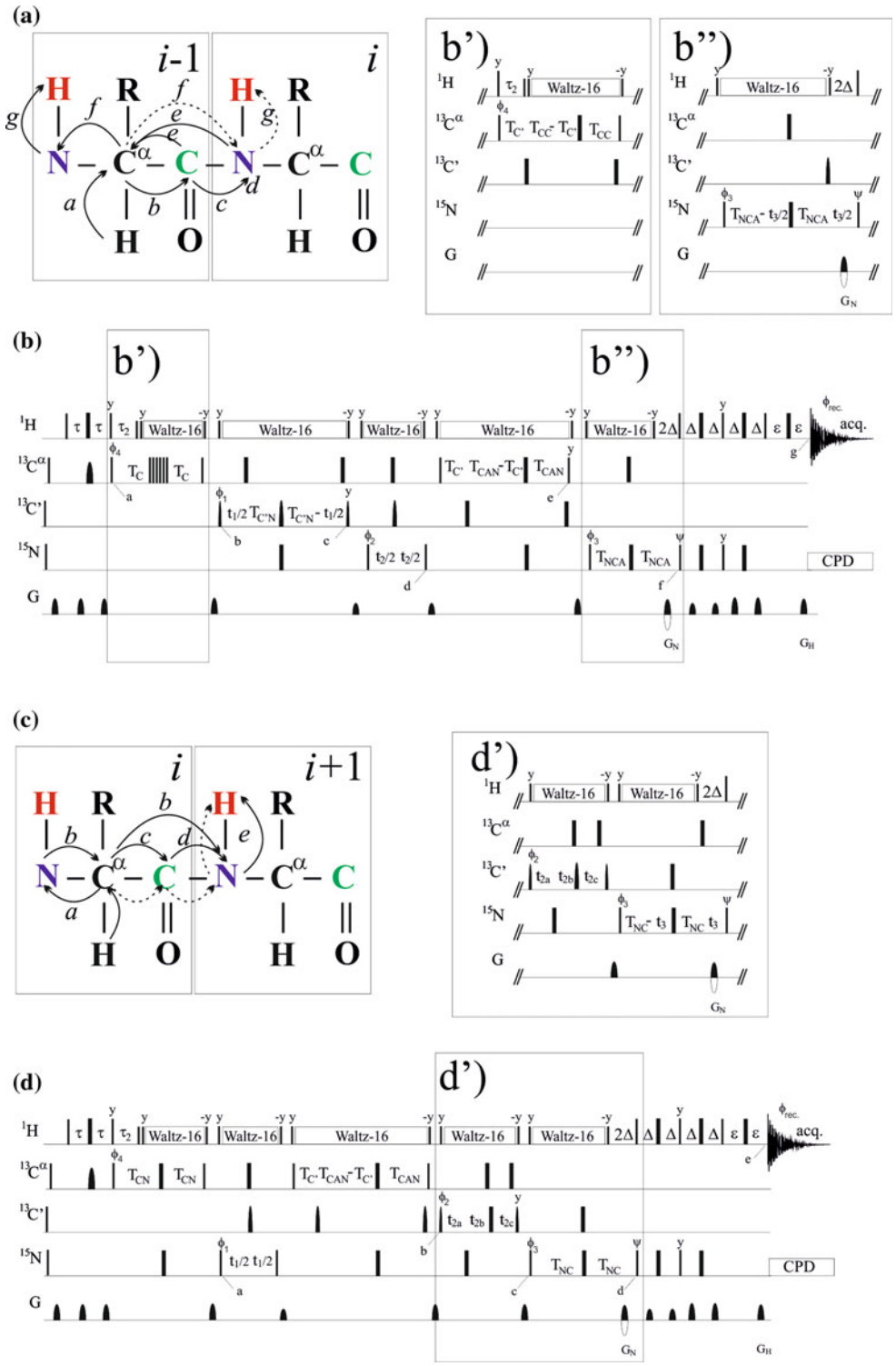
## Results and discussion

### Description of the pulse sequence: (HCA)CON(CAN)H

Figure 2a shows schematically magnetization transfer pathway utilized in the (HCA)CO(CAN)H experiment depicted in Fig. 2b. The experiment is devised for establishing connectivity between the $^{15}$N chemical shift of the proline at the position $i$ and $^{13}$C/$^{1}$H$^{N}$ chemical shifts of the residue $i-1$. Straightforward extension to 4D $^{15}$N($i$)–$^{13}$C′($i-1$)–$^{15}$N($i-1$)–$^{1}$H$^{N}$($i-1$) or 3D $^{15}$N($i$)–$^{15}$N($i-1$)–$^{1}$H$^{N}$($i-1$) correlation experiment is shown in inset Fig. 2b″. The coherence flow of the (HCA)CON(CAN)H experiment can presented as follows:

$$
\begin{aligned}
&{}^{1}H^{\alpha}(i-1)\xrightarrow{2\tau({}^{1}J_{H^{\alpha}C^{\alpha}})}{}^{13}C^{\alpha}(i-1)\xrightarrow{2T_{C},2\tau_{2}({}^{1}J_{C^{\alpha}H^{\alpha}};{}^{1}J_{C^{\alpha}C'})}{}^{13}C'(i-1)\\
&\left[2T_{C'N}-t_{1};{}^{1}J_{C'N}\right]\longrightarrow{}^{15}N(i)[t_{2}]\longrightarrow{}^{13}C^{\alpha}\\
&(i-1)\xrightarrow{2T_{CAN},2T_{C}({}^{1}J_{C^{\alpha}C'},{}^{1}J_{C^{\alpha}N},{}^{2}J_{C^{\alpha}N})}{}^{15}\\
&N(i-1)\longrightarrow{}^{2T_{NCA}({}^{1}J_{C^{\alpha}N},{}^{2}J_{C^{\alpha}N},{}^{1}J_{NH})}{}^{1}H^{H}(i-1)[t_{3}].(1)
\end{aligned}
$$

Active couplings involved in coherence transfer are indicated above the arrows and inside square brackets, whereas $t_i$ ($i$ = 1–3) refers to the acquisition time for the




corresponding spin. The proposed pulse sequence is initiated with coherence transfer from $^{1}$Hα to directly attached $^{13}$Cα spin using the INEPT sequence (Morris and Freeman 1979) and the relevant density operator at time point *a* corresponds to

$$\sigma(a) = H_z^{\alpha}(i-1)C_y^{\alpha}(i-1). \quad (2)$$

During the ensuing delay $2T_C$, that incorporates the $\tau_2$ delay for the refocusing of $^{1}J_{C\alpha H\alpha}$ coupling, the magnetization floats from $^{13}C\alpha(i-1)$ to $^{13}C'(i-1)$. The

**Fig. 2** **a** and **c** Magnetization transfer pathways along the backbone, during the **b** (HCA)CON(CAN)H and **d** (HCA)N(CA)CO(N)H experiments. *Arrows* with *solid line* indicate the actual magnetization transfer pathway, whereas an *arrow* with a *broken line* represents coherence transfer route which can be suppressed with an appropriate delay setting. *Letter codes* above *arrows* refer to time points/steps in the pulse sequence. The (HCA)CON(CAN)H scheme provides correlations with $^{1}H^{N}(i-1)$, $^{15}N(i)$, and $^{13}C'(i-1)$ frequencies. *Inset* b′) depicts alternative implementation that can be utilized in the (HCA)CON(CAN)H experiment for the $^{13}C\alpha \rightarrow {}^{13}C'$ transfer, whereas the *inset* b″) shows modification to the (HCA)CON(CAN)H experiment, allowing to label either $^{15}N(i-1)$ chemical shift in addition to (or instead of) the $^{13}C'(i-1)$ chemical shift in 4D (3D) experiment. The (HCA)N(CA)CO(N)H scheme correlates $^{1}H^{N}(i+1)$, $^{15}N(i)$, and $^{13}C'(i)$ frequencies. *Inset* d′) highlights modification to the (HCA)N(CA)CO(N)H scheme enabling either recording of $^{15}N(i+1)$ chemical shift in addition to (or instead of) the $^{13}C'(i)$ chemical shift in 4D (3D) experiment. *Narrow* and *wide rectangles* correspond to 90° and 180° flip angles, respectively. Pulses are applied with phase × unless otherwise indicated. The $^{1}H$, $^{15}N$, $^{13}C'$, and $^{13}C^{\alpha}$ carrier positions are 4.7 (water), 118 (center of $^{15}N$ spectral region), 175 ppm (center of $^{13}C'$ spectral region), and 57 ppm (center of $^{13}C^{\alpha}$ spectral region). The $^{13}C$ carrier is set initially to the middle of $^{13}C^{\alpha}$ region (57 ppm), shifted to 175 ppm prior to the first 90° $^{13}C'$ pulse (at time point *b* for both experiments), and shifted back to 57 ppm after the time point *c* in (HCA)CON(CAN)H. All rectangular 90° pulses for $^{13}C^{\alpha}$ (57 ppm) and 180° pulses for $^{13}C'$ (175 ppm) were applied with durations of 40.4 μs (90°) and 36.2 μs (180°) at 800 MHz, respectively, in order to provide null mutual excitation (Kay et al. 1990). The cascade of rectangular pulses on $^{13}C^{\alpha}$ denotes a composite pulse for ultra-broadband inversion with durations defined by pwC × ($\beta_i$/90), where $\beta_i$ is a flip angle for individual pulses in the cascade i.e. 158.0, 171.2, 342.8, 145.5, 81.2, 85.3 (Shaka 1985). Selective 90° pulses for $^{13}C'$ have the shape of center lobe of a sinc function and duration of 66.8 μs at 800 MHz. Phase modulated 180° pulses, applied off-resonance for $^{13}C^{\alpha}$, have the shape of one-lobe sinc profile and duration of 60.4 μs. The Waltz-16 sequence (Shaka et al. 1983) with strength of 4.9 kHz was employed to decouple 1H spins. The adiabatic WURST field (Kupće and Wagner 1995) was used to decouple $^{15}N$ during acquisition. Delay durations: $\tau = 1/(4J_{HC}) \sim 1.7$ ms; $\tau_2 = 2.6$ ms (for scheme b) or 2.2 ms (for schemes b′ and d); ε = duration of $G_H$ + field recovery ~0.4 ms; $2T_C = 1/(6J_{C\alpha C'})$ ~6.2 ms; $2T_{C'}$ ~9.4 ms; $2T_{C'N}$ ~32 ms; $2T_{NCA}$ ~29 ms; The $^{13}C\alpha \rightarrow {}^{15}N$ transfer delay $2T_{CAN}$ ~50–56 ms to suppress the auto-correlated pathway, or $2T_{CAN}$ ~25 ms for observing both sequential and auto-correlated cross-peaks. Maximum $t_1$ (and $t_3$) is restrained in scheme b (in scheme b″) $t_{1,max} < 2.0 \times T_{C'N}$ ($t_{3,max} < 2.0 \times T_{NCA}$, whereas $t_{3,max} < 2.0 \times T_{NC}$ in scheme d′. Frequency discrimination in $^{13}C'$ and $^{15}N$ dimensions are obtained using the States-TPPI protocol (Marion et al. 1989) applied to $\phi_1$ and $\phi_2$. In schemes b″ and d′, frequency discrimination in the second $^{15}N$ dimension ($t_3$) is obtained by the sensitivity-enhanced gradient selection (Kay et al. 1992). The echo and antiecho signals in the $^{15}N$ dimension are collected separately by inverting the sign of the $G_N$ gradient pulse together with the inversion of ψ, respectively. Phase cycling: $\phi_1$ = x, −x; $\phi_2$ = 2(x), 2(−x); $\phi_3$ = x; $\phi_4$ = 4(x), 4(−x); ψ = x; $\phi_{rec.}$ = x, 2(−x), x, −x, 2(x), −x. Gradient strengths and durations: $G_N$ = 13 k G/cm (2.5 ms), $G_H$ = 13 k G/cm (0.25 ms). The pulse sequences code and parameter file for Varian Inova system are available from authors website URL http://www.biocenter.helsinki.fi/bi/nmr/permi

coherence is relayed to the $^{13}C'(i)$ spin using the Shaka-6 composite pulse, which simultaneously refocuses $^{13}C^{\alpha}$ resonances and inverts $^{13}C'$ resonances (Shaka 1985). The delay $\tau_2$ can be set to 3.4 ms, corresponding to $1/(2J_{C\alpha H\alpha})$, which is optimal for IS spin moieties. However, a compromise value of 2.6 ms should be used for IDPs to avoid purging of glycines, which are one of the most abundant amino acid residues in IDPs. This results in only modest loss in coherence transfer efficiency. At time point *b*, the desired coherence can be described with the density operator

$$\sigma(b) = C_z^{\alpha}(i-1)C_y'(i-1)\Gamma_1, \tag{3}$$

where $\Gamma_1 = \sin(2\pi J_{C\alpha C'}T_C)\cos^{m}(2\pi J_{C\alpha C\beta}T_C)\sin(\pi J_{C\alpha H\alpha}\tau_2)\cos^{n-1}(\pi J_{C\alpha H\alpha}\tau_2)$ (m equals 0 for glycine and 1 otherwise, and n is the proton multiplicity of α-carbon i.e. 2 for glycine and 1 otherwise). The delay $2T_C$ (~6.2 ms) is a compromise between $^{1}J_{C\alpha C'}$ and $^{1}J_{C\alpha C\beta}$ couplings and $T_2$ relaxation time of $^{13}C\alpha$. Inset b′ shows an alternative $^{13}C\alpha \rightarrow {}^{13}C'$ magnetization transfer scheme, where the delay $2T_{CC}$ is set to $1/2J_{C\alpha C\beta} = 28$ ms so that $2T_{C'}$ can be set to optimal value of 9.4 ms ($1/2J_{C\alpha C'}$). In this case $\Gamma_1$ equals to $\sin(2\pi J_{C\alpha C'}T_{C'})\cos^{m}(2\pi J_{C\alpha C\beta}T_{CC})\sin(2\pi J_{C\alpha H\alpha}\tau_2)\cos^{n-1}(2\pi J_{C\alpha H\alpha}\tau_2)$. Calculations show that this transfer scheme is superior to the shorter $2T_C$ transfer step as long as $^{13}C\alpha$ $T_2 > \sim 50$ ms.

After transfer of magnetization to $^{13}C'(i-1)$ spin, the labeling of its chemical shift takes place during the $t_1$ period, which has been incorporated into $2T_{C'N}$ delay that is used for de-phasing of the coherence with respect to the $^{1}J_{C'(i-1)N(i)}$ couplings. The density operator at time point *c* is

$$\sigma(c) = C_z^{\alpha}(i-1)C_x'(i-1)N_z(i)\cos(\omega_{C'(i-1)}t_1)\Gamma_1\Gamma_2, \tag{4}$$

where $\Gamma_2 = \sin(2\pi^{1}J_{C'N}T_{C'N})$. After $t_1$ labeling, the $^{15}N$ transverse magnetization is generated that evolves during the $t_2$ period under the chemical shift Hamiltonian. Hence, the density operator after the $t_2$ evolution period, at time point *d* is

$$\sigma(d) = C_z^{\alpha}(i-1)C_z'(i-1)N_y(i)\cos(\omega_{C'(i-1)}t_1)\cos(\omega_{N(i)}t_2)\Gamma_1\Gamma_2. \tag{5}$$

After $t_2$ labeling, the magnetization is converted back to the $^{13}C\alpha(i-1)$ coherence. This magnetization evolves under the $^{1}J_{C\alpha N}$ and $^{2}J_{C\alpha N}$ couplings during the following $2T_{CAN}$ delay as well as under the $^{1}J_{C\alpha C'}$ coupling during the $2T_{C'}$ delay, incorporated into the $2T_{CAN}$ delay. The relevant coherences at the time point *e*, can be described with the density operator

$$\sigma(e) = \left[N_z(i-1)C_x^{\alpha}(i-1)\Gamma_3 + C_x^{\alpha}(i-1)N_z(i)\Gamma_4\right]\cos(\omega_{C'(i-1)}t_1)\cos(\omega_{N(i)}t_2)\Gamma_1\Gamma_2, \tag{6}$$

where $\Gamma_3 = \sin(2\pi^{1}J_{NC\alpha}T_{CAN})\sin(2\pi^{2}J_{NC\alpha}T_{CAN})\sin(2\pi J_{C\alpha C'}T_{C'})\cos^{m}(2\pi J_{C\alpha C\beta}T_{CAN})$ and $\Gamma_4 = \cos(2\pi^{1}J_{NC\alpha}T_{CAN})\cos(2\pi^{2}J_{NC\alpha}T_{CAN})\sin(2\pi J_{C\alpha C'}T_{C'})\cos^{m}(2\pi J_{C\alpha C\beta}T_{CAN})$. These

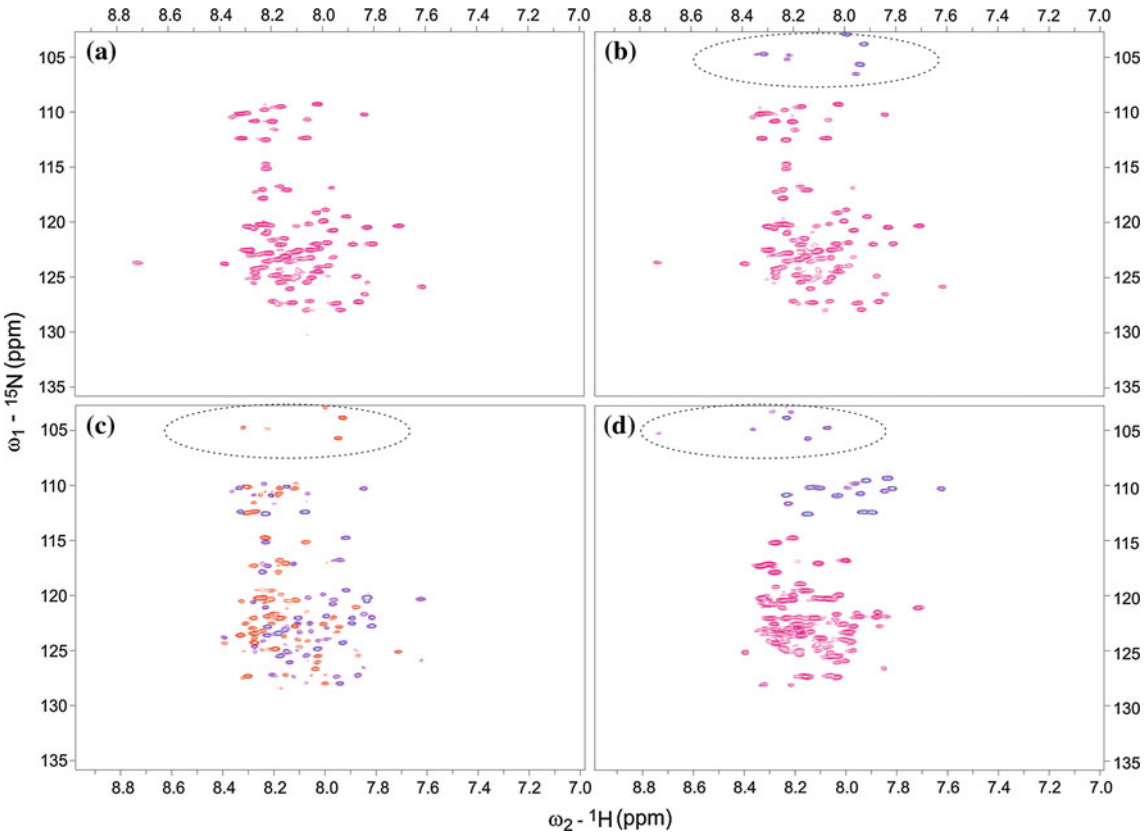



**Fig. 3** Two-dimensional $^{15}$N-$^{1}$H correlation maps from A2A-ct. **a** Displays $^{15}$N($i$)–$^{1}$H$^{N}$($i-1$) correlations in *red*, measured using the hNCOcanH experiment (Kumar and Hosur 2011). The spectrum was recorded using the $^{15}$N → $^{13}$Cα transfer delay set to 56 ms. The spectrum in **b**, measured with the (HCA)CON(CAN)H pulse scheme in Fig. 2b, shows the same $^{15}$N($i$)–$^{1}$H$^{N}$($i-1$) correlations, whereas additional correlations to proline $^{15}$N chemical shifts are shown in *blue*. The spectrum was measured by setting the $^{15}$N → $^{13}$Cα transfer delay ($2T_{CAN}$) to 56 ms i.e. for suppressing the $^{15}$N($i$)–$^{1}$H($i$) cross-peaks. Of note, to minimize $^{15}$N spectral width, the proline $^{15}$N chemical shifts are folded in. Spectrum in **c** both $^{15}$N($i$)–$^{1}$H$^{N}$($i-1$) (*blue*) and $^{15}$N($i$)–$^{1}$H($i$) (*red*) cross-peaks, measured with the (HCA)CON(CAN)H pulse sequence using the delay $2T_{CAN}$ set to 27 ms. Note that most of the missing cross peak can be seen by increasing the noise level. Spectrum in **d** $^{15}$N($i$)–$^{1}$H$^{N}$($i+1$) correlations, measured with the (HCA)N(CA)CO(N)H experiment shown in Fig. 2d. In this spectrum, correlations to glycine and proline residues are shown *blue*. As in the case of (HCA)CON(CAN)H spectrum (**b**), the 180° phase difference for prolines is due to spectral aliasing, whereas in the case of glycines (109–113 $^{15}$N ppm) this is due to transfer delay $2T_{CN}$ set to $1/J_{C\alpha C\beta}$ in (HCA)N(CA)CO(N)H i.e. for non-glycine residues signal will have additional modulation by $\cos(2\pi J_{C\alpha C\beta} T_{CN}) \sim -1$

coherences will be converted to the $^{15}$N magnetization that will be refocused with respect to $^{1}J_{N(i-1)C\alpha(i-1)}$ and $^{2}J_{N(i)C\alpha(i-1)}$ couplings and de-phased with respect to $^{1}J_{NH}$ coupling during the ensuing $2T_{NCA}$ and $2\Delta$ delays (time point *f*).

$$\sigma(f) = \begin{bmatrix} H_z^N N_y(i-1)\Gamma_3\Gamma_5 \\ +H_z^N N_x(i)\Gamma_4\Gamma_6 \end{bmatrix} \cos(\omega_{C'(i-1)}t_1)\cos(\omega_{N(i)}t_2)\sin(2\pi J_{NH}\Delta)\Gamma_1\Gamma_2, \quad (7)$$

where $\Gamma_5 = \sin(2\pi^1 J_{NC\alpha}T_{NCA})\cos(2\pi^2 J_{NC\alpha}T_{NCA})$ and $\Gamma_6 = \cos(2\pi^1 J_{NC\alpha}T_{NCA})\sin(2\pi^2 J_{NC\alpha}T_{NCA})$. Inset b″ shows another implementation of the pulse sequence, enabling labeling of $^{15}$N chemical shift, for instance during $t_3$ period in the 4D experiment utilizing non-linear sampling, or as a replacement for the $^{13}$C′ chemical shift labeling during $t_1$. Finally, gradient selected, sensitivity enhanced back-INEPT transfer is employed to transfer magnetization back to the $^{1}$H$^{N}$($i-1$) single quantum coherence (Kay et al. 1992; Schleucher et al. 1994). Alternatively, a shorter re-INEPT transfer could be used for the $^{15}$N-$^{1}$H$^{N}$ back-transfer, but we prefer more robust water suppression provided by the gradient selection.

The observable magnetization at time point *f* corresponds to a density operator

$$\sigma(g) = H^{N-}\cos(\omega_{C'(i-1)}t_1)\cos(\omega_{N(i)}t_2)\exp(i\omega_{HN(i-1)}t_3) + H^{N-}\cos(\omega_{C'(i-1)}t_1)\cos(\omega_{N(i)}t_2)\exp(i\omega_{HN(i)}t_3), \quad (8)$$

that is frequency modulated by $^{13}$C′($i-1$), $^{15}$N($i$) and $^{1}$H$^{N}$($i-1$) chemical shifts during $t_1$, $t_2$ and $t_3$ periods,

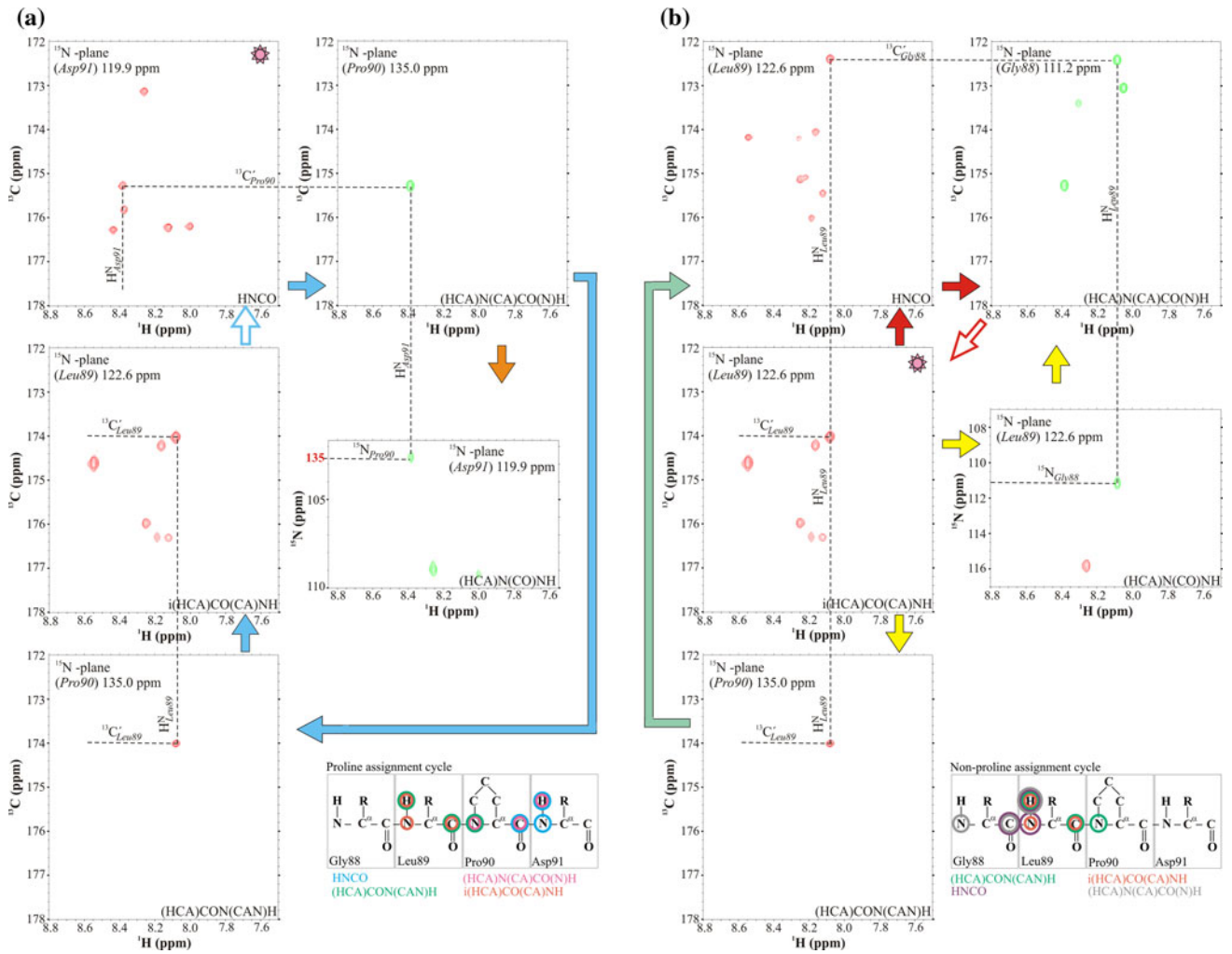



**Fig. 4** Illustration of the 'sequential walk' by means of i(HCA)CO(-CA)NH, HNCO, and novel (HCA)N(CA)CO(N)H, (HCA)CON(-CAN)H and (HCA)N(CACO)NH spectra. $^{15}$N-planes showing $^{13}C'(i)$–$^{1}H(i)$, $^{13}C'(i+1)$–$^{1}H(i)$, $^{13}C'(i-1)$–$^{1}H(i)$, $^{13}C'(i-1)$–$^{1}H(i-1)$ or $^{15}N(i-1)$–$^{1}H(i)$ projections of three-dimensional spectra taken at the chemical shift of $^{15}N(i)$ are shown. Positive and negative peaks appear in the spectra as *red* and *green*, respectively. *Colors of arrows*, indicating assignment routes, are comparable to colors used in Fig. 1. **a** Assignment procedure using proline bridge crossing is exemplified with residue Pro90 of A2A-ct. In the (HCA)N(CA)CO(N)H and (HCA)N(CO)NH spectra resonance peaks of proline $^{15}$N are folded and are negative due to the narrow spectral width in nitrogen dimension. **b** Exemplification of non-proline assignment procedure by continuing with Leu89. Assignment follows the route indicated by *red arrows*, while routes marked with *yellow* and *light green arrow* can complement the assignment. In the (HCA)N(CA)CO(N)H and (HCA)N(CACO)NH spectra, glycine $^{15}$N resonances appear as negative peaks owing to lack of cos ($2\pi J_{C\alpha C\beta} T_{CN}$) dependence during the first $^{13}C\alpha$-$^{15}N$ transfer step

respectively. In addition, depending on the delay setting $2T_{CAN}$, another auto-correlated $^{13}C'(i-1)$, $^{15}N(i)$ and $^{1}H^{N}(i)$ cross peak can emerge, the second term on the right hand side in Eq. (8). This cross peak corresponds to the pathway obtained with the conventional HNCO experiment, albeit with much lower sensitivity. Furthermore, this cross peak will emerge with its sign inverted with respect to the desired $\omega_{C'(i-1)}$, $\omega_{N(i)}$, $\omega_{HN(i-1)}$ correlation and in the case of crowded spectrum, a condition usually met with relatively small IDPs, it may lead to mutual cancellation of cross peaks. Therefore, we prefer maximizing efficiency of the $^{13}C'(i-1)$–$^{15}N(i)$–$^{1}H^{N}(i-1)$ transfer and minimizing the HNCO correlation. This can be obtained using so-called intraresidual transfer (Permi 2002; Mäntylahti et al. 2009; Mäntylahti et al. 2011) that is, by setting to $2T_{CAN}$ delay near to 50–56 ms.

Theoretical transfer efficiency for the desired magnetization, that is the first term on the right hand side in Eq. (8), in the (HCA)CON(CAN)H experiment can be calculated according to Eq. (9)

$$I_{(HCA)CON(CAN)H} \approx \Gamma_1\Gamma_2\Gamma_3\Gamma_5 \exp(-2T_{CAN}/T_{2,C\alpha}) \exp(-2T_{C'N}/T_{2,C'}) \exp(-(2T_{NCA}+t_2)/T_{2,N}). \quad (9)$$

The transfer efficiency in case of an IDP is 0.15, which can be obtained by setting delays $2T_C$, $2T_{C'N}$, $2T_{CAN}$, $2T_{C'}$, and

$2T_{NCA}$ to 6.2 ms, 30 ms, 56 ms, 9.4 ms, and 30 ms, and given that average transverse relaxation times of 200 ms, 100 ms and 200 ms for $^{15}N$, $^{13}C^{\alpha}$ and $^{13}C'$ spins, and coupling constants of 10.6 Hz, 7.5 Hz, 15 Hz, 53 Hz and 35 Hz for $^{1}J_{NC\alpha}$, $^{2}J_{NC\alpha}$, $^{1}J_{NC'}$, $^{1}J_{C\alpha C'}$, $^{1}J_{C\alpha C\beta}$ holds, respectively (Mäntylahti et al. 2010). If alternative implementation, shown in inset b′ for Fig. 2, is employed, the transfer efficiency becomes 0.19. In both cases, the sensitivity loss associated with $^{1}H \rightarrow {}^{13}C$ and $^{15}N \rightarrow {}^{1}H$ transfer steps has been omitted. The coherence transfer efficiency of the (HCA)CON(CAN)H experiment is significantly smaller than for the HNCO owing to longer coherence transfer pathway involved, but is comparable to the i(HCA)CO(CA)NH experiment ($I_{i(HCA)CO(CA)NH} \sim 0.25$) and to the hNCOcanH experiment ($I_{nNCOcanH} \sim 0.23$) given that the delay settings optimized for the *unidirectional* transfer are employed. The hNCOcanH scheme (Kumar and Hosur 2011) then provides the same correlations as the proposed (HCA)CON(CAN)H experiment except for the missing correlations of proline $^{15}N$ chemical shifts at the C-terminal position. As it will be shown later, these proline connectivities are of utmost importance for the assignment of IDP. However, another experiment that links the prolines at the N-terminal position is required and will be described below.

Description of the pulse sequence: (HCA)N(CA)CO(N)H

Figure 2c displays schematically magnetization transfer pathway used in the (HCA)N(CA)CO(N)H experiment shown in Fig. 2d. The experiment is designed for linking the proline $^{15}N$ and $^{13}C'$ chemical shifts at the position $i$ to $^{1}H^{N}$ shift of the residue $i + 1$. Alternative implementation, for extension to 4D $^{15}N(i)$–$^{13}C'(i)$–$^{15}N(i + 1)$–$^{1}H^{N}(i + 1)$ correlation experiment or 3D $^{15}N(i)$–$^{15}N(i + 1)$–$^{1}H^{N}(i + 1)$ experiment is shown in inset Fig. 2d′. The coherence flow of the (HCA)N(CA)CO(N)H scheme is described in Eq. (10):

$$
\begin{aligned}
&{}^{1}H^{\alpha}(i) \xrightarrow{2\tau({}^{1}J_{H^{\alpha}C^{\alpha}})} {}^{13}C^{\alpha}(i) \xrightarrow{2T_{CN},2\tau_2({}^{1}J_{C^{\alpha}H^{\alpha}};{}^{1}J_{C^{\alpha}N})} \\
&{}^{15}N(i)[t_1] \longrightarrow {}^{13}C^{\alpha}(i) \xrightarrow{2T_{CAN},2T_C({}^{1}J_{C^{\alpha}C'},{}^{1}J_{C^{\alpha}N},{}^{2}J_{C^{\alpha}N})} \\
&{}^{13}C'(i)\left[2T_C - t_2; {}^{1}J_{C^{\alpha}C'}\right] \longrightarrow \\
&{}^{15}N(i+1) \xrightarrow{2T_{NCA}({}^{1}J_{C^{\alpha}N},{}^{2}J_{C^{\alpha}N},{}^{1}J_{NH})} {}^{1}H^{H}(i+1)[t_3].
\end{aligned}
\tag{10}
$$

Couplings employed to coherence transfer are indicated above the arrows and inside square brackets, whereas $t_i$ ($i = 1$–3) refers to the acquisition time for the corresponding spin. The (HCA)N(CA)CO(N)H scheme starts magnetization transfer from $^{1}H\alpha(i)$ to $^{13}C\alpha(i)$, followed by transfer delay $2T_{CN}$, during which the magnetization is dephased with respect to the $^{15}N(i)$ spin. Part of the magnetization will be dephased with respect to $^{15}N(i + 1)$ spin as well, due to comparable sizes of $^{1}J_{C\alpha N}$ and $^{2}J_{C\alpha N}$ couplings (Delaglio et al. 1991). However, as in the case of (HCA)CON(CAN)H experiment, this pathway will be suppressed by appropriate delay settings in the later part of the pulse sequence. The coherence of interest prior to $t_1$ ($^{15}N$) evolution period (time point *a*) can be described with the density operator:

$$
\sigma(a) = \left[N_y(i)C_z^{\alpha}(i)\Gamma_2 + C_z^{\alpha}(i)N_y(i+1)\Gamma_3\right]\Gamma_1, \tag{11}
$$

where $\Gamma_1 = \cos^{m}(2\pi^{1}J_{C\alpha C\beta}T_{CN})\sin(\pi^{1}J_{C\alpha H\alpha}\tau_2)\cos^{n-1}(\pi^{1}J_{C\alpha H\alpha}\tau_2)$, $\Gamma_2 = \sin(2\pi^{1}J_{NC\alpha}T_{CN})\cos(2\pi^{2}J_{NC\alpha}T_{CN})$ and $\Gamma_3 = \cos(2\pi^{1}J_{NC\alpha}T_{CN})\sin(2\pi^{2}J_{NC\alpha}T_{CN})$. Of note, here the delay $\tau_2$ should be set to somewhat smaller value than in the case of (HCA)CON(CAN)H experiment, that is 2.2 ms, in order to maximize coherence transfer for both HC and $H_2C$ spin moieties. This is because the delay $2T_{CN}$ is optimally set to $1/J_{C\alpha C\beta} \sim 28$ ms, resulting in uniform transfer efficiency for glycine and non-glycine residues except for the sign inversion. During the $t_1$ period, chemical shift of $^{15}N$ spin will be labeled and the magnetization will be converted back to $^{13}C\alpha$ antiphase coherence. This magnetization will evolve under $^{1}J_{C\alpha N}$ and $^{2}J_{C\alpha N}$ coupling, and $^{1}J_{C\alpha C'}$ coupling during the ensuing delay $2T_{CAN}$ and $2T_{C'}$, respectively. Thus, the coherence of interest at time point *b* is

$$
\sigma(b) = \left[\begin{array}{l}
\left(\begin{array}{l} C_z^{\alpha}(i)C_z^{'}(i)N_z(i+1)\Gamma_4\cos(\omega_{N(i)}t_1) \\ +N_z(i)C_z^{\alpha}(i)C_z^{'}(i)\Gamma_5\cos(\omega_{N(i)}t_1)\end{array}\right)\Gamma_2 \quad + \\
\left(\begin{array}{l} N_z(i)C_z^{\alpha}(i)C_z^{'}(i)\Gamma_4\cos(\omega_{N(i+1)}t_1) \\ +C_z^{\alpha}(i)C_z^{'}(i)N_z(i+1)\Gamma_5\cos(\omega_{N(i+1)}t_1)\end{array}\right)\Gamma_3
\end{array}\right]\Gamma_1, \tag{12}
$$

where $\Gamma_4 = \sin(2\pi^{1}J_{NC\alpha}T_{CAN})\sin(2\pi^{2}J_{NC\alpha}T_{CAN})\sin(2\pi^{1}J_{C\alpha C'}T_{C'})\cos^{m}(2\pi^{1}J_{C\alpha C\beta}T_{CAN})$ and $\Gamma_5 = \cos(2\pi^{1}J_{NC\alpha}T_{CAN})\cos(2\pi^{2}J_{NC\alpha}T_{CAN})\sin(2\pi^{1}J_{C\alpha C'}T_{C'})\cos^{m}(2\pi^{1}J_{C\alpha C\beta}T_{CAN})$. In the next step, the 90° pulse on $^{13}C'$ generates $^{13}C'(i)$ antiphase coherence whose chemical shift will be labeled during the $t_2$ period, implemented into a $2T_{C'}$ delay in a semi-constant time manner (Grzesiek and Bax 1993; Logan et al. 1993). This enables use of $t_{2,max}$ longer than $2T_{C'}$, which would otherwise be limited to $1/2J_{C\alpha C'} \sim 9.5$ ms i.e. high resolution in $^{13}C'$ dimension is obtained using the semi-constant time implementation for $t_2$ incrementation. During the delay $2T_{C'}$, the $^{1}J_{C\alpha C'}$ coupling will be refocused and the density operator at time point *c* can be given:

$$
\sigma(c) = \left[\begin{array}{l}
\left(\begin{array}{l} C_z^{'}(i)N_z(i+1)\Gamma_4\cos(\omega_{N(i)}t_1) \\ +N_z(i)C_z^{'}(i)\Gamma_5\cos(\omega_{N(i)}t_1)\end{array}\right)\Gamma_2 \quad + \\
\left(\begin{array}{l} N_z(i)C_z^{'}(i)\Gamma_4\cos(\omega_{N(i+1)}t_1) \\ +C_z^{'}(i)N_z(i+1)\Gamma_5\cos(\omega_{N(i+1)}t_1)\end{array}\right)\Gamma_3
\end{array}\right]
\cos(\omega_{C'(i)}t_2)\Gamma_1\Gamma_6, \tag{13}
$$

where $\Gamma_6 = \sin(2\pi^1 J_{C\alpha C'} T_{C'})$. After refocusing of the $^1J_{C\alpha C'}$ coupling and labeling of the $^{13}C'$ chemical shift during $t_2$, the $^{15}N$ antiphase coherence is generated and $^1J_{NC'}$ coupling is allowed to refocus during the $2T_{NC'}$ delay while $^1J_{NH}$ coupling de-phases the magnetization during $2\Delta$. Due to very small size of $2J_{NC'}$ coupling, the second and third terms in Eq. (13) will not be converted into observable magnetization at the end of the pulse sequence, and hence the density operator at the time point $d$ is given by:

$$\sigma(d) = \begin{bmatrix} \left(H_z^N N_y (i+1)\Gamma_4 \cos(\omega_{N(i)} t_1)\right)\Gamma_2 \\ +\left(H_z^N N_y (i+1)\Gamma_5 \cos(\omega_{N(i+1)} t_1)\right)\Gamma_3 \end{bmatrix} \cos(\omega_{C'(i)} t_2)\Gamma_1\Gamma_6\Gamma_7 \quad (14)$$

where $\Gamma_7 = \sin(2\pi^1 J_{NC'} T_{NC'})\sin(2\pi^1 J_{NH}\Delta)$. If desired, labeling of $^{15}N$ chemical shift, either during the $t_3$ period in 4D experiment or instead of $^{13}C'$ labeling during $t_2$ in 3D experiment, can be obtained by implementation shown in inset d′. Finally, coherence-order selective coherence transfer utilizing gradient selection, is employed to transfer the $^{15}N$ coherence to amide proton and the density operator prior to the start of acquisition (time point $e$) is:

$$\sigma(e) = \begin{bmatrix} \left(H^{N-}(i+1)\Gamma_4\right)\Gamma_2 \cos(\omega_{N(i)} t_1) \\ +\left(H^{N-}(i+1)\Gamma_5\right)\Gamma_3 \cos(\omega_{N(i+1)} t_1) \end{bmatrix} \cos(\omega_{C'(i)} t_2) \exp(\omega_{H^N(i+1)} t_3)\Gamma_1\Gamma_6\Gamma_7. \quad (15)$$

The signal of interest, the first term on the right hand side of Eq. (15) is therefore labeled by the chemical shift of $^{15}N(i)$, $^{13}C'(i)$ and $^1H^N(i+1)$ in $t_1$, $t_2$ and $t_3$, respectively. As in the case of (HCA)CON(CAN)H experiment (vide infra) another, conventional HNCO-type correlation emerge at the intersection of $\omega_{C'(i)}$, $\omega_{N(i+1)}$, $\omega_{HN(i+1)}$ frequencies if the delay setting $2T_{CAN}$ deviates drastically from ~50 to 56 ms. For IDPs, we prefer using $2T_{CAN}$ value close to 55 ms in order to maximize sensitivity of the desired cross peak and minimize the redundant information in the spectrum.

Analogously to the (HCA)CON(CAN)H experiment above, a theoretical coherence transfer efficiency for the desired magnetization in the (HCA)N(CA)CO(N)H experiment can be calculated according to Eq. (16)

$$I_{N(HCA)CO(N)H} \approx \Gamma_1\Gamma_2\Gamma_6\Gamma_7 \exp(-2(T_{CAN}+T_{CN})/T_{2,C\alpha}) \exp(-(2T_{C'}+t_2)/T_{2,C'}) \exp(-(2T_{NC'}+t_1)/T_{2,N}). \quad (16)$$

Considering identical conditions as in the case of (HCA)CON(CAN)H scheme (vide infra), the transfer efficiency reaches 0.2, obtained by setting $2T_{CN}$, $2T_{CAN}$, $2T_{C'}$, and $2T_{NC'}$ to 28 ms, 56 ms, 9.4 ms, and 32 ms, respectively. Linkage of $^{15}N$ chemical shift of proline in $i-1$ position can be also obtained with the HNCAN experiment, originally targeted for the assignment of perdeuterated globular proteins (Löhr et al. 2000). Calculations made under identical experimental conditions, that is transverse relaxation rates and J couplings, indicate that the transfer efficiency of the HNCAN experiment ($I \sim 0.02$) is drastically smaller in comparison to the proposed (HCA)N(CA)CO(N)H experiment. Recently proposed hNCOcanCO and hnCOcaNCO experiments (Pantoja-Uceda and Santoro 2013) offer similar transfer efficiencies ($I \sim 0.19$) and connectivities over proline residues. However, due to the $^{13}C'$ detection an inherent sensitivity loss by a factor of 8 results in, which can be only partially compensated by $^{13}C$ detection optimized probes.

## Assignment strategies for non-proline and proline residues using HN detected experiments

The assignment protocol utilizing a set of unidirectional 3D experiments e.g., HNCO, i(HCA)CO(CA)NH, (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H is demonstrated in Figs. 1 and 4. Complementary information can be obtained through the (HCA)N(CO)NH experiment. In case of non-proline residues (Figs. 1b, 4b), assignment procedure can be started from the i(HCA)CO(CA)NH spectrum in which cross peak stems from the intraresidual frequencies of $^1H^N(i)$, $^{13}C'(i)$ and $^{15}N(i)$. Next, inspection of the HNCO spectrum at the $^{15}N(i)$ plane yields $^1H^N(i)$ and $^{13}C'(i-1)$ frequencies. Furthermore, at the $^{15}N(i-1)$ plane of (HCA)N(CA)CO(N)H spectrum, cross peak between $^{13}C'(i-1)$ and $^1H^N(i)$ can be found. Turning back to the i(HCA)CO(CA)NH spectrum at the $^{15}N(i-1)$ plane provides matching of $^{13}C'(i-1)$ and $^1H^N(i-1)$ frequencies and assignment cycle can continue further. Typically, NMR spectra of IDP contain several overlapping signals. In this procedure (HCA)CON(CAN)H and N(HCACO)NH spectra can be used for acquiring complementary data for assignments, i.e. from (HCA)CON(CAN)H cross peak for $^1H^N(i)$, $^{13}C'(i)$ and $^{15}N(i+1)$ and from (HCA)N(CO)NH for $^1H^N(i)$, $^{15}N(i-1)$ and $^{15}N(i)$.

In the case of proline residues, the assignment procedure for crossing the proline bridge differs from assignment of non-proline residues (Figs. 1a, 4a). The assignment cycle can be started for example from the HNCO spectrum, where the cross peak appears between $^1H^N(Pro+1)$, $^{13}C'(Pro)$ and $^{15}N(Pro+1)$. Moving to the (HCA)N(CA)CO(N)H spectrum yields the cross peak that appears at $^{15}N(Pro)$ plane at frequencies of $^1H^N(Pro+1)$ and $^{13}C'(Pro)$. Furthermore, at the $^{15}N(Pro)$ plane of

(HCA)CON(CAN)H spectrum, the cross peak between $^{13}C'(Pro\text{-}1)$ and $^{1}H^{N}(Pro\text{-}1)$ is found. Jumping to the i(HCA)CO(CA)NH spectrum at the $^{15}N(Pro\text{-}1)$ plane provides intraresidual cross peak at the frequencies of $^{13}C'(Pro\text{-}1)$ and $^{1}H^{N}(Pro\text{-}1)$ and assignment cycle can continue further, either by following the non-proline or proline route (Fig. 1). From the (HCA)N(CACO)NH spectrum complementary data can be acquired for assignments, that is $^{1}H^{N}(Pro + 1)$, $^{15}N(Pro)$ and $^{15}N(Pro + 1)$.

The proposed approach enables linking of amino acid stretches that are flanking a proline residue. This can be readily accomplished with the convenient HN detected approach. Moreover, the proposed experiments can be readily converted into 4D (HCA)CON(CA)NH and 4D (HCA)N(CA)CONH schemes, without any sensitivity loss, by adding another $^{15}N$ dimension as shown in Fig. 2b″ and d′. However, in case of two or more consecutive prolines, alternative approaches, for instance $^{1}H\alpha$ or $^{13}C'$ detected experiments, need to be applied (Mäntylahti et al. 2010, 2011; Bermel et al. 2009, 2012; Nováček et al. 2011; Permi and Hellman 2012).

Exchange broadening of amide proton resonances is much more severe problem in IDPs in comparison to globular due to lack of stable hydrogen bonding network. Therefore, any method where polarization transfer includes amide protons is susceptible to this effect. The pulse sequences proposed in this manuscript are no exceptions. However, as Hα protons are utilized for the initial polarization transfer, the (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H experiments are less sensitive to amide proton associated exchange broadening. In case of severe exchange broadening i.e. if measurements need to be carried out under alkali pH and/or at elevated temperature, either $^{1}H\alpha$ detected (Mäntylahti et al. 2010, 2011; Permi and Hellman 2012) or $^{13}C'$ detected approaches utilizing non-linear sampling, should be employed (Bermel et al. 2009, 2012; Nováček et al. 2011, 2012).

For those IDPs having poor solubility and/or poor relaxation properties i.e. being in so-called sensitivity limited regime, traditional HN detected approaches such as HNN or HN(C)N should offer adequate sensitivity (Panchal et al. 2001). This, however, necessitates at the same time small number of prolines and slow amide proton exchange rates.

## Application of the pulse sequences for assignment of A2A cytoplasmic tail

We applied the novel (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H experiments for the assignment of 134-residue construct of intrinsically disordered C-terminal cytoplasmic tail of adenosine receptor A2A-ct. Figure 3 shows two-dimensional $^{15}N$, $^{1}H^{N}$ correlation maps of the spectra acquired with the hNCOcanH experiment (a), the proposed (HCA)CON(CAN)H experiment, depicted in Fig. 3b, with two $T_{CAN}$ delay settings 56 ms (b) and 27 ms (c), and the novel (HCA)N(CA)CO(N)H experiment shown in Fig. 3d (d). Comparison of these spectra highlights their distinct features in terms of sensitivity, resolution and information content. The overall sensitivities of original hNCOcanH (Fig. 3a) and new (HCA)CON(CAN)H (Fig. 3b) experiments are comparable when both are optimized for the selection of desired $\omega_{C'(i-1)}$, $\omega_{N(i)}$, $\omega_{HN(i-1)}$ pathway. For (HCA)CON(CAN)H, when both $\omega_{C'(i-1)}$, $\omega_{N(i)}$, $\omega_{HN(i-1)}$ and $\omega_{C'(i-1)}$, $\omega_{N(i)}$, $\omega_{HN(i)}$ pathways are selected, the sensitivity is significantly compromised (Fig. 3c), yet cancellation of several cross peaks is also evident due to 180° phase difference between $\omega_{N(i)}$, $\omega_{HN(i-1)}$ and $\omega_{N(i)}$, $\omega_{HN(i)}$ correlations although the sign information can be useful during the assignment procedure. Moreover, the $\omega_{C'(i-1)}$, $\omega_{N(i)}$, $\omega_{HN(i)}$ correlations can be obtained using the conventional HNCO scheme with much higher sensitivity and we thus propose using the delay setting ($2T_{CAN} \sim 55$ ms) that maximizes transfer efficiency for the non-redundant $\omega_{C'(i-1)}$, $\omega_{N(i)}$, $\omega_{HN(i-1)}$ cross peaks (Fig. 3b). Figure 3d shows a complementary $^{15}N$-$^{1}H$ correlation map, obtained with the proposed (HCA)N(CA)CO(N)H experiment with the $2T_{CAN}$ delay set to 56 ms in order to minimize emergence of the undesired $\omega_{C'(i)}$, $\omega_{N(i+1)}$, $\omega_{HN(i+1)}$ (that is HNCO) pathway and maximize transfer efficiency of the sequential $\omega_{N(i)}$, $\omega_{C'(i)}$, $\omega_{HN(i+1)}$ correlations. The most salient feature of novel (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H pulse schemes is their ability to establish correlations to proline residues in C-terminal and N-terminal positions, respectively. The correlations to $^{15}N$ frequencies of prolines are encircled with a dashed ellipsoid in the panels b-d of Fig. 3. As can be appreciated, these connectivities are missing in the corresponding hNCOcanH spectrum (Fig. 3a). Although prolines can facilitate HN detection based assignment in globular proteins where their relatively low abundance provides useful check points during the assignment procedure, high occurrence of prolines in IDPs render traditional approaches ambiguous. Therefore ability to bridge proline flanking residues is very helpful for the backbone resonance assignment in IDPs.

Initially, an attempt for the backbone assignment of A2A-ct with the established HNCO and i(HCA)CO(CA)NH experiments was made, supplemented with data from CBCA(CO)NH and iHNCACB experiments (Rios et al. 1996; Tossavainen and Permi 2004). Previously we have successfully employed this approach for assignment of 110-residue Prostate-Associated Gene 5 Protein, PAGE5, where over 90 % of backbone $^{1}H^{N}$, $^{15}N$, and $^{13}C$ chemical shifts were assigned (Hellman et al. 2011). In the case of A2A-ct, we were only able to assign 75.1 % of

backbone $^{1}H^{N}$, $^{15}N$, $^{13}C'$, and $^{13}C\alpha$ resonances by utilizing $^{13}C'$ and $^{13}C\alpha/^{13}C\beta$ pathways for the assignment. Difficulties were especially associated with the linkage of numerous XPX stretches, where X stands for any amino acid residue except for proline. For the assignment of A2A-ct, new (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H experiments turned out to be invaluable. Their concomitant usage with the HNCO and i(HCA)CO(CA)NH experiments not only complemented the $^{13}C'$ chemical shift information based assignment with the corresponding $^{15}N$ frequencies but also provided links across the proline residues. Figure 4 illustrates representative example of the proposed approach for the assignment of Gly88-Leu89-Pro90-Asp91 stretch in A2A-ct. In overall, 95.6 % of $^{1}H^{N}$, $^{15}N$, $^{13}C'$ and $^{13}C\alpha$ resonances were assigned with the proposed NH-detected approach utilizing (HCA)CON(CAN)H, (HCA)N(CA)-CO(N)H, (HCA)CO(CA)NH and HNCO experiments.

## Conclusions

In this manuscript, we have introduced two new NH-detected pulse sequences, (HCA)CON(CAN)H and (HCA)N(CA)CO(N)H. Both experiments provide correlations to $^{15}N$ frequencies of proline residues either N- or C-terminal to a non-proline residue and hence can be employed to establish sequential links across the proline residues. This is of special interest in the case of IDPs where the proline is one of the most abundant residues and numerous 'proline stops' become hurdle rather than a remedy during the assignment procedure. We have also shown with the new assignment protocol that experiments can be utilized in an efficient way concomitantly with the conventional HNCO as well as intraresidual i(HCA)CO(CA)NH schemes. We reckon that this approach offers convenient approach for the assignment of IDPs using solely NH-detected experiments, which is of general interest especially for those who do not have $^{13}C$ detection optimized probes.

**Acknowledgments** This work was financially supported by the Grants 259447 (to P. P.) and 132138 (to V. P. J.) from the Academy of Finland, and by the IRG 249081 from FP7 Marie Curie European Reintegration Grant and Biocenter Oulu (to V. P. J.). Biocenter Finland and Biocentrum Helsinki are acknowledged for the infrastructure support.

## References

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

Bermel W, Bertini I, Felli IC, Piccioli M, Pierattelli R (2006) $^{13}C$-detected protonless NMR spectroscopy of proteins in solution. Prog Nucl Magn Reson Spectr 48:25–45

Bermel W, Bertini I, Csizmok V, Felli IC, Pierattelli R, Tompa P (2009) H-start for exclusively heteronuclear NMR spectroscopy: the case of intrinsically disordered protein. J Magn Reson 198:275–281

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

Bottomley MJ, Macias MJ, Liu Z, Sattler M (1999) A novel NMR experiment for the sequential assignment of proline residues and proline stretches in $^{13}C/^{15}N$-labeled proteins. J Biomol NMR 13:381–385

Brutscher B (2002) Intraresidue HNCA and COHNCA experiments for protein backbone resonance assignment. J Magn Reson 156:155–159

Delaglio F, Torchia DA, Bax A (1991) Measurement of $^{15}N$-$^{13}C$ J couplings in staphylococcal nuclease. J Biomol NMR 1:439–446

Dyson HJ, Wright PE (2001) Nuclear magnetic resonance methods for elucidation of structure and dynamics in disordered states. Methods Enzymol 339:258–270

Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol 6:197–208

Fiorito F, Hiller S, Wider G, Wüthrich K (2006) Automated resonance assignment of proteins: 6D APSY-NMR. J Biomol NMR 35:27–37

Grzesiek S, Bax A (1993) Amino acid type determination in the sequential assignment procedure of uniformly 13C/15N-enriched proteins. J Biomol NMR 3:185–204

Grzesiek S, Bax A, Hu J-S, Kaufman J, Palmer I, Stahl SJ, Tjandra N, Wingfield PT (1997) Refined solution structure and backbone dynamics of HIV-1 Nef. Protein Sci 6:1248–1263

Hellman M, Tossavainen H, Rappu P, Heino J, Permi P (2011) Characterization of intrinsically disordered prostate associated gene (PAGE5) at single residue resolution by NMR spectroscopy. PLoS One 6:e26633

Hu K, Vögeli B, Clore GM (2007) Spin-state selective carbon-detected HNCO with TROSY optimization in all dimensions and double echo-antiecho sensitivity enhancement in both indirect dimensions. J Am Chem Soc 129:5484–5491

Kanelis V, Donaldson L, Muhandiram DR, Rotin D, Forman-Kay JD, Kay LE (2000) Sequential assignment of proline-rich regions in proteins: application to modular binding domain complexes. J Biomol NMR 16:253–259

Kay LE, Keifer P, Saarinen T (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. J Am Chem Soc 114: 10663–10665

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

Kosol S, Contreras-Martos S, Cedeno C, Tompa P (2013) Structural characterization of intrinsically disordered proteins by NMR spectroscopy. Molecules 18:10802–10828

Kumar D, Hosur RV (2011) hNCOcanH pulse sequence and a robust protocol for rapid and unambiguous assignment of backbone ($^{1}HN$, $^{15}N$ and $^{13}C'$) resonances in $^{15}N/^{13}C$-labeled proteins. Magn Reson Chem 49:575–583

Kupće E, Wagner G (1995) Wideband homonuclear decoupling in protein spectra. J Magn Reson 109A:329–333

Logan TM, Olejniczak ET, Xu RX, Fesik SW (1993) A general method for assigning NMR spectra of denatured proteins using 3D HC(CO)NH-TOCSY triple resonance experiments. J Biomol NMR 3:225–231

Löhr F, Pfeiffer S, Lin Y-J, Hartleib J, Klimmek O, Rüterjans H (2000) HNCAN pulse sequences for sequential backbone resonance assignment across proline residues in perdeuterated proteins. J Biomol NMR 18:337–346

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

Mäntylahti S, Tossavainen H, Hellman M, Permi P (2009) An intraresidual i(HCA)CO(CA)NH experiment for the assignment of main-chain resonances in $^{15}$N, $^{13}$C labeled proteins. J Biomol NMR 45:301–310

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

Marion D, Ikura M, Tschudin R, Bax A (1989) Rapid recording of 2D NMR-spectra without phase cycling—application to the study of hydrogen-exchange in proteins. J Magn Reson 85:393–399

Morris GA, Freeman R (1979) Enhancement of nuclear magnetic resonance signals by polarization transfer. J Am Chem Soc 101:760–762

Muhandiram DR, Kay LE (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. J Magn Reson 103:203–216

Nietlispach D (2002) A novel approach for the sequential backbone assignment of larger proteins: selective intra-HNCA and DQ-HNCA. J Am Chem Soc 124:11199–11207

Nováček J, Zawadzka-Kazimierczuk A, Papoušková V, Zídek 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, Haba NY, Chill JH, Zídek L, Sklenář V (2012) 4D non-uniformly sampled HCBCACON and $^{1}$J(NCα)-selective HCB-CANCO experiments for the sequential assignment and chemical shift analysis of intrinsically disordered proteins. J Biomol NMR 53:139–148

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

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

Permi P (2002) Intraresidual HNCA: an experiment for correlating only intraresidual backbone resonances. J Biomol NMR 23:201–209

Permi P, Annila A (2004) Coherence transfer in proteins. Prog Nucl Magn Reson Spectr 44:97–137

Permi P, Hellman M (2012) Alpha proton detection based backbone assignment of intrinsically disordered proteins. Methods Mol Biol 895:211–226

Rios CB, Feng W, Tashiro M, Shang Z, Montelione GT (1996) Phase labeling of C–H and C–C spin-system topologies: application in constant-time PFG-CBCA(CO)NH experiments for discriminating amino acid spin-system types. J Biomol NMR 8:345–350

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

Schleucher J, Schwendinger MG, Sattler M, Schmidt P, Glaser SJ, Sørensen OW, Griesinger C (1994) A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed-field gradients. J Biomol NMR 4:301–306

Shaka AJ (1985) Composite pulses for ultra-broadband spin inversion. Chem Phys Lett 120:201–205

Shaka AJ, Keeler J, Frenkiel T, Freeman R (1983) An improved sequence for broad-band decoupling—Waltz-16. J Magn Reson 52:335–338

Tossavainen H, Permi P (2004) Optimized pathway selection in intraresidual triple-resonance experiments. J Magn Reson 170:244–251

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

Yao J, Dyson HJ, Wright PE (1997) Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins. FEBS Lett 419:285–289