

## Communication

# iHADAMAC: A complementary tool for sequential resonance assignment of globular and highly disordered proteins

Sophie Feuerstein<sup>a,b,c</sup>, Michael J. Plevin<sup>a,b,c</sup>, Dieter Willbold<sup>b,d,e</sup>, Bernhard Brutscher<sup>a,b,c,\*</sup>

<sup>a</sup>Institut de Biologie Structurale, Université Grenoble 1, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France

<sup>b</sup>Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Grenoble, France

<sup>c</sup>Centre National de Recherche Scientifique (CNRS), Grenoble, France

<sup>d</sup>Institute of Complex Systems (ICS-6) Structural Biochemistry, Forschungszentrum Jülich, 52425 Jülich, Germany

<sup>e</sup>Institut für Physikalische Biologie, Heinrich-Heine-Universität, 40225 Düsseldorf, Germany

## ARTICLE INFO

## Article history:

Received 5 September 2011

Revised 26 October 2011

Available online 9 November 2011

## Keywords:

NMR

Protein

IDP

Sequential resonance assignment

HADAMAC

## ABSTRACT

An experiment, iHADAMAC, is presented that yields information on the amino-acid type of individual residues in a protein by editing the  $^1\text{H}$ – $^{15}\text{N}$  correlations into seven different 2D spectra, each corresponding to a different class of amino-acid types. Amino-acid type discrimination is realized via a Hadamard encoding scheme based on four different spin manipulations as recently introduced in the context of the sequential HADAMAC experiment. Both sequential and intra-residue HADAMAC experiments yield highly complementary information that greatly facilitate resonance assignment of proteins with high frequency degeneracy, as demonstrated here for a 188-residue intrinsically disordered protein fragment of the hepatitis C virus protein NS5A.

© 2011 Elsevier Inc. All rights reserved.

## 1. Introduction

Sequence-specific resonance assignment of  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled proteins, a prerequisite for further NMR investigations, relies on heteronuclear correlation experiments that allow building chains of sequentially linked spin systems (peptide segments). This sequential correlation data needs to be complemented by site-specific amino-acid-type information in order to match these segments to a particular position in the amino-acid sequence of the protein. The importance of additional amino-acid type information increases for polypeptides with highly degenerate NMR frequencies that makes the assembly of long fragments of connected spin systems difficult. This problem is particularly pronounced for intrinsically disordered proteins (IDPs) as the lack of persistent tertiary structure results in similar chemical environments along the polypeptide chain, and thus reduced chemical shift dispersion. Despite the absence of a global fold such IDPs carry out important cellular functions such as signal transduction, transcription regulation, and other regulatory processes by interacting with a variety of binding partners, often via transiently structured peptide segments called molecular recognition elements [1]. Multidimensional NMR spectroscopy is currently the only available technique that provides atomic-resolution

information on the conformational dynamics of IDPs, and on the mode of interaction with their binding partners. Therefore it is of interest to develop new NMR tools that facilitate sequence-specific resonance assignment of polypeptide chains with highly degenerate NMR frequencies.

Amino-acid type information can be obtained from NMR data (i) by the selective incorporation of labeled (or unlabeled) amino acids in an otherwise unlabeled (or labeled) protein [2–4], (ii) from  $^{13}\text{C}^\alpha$  and  $^{13}\text{C}^\beta$  chemical shift data [5], and (iii) from specifically tailored correlation experiments that exploit the particular spin coupling topologies of the different amino-acid side chains [6–12]. While the first approach provides the most direct amino-acid-type information, it requires the preparation of several differently labeled protein samples that can be a severe practical drawback. The use of  $^{13}\text{C}$  chemical shifts, the most prominent approach today, only provides unambiguous amino-acid type identification for a few amino-acid types, namely G, A, and the S/T pair. For all other amino acids, only probability scores can be computed as they have overlapping  $^{13}\text{C}$  chemical shift ranges. The third approach has the advantage that a single uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labeled sample is sufficient and that, in principle, each amino-acid type can be unambiguously identified by a correlation experiment designed to select a coherence transfer pathway that is specific to the spin topology of this particular amino-acid. Although this approach has been applied successfully in the past, it has several disadvantages: many different spectra need to be recorded making the approach time

\* Corresponding author. Address: Biomolecular NMR spectroscopy group, Institut de Biologie Structurale – Jean Pierre Ebel, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France. Fax: +33 4 76 88 54 94.

E-mail address: [bernhard.brutscher@ibs.fr](mailto:bernhard.brutscher@ibs.fr) (B. Brutscher).

consuming, while the multiple transfer steps required for the more complicated pulse sequences result in low experimental sensitivity.

In order to overcome some of these drawbacks, we have recently introduced the HADAMAC experiment [13,14] that has been setup for a high level of amino-acid-type discrimination in a single experiment, and thus in minimal experimental time. HADAMAC is a pseudo-3D experiment using Hadamard encoding [15] in the third dimension to separate the  $^1\text{H}$ - $^{15}\text{N}$  correlation peaks into seven 2D subspectra according to the amino-acid type classes: (1) *Val-Ile-Ala* (2) *Gly*, (3) *Ser*, (4) *Thr*, (5) *Asn-Asp*, (6) *Arom* (*Phe*, *His*, *Trp*, *Tyr*)-*Cys*, and (7) *Arg*, *Glu*, *Lys*, *Pro*, *Gln*, *Met*, *Leu*. Hadamard encoding ensures highest possible sensitivity as in each experiment signals from all amino-acid types are detected, and only their sign ( $180^\circ$  phase shift) is changed in between successive experiments according to a given Hadamard matrix [16]. The HADAMAC experiment is based on a sequential HBCBCACONH coherence transfer pathway [17,18] thus providing information on the amino-acid type of the residue preceding the detected amide  $^1\text{H}$ - $^{15}\text{N}$  frequency pair. This approach has later on been extended to bi-directional coherence transfer [19] where two correlation peaks are detected per residue in the subspectra corresponding to the amino-acid type of the preceding and the detected residue. Here we introduce iHADAMAC, an experiment that yields a single correlation peak per residue (amide  $^1\text{H}$ ,  $^{15}\text{N}$  pair) providing direct intra-residue amino-acid-type information. The performance of this new experiment is demonstrated for native and acid-denatured ubiquitin, as well as for a 188-residue IDP, a highly disordered fragment of the NS5A protein from hepatitis C virus.

## 2. The iHADAMAC experiment

The pulse sequence for the intra-residue iHADAMAC experiment is shown in Fig. 1a. The sequence performs the following intra-residue coherence transfer pathway:  $\text{HB}(i) \rightarrow \text{CB}(i) \rightarrow \text{CA}(i) \rightarrow \text{N}(i) \rightarrow \text{H}^{\text{N}}(i)$ . In the past, pulse sequence building blocks have been proposed [20–23] that allow N-CA and CA-N coherence transfer solely within a given residue yielding pure *intra-residue* correlation experiments. The pulse sequence element used here for the intra-residue  $\text{CA}(i) \rightarrow \text{N}(i)$  transfer is highlighted in gray in Fig. 1a. This sequence block is similar to the one previously introduced and described in the context of an intra-residue  $\text{HN}(\text{CA})\text{CO}$ -type experiment,  $\text{i}(\text{HCA})\text{CO}(\text{CA})\text{NH}$  [24]. Therefore, in the following, only the principal features will be briefly reviewed. Three spin couplings  $^1\text{J}_{\text{CAN}}$ ,  $^2\text{J}_{\text{CAN}}$ , and  $^1\text{J}_{\text{CAB}}$  are active during the whole transfer element of duration  $2(\eta + \gamma) \approx 45$  ms. In addition,  $^1\text{J}_{\text{COCA}}$  and  $^1\text{J}_{\text{CON}}$  couplings are active during the delays  $\gamma$  and  $\eta$ , respectively. Chemical shift evolution during these transfer delays is refocused by means of  $180^\circ$  pulses. As a result, this building block performs an intra-residue coherence transfer  $2\text{CA}_y\text{CB}_z \rightarrow 2\text{CA}_x\text{N}_z$  with a transfer amplitude (efficiency) given by the following expression:

$$I(\gamma, \eta) = \sin(\pi^1\text{J}_{\text{NCA}}2(\gamma + \eta)) \sin(\pi^2\text{J}_{\text{NCA}}2(\gamma + \eta)) \\ \times \sin^2(\pi^1\text{J}_{\text{COCA}}\gamma) \times \sin(\pi^1\text{J}_{\text{NCO}}2\eta) \sin(\pi^1\text{J}_{\text{CAB}}2(\gamma \\ + \eta)) \exp\{-2(\gamma + \eta)/T_2^{\text{eff}}\} \quad (1)$$

In Fig. 1b, the transfer efficiency has been computed as a function of the delay  $\eta$  in the absence of spin relaxation, and assuming different effective transverse relaxation times  $T_2^{\text{eff}}$  ranging from 30 to 90 ms. For the calculation, the spin coupling constants were fixed to  $^1\text{J}_{\text{CAN}} = 12$  Hz,  $^2\text{J}_{\text{CAN}} = 7$  Hz,  $^1\text{J}_{\text{CACO}} = 55$  Hz,  $^1\text{J}_{\text{CAB}} = 35$  Hz, and  $^1\text{J}_{\text{NCO}} = 15$  Hz, and the transfer delay  $\gamma$  was set to 8.5 ms. Neglecting relaxation-induced signal loss, a maximal transfer amplitude of  $\sim 0.8$  is obtained for  $\eta = 14$  ms, while the transfer efficiency drops to  $\sim 0.2$  for  $T_2^{\text{eff}} = 30$  ms. The calculations also show that for

increasing relaxation rates (shorter  $T_2^{\text{eff}}$ ) the  $\eta$  delay should be slightly reduced to achieve optimal sensitivity. Note that no cross peak is expected in the iHADAMAC spectrum for the C-terminal residue of the peptide chain that lacks a  $^2\text{J}_{\text{CAN}}$  coupling.

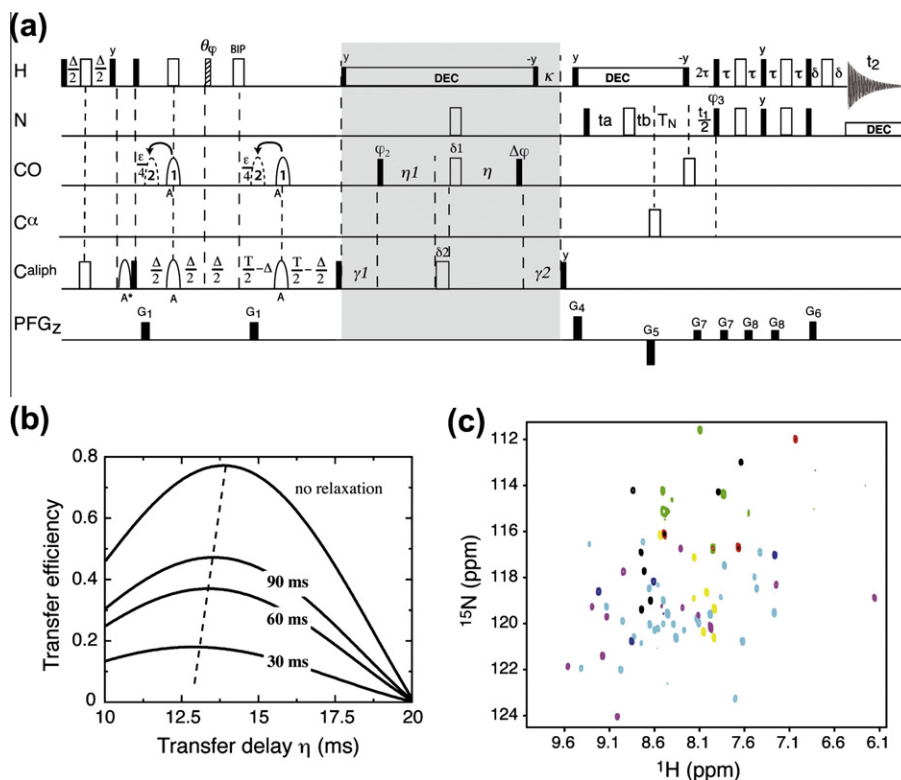
The amino-acid type encoding is based on four spin manipulations (see Table 1) exploiting different properties of the amino acid side chains to selectively invert the signals from certain amino-acid types: (i) changing the flip angle  $\theta$  of the DEPT sequence affects residues with a CH or  $\text{CH}_3$  group at the CB carbon position; (ii) changing the delay  $\kappa$  from 0 to  $1/\text{J}_{\text{CH}}$  inverts NMR signals of residues with a CH at CA carbon position; (iii) the presence/absence of a CO attached to CB is exploited by shifting CO refocusing pulses from position 1 to 2, and (iv) Ser and Thr resonances are inverted by a selective inversion pulse taking advantage of the particular  $^{13}\text{C}$  chemical shift range of these amino acids. The 8-step Hadamard encoding used for iHADAMAC, shown in Table 1, is identical to the one described recently for the HADAMAC-2 experiment [13]. This encoding results in eight different 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra (Table 2); six of them containing the desired amino-acid-type information, while the two remaining ones are used to separate signals from the undesired  $\text{HA}(i) \rightarrow \text{CA}(i) \rightarrow \text{N}(i) \rightarrow \text{H}^{\text{N}}(i)$  coherence transfer pathway that may also be detected in the iHADAMAC experiment. Note that the amino-acid type classes (6) *Arom* (*Phe*, *His*, *Trp*, *Tyr*)-*Cys*, and (7) *Arg*, *Glu*, *Lys*, *Pro*, *Gln*, *Met*, *Leu* are detected in the same Hadamard subspectrum, but with opposite signs.

In addition, we have implemented semi-CT  $^{15}\text{N}$  editing in the iHADAMAC sequence of Fig. 1a in order to allow for high spectral resolution in the  $^{15}\text{N}$  dimension as required for NMR studies of IDPs in order to distinguish most of the individual  $^1\text{H}$ - $^{15}\text{N}$  correlation peaks in the 2D map.

## 3. Results and discussion

As an application of the iHADAMAC experiment to small globular proteins, the spectrum obtained for a 2 mM sample of the 76-residue protein ubiquitin is shown in Fig. 1c. The seven  $^1\text{H}$ - $^{15}\text{N}$  subspectra corresponding to the seven amino-acid type classes are color coded, and superposed on the same graph. This iHADAMAC spectrum has been recorded at  $25^\circ\text{C}$  in 1 h on a 600 MHz spectrometer equipped with a cryogenic probe, illustrating the high spectral quality that can be obtained in short overall time for small fast tumbling proteins. Of course, as shown in Fig. 1b the experimental sensitivity will drop significantly for proteins with slower molecular tumbling, thus limiting the application to globular proteins in the  $<15$  kDa range for which high quality spectra can be recorded in a reasonable amount of acquisition time ( $<\sim 48$  h).

In the following we will focus on applications to highly disordered proteins. The dynamic behavior of such molecular systems is generally characterized by fast effective local rotational correlation times, resulting in favorable spin relaxation properties that make the iHADAMAC experiment amenable to the study of larger protein systems. As a first example we show the results obtained for a small amount ( $<150$   $\mu\text{g}$ ) of urea-denatured ubiquitin (8 M urea, pH 2.5) as a model system for a protein with high intrinsic disorder and almost no residual transient structure. The protein was dissolved in 150  $\mu\text{L}$  buffer solution, resulting in a final protein concentration of 90  $\mu\text{M}$ , and placed in a 3 mm NMR tube. The protein containing tube was then put into a standard 5 mm NMR tube filled with  $\text{D}_2\text{O}$  amenable to NMR investigation in a standard 5 mm triple-resonance cryogenic probe. The iHADAMAC spectrum, shown in Fig. 2, was recorded in 24 h using an 800 MHz spectrometer. Despite the small amount of protein sample used for recording this spectrum, all expected  $^1\text{H}$ - $^{15}\text{N}$  correlations could be detected and unambiguously assigned to one of the seven amino-acid type classes.



**Fig. 1.** (a) Pulse sequence for the intra-residue Hadamard-encoded amino-acid-type-editing (iHADAMAC) experiment. All radio-frequency (rf) pulses are applied along the x-axis unless indicated.  $90^\circ$  and  $180^\circ$  rf pulses are represented by filled and open pulse symbols, respectively. The dashed  $^1\text{H}$  pulse corresponds to the DEPT pulse with variable flip angle  $\theta$  and phase  $\phi$ . The  $^1\text{H}$  pulse labeled “BIP” is a broadband inversion pulse with a BIP-720–50–20 phase modulation [28]. The  $^{13}\text{C}$  shaped pulses indicated by (A) and (A\*) have an isnob5 shape [29] covering band widths of  $46 \pm 40$  ppm ( $\text{C}^{\text{aliph}}$ , A),  $70 \pm 15$  ppm ( $\text{C}^{\text{aliph}}$ , A\*), and  $196 \pm 40$  ppm (CO). All other  $\text{C}^{\text{aliph}}$  and  $\text{C}^\alpha$  pulses are applied with a rectangular shape and field strengths of  $\Delta/\sqrt{15}$  ( $90^\circ$ ) and  $\Delta/\sqrt{3}$  ( $180^\circ$ ), where  $\Delta$  is the separation in Hz between the centers of the  $\text{C}^\alpha$  (56 ppm) or  $\text{C}^{\text{aliph}}$  (46 ppm) and CO (175 ppm) chemical shift regions, while the other CO pulses have the shape of the center lobe of a  $\sin x/x$  function. The transfer delays are adjusted to  $\Delta = 3.7$  ms,  $T = 20$  ms,  $\varepsilon = 2.5$  ms,  $\gamma = 8.5$  ms,  $\gamma_1 = \gamma + \delta_2$ ,  $\gamma_2 = \gamma - \delta_1$ ,  $\eta = 12.5 - 14$  ms,  $\eta_1 = \eta - \delta_2$ ,  $\kappa =$  (see Table 1),  $T_N = 12$  ms,  $\tau = 2.7$  ms,  $\text{ta} = T_N - t_1/2 + \text{tb}$ , and  $\text{tb} = 0$ . The delay  $\text{tb}$  is incremented together with  $t_1$  for semi-CT  $^{15}\text{N}$  editing using the following time increment:  $\Delta \text{tb} = (T_1^{\text{max}}/2 - T_N)/N_1$  with  $N_1$  the number of total increments in the  $t_1$  dimension. The phase  $\Delta\phi$  ( $\sim 70^\circ$ ) accounts for Bloch–Siegert phase shift and needs to be optimized experimentally for maximal sensitivity. Pulsed field gradients  $G_1$ – $G_8$  are applied along the z-axis (PFGz) with a gradient strength of approximately 20 G/cm and lengths ranging from 100 to 2000  $\mu\text{s}$ . The relative durations of  $G_5$  and  $G_6$  are given by the gyromagnetic ratios of  $^1\text{H}$  and  $^{15}\text{N}$  as  $G_5/G_6 = \gamma_{\text{H}}/\gamma_{\text{N}}$ . A four-step phase cycle is used with  $\phi = y, y, -y, -y$ ;  $\phi_2 = x, -x$ , and the receiver phase  $\phi_{\text{rec}} = x, -x, (-x, x)$ . Quadrature detection in  $t_1$  is achieved according to the echo-antiecho scheme by inverting the sign of gradient  $G_5$  and phase  $\phi_3$ . In addition, phase  $\phi_3$  is inverted for every second  $t_1$  increment. Eight repetitions of the experiments are recorded for Hadamard-based amino-acid-type encoding with the parameter settings given in Table 1. The HADAMAC pulse sequence in Varian language, and processing protocols for NMRPipe software [30] are available from the authors upon request. (b) Calculation of transfer efficiency of the pulse sequence block highlighted in gray as a function of the delay  $\eta$  assuming different relaxation times. (c) iHADAMAC spectrum of native ubiquitin recorded at  $25^\circ\text{C}$  and 600 MHz  $^1\text{H}$  frequency. The seven subspectra corresponding to the different amino-acid type classes are color-coded and superposed on the same graph. The color-coding used throughout this manuscript is the following: Val-Ile-Ala (magenta), Gly (green), Ser (red), Thr (black), Asn-Asp (yellow), Phe, His, Trp, Tyr, Cys (blue), and Arg, Glu, Lys, Pro, Gln, Met, Leu (cyan). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 1**

Experimental settings used for iHADAMAC (Fig. 1a) in order to realize 8-step Hadamard amino-acid type encoding.

	CO inversion pulses A at position (1 or 2)	$\text{C}^{\text{aliph}}$ inversion pulse A* applied (yes/no)	Flip angle $\theta$	Delay $\kappa$
Exp 1	1	no	$120^\circ$	0
Exp 2	2	no	$120^\circ$	0
Exp 3	1	yes	$120^\circ$	7ms
Exp 4	2	yes	$120^\circ$	7ms
Exp 5	1	no	$60^\circ$	0
Exp 6	2	no	$60^\circ$	0
Exp 7	1	yes	$60^\circ$	7ms
Exp 8	2	yes	$60^\circ$	7ms

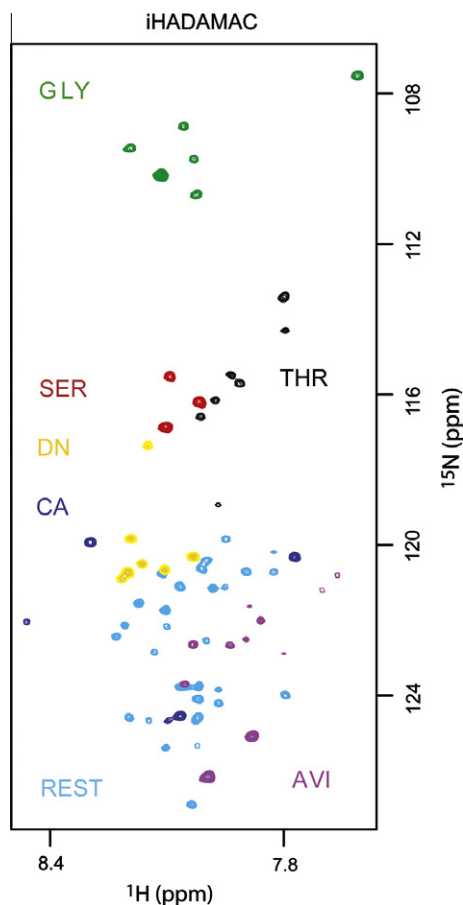
Finally, we have applied the iHADAMAC experiment to a *real* IDP system, a 188-residue fragment of the hepatitis C virus protein NS5A [25]. This fragment comprises the entire domain two as well as elongations in both, C- and N-terminal directions that contain crucial binding motifs necessary for interaction with other viral and host proteins [26]. This NS5A fragment presents a challenge for NMR studies, not only because of its length and high NMR frequency degeneracy, but also because of the heterogeneous spin

relaxation behavior observed for this protein, translating to significant intensity variations even in a simple  $^1\text{H}$ – $^{15}\text{N}$  correlation experiment. The measured  $^{15}\text{N}$  transverse relaxation rates cover a range of almost one order of magnitude (data not shown). This peculiar behavior is explained by the structural model emerging from our NMR data pointing toward the presence of three transiently populated  $\alpha$ -helical segments that are connected by flexible linkers [27]. The iHADAMAC spectrum recorded on a

**Table 2**  
Hadamard matrix resulting from the spin manipulation of Table 1, used for encoding and decoding in iHADAMAC.

	Gly	AVI	Ser	Thr	Asx	Cys-Arom/REST	CA pathway I (<55 ppm) <sup>a</sup>	CA pathway II (55–85 ppm) <sup>a</sup>
Exp 1	+	+	+	+	+	+	+	+
Exp 2	–	+	+	+	–	+	–	–
Exp 3	–	+	–	–	+	+	+	–
Exp 4	+	+	–	–	–	+	–	+
Exp 5	–	+	–	+	–	–	+	+
Exp 6	+	+	–	+	+	–	–	–
Exp 7	+	+	+	–	–	–	+	–
Exp 8	–	+	+	–	+	–	–	+

<sup>a</sup> The CA pathways I and II correspond to the  $H_i^z - C_i^z - CO_i - N_i - HN_i$  transfer. The Ser/Thr  $C^{\beta}$  inversion pulse also affects sign encoding for residues with  $C^z$  chemical shift values in the 55–85 ppm range leading to the creation of the “CA pathway” groups I and II.



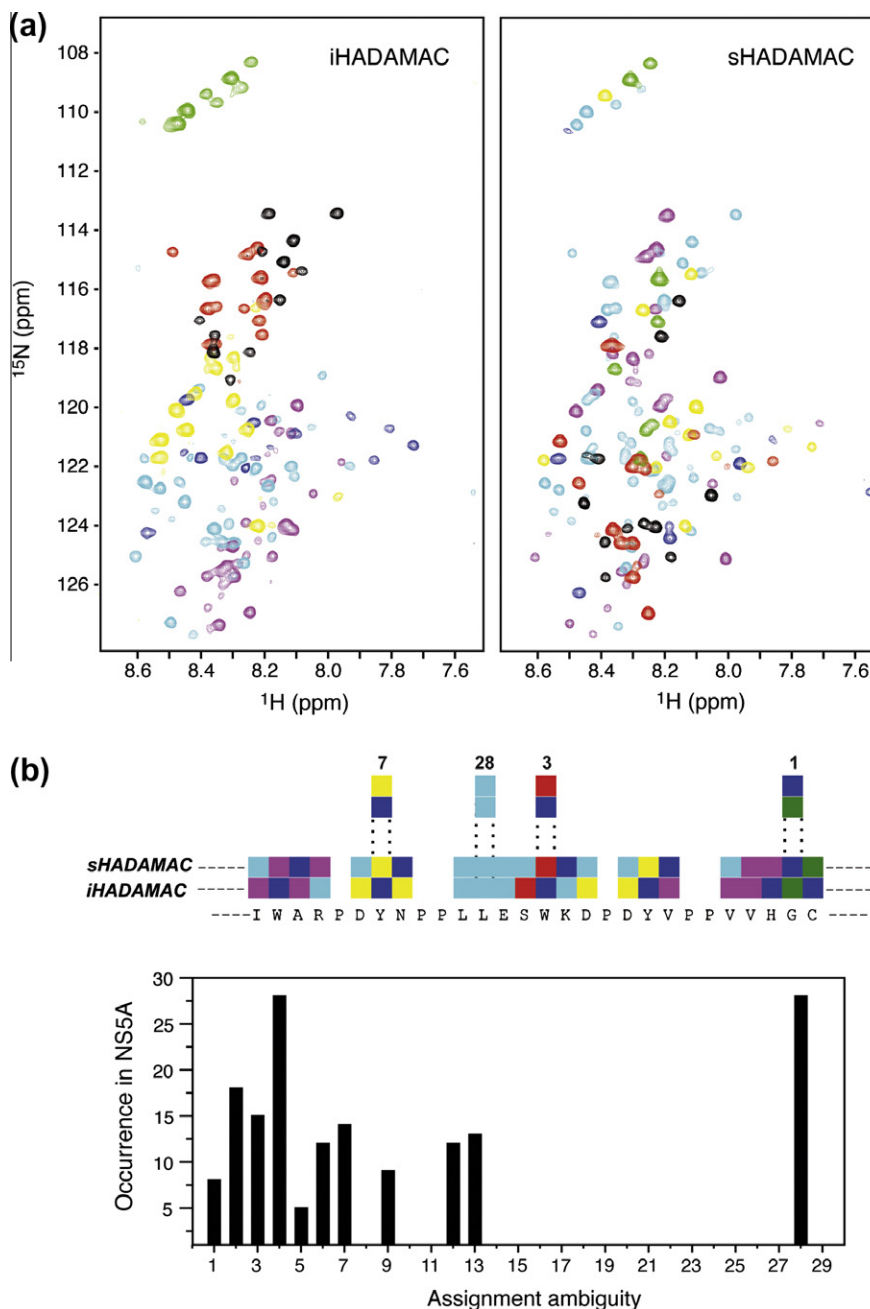
**Fig. 2.** iHADAMAC spectrum recorded on a 90  $\mu$ M sample of acid-denatured ubiquitin (20 mM Glycine, pH2.5, 8 M Urea) at 25 °C and 800 MHz  $^1$ H frequency. 150  $\mu$ L of sample was placed in a 3 mm NMR tube sitting inside a standard 5 mm tube filled with  $D_2O$  [31]. 100 complex points were recorded in the  $t_1$  ( $^{15}N$ ) dimension for a spectral width of 2000 Hz using semi-CT editing. The total acquisition time was 24 h. The seven subspectra corresponding to the different amino-acid type classes are color-coded and superposed on the same graph. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

90  $\mu$ M sample (500  $\mu$ L solution in standard 5 mm NMR tube) is shown in Fig. 3a (left spectrum). The different 2D planes are color-coded and superposed on the same graph. In addition, a plane-by-plane representation with peak assignments is provided as Supporting Information. This iHADAMAC data set has been recorded at 20 °C on a 800 MHz machine equipped with a cryogenic probe in a total experimental time of 12 h. About 90% of the correlation peaks that are observed in a  $^1H-^{15}N$  correlation spectrum are also detected in this iHADAMAC spectrum, although some of them

with very low intensity. The missing peaks correspond to residues within transiently structured segments of the IDP that are characterized by slower effective rotational correlation times, and thus more efficient transverse relaxation of the nuclear spins.

The intra-residue amino-acid type information nicely complements the sequential amino-acid type information obtained from a standard HADAMAC experiment (from here on called sHADAMAC). The sHADAMAC spectrum recorded in 2 h on the same NS5A sample under identical experimental conditions is shown in Fig. 3a (right spectrum). The different experimental times chosen for the iHADAMAC (12 h) and sHADAMAC (2 h) reflect the relative sensitivity of the respective pulse sequences for this protein. The combined amino-acid type information extracted from the two HADAMAC spectra (iHADAMAC and sHADAMAC) drastically reduces the assignment possibilities of a given  $^1H-^{15}N$  frequency pair (residue) to a specific position in the primary sequence of the protein that is now constrained by the amino-acid type of two neighboring residues (Fig. 3b). For the 188-residue NS5A fragment, this combined HADAMAC information reduces the assignment ambiguities for 60% of the  $^1H-^{15}N$  correlation peaks to less than eight positions in the peptide sequence, and allows unambiguous sequence-specific assignment of eight  $^1H-^{15}N$  frequencies without any additional correlation information. As a consequence, it is no longer necessary to build up long fragments of sequentially correlated residues (NMR frequencies) from the usual 3D (or 4D) H–N–C type correlation experiments in order to achieve complete sequential assignment of the IDP. This overcomes most of the problems related to missing  $^1H-^{15}N$  correlations due to Pro residues or exchange broadened amide protons that are not observable in the amide  $^1H$  detected experiments, and therefore present break points in the sequential assignment walk. In addition, the assignment procedure becomes more robust with respect to errors in the connected residue fragments due to missing correlation peaks, frequency degeneracy, and other spectral imperfections. Moreover, the combination of iHADAMAC and sHADAMAC provides a convenient tool for re-assigning proteins for which an NMR assignment is already available, but in different experimental sample conditions (T, pH, buffer, ...), of a homologous protein containing a few point mutations, or any other situation inducing small, often localized peak shifts in the NMR spectrum. As long as spectral resolution is sufficient, these peak shifts can be conveniently identified from simple inspection of the set of 2D HADAMAC planes, instead of a more cumbersome analysis of 3D HNC correlation maps.

In summary, we have introduced iHADAMAC, a new pulse sequence that yields intra-residue amino-acid-type information complementary to its sequential counterpart sHADAMAC. Together these HADAMAC data drastically decrease the sequential assignment ambiguities of individual amide groups as demonstrated here for a 188-residue IDP. We believe that this experiment will prove especially useful for NMR studies of proteins with a high degree of frequency degeneracy and/or proteins with many missing



**Fig. 3.** (a) iHADAMAC (left) and sHADAMAC (right) spectra recorded on a 90  $\mu\text{M}$  sample of the 188-residue NS5A fragment at 20  $^\circ\text{C}$  and 800 MHz  $^1\text{H}$  frequency. For both spectra, 100 complex points were recorded in the  $t_1$  ( $^{15}\text{N}$ ) dimension for a spectral width of 2000 Hz using semi-CT editing. The total acquisition times were set to 12 h (iHADAMAC) and 2 h (sHADAMAC). The seven subspectra corresponding to the different amino-acid type classes are color-coded and superposed on the same graph. (b) Graphical representation of the information content contained in the two HADAMAC spectra. A pair of colored boxes represents the amino-acid type information at each residue position, illustrated for a small part of the NS5A sequence. The assignment ambiguity is given by the number of times a particular pair of sequential amino-acid types (pair of two colors) is found in the primary sequence of the protein, as indicated here for four representative examples of color pairs. A statistical analysis of the remaining assignment ambiguity for the 188-residue NS5A fragment is shown at the bottom. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$^1\text{H}$ - $^{15}\text{N}$  correlation peaks, either due to Pro residues or to unfavorable local dynamics.

#### Acknowledgments

We thank I. Ayala (IBS Grenoble) for help with the production of the protein samples used for this study, and S. Hoffmann (FZ Jülich) for stimulating discussion on the NS5A project. This work has been supported in part by grants from the European Commission (FP7-

I3-BIO-NMR contract No. 261862, FP7-ITN-IDPbyNMR contract No. 264257, and FPR-IRG-2008 contract No. 231082), and financial support from ARC to M.J.P. and from CEA to S.F.

#### Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmr.2011.10.019](https://doi.org/10.1016/j.jmr.2011.10.019).



## References

- [1] P. Tompa, Structure and function of intrinsically disordered proteins, Chapman & Hall/CRC Press, 2009.
- [2] L.P. McIntosh, F.W. Dahlquist, Biosynthetic incorporation of  $^{15}\text{N}$  and  $^{13}\text{C}$  for assignment and interpretation of nuclear magnetic resonance spectra of proteins, *Q. Rev. Biophys.* 23 (1990) 1–38.
- [3] B. Krishnarajana, G. Jaipuria, A. Thakur, P. D'Silva, H.S. Atreya, Amino acid selective unlabeled for sequence specific resonance assignments in proteins, *J. Biomol. NMR* 49 (2011) 39–51.
- [4] K.I. Tong, M. Yamamoto, T. Tanaka, A simple method for amino acid selective isotope labeling of recombinant proteins in *E. coli*, *J. Biomol. NMR* 42 (2008) 59–67.
- [5] S. Grzesiek, A. Bax, Amino-acid type determination in the sequential assignment of uniformly  $^{13}\text{C}/^{15}\text{N}$  enriched proteins, *J. Biomol. NMR* 3 (1993) 185–204.
- [6] V. Dötsch, R.E. Oswald, G. Wagner, Selective identification of threonine, valine, and isoleucine sequential connectivities with a TVI-CBCACONH experiment, *J. Magn. Reson. B* 110 (1996) 304–308.
- [7] V. Dötsch, R.E. Oswald, G. Wagner, Amino-acid-type-selective triple-resonance experiments, *J. Magn. Reson. B* 110 (1996) 107–111.
- [8] M. Schubert, H. Oschkinat, P. Schmieder, MUSIC and aromatic residues: amino acid type-selective H-1-N-15 correlations, III, *J. Magn. Reson.* 153 (2001) 186–192.
- [9] M. Schubert, H. Oschkinat, P. Schmieder, MUSIC, selective pulses, and tuned delays: amino acid type-selective H-1-N-15 correlations, II, *J. Magn. Reson.* 148 (2001) 61–72.
- [10] M. Schubert, M. Smalla, P. Schmieder, H. Oschkinat, MUSIC in triple-resonance experiments: Amino acid type-selective H-1-N-15 correlations, *J. Magn. Reson.* 141 (1999) 34–43.
- [11] J. Chugh, D. Kumar, R.V. Hosur, Tuning the HNN experiment: generation of serine-threonine check points, *J. Biomol. NMR* 40 (2008) 145–152.
- [12] D. Kumar, J.G. Reddy, R.V. Hosur, hnCOcaNH and hncoCANH pulse sequences for rapid and unambiguous backbone assignment in (( $^{13}\text{C}$ ), ( $^{15}\text{N}$ )) labeled proteins, *J. Magn. Reson.* 206 (2010) 134–138.
- [13] E. Lescop, B. Brutscher, Highly automated protein backbone resonance assignment within a few hours: the BATCH strategy and software package, *J. Biomol. NMR* 44 (2009) 43–57.
- [14] E. Lescop, R. Rasia, B. Brutscher, Hadamard amino-acid-type edited NMR experiment for fast protein resonance assignment, *J. Am. Chem. Soc.* 130 (2008) 5014–5015.
- [15] E. Kupce, T. Nishida, R. Freeman, Hadamard NMR spectroscopy, *Prog. NMR Spectrosc.* 42 (2003) 95–122.
- [16] J. Hadamard, Resolution d'une question relative aux determinants, *Bull. Sci. Math.* 17 (1893) 240–248.
- [17] B. Brutscher, DEPT spectral editing in HCCONH-type experiments. Application to fast protein backbone and side chain assignment, *J. Magn. Reson.* 167 (2004) 178–184.
- [18] S. Grzesiek, A. Bax, Correlating backbone amide and side-chain resonances in larger proteins by multiple relayed triple resonance Nmr, *J. Am. Chem. Soc.* 114 (1992) 6291–6293.
- [19] D. Pantoja-Uceda, J. Santoro, Amino acid type identification in NMR spectra of proteins via beta- and gamma-carbon edited experiments, *J. Magn. Reson.* 195 (2008) 187–195.
- [20] B. Brutscher, Intraresidue HNCA and COHNCA experiments for protein backbone resonance assignment, *J. Magn. Reson.* 156 (2002) 155–159.
- [21] D. Nietlispach, A selective intra-HN(CA)CO experiment for the backbone assignment of deuterated proteins, *J. Biomol. NMR* 28 (2004) 131–136.
- [22] D. Nietlispach, Y. Ito, E.D. Laue, A novel approach for the sequential backbone assignment of larger proteins: selective intra-HNCA and DQ-HNCA, *J. Am. Chem. Soc.* 124 (2002) 11199–11207.
- [23] P. Permi, Intraresidual HNCA: an experiment for correlating only intraresidual backbone resonances, *J. Biomol. NMR* 23 (2002) 201–209.
- [24] S. Mantylahti, H. Tossavainen, M. Hellman, P. Permi, An intraresidual i(HCA)CO(CA)NH experiment for the assignment of main-chain resonances in N-15, C-13 labeled proteins, *J. Biomol. NMR* 45 (2009) 301–310.
- [25] D. Moradpour, F. Penin, C.M. Rice, Replication of hepatitis C virus, *Nat. Rev. Microbiol.* 5 (2007) 453–463.
- [26] T.L. Tellinghuisen, J. Marcotrigiano, A.E. Gorbalenya, C.M. Rice, The NS5A protein of hepatitis C virus is a zinc metalloprotein, *J. Biol. Chem.* 279 (2004) 48576–48587.
- [27] S. Feuerstein, Z. Solyom, A. Aladag, S. Hoffmann, D. Willbold, B. Brutscher,  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{15}\text{N}$  resonance assignment of a 179 residue fragment of hepatitis C virus non-structural protein 5A, *Biomol NMR Assign* 5 (2011) 241–243.
- [28] M.A. Smith, H. Hu, A.J. Shaka, Improved broadband inversion performance for NMR in liquids, *J. Magn. Reson.* 151 (2001) 269–283.
- [29] E. Kupce, J. Boyd, I.D. Campbell, Short selective pulses for biochemical applications, *J. Magn. Reson. B* 106 (1995) 300–303.
- [30] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, Nmrpipe – A multidimensional spectral processing system based on unix pipes, *J. Biomol. NMR* 6 (1995) 277–293.
- [31] R.M. Rasia, M. Noirclerc-Savoye, N.G. Bologna, B. Gallet, M.J. Plevin, L. Blanchard, J.F. Palatnik, B. Brutscher, T. Vernet, J. Boisbouvier, Parallel screening and optimization of protein constructs for structural studies, *Protein Sci.* 18 (2009) 434–439.