J Biomol NMR (2013) 55:311–321
DOI 10.1007/s10858-013-9715-0

ARTICLE

BEST-TROSY experiments for time-efficient sequential resonance assignment of large disordered proteins

Zsofia Solyom · Melanie Schwarten · Leonhard Geist · Robert Konrat · Dieter Willbold · Bernhard Brutscher

Received: 17 January 2013 / Accepted: 15 February 2013 / Published online: 24 February 2013

Abstract The characterization of the conformational properties of intrinsically disordered proteins (IDPs), and their interaction modes with physiological partners has recently become a major research topic for understanding biological function on the molecular level. Although multidimensional NMR spectroscopy is the technique of choice for the study of IDPs at atomic resolution, the intrinsically low resolution, and the large peak intensity variations often observed in NMR spectra of IDPs call for resolution- and sensitivity-optimized pulse schemes. We present here a set of amide proton-detected 3D BEST-TROSY correlation experiments that yield the required sensitivity and spectral resolution for time-efficient sequential resonance assignment of large IDPs. In addition, we introduce two proline-edited 2D experiments that allow unambiguous identification of residues adjacent to proline that is one of the most abundant amino acids in IDPs. The performance of these experiments, and the advantages of BEST-TROSY pulse schemes are discussed and illustrated for two IDPs of similar length (~270 residues) but with different conformational sampling properties.

Keywords BEST · TROSY · IDP · Viral protein · Amino-acid-type editing · Longitudinal-relaxation enhancement

Electronic supplementary material The online version of this article (doi:10.1007/s10858-013-9715-0) contains supplementary material, which is available to authorized users.

Z. Solyom · M. Schwarten · B. Brutscher (✉)
Institut de Biologie Structurale, Université Grenoble 1, 41 Rue Jules Horowitz, 38027 Grenoble Cedex 1, France
e-mail: bernhard.brutscher@ibs.fr

Z. Solyom · M. Schwarten · D. Willbold · B. Brutscher
Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Grenoble, France

Z. Solyom · M. Schwarten · B. Brutscher
Centre National de Recherche Scientifique (CNRS), Grenoble, France

L. Geist · R. Konrat
Department of Computational and Structural Biology, Max F. Perutz Laboratories, Campus Vienna Biocenter 5, A-1030 Vienna, Austria

D. Willbold
Institute of Complex Systems (ICS-6) Structural Biochemistry, Forschungszentrum Jülich, 52425 Jülich, Germany

D. Willbold
Institut für Physikalische Biologie, Heinrich-Heine-Universität, 40225 Düsseldorf, Germany

Introduction

The structure–function paradigm stating that a well-defined structure is required for a protein's function has been challenged by the discovery of a large number of highly flexible proteins or protein segments that exist as ensembles of partly collapsed or extended structural conformers in the cell and that are functional as such (Dunker et al. 2001; Tompa 2002; Wright and Dyson 1999; Uversky and Dunker 2010). Intrinsically disordered proteins (IDPs) or protein regions (IDRs) have been shown to play important roles in regulatory and signaling processes where the structural flexibility allows the protein to adapt to and interact with a large number of distinct molecular partners (Tompa 2012). Similarly, structural disorder is also abundant in viral proteins (Davey et al. 2011; Xue et al. 2010). Viruses are often characterized by a small genome, only coding for a few proteins. High mutation rates in these

genomes allow to adapt to changing environments and to escape the defense mechanisms of the host cell. Again, structural flexibility presents a functional advantage in terms of binding promiscuity, as well as a high tolerance to mutations.

During recent years, NMR spectroscopy has become the technique of choice to obtain atomic-resolution information for IDPs, and to extract useful information on the structural ensemble that the IDP forms in solution. In particular, NMR allows the identification of peptide regions with increased propensity to form α-helical or extended (β-strand) structures that often play a role in molecular recognition events, or the characterization of transient long-range interactions. Furthermore, NMR is a powerful technique to characterize binding events in terms of interaction surfaces, and to study eventual conformational transitions of the IDP upon binding to its partner(s). The rapid interconversion between the large number of conformations sampled by the ensemble typically results in a single set of sharp NMR signals. However, as a consequence of the lack of a stable structure, NMR spectra of IDPs are characterized by low chemical shift dispersion as compared to well-structured globular proteins, which makes NMR studies of large IDPs a challenging task. In addition, fast solvent exchange of the solvent-exposed labile protons with water protons may result in extensive line broadening of amide proton resonances. Therefore, amide ^{1}H-detected NMR experiments often need to be performed at lower than ambient sample temperature to slow down hydrogen exchange. Alternatively, H^{α}-detected (Mantylahti et al. 2010) or ^{13}C-detected experiments (Felli and Brutscher 2009; Csizmok et al. 2008; Bermel et al. 2012) can be used to circumvent the problem of unobservable amide proton resonances. However, H^{α}-detected experiments require protein samples in deuterated solvents, while ^{13}C-detected experiments suffer from reduced sensitivity due to the lower gyromagnetic ratio of ^{13}C with respect to ^{1}H. Other concerns for NMR studies of IDPs are often limited sample stability, low sample concentrations to avoid protein aggregation, and substantial peak intensity heterogeneities in the NMR spectra. Therefore, sensitive NMR pulse schemes are required to detect correlation peaks also for the sites with the lowest signal intensity; high dimensional (≥3D) NMR experiments with long acquisition times in all dimensions are needed in order to resolve overlapping correlation peaks; and last but not least fast acquisition techniques are mandatory to enable multidimensional data acquisition in a reasonable amount of time.

Herein, we present BEST-TROSY pulse sequences that yield the required sensitivity and spectral resolution for time-efficient sequential resonance assignment of large IDPs from a set of typically four 3D correlation spectra. These experiments are particularly attractive for NMR studies performed at high magnetic field strengths. Compared to other ^{1}H-detected pulse schemes, BEST-TROSY experiments provide significant advantages in terms of experimental sensitivity and spectral resolution as demonstrated for two IDPs of about 270 residues in length, the C-terminal part of the NS5A protein from hepatitis C virus (268 residues including tag) that acts as a multifunctional regulator of cellular pathways, and the chicken BASP1 protein (270 residues including tag) involved in transcription regulation.

Materials and methods

NMR measurements were performed on an Agilent VNMRS 800 MHz spectrometer at 278 K equipped with a cryogenically cooled triple-resonance (HCN) probe, and pulsed z-field gradients.

The presented pulse sequences were tested on two NMR samples: the first contains 0.12 mM [*U*–^{13}C, *U*–^{15}N] enriched NS5A (Non-Structural Protein 5A, residues 191–447 plus an 11-residues tag remaining after cleavage of a His tag) from hepatitis C virus in 50 mM potassium phosphate buffer (pH 6.5), 20 mM NaCl, 2 mM β-mercaptoethanol and 5 % (v/v) D_2O, the second sample contains 0.6 mM [*U*–^{13}C, *U*–^{15}N] chicken BASP1 (Brain Acid-Soluble Protein 1, 244 residues) in 20 mM citrate buffer (pH 2.0), 0.2 % NaN_3 and 10 % (v/v) D_2O. The BASP1 construct includes a 26-residue cleavable His tag.

To characterize the conformational dynamics of the two proteins, ^{15}N relaxation experiments (T_1, T_2, HETNOE) were performed using standard pulse sequences (Farrow et al. 1994). The relaxation curves were fitted onto the extracted peak intensities from the subspectra measured with 10 relaxation delays, ranging from 0.01 to 1.6 s in the case of T_1 and from 0.01 to 0.25 s for T_2.

Proton T_1 relaxation time constants were measured by inversion recovery experiments, with amide selective inversion, water-flip-back (WFB) and non-selective inversion using the pulse sequence elements of Figure S1. The spectra were recorded with 10 relaxation delays ranging from 0 to 1.4 s in the non-selective and WFB case and 0–0.9 s in the amide selective case. All relaxation curves were fitted to a mono-exponential function using a Python program.

3D HNcoCACB, iHNCACB, hNcocaNH, and hnCOcaNH correlation spectra (Fig. 3) were recorded for NMR assignment of NS5A with the recycle delay set to $T_{rec} = 0.15$ s, the acquisition time to 70 ms with 4 scans per (t_1, t_2) increment. In the ^{15}N dimension, 150 complex points were acquired for a spectral width of 2,000 Hz resulting in a maximal evolution time of t_2^{max} = 75 ms. The number of recorded data points (spectral widths) for the other indirect dimensions are: 110 (10,000 Hz) for CACB, 100 (2,000 Hz) for N, and 90 (2,000 Hz) for CO, resulting in total data acquisition times of about 20 h per experiment.

Results and discussion

Amide ^{1}H-detected BEST-TROSY pulse schemes provide increased sensitivity and spectral resolution, while allowing for short overall experimental times, by exploiting three complementary effects: (i) The selective manipulation of amide ^{1}H by means of band-selective radio-frequency (RF) pulses enhances longitudinal ^{1}H relaxation (Pervushin et al. 2002; Schanda 2009), and thus provides higher sensitivity due to an increased ^{1}H steady-state polarization (BEST principle) (Schanda et al. 2006). (ii) Single-transition spin-state selection results in favorable relaxation properties for ^{1}H and ^{15}N due to CSA-dipolar cross-correlation (TROSY principle) (Pervushin et al. 1997). (iii) BEST-TROSY contains a build-in module for the conversion of (undetected) ^{1}H polarization that builds up during the pulse sequence due to spin relaxation, into enhanced ^{15}N polarization that contributes to the detected signal in the subsequent scan. The efficiency of this ^{15}N polarization enhancement mechanism depends on the longitudinal relaxation times (T_1) of both ^{1}H and ^{15}N, with the highest gains obtained for short ^{1}H T_1 and long ^{15}N T_1 (Favier and Brutscher 2011).

3D BEST-TROSY HNC correlation experiments

So far, BEST-TROSY (BT) techniques have been successfully applied to the study of globular proteins (Favier and Brutscher 2011) and nucleic acids (Farjon et al. 2009), but not yet to long, highly disordered polypeptides which is the subject of the present report. In order to illustrate the performance of BEST-TROSY for large IDPs we have chosen two proteins (or protein fragments), NS5A and BASP1 that are similar in size ($\sim$270 residues), but differ in their amino-acid-type composition and their structural compactness, resulting in different NMR properties (chemical shift dispersion, relaxation times, …) as shown in Figs. 1 and 5c. In particular, the structural dynamics in the BASP1 protein are quite uniform along the polypeptide chain, while NS5A shows a more heterogeneous behavior with several peptide regions that are characterized by increased local tumbling correlation times (τ_c) and higher local order (larger HET-NOE values). The 2D ^{1}H–^{15}N BEST-TROSY correlation spectra shown in Fig. 1 for NS5A (left) and BASP1 (right) show a large number of resolved resonances (especially for NS5A) despite the intrinsically low frequency dispersion. This high resolution can be maintained in 3D H–N–C correlation spectra, required for NMR assignment, by using semi-constant-time (semi-CT) editing (Grzesiek and Bax 1993) in the ^{15}N dimension allowing for long maximal evolution times. A series of such BT-optimized triple-resonance pulse sequences is displayed in Fig. 2. We used these experiments for the assignment of NS5A, the BASP1 protein had been assigned previously by other means. These experiments correlate the backbone amide ^{1}H and ^{15}N either with the ^{13}CA, ^{13}CB, ^{13}CO, or ^{15}N of the same (intra-residue correlation) or a sequentially adjacent residue (sequential correlation). A 2-step phase cycle is sufficient to obtain artifact-free correlation spectra. As discussed in more detail below, BT-optimization results in accelerated longitudinal relaxation of amide protons, yielding significantly reduced inter-scan delays required for optimal sensitivity, thus enhancing the overall sensitivity and reducing the total experimental time. As a consequence, even for large IDPs such as NS5A or BASP1 high-resolution 3D data sets can be recorded in only a few hours using a regular uniform sampling grid and Fourier-transform (FT) processing. Example strip extracted from NS5A spectra are shown in Fig. 3.

Particularly useful for sequential resonance assignment of large IDPs are the 3D hNcocaNH (Fig. 3a) and 3D hnCOcaNH (Fig. 3b) spectra that allow building sequential connectivities on the basis of backbone ^{15}N and ^{13}CO chemical shifts, characterized by a higher chemical shift dispersion than ^{13}CA or ^{13}CB in highly flexible proteins (Panchal et al. 2001, Kumar et al. 2010). Note that in these experiments the signal originating from ^{15}N polarization only enhances the "out-and-back" coherence transfer pathways resulting in the diagonal peaks.

BEST versus conventional techniques

In the following we will discuss the advantages of BEST-TROSY techniques (Fig. 2) with respect to alternative conventional pulse schemes. Longitudinal ^{1}H relaxation

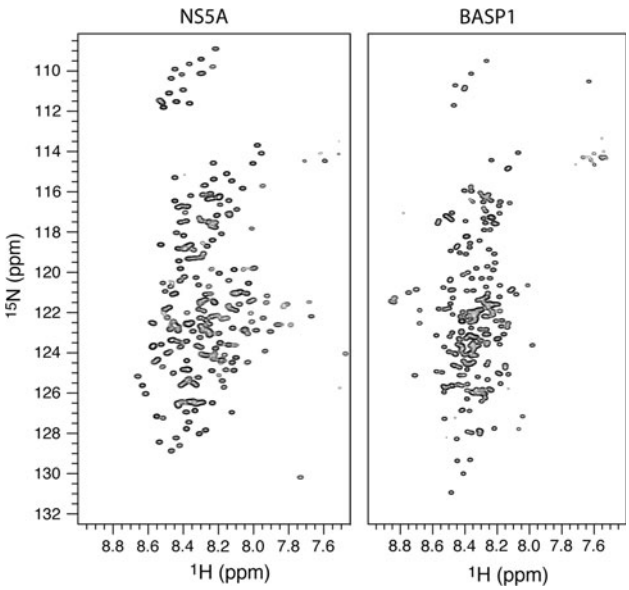
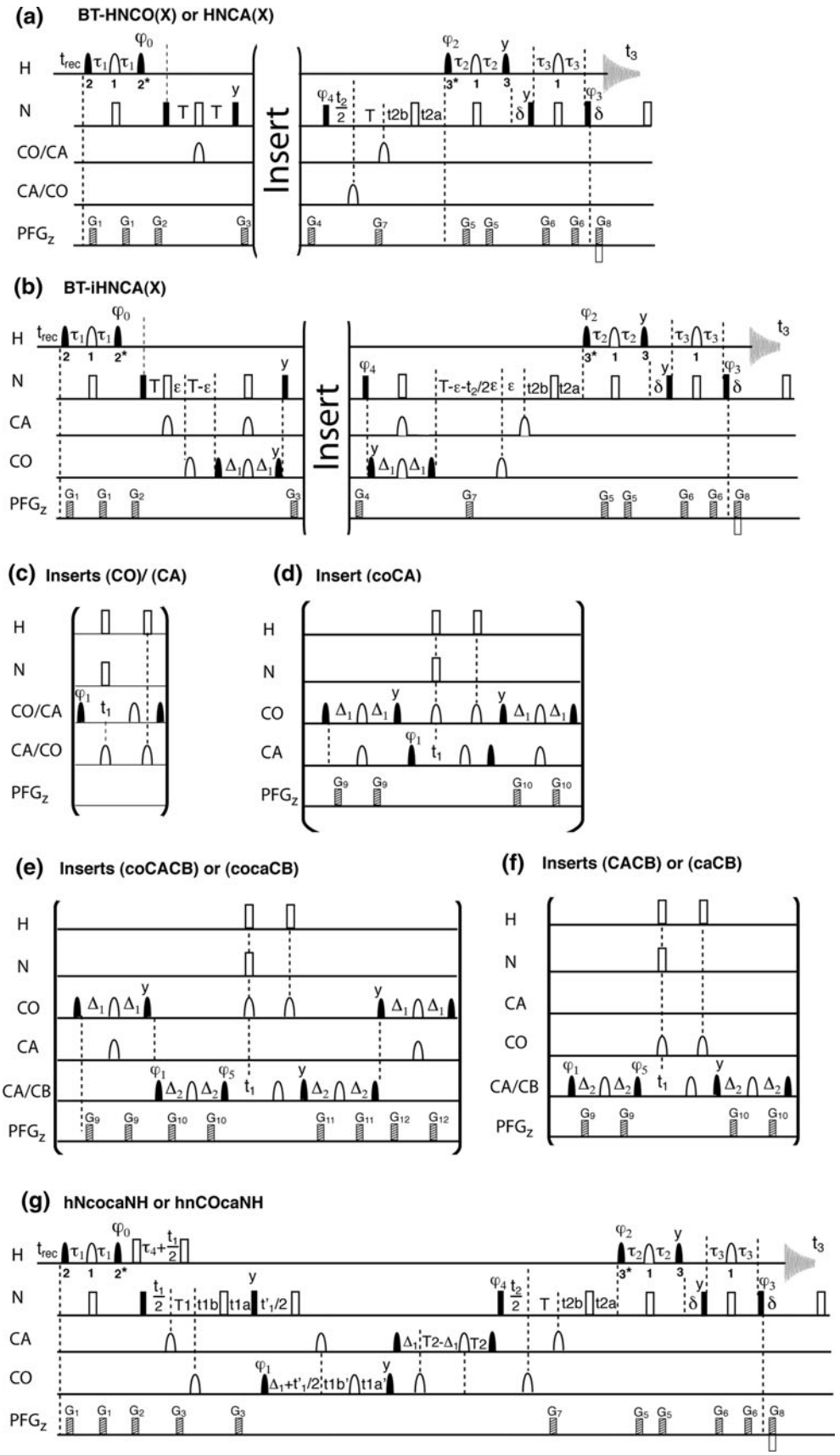


Fig. 1 2D ^{1}H–^{15}N BEST-TROSY correlation spectra of NS5A(191–447) (*left*) and BASP1 (*right*)


(a) BT-HNCO(X) or HNCA(X)
(b) BT-iHNCA(X)
(c) Inserts (CO)/ (CA)
(d) Insert (coCA)
(e) Inserts (coCACB) or (cocaCB)
(f) Inserts (CACB) or (caCB)
(g) hNcocaNH or hnCOcaNH
Insert
H
N
CO/CA
CA/CO
CA
CO
CA/CB
PFGz

◀**Fig. 2** BEST-TROSY sequences: (a+c) HNCO or HNCA, (a+d) HNcoCA, (a+e) HNcoCACB or HNcocaCB, (a+f) HNCACB or HNcaCB, (b+c) iHNCA, (b+f) iHNCACB, (g) hNcocaNH and hnCOcaNH. Filled and open pulse symbols indicate 90° and 180° rf pulses. Unless indicated, all pulses are applied with phase x. All selective ^{1}H pulses are centered at 8.5 ppm, covering a bandwidth of 3.5 ppm, with the following shapes: [1] REBURP (Geen and Freeman 1991), [2] PC9 (Kupce and Freeman 1994) and [3] E-BURP2 (Geen and Freeman 1991). A star indicates a flip-back pulse obtained by time inversion of the excitation pulse shape. Open squares on ^{1}H indicate BIP-720-50-20 broadband inversion pulses (Smith et al. 2001). CO pulses have the shape of the center lobe of a sinx/x function, whereas CA and CA/CB pulses are applied with a rectangular shape and zero excitation at the CO frequency. The transfer delays common to all BEST-TROSY sequences are adjusted to $\tau_1 = 1\backslash(4J_{NH}) - 0.5\delta_1 - 0.5\delta_2$, $\tau_2 = 1\backslash(4J_{NH}) - 0.5\delta_1 - k\delta_3$, $\tau_3 = 1\backslash(4J_{NH})$ with $1\backslash(4J_{NH}) \approx 2.7$ ms, $\tau_4 = 1\backslash(2J_{NH})$, t2a = T-t$_2$/2 + t2b, and t2b = 0. The delays δ_1, δ_2, and δ_3 correspond to the lengths of the REBURP, PC9, and E-BURP2 pulses, respectively, and the parameter $k \approx 0.7$ can be fine-tuned to equilibrate the transfer amplitudes of the different coherence transfer pathways (Clean-TROSY) for optimal suppression of the unwanted quadruplet components in the spectrum (Schulte-Herbruggen and Sorensen 2000). For semi-CT ^{15}N editing, the delay t2b is incremented together with t$_2$ using the following time increment: $\Delta t2b = (t_2^{max}/2\text{-}T)/N_2$ with N_2 the number of total increments in the t$_2$ dimension. The additional delays are set to **a** T = 15 ms; **b** T = 17.5 ms; ε = 0.5 ms; $\Delta_1 = 4.5$ ms; **d** $\Delta_1 = 4.5$ ms; **e** and **f** $\Delta_1 = 4.5$ ms, $\Delta_2 = 3.5$ ms (for HNcoCACB) and $\Delta_2 = 7$ ms (for HNcocaCB); **g** T1 = 15 ms; T2 = 13.5 ms; $\Delta_1 = 4.5$ ms; t1a = T-t$_1$/2 + t1b; t1b = 0; t1a' = Δ_1-t'$_1$/2 + t1b'; t1b' = 0. For semi-CT ^{15}N (^{13}CO) editing, the delay t1b (t1b') is incremented together with t$_1$ (t$_1$′) using the following time increment: $\Delta t1b = (t_1^{max}/2\text{-}T)/N_1$ ($\Delta t1b = (t_1^{max}/2\text{-}\Delta_1)/N_1$) with N_1 the number of total increments in the t$_1$ dimension. Pulsed field gradients G_1–G_8 are applied along the z-axis (PFG_z) with durations of 200 μs to 2 ms and field strengths ranging from 5 to 40 G/cm. The 2-step phase cycling is: $\varphi_1 = -x, x$; $\varphi_2 = -y$ $\varphi_3 = -x$; $\varphi_4 = x, x$; $\varphi_5 = y$; $\varphi_{rec} = -x, x$. The relative durations of G_7 and G_8 are given by the gyromagnetic ratios $G_7/G_8 = \gamma_H/\gamma_N$. Quadrature detection in t$_1$ is obtained by time-proportional phase incrementation of φ_1 (and φ_5) according to TPPI-States. For quadrature detection in t$_2$, echo-antiecho data are recorded by inverting the sign of gradient G_8 together with phases φ_2 and φ_3. All pulse sequences (in Agilent pulse program language) are available from the authors upon request

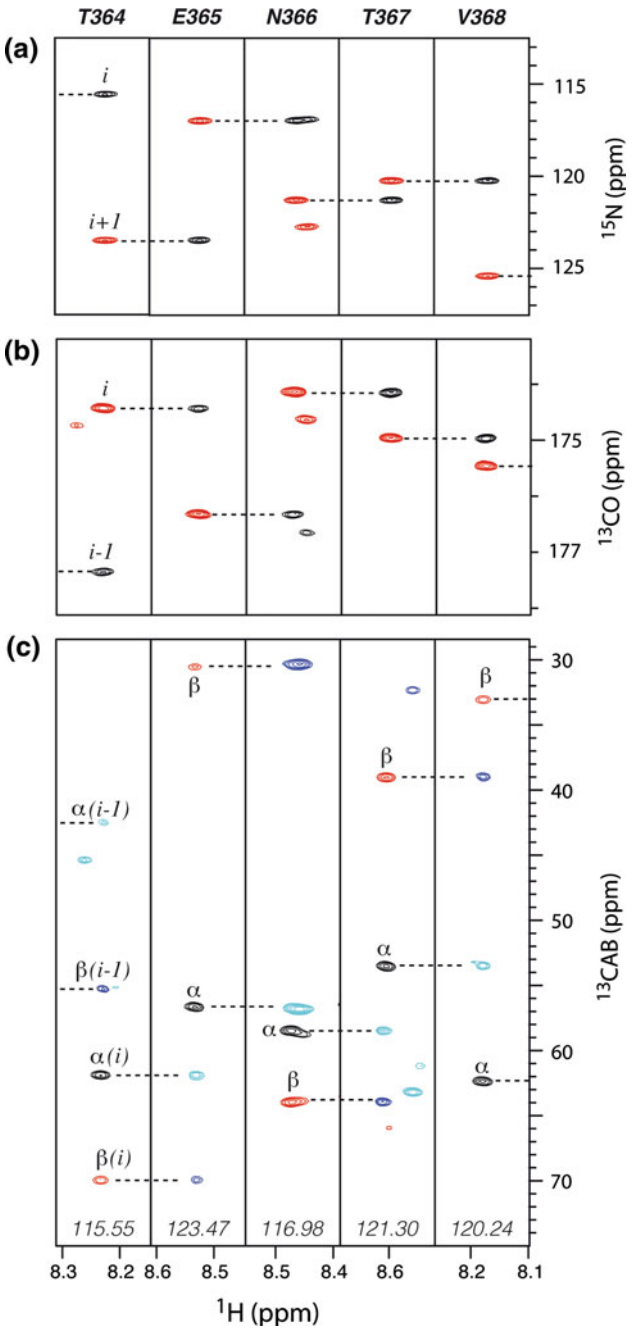


Fig. 3 Sequential resonance assignment of the NS5A segment 364-368 based on **a** 3D hNcocaNH, **b** 3D hnCOcaNH, and **c** 3D HNcoCACB and iHNCACB spectra. In **a** and **b** 2 correlation *peaks* of opposite sign (except if Gly residues are involved) are detected per amide group, one corresponding to the intra-residue, one to a sequential correlation. In **c** two spectra, each containing only intra-residue or sequential correlations, are superposed on the same graph

enhancement as achieved by BEST-type sequences exploits the fact that unperturbed proton spins that are coupled to the excited amide protons either by dipolar interactions with other aliphatic or aromatic protons, or chemical exchange with water protons take up some of the energy put into the spin system, thus accelerating spin–lattice (longitudinal) relaxation. While NS5A was studied at pH 6.5, for the BASP1 sample the pH was adjusted to two in order to investigate the relative contributions of ^{1}H–^{1}H dipolar interactions and amide-solvent ^{1}H exchange (rendered inefficient at low pH) on longitudinal relaxation enhancement. In order to quantify these effects for the two IDPs studied here, we have performed inversion-recovery experiments using three different inversion pulse schemes to simulate the proton relaxation behavior encountered in different types of experiments: (A) selective inversion of amide protons (BEST-type experiments), (B) inversion of all but water protons (water-flip-back (WFB) experiments), (C) inversion of the entire proton spectrum (conventional non-selective experiments). The inversion sequence is followed by a variable relaxation delay and a ^{1}H-^{15}N BEST-HSQC sequence for readout of the ^{1}H polarization

(see Figure S1). The measured ^{1}H T_1 values are displayed in Fig. 4a as a function of the protein sequences of NS5A and BASP1, respectively. The average ^{1}H T_1 values measured for the two proteins are 0.92 ± 0.11 s (NS5A) and 0.91 ± 0.08 s (BASP1) for the non-selective inversion (black bars), 0.70 ± 0.15 s (NS5A) and 0.89 ± 0.07 s (BASP1) for the WFB situation (green bars), and 0.21 ± 0.06 s (NS5A) and 0.31 ± 0.06 s (BASP1) for the amide-selective (BEST-type) experiment (red bars). The measured relaxation times are quite uniform along the peptide sequence for the different experimental scenarios, with local differences mainly reflecting the variation in local structure (water accessibility and effective tumbling correlation time). The first interesting conclusion from these measurements is that for both proteins the non-selective ^{1}H T_1 of about 900 ms is reduced to 200–300 ms in BEST-type experiments. A second observation is that for both proteins the major relaxation enhancement mechanisms are dipolar interactions of the amide proton with surrounding protons, rather than chemical exchange with water protons. This conclusion, however, may be different for samples studied at higher pH and higher temperature. Finally, a third conclusion from these data is that longitudinal relaxation enhancement is more efficient for NS5A than for BASP1. This is mainly explained by the more compact conformers in the structural ensemble of NS5A resulting in more efficient proton–proton spin diffusion within the molecule, and some additional contribution from hydrogen-exchange mediated polarization transfer in NS5A.

The average ^{1}H T_1 values can be used to compute the expected signal-to-noise ratio per unit time (SNR_{UT}) as a function of the recycle delay (T_{rec}) between scans according to the analytical relation.

$$SNR_{UT} = (1 - \exp(-T_{rec}/T_1))/\sqrt{T_{scan}}. \quad (1)$$

Note that T_{rec} includes the data acquisition time, and that $T_{scan} = T_{rec} + T_{seq}$ with the sequence length arbitrarily set to $T_{seq} = 100$ ms. The resulting sensitivity curves for NS5A and BASP1, shown in Fig. 4b, indicate sensitivity gains for BEST-type optimized sequences of about a factor of 2 with the maximal sensitivity obtained for a recycle delay $T_{rec} \approx 1.25\ T_1$.

BEST-TROSY versus BEST-HSQC

Having shown in the previous paragraph that BEST sequences provide a sensitivity (and time) advantage with respect to non-selective or WFB ^{1}H-^{15}N correlation experiments, we now want to address the question how

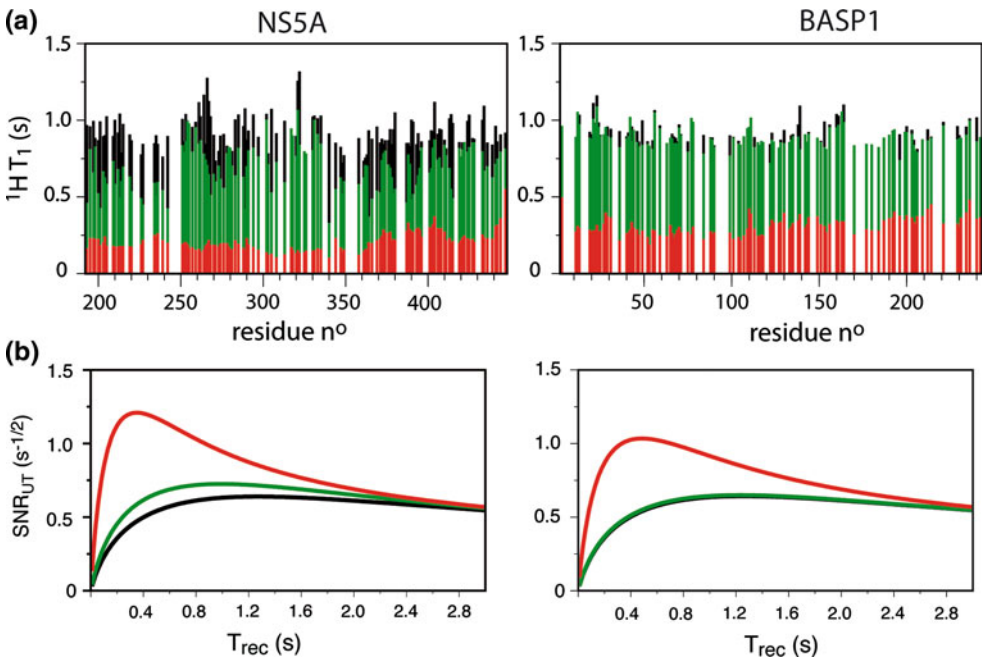


Fig. 4 Backbone amide ^{1}H T_1 relaxation time constants measured for NS5A (*left*) and BASP1 (*right*) are shown in **a** as function of the peptide sequence. The residue numbering corresponds to the real protein sequence, and the values measured for the tag-residues are not shown. Relaxation rates were measured by inversion-recovery using different inversion sequences (see Figure S1): non-selective ^{1}H inversion (*black bars*), water-flip-back ^{1}H inversion (*green bars*), and selective amide ^{1}H inversion (*red bars*). In **b** the expected average signal-to noise-ratio (SNR) per unit time (UT) has been computed from equation [1] as a function of the inter-scan delay (T_{rec}) for the three experimental scenarios: non-selective (*black*), water-flip-back (*green*), and amide-selective (*red*) BEST-type pulse sequences. The residue numbering corresponds to the real protein sequence; values measured for the tag-residues are not shown

BEST-TROSY (BT) compares to BEST-HSQC (BH) sequences. In order to do so, we have recorded 2D HNco planes of both proteins using either BEST-TROSY (Fig. 2a, c) or BEST-HSQC (Lescop et al. 2007) versions of the experiment. Except for the pulse sequence, all other acquisition and processing parameters were chosen identical. As can be appreciated from the spectra shown in Fig. 5a, the BEST-TROSY-optimized sequence yields higher spectral resolution due to the narrower single-transition line widths, and the signal contributed from ^{15}N polarization. Actually, as the ^{15}N labeling period t_2 becomes longer, more ^{1}H polarization builds up as a consequence of longitudinal spin relaxation during t_2, that is converted into enhanced ^{15}N polarization by the final coherence transfer step (Favier and Brutscher, 2011). Furthermore, also the peak intensities are increased in the BEST-TROSY spectra, on average by 80 % for NS5A and by 20 % for BASP1 despite an intrinsic factor-of-2 signal loss due to the single-transition selection. This signal enhancement reaches a factor of 3 (NS5A) and 1.6 (BASP1) for individual residues (Fig. 5b). Note however, that despite this overall increase in signal intensity, some residues in BASP1 have a reduced intensity in BEST-TROSY with respect to BEST-HSQC spectra. The differences observed for the two proteins are most likely ascribed to the shorter ^{1}H T_1 and longer ^{15}N T_1 values in NS5A, resulting in increased ^{15}N polarization enhancement, as well as more efficient CSA-dipolar cross-correlation (TROSY line narrowing effect) due to a more compact average structure in the N-terminal half of the NS5A protein (Feuerstein et al. 2012b). The highest intensity gains observed correlate well with regions of increased structural rigidity, as revealed by the larger τ_c and HETNOE values (Fig. 5c). Finally, BEST-TROSY also leads to a more uniform intensity distribution in the NMR spectra of these IDPs, as the weakest NMR signals become enhanced most (Fig. 5d).

Proline-selective ^{1}H–^{15}N BEST-TROSY correlation experiments

A specific feature of IDPs is that they often have repetitive sequence parts, and that they are rich in proline (Pro) residues (Tompa 2002). Because Pro residues are break points in the sequential assignment walk, this presents an additional complication for sequence-specific resonance assignment of IDPs. Therefore, NMR tools that allow identifying Pro-neighboring residues are of particular interest for IDP resonance assignment. Such information can be either obtained by 3D HNCAN-type pulse sequences (Lohr et al. 2000) that allow coherence transfer across Pro residues, or by Pro-edited 2D ^{1}H–^{15}N correlation experiments (Schubert et al. 2000) that exploit the particular spin-coupling topology of the Pro-side-chain to selectively detect coherence-transfer pathways from residues preceding or following Pro.

Here we present new pulse sequences for recording Pro-edited 2D ^{1}H–^{15}N correlation spectra that are of the BEST-TROSY type and therefore benefit from all the advantages discussed in the previous paragraphs. The Pro-HNcocan experiment selectively detects ^{1}H–^{15}N correlations of residues following a proline in the peptide sequence, while the Pro-iHNcan is selective to residues preceding a proline. The pulse sequences are obtained by inserting the pulse sequence element depicted in Fig. 6a into either the BT-HNCO(X) building block of Fig. 2a (Pro-HNcocan) or the BT-iHN-CA(X) building block of Fig. 2b (Pro-iHNcan)—see also Figure S2 of the Supporting Information. The experiments exploit the fact that the ^{15}N chemical shift range of Pro in IDPs is well separated from all other amino-acid types (Fig. 6b). This makes it possible to selectively dephase C^{α} coherence during a constant-time delay T with respect to either the one-bond (Pro-HNcocan) or two-bond (Pro-iHNcan) J_{CAN} coupling, similar to a technique presented recently for the selective detection of phsophoserine and phosphothreonine residues (McIntosh et al. 2009). Two experiments need to be performed either with (*transfer experiment*) or without (*reference experiment*) application of the Pro-selective ^{15}N shaped pulse. In the transfer experiment the ^{13}C–^{15}N coupling evolution is active during the constant time delay T while in the reference experiment, coupling evolution is refocused. Subtraction of the 2 data sets (*difference spectrum*) then yields the desired Pro-selective ^{1}H–^{15}N correlation maps. The relative sensitivity of these experiments with respect to either HNcoca or iHNca experiments can be estimated by computing the transfer amplitude.

$$\mathrm{TA}_{pro} = 0.5(1 - \cos(\pi J_{NCA}\mathrm{T}))\cos(\pi J_{CC}\mathrm{T})\exp(-T/T_2), \quad (2)$$

with T_2 the transverse relaxation time constant of the C^{α} coherence present during the time T, $J_{cc} = \approx 35$ Hz the one-bond C^{α}–C^{β} coupling constant, and the factor 0.5 accounting for the sensitivity loss induced by difference spectroscopy. In Fig. 6c, the computed transfer amplitudes are plotted as a function of the delay T for the two experiments assuming relaxation time constants T_2 in the range from 30 to 100 ms as typical for highly flexible proteins. For both experiments, highest sensitivity (largest transfer amplitude) is obtained for T values of about 58 ms, with the overall sensitivity about a factor of 2 higher in the Pro-HNcocan with respect to Pro-iHNcan experiment which is explained by the larger one-bond ^{13}C-^{15}N coupling 12 Hz active in Pro-HNcocan with respect to the smaller two-bond coupling $J_{NCA} \approx 7$ Hz that is active in Pro-iHNcan. Pro-HNcocan and Pro-iHNcan spectra of BASP1, recorded in 15 min and 30 min, respectively, are

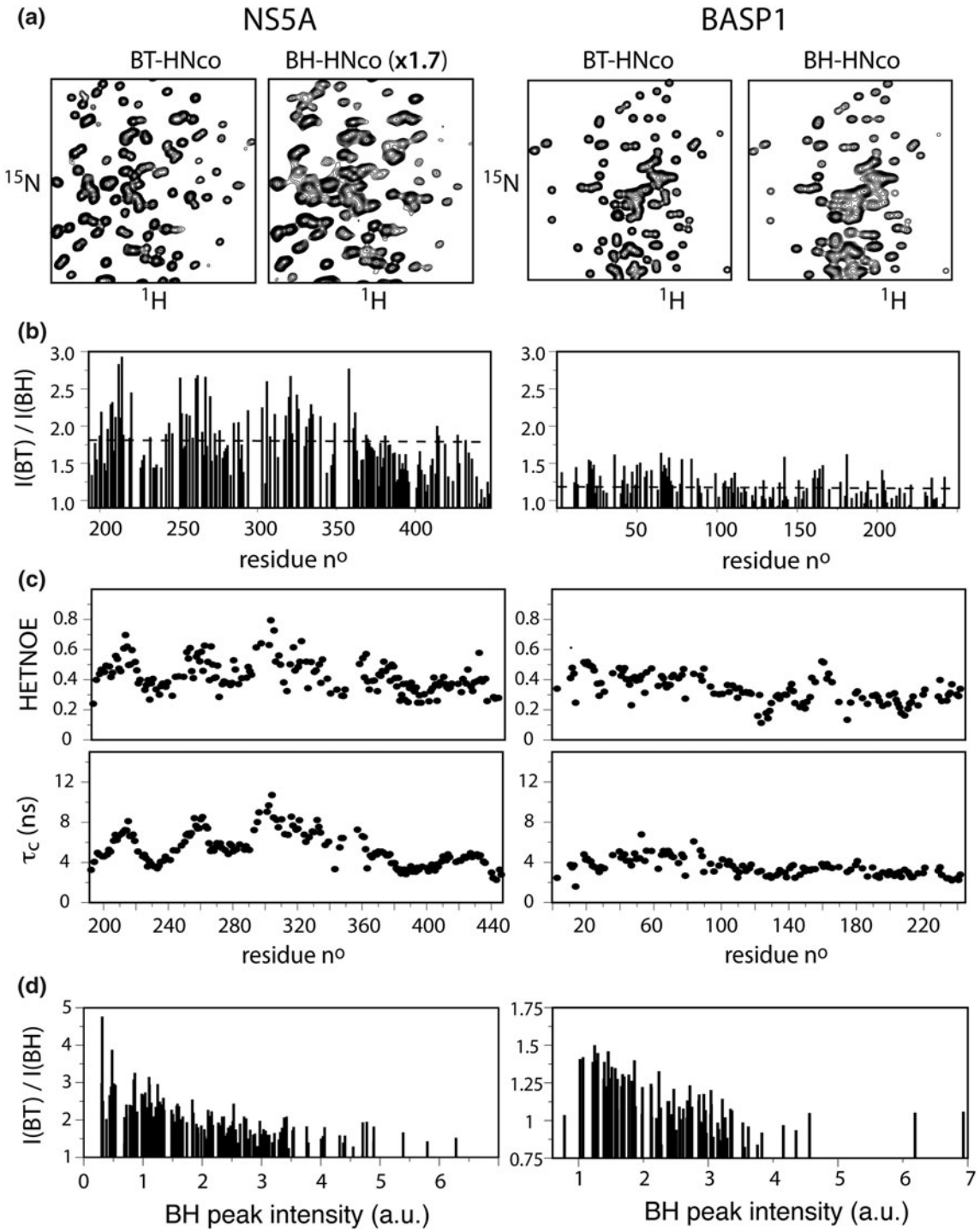


Fig. 5 Experimental comparison of the performance of BEST-HSQC (BH) and BEST-TROSY (BT)—type HNC correlation experiments for the two IDPs NS5A (*left*) and BASP1 (*right*). Extracts from 2D BT-HNco and BH-HNco spectra recorded with identical experimental parameters (except for the pulse sequence) are shown in **a**. For NS5A the BH-HNco spectrum was scaled by a factor 1.7 to obtain similar *peak* intensities for a better appreciation of the differences in spectral resolution in the two spectra. The *peak* intensity ratios (BT over BH) measured for the two proteins are plotted in **b** as a function of the peptide sequence. **c** ^{15}N relaxation data of NS5A(191–447) (*left*) and BASP1 (*right*). Local tumbling correlation times (τ_c) were obtained from the measured ^{15}N T_1 and T_2 relaxation time constants by computing the following expression $\tau_c \cong \sqrt{(6T_1/T_2 - 7)}/(4\pi\nu_N)$. **d** Same data as shown in **b**, but here the intensity gain (BT/BH) is plotted as a function of the peak intensity measured in the BEST-HSQC spectrum, in order to illustrate that highest signal enhancement is obtained for the weakest peaks leading to a more uniform spectral intensity distribution

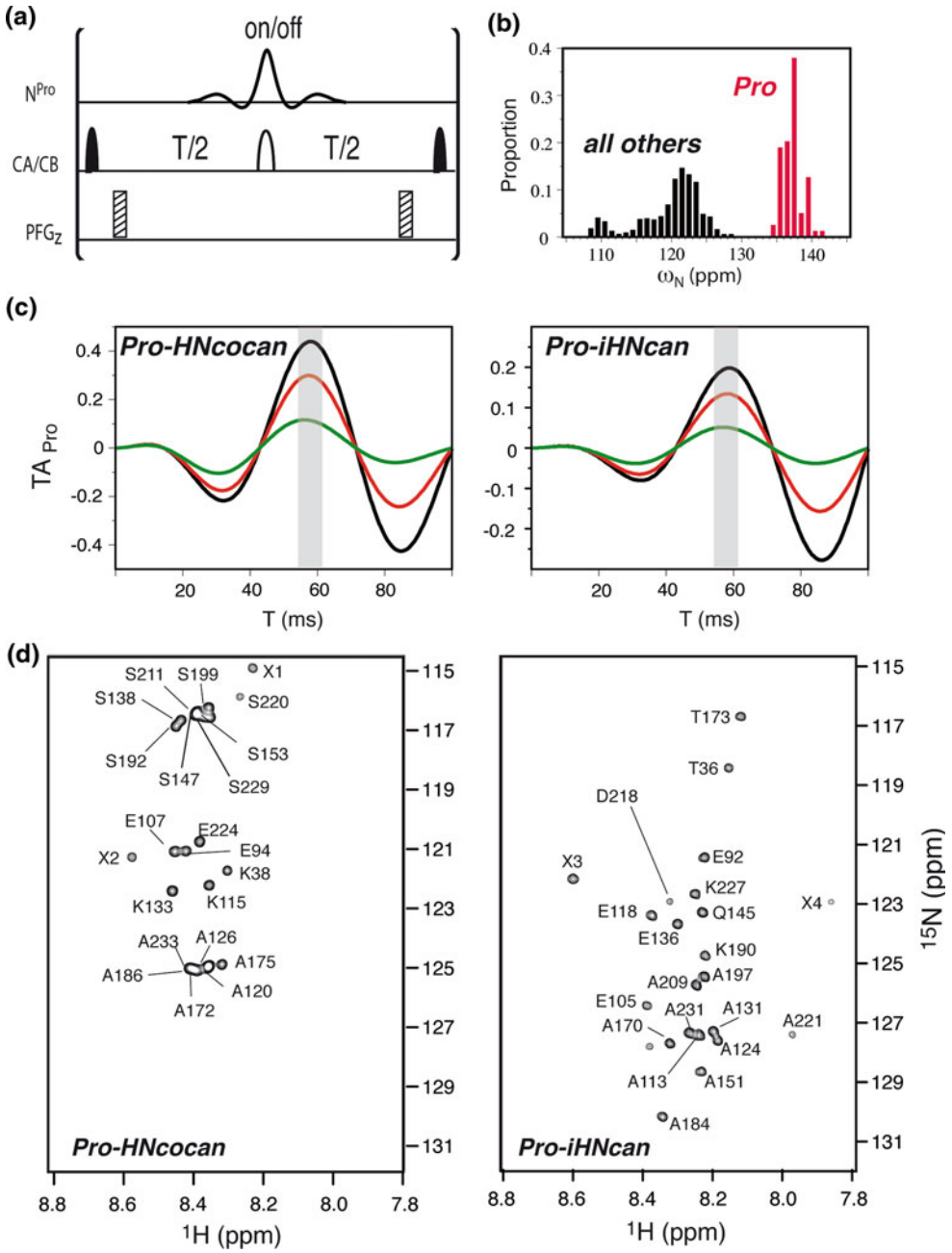


Fig. 6 Pro-selective ^{1}H–^{15}N spin-echo difference experiments. The pulse sequence element shown in **a** can be inserted into the BEST-TROSY sequence of Fig. 2a or 2b to obtain the pulse sequence of Pro-HNcocan or Pro-iHNcan, respectively. Pro-selective correlation spectra are obtained by subtracting two data sets recorded with and without the Pro ^{15}N refocusing pulse, centered at 138 ppm and covering a band width of 7 ppm (corresponding to a pulse length of 8.6 ms at 80 MHz). Use of a REBURP pulse shape (Geen and Freeman 1991) for Pro ^{15}N refocusing, makes sure that C^{α}–N coupling evolution remains active during the entire delay T (Lescop et al. 2010), thus ensuring optimal sensitivity. **b** ^{15}N chemical shift statistics of IDPs obtained from 14 data sets (1,562 chemical shifts) deposited with the BMRB for non-Pro residues (*black*), and 6 data sets (79 chemical shifts) for Pro residues (*red*), including NS5A and BASP1. **c** Transfer amplitudes (TA_{Pro}) computed as a function of the filter delay (T) using equation [2] assuming different relaxation time constants: $T_2 = 100$ ms (*black*), 60 ms (*red*) and 30 ms (*green*). Highest transfer efficiency is obtained for a filter delay $T \cong 58$ ms. (d) Pro-HNcocan (*left*) and Pro-iHNcan (*right*) spectra recorded for BASP1 at 25 °C in 15 min and 30 min, respectively. All residues adjacent to prolines can be identified unambiguously from these spectra. Peaks are annotated by the residue number (and amino-acid type) of the detected amide ^{1}H–^{15}N correlations. An annotation 'X' refers to residues within the N-terminal His-tag extension of the BASP1 protein

displayed in Fig. 6d. All residues preceding or following Pro can be unambiguously identified from these spectra providing valuable starting points for the assignment of this 270-residue IDP. The corresponding spectra recorded for NS5A are shown in Figure S2 of the Supporting Information.

Recently we have introduced HADAMAC (Lescop et al. 2008), a sensitive NMR technique that allows distinguishing between seven different classes of amino-acid types from a set of amino-acid-type-edited 2D ^{1}H–^{15}N correlation spectra. HADAMAC is particularly useful for sequential resonance assignment of IDPs (Feuerstein et al. 2012a), but Pro residues cannot be distinguished from Arg, Glu, Lys, Gln, Met, and Leu residues, because they all fall within the so-called 'rest' class, and are detected in the same 2D spectrum. Thus, the Pro-edited ^{1}H–^{15}N BEST-TROSY correlation experiments introduced here, nicely complement HADAMAC for efficient amino-acid-type discrimination in IDPs.

In summary, we have proposed a set of sensitivity- and resolution-enhanced correlation experiments for sequential resonance assignment of highly flexible protein systems such as IDPs. We have demonstrated the performance of these BEST-TROSY experiments for two IDPs of $\sim$270 residues in length. Short inter-scan delays of about 200 ms provide high sensitivity and allow to perform 3D data acquisition with long maximal evolution times in indirect dimensions (especially ^{15}N) in a total experimental time of only a few hours. If the resolution in 3D spectral space is not sufficient, the proposed experiments can be easily extended to 4D versions by additionally editing the chemical shifts of ^{13}CO that is involved in the coherence transfer pathways of all sequential and intra-residue correlation experiments. In addition, Pro-edited 2D ^{1}H-^{15}N correlation experiments have been introduced that provide complementary amino-acid-type information for resonance assignment of IDPs.

Acknowledgments We are grateful to Isabel Ayala and Adrien Favier for help in protein production and technical support. This work has been supported by grants from the European Commission (FP7-ITN IDPbyNMR contract No. 264257 and FP7-I3 BIO-NMR contract No. 261863), from the DFG (SFB974, A11), and from the Austrian Science Foundation FWF (W1221-B03 and P 20549-N19).

References

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

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

Davey NE, Trave G, Gibson TJ (2011) How viruses hijack cell regulation. Trends Biochem Sci 36:159–169

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

Farjon J, Boisbouvier J, Schanda P, Pardi A, Simorre JP, Brutscher B (2009) Longitudinal relaxation enhanced NMR experiments for the study of nucleic acids in solution. J Am Chem Soc 131:8571–8577

Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Formankay JD, Kay LE (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology-2 domain studied by 15 N NMR relaxation. Biochemistry 33:5984–6003

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 advances in solution NMR: fast methods and heteronuclear direct detection. Chem Phys Chem 10:1356–1368

Feuerstein S, Plevin MJ, Willbold D, Brutscher B (2012a) iHADAMAC: a complementary tool for sequential resonance assignment of globular and highly disordered proteins. J Magn Reson 214:329–334

Feuerstein S, Solyom Z, Aladag A, Favier A, Schwarten M, Hoffmann S, Willbold D, Brutscher B (2012b) Transient structure and SH3 interaction sites in an intrinsically disordered fragment of the hepatitis C virus protein NS5A. J Mol Biol 420:310–323

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

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

Kumar D, Paul S, Hosur RV (2010) BEST-HNN and 2D-(HN)NH experiments for rapid backbone assignment in proteins. J Magn Reson 204:111–117

Kupce E, Freeman R (1994) Wide-band excitation with polychromatic pulses. J Magn Reson A 108:268–273

Lescop E, Schanda P, Brutscher B (2007) A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. J Magn Reson 187:163–169

Lescop E, Rasia R, Brutscher B (2008) Hadamard amino-acid-type edited NMR experiment for fast protein resonance assignment. J Am Chem Soc 130:5014–5015

Lescop E, Kern T, Brutscher B (2010) Guidelines for the use of band-selective radiofrequency pulses in hetero-nuclear NMR: example of longitudinal-relaxation-enhanced BEST-type H-1-N-15 correlation experiments. J Magn Reson 203:190–198

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

Mantylahti 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

Mcintosh LP, Kang HS, Okon M, Nelson ML, Graves BJ, Brutscher B (2009) Detection and assignment of phosphoserine and phosphothreonine residues by (13)C-(31)P spin-echo difference NMR spectroscopy. J Biomol NMR 43:31–37

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

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

Schanda P (2009) Fast-pulsing longitudinal relaxation optimized techniques: enriching the toolbox of fast biomolecular NMR spectroscopy. Prog NMR Scpectrosc 55:238–265

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

Schubert M, Ball LJ, Oschkinat H, Schmieder P (2000) Bridging the gap: a set of selective H-1-N-15-correlations to link sequential neighbors of prolines. J Biomol NMR 17:331–335

Schulte-Herbruggen T, Sorensen OW (2000) Clean TROSY: compensation for relaxation-induced artifacts. J Magn Reson 144:123–128

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

Tompa P (2002) Intrinsically unstructured proteins. Trends Biochem Sci 27:527–533

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

Uversky VN, Dunker AK (2010) Understanding protein non-folding. BBA-proteins proteom 1804:1231–1264

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

Xue B, Williams RW, Oldfield CJ, Goh GKM, Dunker AK, Uversky VN (2010) Viral disorder or disordered viruses: do viral proteins possess unique features? Protein Peptide Lett 17:932–951