Journal of Magnetic Resonance 302 (2019) 50–63

Contents lists available at ScienceDirect

Journal of Magnetic Resonance

journal homepage: www.elsevier.com/locate/jmr

Towards complete polypeptide backbone NH assignment via combinatorial labeling

Frank Löhr, Jakob Gebel, Erik Henrich, Christopher Hein, Volker Dötsch *

Institute of Biophysical Chemistry & Center for Biomolecular Magnetic Resonance, Goethe University, Max-von-Laue-Str. 9, 60438 Frankfurt, Germany

ARTICLE INFO

Article history:
Received 19 February 2019
Revised 27 March 2019
Accepted 28 March 2019
Available online 29 March 2019

Keywords:
BEST-TROSY
Cell-free expression
Cyclophilin D
Isotope editing/filtering
Phosphorylation activation domain
Selective labeling
Sequence repeats

ABSTRACT

Combinatorial selective isotope labeling is a valuable tool to facilitate polypeptide backbone resonance assignment in cases of low sensitivity or extensive chemical shift degeneracy. It involves recording of ^{15}N-HSQC and 2D HN-projections of triple-resonance spectra on a limited set of samples containing different combinations of labeled and unlabeled amino acid types. Using labeling schemes in which the three backbone heteronuclei (amide nitrogen, α-carbon and carbonyl carbon) are enriched in ^{15}N or ^{13}C isotopes – individually as well as simultaneously – usually yields abundant amino-acid type information of consecutive residues i and $i-1$. Although this results in a large number of anchor points that can be used in the sequential assignment process, for most amide signals the exact positioning of the corresponding residue the polypeptide sequence still relies on matching intra- and interresidual ^{13}C chemical shifts obtained from 3D spectra. An obvious way to obtain more sequence-specific assignments directly with combinatorial labeling would be to increase the number of samples. This is, however, undesirable because of increased sample preparation efforts and costs. Irrespective of the number of samples, unambiguous assignments cannot be accomplished for $i-1/i$ pairs that are not unique in the sequence. Here we show that the ambiguity for non-unique pairs can be resolved by including information about the labeling state of residues $i+1$ and $i-2$. Application to a 35-residue peptide resulted in complete assignments of all detectable signals in the ^{15}N HSQC which, due to its repetitive sequence and ^{13}C chemical shift degeneracies, was difficult to achieve by other means. For a medium-sized protein (165 residues, rotational correlation time 8.2 ns) the improved protocol allowed the extent of backbone amide assignment to be expanded to 88% solely using a suite of 2D ^{1}H-^{15}N correlated spectra.

1. Introduction

Sequence-specific assignments of backbone amide resonances enable interaction studies of proteins with their binding partners, investigations of protein dynamics and proton exchange with the solvent, measurements of, e.g., paramagnetic relaxation enhancements, pseudocontact shifts, or residual dipolar couplings, and form the first step in protein structure determinations by NMR. In most cases this is achieved by correlating $^{1}H^{N}$ and ^{15}N resonances with those of neighboring nuclei using uniformly ^{13}C and/or ^{15}N enriched samples in combination with three- or higher dimensional pulse sequences. In an independent approach, which proved to be beneficial in situations of limited sensitivity and extensive chemical shift degeneracy, valuable assignment information can be obtained with the help of amino-acid specific labeling, usually requiring only one- or two-dimensional NMR experiments. Substantial spectral simplification along with identification of amino acid types can be achieved by specific ^{15}N labeling [1–6]. Dual-selective labeling with ^{15}N and 1-^{13}C labeled amino acids, pioneered by Kainosho [7], allows site-specific information to be attained by exploiting the one-bond scalar coupling across the peptide bond, either observed as J splittings in the ^{15}N dimensions of HMQC/HSQC spectra [8–12], in 2D carbon-detected ^{13}C-^{15}N correlations [13], or using 2D HN(CO) experiments [14–21]. Preparation of up to 19 NMR samples would be necessary in order to obtain amino-acid type information of all non-proline residue types present in a particular target protein by means of individual ^{15}N labeling. Even more samples – in the dimension of the number of residues – would be required when aiming at sequence-specific assignments of all distinguishable amino acid pairs via dual-selective ^{15}N/1-^{13}C labeling. Considering that only sensitive experiments such as ^{15}N-HSQC and 2D HN(CO) need to be recorded and that cell-free expression systems allow a fast and cost-efficient production of selectively labeled protein samples while to a large extent avoiding problems due to isotope scrambling [15,18,21–

* Corresponding author.
E-mail address: vdoetsch@em.uni-frankfurt.de (V. Dötsch).

https://doi.org/10.1016/j.jmr.2019.03.010
1090-7807/

26], this approach may be feasible in principle but would be quite laborious.

Recently reviewed [27–29] combinatorial selective labeling and unlabeling methods take advantage of multiplexing, resulting in a tremendous reduction of the number of samples to achieve the same task in a more efficient manner [23,25,30–33]. For instance, identification of all 19 amino acid types in ^{1}H-^{15}N correlation maps can be accomplished with as few as three specifically labeled samples [34,35]. The systematic analysis of ^{15}N HSQC and 2D HN(CO) spectra recorded on multiple protein samples, each containing different subsets of ^{13}C and ^{15}N labeled amino acids rather than a single of each labeling type, was introduced by Parker and co-workers [31]. Subsequently, similar combinatorial dual-selective labeling schemes were employed to facilitate the backbone assignment of membrane proteins [36–38], which is severely complicated by a lack of spectral dispersion, further compounded by broad lines, when otherwise relying on uniformly labeled samples. Including ^{13}C labeling at the α-carbon (2-^{13}C) in addition to selective ^{15}N and 1-^{13}C labeling increases the number of single amino acid types and amino acid pairs that can be identified with a given number of samples [39–41]. On the other hand, to differentiate between all resulting isotopomeric species involving the observed amide group, additional 2D HN(CX) type triple resonance experiments have to be acquired. Accordingly, combinatorial triple-selective labeling protocols [35,40–42] feature a "sample dimension" to identify amino acid types from the occurrence of cross peaks in the spectra of n different samples and an "experiment dimension" in which sequential information is derived from the pattern of correlations detected in a series of m 2D HN(CX) spectra. This results in an $m \times n$ grid of spectra where the presence/absence pattern is characteristic for a specific residue in the polypeptide sequence. Owing to the combinatorial nature of both dimensions the number of distinguishable $i-1/i$ pairs highly exceeds the total number of spectra to be acquired.

So far, combinatorial selective labeling of proteins has been utilized to disambiguate crowded regions of the spectra or, conversely, to provide amino acid types or selected sequence-specifically assigned peaks as starting points for the conventional assignment procedure [18,19,21,32,35–42]. In any case, preparation of a uniformly labeled sample, followed by acquisition of a reduced set of 3D triple-resonance experiments is necessary to fill the gaps between the residues assigned with the help of selectively labeled samples. One major issue preventing full backbone NH assignments exclusively with combinatorial selective labeling is its restriction to unique $i-1/i$ pairs. Unambiguous assignments cannot be obtained whenever an amino-acid type combination of two consecutive residues occurs more than once in a protein sequence. Recently, Atreya and co-workers demonstrated that this type of degeneracy can be lifted by application of a pulse sequence that senses whether the amide nitrogen of residue $i+1$ is ^{15}N labeled or not [43]. Using a single sample in which six residue types were unlabeled in a ^{13}C/^{15}N labeled background, together with $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shift information taken from 3D spectra of a uniformly labeled sample, these authors were able to assign large portions of the signals in the ^{15}N HSQC of two challenging targets. Here we use a closely related experiment to expand the information content of previous combinatorial triple-selective labeling schemes. Despite this improvement ambiguities can remain due to a degeneracy of $i-1/i/i+1$ tripeptides or because only two cases, i.e. $i-1/i/^{14}N_{i+1}$ and $i-1/i/^{15}N_{i+1}$, are differentiated for identical $i-1/i$ pairs. As a remedy, use of an experiment that probes the labeling state of the carbonyl carbon of residue $i-2$ is suggested, finally resulting in a high coverage of residues assigned via combinatorial labeling.

2. Materials and methods

2.1. Sample preparations

The production of selectively labeled NMR samples was carried out using a continuous-exchange cell-free expression system based on an *E. coli* S30 extract [44]. The p63 PAD peptide (amino acids 569 to 598 of p63) was cloned into a pET28b vector containing sfGFP as an expression tag for cell free protein synthesis. In between the sfGFP protein and peptide sequence a 3C protease cleavable sequence and a His-tag were introduced, to allow for standard metal affinity chromatography purification. For NMR samples of p63 PAD 6 ml of reaction mixture (containing ~0.5 mM of each amino acid) were dialyzed overnight against 100 ml of feeding mixture (containing ~0.6 mM of each amino acid) at 30 °C under 150 rpm shaking. Amino acid type compositions of selectively labelled samples are summarized in Table 1. The remaining amino acids were supplied in natural abundance.

After expression the reaction mixture was centrifuged at 16,000 g to remove precipitated protein. The supernatant was diluted 1:3 with buffer A (25 mM Tris, pH 8, 200 mM NaCl, 30 mM imidazole) to dilute the EDTA and DTT in the cell free expression mixture. The protein was loaded to a 5 ml HisTrap column (GE life sciences) at 1 ml/min. The column was washed with 5 CV of buffer A prior to elution in buffer B (25 mM Tris, pH 8, 200 mM NaCl, 500 mM imidazole). The peptide was cleaved from the sfGFP-expression tag by overnight incubation with 1:10 (w/w) of 3C protease at 4 °C. On the next day the mixture was concentrated to ~350 µl and subjected to gel filtration to separate sfGFP and 3C protease from the peptide as well as buffer exchange to NMR measurement buffer.

Table 1
Combinatorial selective labeling scheme employed to obtain complete backbone NH assignment of p63 PAD.

Amino acid type[a]	Labeling type[b]		
	Sample #		
	1	2	3
Alanine	^{15}N	2,3-$^{13}C_2$[c]	
Arginine	^{13}C/^{15}N	^{15}N	
Glutamine[d]		1-^{13}C	^{15}N
Glutamate	^{13}C/^{15}N	1-^{13}C	^{15}N
Glycine	^{13}C/^{15}N		^{13}C/^{15}N
Histidine	^{15}N	^{15}N	
Leucine	2-^{13}C		^{13}C/^{15}N
Phenylalanine		^{15}N	2-^{13}C
Proline		1-^{13}C	1-^{13}C
Serine	1-^{13}C	^{13}C/^{15}N	^{13}C/^{15}N
Threonine		^{13}C/^{15}N	
Tryptophan			^{15}N
Valine	^{15}N	^{13}C/^{15}N	

[a]Amino acid types not listed do not occur in the sequence of p63 PAD and were supplied at natural isotopic abundance to enable expression of the N-terminal sfGFP-tagged peptide. [b] ^{15}N and ^{13}C labels are color coded in blue and red, respectively. Amino acids selectively labeled at the backbone nitrogen, carbonyl and α-carbon are denoted ^{15}N, 1-^{13}C and 2-^{13}C, respectively. ^{13}C/^{15}N refers to fully ^{15}N and ^{13}C enriched species. Missing entries imply non-labeled amino acids. [c] ^{13}C labeled at the α- and β-positions. The additional label at C^{β} is not required and is of no experimental relevance. [d]Although the amino acid type does not occur in the sequence of p63 PAD labeled glutamine was added to samples 2 and 3 to avoid scrambling of unlabeled glutamine into glutamate.

Final peptide concentrations varied between 0.15 mM and 0.45 mM, as detailed in Table S1, in deuterated sodium acetate buffer (pH 4.7), containing 5% D_2O and 0.15 mM 4, 4-dimethyl-4-sila pentane-1-sulfonic acid (DSS) as internal chemical shift standard.

A set of four cyclophilin D (CypD) samples based on the combinatorial labeling scheme in Table 2 was prepared as described previously [35]. Protein concentrations varied between 0.5 and 0.6 mM (specified in Table S2) in a 50-mM sodium phosphate buffer (pH 7.0) containing 1 mM DTT, 5 mM EDTA, 0.15 mM DSS and 4% D_2O. All p63 PAD and CypD samples had a volume of 0.32 ml and were placed in 5-mm susceptibility-matched Shigemi tubes.

2.2. NMR spectroscopy

All spectra of p63 PAD were recorded at a sample temperature of 15 °C using a Bruker AV II 500 MHz spectrometer equipped with a room-temperature, three-axis gradient triple-resonance probe. For CypD the majority of experiments employed the same spectrometer, while HN(CACON) spectra of three of the four samples and all (H)N(COCO)NH spectra were recorded at Bruker AV II or AV III 600 MHz spectrometers equipped with cryogenic ^1H {$^{13}C/^{15}N$} triple-resonance z-gradient probes at a sample temperature of 30 °C.

Pulse sequences applied in this study were of the [^{15}N, ^1H]-TROSY [45–47] type and employed sensitivity-enhanced gradient echo/antiecho coherence selection [48–51]. Acceleration of longitudinal ^1H relaxation between scans was achieved in the Band-Selective Excitation Short-Transient (BEST) [52–54] manner using exclusively shaped proton pulses with bandwidths/offsets of 4.2/8.7 ppm for p63 PAD and 4.8/8.6 ppm for CypD. The delay between scans was set to 0.3 s in all experiments. The acquisition period was immediately followed by a ^{15}N 180° pulse in order to constructively add polarization from proton magnetization longitudinally relaxed during the pulse sequence and transferred to ^{15}N in the final single-transition-to-single-transition (ST2-PT) element [55].

Two-dimensional time-shared (ts) HN(CO)/HN(COCA) and SQ/DQ-HN(CA) experiments were performed essentially as described previously [35,56]. Linear combinations of transients acquired with different 90° ^{13}C pulse phases in an interleaved manner resulted in CO-filtered ^{15}N-HSQC, C^{α}-filtered HN(CO) and HN(COCA) subspectra from the ts-HN(CO)/HN(COCA), and in HN(CA) and double-quantum (DQ) HN(CA) [57] subspectra from the ts-SQ/DQ-HN (CA) data set. In addition, a pair of ^{15}N single-quantum (SQ) and double-quantum/zero-quantum (DQ/ZQ) HN(CACON) spectra was obtained with the pulse sequence shown in Fig. S1, and (H)N (COCON)H [58–60] spectra resulted from applications of the pulse sequence in Fig. S2 using 32 cycles of MOCCA-XY16 [61,62] which corresponds to ^{13}CO TOCSY mixing times of 285.4 ms for p63 PAD and 280.5 ms for CypD (the small discrepancy is due to the difference in ^{13}CO 180° pulse width at 500 and 600 MHz spectrometer frequencies, see legend to Fig. S2).

Experiments on p63 PAD were carried out with spectral widths of 11 and 20 ppm along the ^1H and ^{15}N dimensions, respectively, and an ^1H acquisition time of 81 ms. Spectral widths in experiments of CypD samples were 13 ppm (^1H) and 50 (^{15}N) ppm, with ^1H acquisition times of 69 ms (at 500 MHz) or 57 ms (at 600 MHz). Further experimental details regarding acquisition of all 2D spectra are listed in Tables S1 and S2. Three-dimensional HNCACB spectra were recorded of all three selectively labelled samples of p63 PAD to demonstrate the lack of ^{13}C chemical shift dispersion. Each spectrum was acquired in 18 h (8 scans/FID) using acquisition times of

Table 2
Four-sample combinatorial selective labeling scheme employed for CypD.

Amino acid type	Labeling type[a]			
	Sample #			
	1	2	3	4
Alanine	2,3-$^{13}C_2$[b]	$^{13}C/^{15}N$	^{15}N	$^{13}C/^{15}N$
Arginine	^{15}N	^{15}N	-	$^{13}C/^{15}N$
Asparagine	^{15}N	-	^{15}N	$^{13}C/^{15}N$
Aspartate	$^{13}C/^{15}N$	1-^{13}C	^{15}N	$^{13}C/^{15}N$
Cysteine	-	1-^{13}C	1-$^{13}C/^{15}N$[c]	1-$^{13}C/^{15}N$[c]
Glutamine	-	^{15}N	1-^{13}C	1-^{13}C
Glutamate	^{15}N	^{15}N	1-^{13}C	^{15}N
Glycine	2-^{13}C	2-^{13}C	$^{13}C/^{15}N$	$^{13}C/^{15}N$
Histidine	1-^{13}C	^{15}N	^{15}N	-
Isoleucine	-	$^{13}C/^{15}N$	^{15}N	^{15}N
Leucine	$^{13}C/^{15}N$	2-^{13}C	$^{13}C/^{15}N$	2-^{13}C
Lysine	^{15}N	^{15}N	$^{13}C/^{15}N$	1-^{13}C
Methionine	$^{13}C/^{15}N$	^{15}N	1-^{13}C	^{15}N
Phenylalanine	$^{13}C/^{15}N$	$^{13}C/^{15}N$	$^{13}C/^{15}N$	2-^{13}C
Proline	1-^{13}C	1-^{13}C	-	1-^{13}C
Serine	1-^{13}C	$^{13}C/^{15}N$	2-^{13}C	$^{13}C/^{15}N$
Threonine	-	$^{13}C/^{15}N$	$^{13}C/^{15}N$	-
Tryptophan	^{15}N	-	-	-
Tyrosine	$^{13}C/^{15}N$	-	2-^{13}C	-
Valine	$^{13}C/^{15}N$	$^{13}C/^{15}N$	-	^{15}N

^{15}N and ^{13}C labels are color coded in blue and red, respectively. Amino acids selectively labeled at the backbone nitrogen, carbonyl and α-carbon are denoted ^{15}N, 1-^{13}C and 2-^{13}C, respectively. $^{13}C/^{15}N$ refers to fully ^{15}N and ^{13}C enriched species. Where no labeling type is indicated, amino acids were supplied at natural isotopic abundance. [b]Alanine was supplied as ^{13}C labeled at both α- and β-positions instead of selectively $^{13}C^{\alpha}$ labeled. The label at C^{β} is not required and has no consequences for cross peak occurrence in any of the spectra recorded here. [c]In samples 3 and 4 cysteine was carbonyl-selectively ^{13}C and ^{15}N doubly labeled.

81 ms, 66 ms, and 10.5 ms in the ^{1}H, ^{15}N, and ^{13}C dimensions, respectively.

The rotational correlation of CypD was determined by the TRACT method [63], using a two-dimensional version to allow exclusion of very mobile residues.

Spectra processing and analysis was performed with TopSpin 3.5. Cosine-squared window functions were applied for apodization in all dimensions. Spectra were referenced with respect to internal DSS using consensus Ξ values for ^{13}C and ^{15}N [64]. Contour levels were drawn on an exponential scale using a factor of $2^{1/2}$ for preparation of all plots shown in the following.

3. Results and discussion

3.1. Purely combinatorial labeling based backbone amide assignment: Application to an unstructured peptide

The established protocol for obtaining backbone resonance assignments of isotopically labeled polypeptides relies on matching $^{13}C^{\alpha}$ [65], ^{13}CO [66], or $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ [67] chemical shifts in intraresidual and sequential correlations of three-dimensional HN-detected triple-resonance spectra. This approach might fail as a result of extensive ^{13}C chemical shift degeneracy, as often encountered for unstructured peptides, especially in the presence of sequence repeats. A typical example is the Phosphorylation Activation Domain of TAp63α (herein referred to as p63 PAD) comprising 35 amino acid residues, ten of which are serines. The full amino acid sequence is given in Fig. S3a. This sequence is crucial for activation of the tumor suppressor protein TAp63α in oocytes upon DNA damage. TAp63α is a transcriptionally inactive dimer, but upon DNA damage the PAD sequence becomes highly phosphorylated by the kinases Chk2 and CK1, leading to disruption of a charge-charge interaction network and subsequent tetramerization. The tetrameric protein is then able to trigger a DNA damage response similar to that of p53 [68]. While most of the 31 signals of the peptide in a ^{15}N-HSQC (three prolines and the N-terminal glycine are undetectable) are resolved, their sequence specific assignment is hampered by the lack of ^{13}C chemical shift dispersion. This becomes apparent from the ^{1}H-^{13}C strips of 3D HNCACB spectra of the three selectively labeled samples of p63 PAD shown in Fig. S3b.

Except for Ser9 and Thr16, that sequentially precede proline residues, all $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts are almost degenerate for a given amino acid type, implying that unambiguous assignments will be difficult to obtain using the classical approach. It is not unlikely that a distinction of individual amino acids based on ^{13}C chemical shifts would be possible using constant-time ^{13}C evolution periods and/or stronger magnetic fields. This would, however, require relatively long periods of high-field measurement time and was not attempted here. Also, alternative methods are available to avoid the problem of low ^{13}C chemical shift dispersion, most of which were successfully applied to the assignment of intrinsically disordered proteins where signal overlap is a major concern. Examples include the direct correlation of neighboring amide groups [69–72], the use of $^{1}H^{\alpha}$-detected triple-resonance experiments [73,74] or high-dimensional (≥4D) pulse sequences in combination with either non-uniform sampling [75–82] or automated projection spectroscopy [74,83]. Conversely, the aim of the current study is to achieve complete sequence-specific amide $^{1}H/^{15}N$ resonance assignments by exclusively relying on 2D proton-nitrogen correlation spectra acquired for a set of selectively labeled samples.

The objective of previous multiple-selective labeling methods [10,12,18,21,25,31,34,36,37,39–43,84] was the identification of amino acid types (providing information about a residue i) along with unique sequential pairs (residues i and $i-1$) for as many sites in a polypeptide chain as possible. Subsequently, these residues were used as anchor points to enable full assignments from standard 3D triple-resonance spectra in cases where the latter were difficult to analyze due to low sensitivity and/or resolution. While this complementary approach was successful in a variety of applications, it requires preparation of an additional uniformly $^{13}C/^{15}N$ labeled sample and time-consuming acquisition of 3D spectra. Recently we have shown that all 19 non-proline amino acid types can be distinguished with three samples containing amino acids that are isotope labeled at the backbone nitrogen, α-carbon and carbonyl carbon positions (either individually or simultaneously) in a combinatorial manner [35]. In that study, the spectroscopic deconvolution of the various isotopomeric species occurring in each sample was achieved with the help of six 2D HN(CX) type triple-resonance spectra resulting from one stand-alone and two time-shared [56] experiments. Considering that the p63 PAD peptide comprises only a subset of twelve amino acid types with 31 backbone amides detectable in a ^{15}N-HSQC, it might appear that its complete assignment can be readily achieved via combinatorial labeling. However, due to the occurrence of several identical amino acid pairs in the sequence, only 22 signals, which include no more than three of the ten serine residues, could in fact be assigned using our previous protocol.

A suitable triple-selective labeling scheme for sequence specific assignment of the 22 residues involved in unique pairs is given in Table 1. Two-dimensional spectra for each of the three samples are plotted in Fig. 1 to demonstrate the problem of sequence degeneracy. Shown are the entire spectra resulting from applications of ts-HN(CO)/HN(COCA) and ts-SQ/DQ-HN(CA) pulse sequences. The former generates cross peaks for all amino acids that have a ^{15}N label at the backbone nitrogen and edits them into three subspectra depending on whether the preceding residue is not ^{13}C labeled at the carbonyl group (CO-filtered HSQC), exclusively labeled at the carbonyl group (C^{α}-filtered HN(CO)) or labeled at both the carbonyl and C^{α} positions (HN(COCA)). The second experiment yields cross peaks in the HN(CA) subspectrum for all ^{15}N labeled amino acids with a ^{13}C labeled α-carbon in the same or the preceding residue or both, and additionally in the DQ-HN(CA) subspectrum when both residues are $^{13}C^{\alpha}$ labeled.

Of the total of 31 residues observed, 22 show a distinct pattern in the 15 spectra, immediately resulting in their sequential assignment. Remaining ambiguities arise from degeneracies of $i-1/i$ pairs Gly3/Ser4 and Gly26/Ser27, Ser4/His5 and Ser11/His12, Pro10/Ser11 and Pro17/Ser18 and the three Ser/Ser pairs in positions 8/9, 18/19 and 27/28. It should be noted that combinations of selectively labeled amino acids other than that of Table 1 would also be conceivable to obtain the same information, but none of them would provide further assignments if only the identities of amino acid types in positions i and $i-1$ are established. Likewise, addition of further samples would not solve the problem. Rather, supplemental experiments are required to gain information about adjacent residues, i.e. in positions $i+1$ and $i-2$. The corresponding magnetization transfer pathways are depicted in Fig. 2.

In the first experiment (Fig. 2a), degenerate $i-1/i$ pairs may be distinguished based on the presence or absence of a ^{15}N label in the sequentially following residue. A suitable pulse sequence, termed HN-XU, has recently been introduced by Atreya and co-workers [43] and employed in conjunction with a selective unlabeling scheme to selectively observe $^{13}C/^{15}N$-labeled residues followed by unlabeled ones. Apart from its BEST-TROSY implementation the main difference of the pulse sequence version detailed in Fig. S1 is that signals from $^{15}N_i$-$^{13}C^{\alpha}_i$-$^{13}CO_i$-$^{15}N_{i+1}$ and $^{15}N_i$-$^{13}C^{\alpha}_i$-$^{13}N_i$-$^{13}C^{\alpha}_i$-$^{13}CO_i$-$^{14}N_{i+1}$ moieties are both retained and edited into different subspectra. Following the relay of magnetization from the amide proton of residue i to its own carbonyl, double antiphase

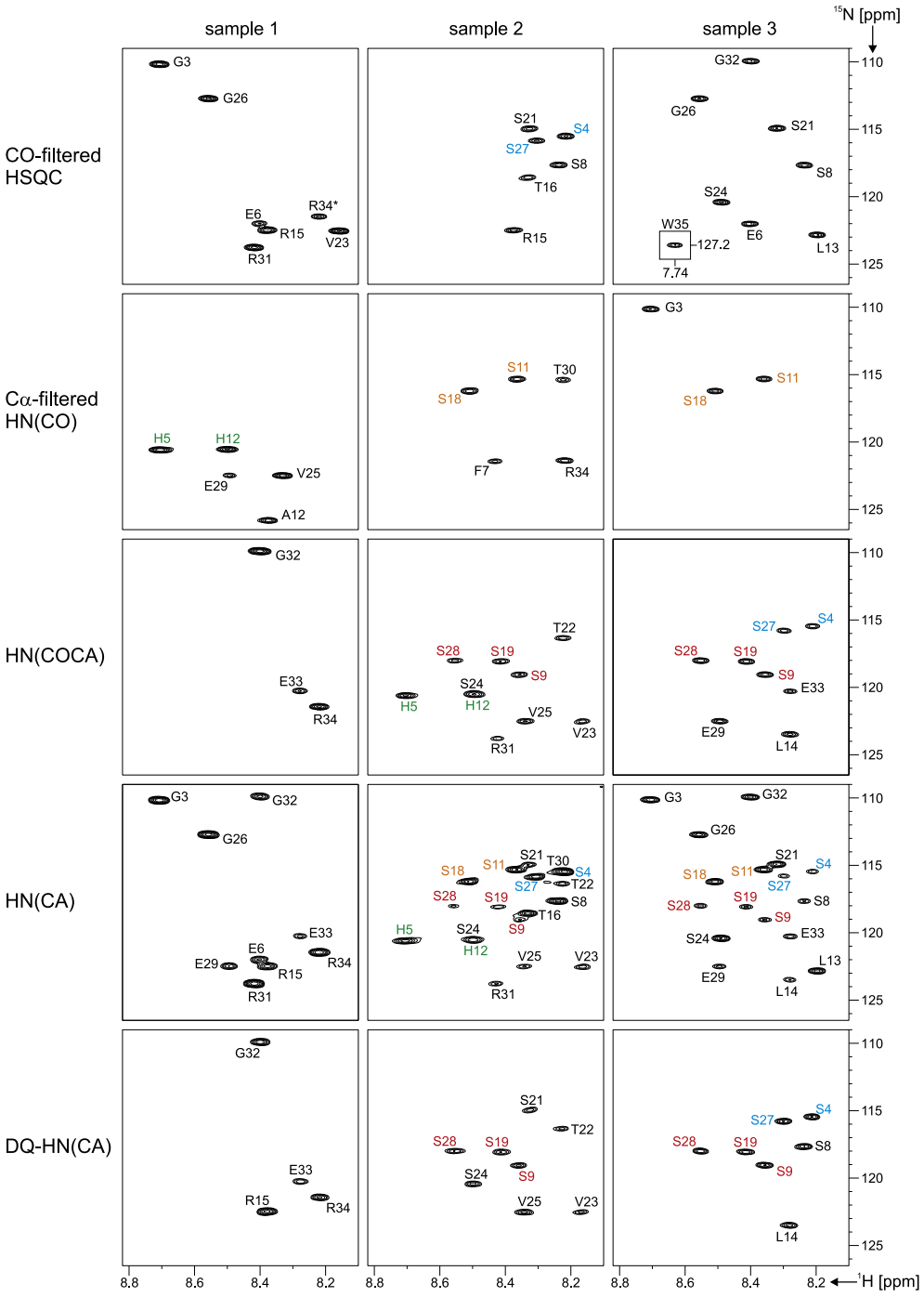

Fig. 1. Two-dimensional ^{1}H-^{15}N correlation spectra obtained by application of ts-HN(CO)/HN(COCA) (upper three panels) and ts-SQ/DQ-HN(CA) experiments to the three combinatorial selectively labeled samples (Table 1) of p63 PAD. Peak annotations (residue *i*) in black indicate unique *i*, *i* − 1 pairs, resulting in unambiguous assignments. Mutually degenerate pairs, showing identical peak presence/absence patterns in the current set of spectra, are highlighted by colored labels. The inset in the CO-filtered HSQC of sample 3 shows a cross peak for the backbone of W35, which resonates outside the plotted region. An asterisk marks a cross peak of R34 in the CO-filtered HSQC of sample 1 which appears due to scrambling of unlabeled glutamine into glutamate, such that the preceding residue is a mixture of unlabeled and fully ^{13}C/^{15}N labeled glutamate. Note that this undesired signal would be suppressed by use of ^{13}C/^{15}N labelled glutamine during in-vitro expression of the peptide. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

^{13}CO magnetization with respect to ^{15}N$_i$ and ^{15}N$_{i+1}$ builds up during the central ^{13}CO-^{15}N HMQC-type element if residue *i* + 1 is ^{15}N labeled, whereas a single ^{13}CO antiphase state with respect to ^{15}N$_i$ will be present otherwise. At this time point a pair of 90° ^{15}N pulses is applied where the phase of second is always x whereas the phase of the first is alternated between x and −x and transients are stored separately. Addition and subtraction of the two experiments selects ^{15}N$_i$-^{15}N$_{i+1}$ double/zero-quantum (DQ/ZQ) and ^{15}N$_i$ single-quantum (SQ) coherence, respectively, during the short time between the two ^{15}N pulses. Therefore, the sum (DQ/ZQ) spectrum contains NH$_i$ cross peaks for those residues where residue *i* + 1 is ^{15}N labeled and the difference (SQ) spectrum contains NH$_i$ cross peaks exclusively for those residues where residue *i* + 1 is not ^{15}N labeled. The advantage of retaining signals of all

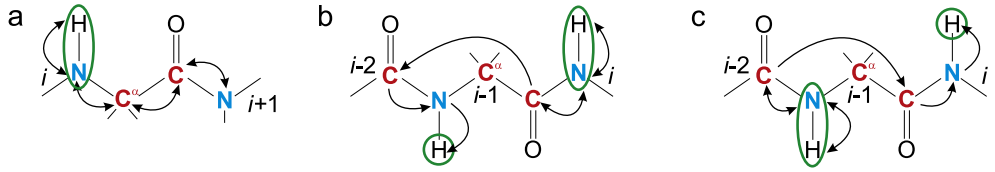

Fig. 2. Magnetization transfer pathways of experiments to determine whether an amino acid in position $i+1$ is ^{15}N labeled (a: HN(CACON)) and whether an amino acid in position $i-2$ is 1-^{13}C labeled (b, c: (H)N(COCON)H). Isotope labeled carbons and nitrogens are drawn in red and blue, respectively. Note that labeling of the α-carbon in residue $i-1$ in schemes a and b is not mandatory, i.e. fully ^{13}C/^{15}N labeled amino acids at this position may be replaced by selectively ^{15}N and 1-^{13}C labeled amino acids. Green ellipses and circles indicate the nuclei whose frequencies are correlated in the 2D spectra. The (H)N(COCON)H experiment is bidirectional and therefore provides cross peaks between $^{1}H_{i-1}$ and $^{15}N_i$ (scheme b) as well as between $^{1}H_i$ and $^{15}N_{i-1}$ (scheme c) in the presence of a ^{13}C label at the carbonyl of residue $i-2$. In addition, $^{1}H_i/^{15}N_i$ and $^{1}H_{i-1}/^{15}N_{i-1}$ auto-correlation peaks are detected via the HN(CO) out-and-back type pathway. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

^{13}C/^{15}N labeled residues in either of the subspectra rather than filtering out those with ^{15}N labeled C-terminal neighbors, is that the HN(COCAN) experiment can thus serve to distinguish $^{12}C^{\alpha}_{i-1}$-$^{12}CO_{i-1}$-$^{15}N_i$-$^{13}C^{\alpha}_{i}$-$^{13}CO_i$ from $^{13}C^{\alpha}_{i-1}$-$^{12}CO_{i-1}$-$^{15}N_i$-$^{12}C^{\alpha}_{i}$-$^{12}CO_i$ dipeptides. This distinction would otherwise require recording of an extra CO-filtered HN(CA) experiment [41] which now becomes redundant.

Pairs of SQ and DQ/ZQ HN(CACON) subspectra obtained for each of the three samples of p63 PAD are shown in Fig. 3. They contain all expected signals for the particular fully ^{13}C/^{15}N labeled residue types, sorted according to whether the nitrogen nucleus of the sequentially following residue is ^{14}N (SQ subspectrum) or ^{15}N (DQ/ZQ subspectrum). This immediately allows the unambiguous assignment of residues Ser4 and Ser27 that were indistinguishable in the previously recorded set of spectra [35,41] (Fig. 2), as well as residues Ser11 and Ser18 from the HN(CACON) spectra of sample 3. Furthermore, Ser28 can now be assigned as it has a ^{15}N labeled C-terminal neighbor, contrasting Ser9 and Ser19.

Using the present labeling scheme, additional information will be required to obtain assignments in the case of higher degrees of degeneracy as, for instance, encountered for the three Ser/Ser pairs in p63 PAD. This is accounted for by incorporating the residue types in positions $i-2$ of Ser9 and Ser19, respectively, without (Phe) and with (Pro) a ^{13}C label at the carbonyl group. Probing whether a remote residue carries a ^{13}CO label is made possible by the (H)N(COCO)NH experiment [58–60] which involves ^{13}C isotropic mixing via small three-bond ^{13}CO-^{13}CO scalar couplings in the polypeptide backbone. Magnetization can thus be relayed from the amide of residue i via the carbonyl of residue $i-1$ to the carbonyl of residue $i-2$ and finally detected on the amide of residue $i-1$ (Fig. 2b), provided that residue $i-1$ is fully ^{13}C/^{15}N labeled and residue $i-2$ is ^{13}CO labeled. The two-dimensional (H)N(COCON)H BEST-TROSY version of the pulse sequence used in the current study is shown in Fig. S2. It gives rise to rectangular cross peak patterns spanned by HN(CO)-like auto peaks at the usual $^{15}N_i$-$^{1}H_i$ and $^{15}N_{i-1}$-$^{1}H_{i-1}$ resonance positions and additional $^{15}N_i$-$^{1}H_{i-1}$ (Fig. 3b) and $^{15}N_{i-1}$-$^{1}H_i$ (Fig. 2c) correlations indicating the presence of a $^{13}CO_{i-2}$-$^{15}N_{i-1}$-$^{13}C^{\alpha}_{i-1}$-$^{13}CO_{i-1}$-$^{15}N_i$ fragment.

The application of the (H)N(COCON)H experiment to p63 PAD is illustrated in Fig. 4. In the samples shown here, serines are ^{13}C/^{15}N

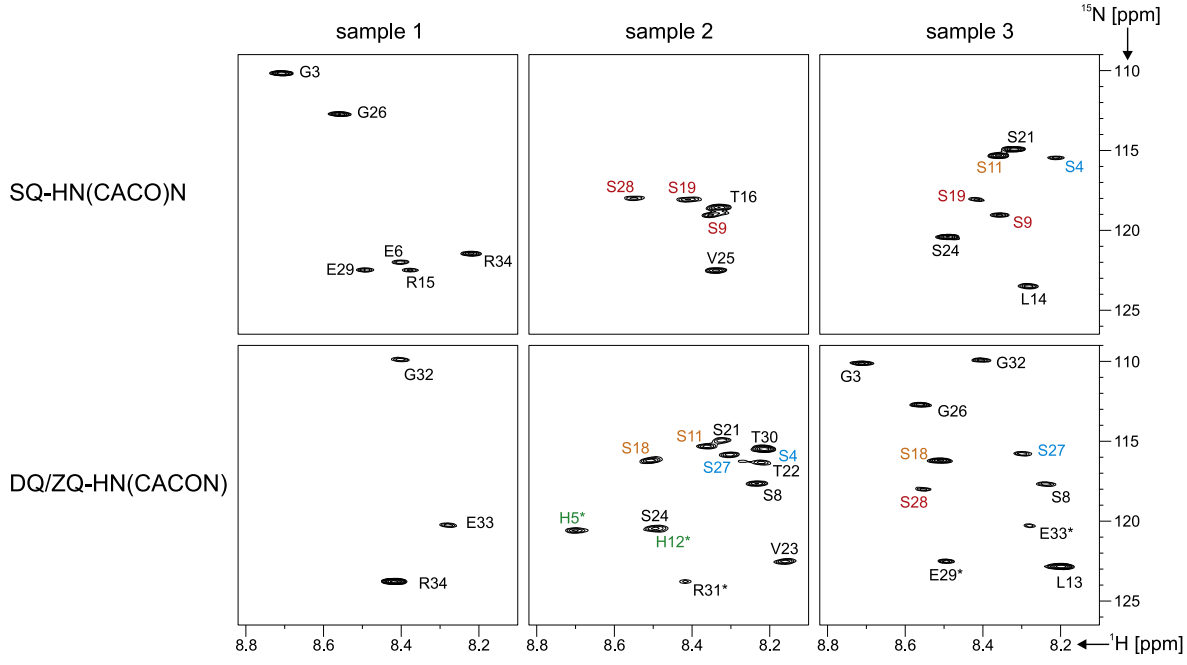

Fig. 3. Differentiation of $^{15}N_i$-$^{13}C^{\alpha}_{i}$-$^{13}CO_i$-$^{14}N_{i+1}$ and $^{15}N_i$-$^{13}C^{\alpha}_{i}$-$^{13}CO_i$-$^{15}N_{i+1}$ moieties in p63 PAD using the HN(CACON) experiment. The former and latter are exclusively observed in SQ and DQ/ZQ subspectra, respectively. Asterisks mark additional sequential cross peaks of ^{15}N labelled residues preceded by fully ^{13}C/^{15}N labelled ones, for which $^{13}CO_{i-1}$ antiphase magnetization with respect to $^{15}N_i$ is converted into in-phase magnetization, which is not separated from double antiphase ^{13}CO magnetization (giving rise to cross peaks in DQ/ZQ subspectra) by the phase cycling employed here. Their appearance does not obstruct the analysis of the HN(CACO)N as these signals are unambiguously identified by the presence/absence of cross peaks in HN(COCA)/DQ-HN(CA) spectra (see Fig. 1). Degenerate $i, i-1$ pairs are annotated with the same colors as in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

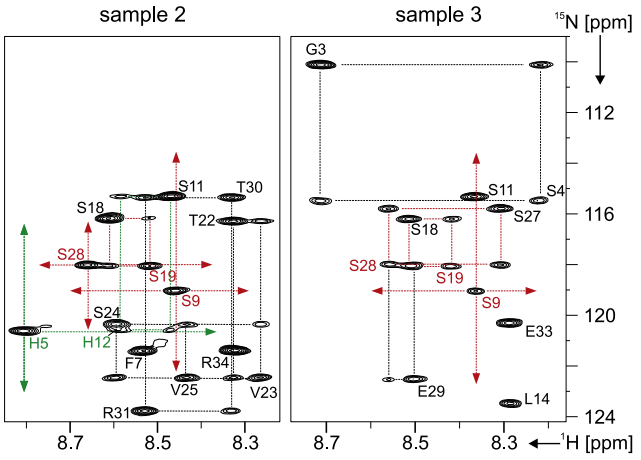

Fig. 4. Identification of $^{13}CO_{i-2}$-$^{15}N_{i-1}$-$^{13}C^{\alpha}_{i-1}$-$^{13}CO_{i-1}$-$^{15}N_i$ fragments in p63 PAD. Plotted are 2D (H)N(COCON)H spectra recorded with 285.4-ms MOCCA-XY16 mixing periods at 500 MHz ^{1}H frequency. Assignments are only provided for the HN (CO) auto-correlation peaks. These are connected with $^{15}N_i$-$^{1}H_{i-1}$ and $^{15}N_{i-1}$-$^{1}H_i$ cross peaks by dashed lines, indicating the presence of a ^{13}C labeled carbonyl group in the N-terminal neighbor. If residue $i-2$ is not labeled at the carbonyl, no such correlations are observed, as visualized by dashed arrows. The differentiation of the three Ser_{i-1}/Ser_i pairs at positions i = 9, 19 and 28, and of the two Ser_{i-1}/His_i pairs at positions i = 5 and 12 is highlighted in red and green, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

labeled and the phenylalanine residue at position $i-2$ from Ser9 is not carbonyl labeled. Consequently, no additional correlations are observed at the F_1 (^{15}N) or F_2 (^{1}H) chemical shifts of Ser9. Both proline ($i-2$ form Ser19) and glycine ($i-2$ from Ser28) contain a ^{13}C label at the carbonyl in sample 3, resulting in the expected rectangular cross peak patterns that connect the two serine residues with the respective sequentially preceding ones. In sample 2, only proline is 1-^{13}C labeled such that Ser19 shows additional cross peaks sharing the F_1 and F_2 coordinates with the preceding residue, whereas Ser9 and Ser28 do not. Individual assignments for each of the three degenerate Ser/Ser pairs are thus obtained with the help of the (H)N(COCON)H method. Likewise ambiguous from the previous set of 2D spectra was the assignment of the degenerate pairs Ser4/His5 and Ser11/His12. It should be noted that a distinction would have been possible based on the HN(CACON) experiment, provided that ^{15}N labeled histidine had been replaced by ^{13}C/^{15}N labeled histidine in sample 1, where only one of the following residues, i.e. Glu6, but not Leu13 is ^{15}N labeled. Incorporation of expensive doubly labeled histidine was however avoided here. Instead, the (H)N(COCON)H spectrum of sample 2 allows straightforward assignment of the two histidine signals based on the fact that Pro10 is 1-^{13}C labeled whereas Gly3 is not labeled in this sample, such that only His12 shows the corresponding connectivities with its N-terminal neighbor.

To summarize, exclusively using combinatorial labeling complete NH assignments were obtained for the p63 PAD peptide which features a repetitive sequence due to its 30% composition of serine residues. The general assignment strategy is depicted schematically in Fig. 5. For each backbone NH_i of a given amino acid sequence (top) the expected cross peak occurrence pattern (bottom right) can be compiled from the list of experiments (left side) and the chosen labeling scheme (top right). Using four different labeling types, i.e. ^{15}N, 1-^{13}C, 2-^{13}C and fully ^{13}C/^{15}N, a total of 15 isotopomers involving the detected residue i, the adjacent residues $i-1$ and $i+1$ as well as the remote residue $i-2$, can be identified from a suite of eight proton-nitrogen correlated 2D spectra, representing the vertical dimension of the grid shown at the bottom right. Seven tetrapeptide fragments are exemplarily shown here; an overview of all isotopomers and the corresponding spectral signature in terms of presence or absence in these spectra is provided in the supplementary material (Table S3). Since the number of amino acid types highly exceeds the number of labelling types employed and at the same time the number of residues highly exceeds the number of isotopomeric tetrapeptides that can be distinguished, this type of information must be multiplexed by the use of several amino-acid specifically labeled samples to provide the horizontal dimension of the grid. For instance, residues C32, G42, and K53 reveal the same pattern in the first column of the table at the bottom right (sample 1) and can only be differentiated from the set of spectra recorded on samples 2 and 3. In order to avoid preparation and analysis of an excessive number of samples selective labeling is implemented in a combinatorial manner. Thus, the experimental cross peak occurrence at each ^{1}H-^{15}N position in the set of spectra of all samples must be compared with the expected patterns of individual residues of a polypeptide. Whenever a match is found and the pattern is unique, the corresponding NH cross peak is sequence-specifically assigned.

3.2. Sequence-specific NH assignment exclusively by combinatorial selective labeling – a feasible approach for folded proteins?

Having demonstrated how information about sequence positions $i+1$ and $i-2$ complements our previous combinatorial labeling protocol to achieve complete amide group resonance assignments of a peptide it shall be explored in the following to what extent the results are transferable to a protein. Two aspects are relevant here: (1) the much higher number of amino-acid type combinations that need to be distinguished with a given labeling scheme and (2) the increased transverse relaxation rates that might render application of the two additional, lengthy pulse sequences to a folded protein inefficient. As a test case we have chosen the 165-residue protein cyclophilin D, a peptidyl-prolyl *cis/trans* isomerase for which inhibitor binding was previously investigated in our laboratory [35]. In that study, assignments were obtained with the aid of a three-sample combinatorial labeling scheme, directly resulting in sequence specific assignments for 37% of the 152 residues detectable in a ^{15}N HSQC, and the identification of all amino acid types. Assignments were completed by recording 3D HNCA and HN(CO)CA spectra on a uniformly ^{13}C/^{15}N labeled sample.

Use of a uniformly labeled sample and acquisition of three-dimensional data sets was dispensed with in the current work. Instead, a combinatorial selective labeling scheme was devised which comprises four samples (Table 2) and represents a compromise to allow collection of a sufficient amount of information while keeping sample preparation efforts at a manageable level. Analysis of six 2D ^{1}H-^{15}N correlated spectra, i.e. CO-filtered HSQC, C^{α}-filtered HN(CO), HN(COCA), HN(CA), DQ-HN(CA), and CO-filtered HN(CA) – in analogy to previous applications [35,41]– potentially provides sequence specific assignments of 70 residues (46%) of CypD. For a further 52 residues (34%) two possible assignments exist, while higher degrees of ambiguity (up to six-fold) apply for the remaining residues (see Fig. 6, red bars, for statistics). Among the 82 unassigned residues 73 are involved in non-unique i, $i-1$ pairs, a priori ruling out their assignment with the above set of experiments, irrespective of the labeling scheme and the number of samples employed.

In order to assess to what extent ambiguities can be resolved by the discrimination of ^{14}N/^{15}N nuclei in residues $i+1$, the HN(CACON) experiment is included in the next step. Note that this allows abandoning the CO-filtered HN(CA) experiment because $^{13}C^{\alpha}_{i-1}$-$^{12}CO_{i-1}$-$^{15}N_i$-$^{12}C^{\alpha}_{i}$-$^{12}CO_i$ moieties are now identified by their absence in both HN(CACON) subspectra. Compared to the unstructured p63 PAD peptide, application of the HN(CACON)

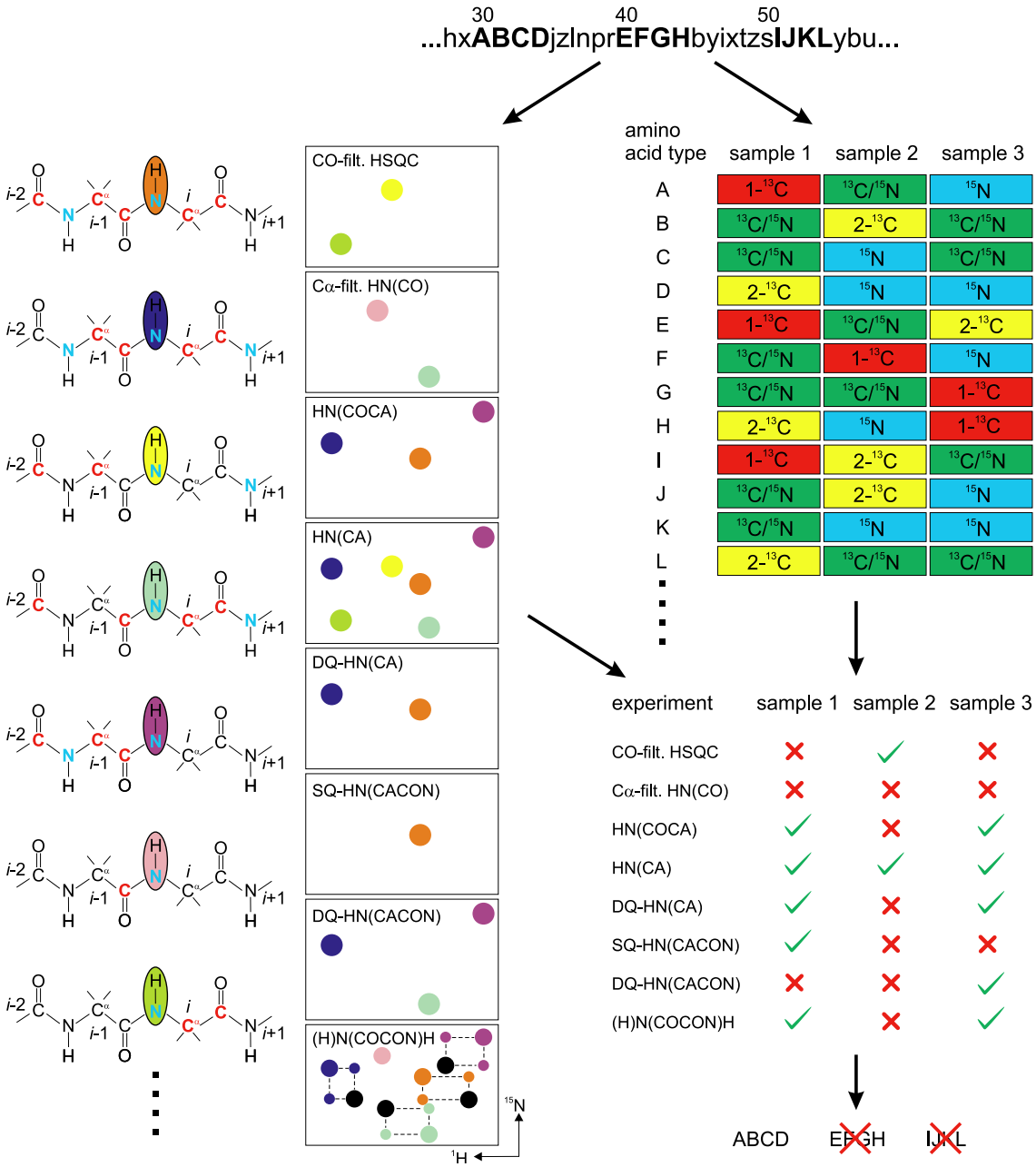

Fig. 5. Summary of assignment strategy employed here. Top: Section from a hypothetical amino acid sequence. Arbitrary letters, not corresponding to one-letter amino acid codes, are used. Left: Examples of polypeptide stretches comprising residues $i-2$, $i-1$, i, $i+1$ and their fingerprint in the set of eight 2D ^{1}H-^{15}N correlation spectra used for identification of the tetrapeptides. ^{13}C and ^{15}N nuclei are indicated in red and blue, respectively. Black dots in the (H)N(COCON)H spectrum are auto peaks from sequential neighbors of the colored amide groups, connected to the latter by (smaller) $^{15}N_i$-$^{1}H_{i\pm1}$ and $^{15}N_{i\pm1}$-$^{1}H_i$ cross peaks. Top right: Hypothetical three-sample combinatorial labeling scheme. Bottom right: Cross peak occurrence pattern for a residue i derived from the array of experiments (vertical dimension) and the labeling type in each sample (horizontal dimension). Of the three $i-2$, $i-1$, i, $i+1$ tetrapeptides marked in the sequence by upper case letters, only one matches the full grid, resulting in the assignment of the amide of C32 (residue i in the tetrapeptide ABCD).

experiment to a folded protein resulted in a larger variation of signal intensities due to differential transverse relaxation of its backbone nuclei and required longer accumulation times to observe the weaker cross peaks. Nevertheless, for CypD assignments could be extended to 122 residues (80%) and the degree of ambiguity for the remaining residues reduced (Fig. 6, green bars).

Three cases where the pair of HN(CACON) subspectra resolved ambiguous assignments in CypD are highlighted in Fig. 7. (1) Threonine residues 41 and 116 are both preceded by a cysteine and therefore indistinguishable with the previously applied set of experiments (data not shown). However, Thr41 is sequentially followed by a glycine, which is 2-^{13}C-labeled in sample 2, while Thr116 is sequentially followed by a fully ^{13}C/^{15}N-labeled isoleucine in the same sample. Consequently, cross peaks for Thr41 and Thr116 are exclusively observed in the SQ and DQ/ZQ subspectra, respectively, of the HN(CACON) (peak annotations printed in blue in Fig. 7a). (2) Individual assignments for the two degenerate $i-1/i$ pairs Ser59/Phe60 and Ser144/Phe145 (red peak annotations) can be obtained with the help of the HN(CACON) spectra of either sample 1 or sample 2, since Phe60 is followed by a

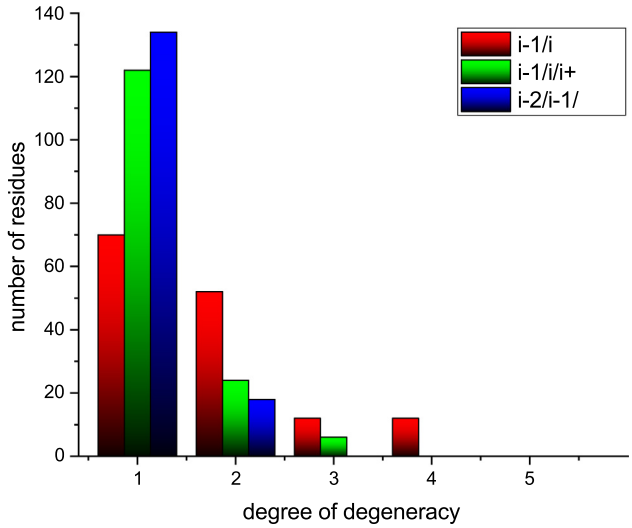

Fig. 6. Uniqueness of CypD NH assignments accomplished with the four-sample combinatorial labeling scheme of Table 2 using different sets of experiments. The height of bars in the histogram indicates the distribution of residues for which either unambiguous sequential information is obtained or up to six possible assignments exist. Red bars: using only information about the amino acid type of residues i and $i-1$; green bars: after including information about residue $i+1$; blue bars: additionally including information about residue $i-2$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

methionine which carries a ^{15}N label in both samples, whereas Phe145 is followed by a glycine, which is not ^{15}N labeled. (3) Due to the binary information about the labeling state of residue $i+1$ provided by the HN(CACON) experiment, combined information from both samples is required when more than two $i-1/i$ pairs are indistinguishable, as found for the three phenylalanine residues in positions 36, 83, and 88 (green peak annotations). Although only Phe36 and Phe88 are preceded by the same residue type, i.e. asparagine, the degree of ambiguity further increases because Phe83 is preceded by arginine, which cannot be distinguished from asparagine in position $i-1$ with the current labeling scheme (Table 2). The reason is that both residue types carry no ^{13}C label in samples 1–3, where phenylalanine cross peaks can be observed (spectra of sample 3 not shown here). For clarity, the labeling pattern of the three phenylalanine residues and their N- and C-terminal neighbors is depicted in Fig. 7b. In both samples 1 and 2, Phe36 and Phe83 form peptide bonds with ^{15}N and ^{14}N nuclei, respectively, and are therefore exclusively detected in either the DQ/ZQ (Phe36) or SQ (Phe83) subspectra. In contrast, Phe88 is observed in the SQ subspectrum of sample 1 and the DQ/ZQ subspectrum of sample 2, revealing that its C-terminal neighbor is ^{15}N labeled in sample 2 but not in sample 1. The three phenylalanine residues are thus readily differentiated by their fingerprints in the HN(CACON) experiment.

Expanding the combinatorial triple-selective labeling approach with the HN(CACON) technique involves analysis of seven ^{1}H-^{15}N correlated spectra that result from three time-shared pulse sequences. In the case of the four-sample labeling scheme of CypD amide resonance assignments are therefore obtained by evaluating the presence or absence of cross peaks at the corresponding ^{1}H and ^{15}N resonance positions in a total of 28 2D spectra. Each finding of a unique pattern in a 7×4 matrix of the type shown in Fig. 8 contributes an unambiguous sequence specific assignment. Representative examples are Leu39 (Fig. 8a) and Ile114 (Fig. 8b). For 30 residues of CypD, however, non-unique patterns were observed, as for asparagines 13 and 102 depicted in Fig. S4. Of these 30 residues, eight are at the central position of degenerate $i-1/i/i+1$ amino–acid type triplets, implying that their indistinguishability is labeling-scheme independent. As demonstrated in the following, the (H)N(COCON)H method is suitable to resolve such ambiguities not only for an unstructured polypeptide but for a folded protein such as CypD, too.

The 2D BEST-[^{15}N,^{1}H]-TROSY-(H)N(COCON)H pulse sequence was applied at 600 MHz to samples 3 and 4 of CypD, where most of the above mentioned ambiguities occurred. An expansion of the spectrum of sample 4 is shown in Fig. 9. Serine residues 110 and 147 are both preceded by a glycine and followed by a non-^{15}N labeled residue and therefore not assigned individually thus far. However, as illustrated below the spectrum, the residue in $i-2$ position of Ser110 is ^{13}C labeled at the carbonyl group, whereas the amino acid in the corresponding position of Ser147 is not. Consequently, no $^{15}N_i$-$^{1}H_{i-1}$ or $^{15}N_{i-1}$-$^{1}H_i$ cross peaks are detected for Ser147, contrasting Ser110 where the expected connectivities to Gly109 are observed. A slightly different situation is encountered for the Asn13/Asn102 pair, which revealed identical patterns in the experiment vs. sample matrices of Fig S4. Again, amino acids in position $i-2$ are carbonyl labeled in one case and non-carbonyl labeled in the other, in principle allowing a differentiation. However, both asparagines are sequentially followed by a ^{13}C/^{15}N labeled residue, resulting in additional C-terminal correlations in the (H)N(COCON)H spectrum at the ^{1}H and ^{15}N coordinates of Asn13 and Asn102. Although it cannot a priori be decided whether these correlations point into the N- or the C-terminal direction, connectivities involving Ala12, Gly14, Ala103, and Gly104 immediately result in sequence specific assignments for Asn13 and Asn102 because the resonance positions of the adjacent residues are known from the previously applied set of experiments. This example illustrates an attractive feature of the (H)N(COCON)H experiment not mentioned so far: assignment information is not only derived from the presence or absence of signals, but the observed cross peaks provide direct sequential links to already assigned residues. An example from sample 3 of CypD is presented in the supplementary material (Fig. S5).

Including information from the (H)N(COCON)H experiment raises the fraction of unambiguous assignments in CypD to 88% with the remaining 18 residues subject to two-fold degeneracies (Fig. 6, blue bars). Reasons for the incomplete assignment are (1) the fact that (H)N(COCON)H spectra were only available from two of the four samples, (2) the lack of suitable isotopomers to resolve ambiguities in the labeling scheme used in this study, and (3) the absence of expected correlations in (H)N(COCON)H experiments. The latter issue is of general concern in the (H)N(COCON)H method as it critically relies on a magnetization transfer via three-bond ^{13}CO-^{13}CO scalar couplings that show a Karplus-type dependence on the backbone torsion angle φ [85,86]. Applications to intrinsically disordered proteins took advantage of conformationally averaged $^{3}J_{C'C'}$ coupling constants resulting in uniform transfer efficiencies along the backbone [59,60]. In folded proteins, however, this scalar interaction can approach zero in α-helical regions, preventing the use of the (H)N(COCON)H experiment for distinction of certain $i-2$, $i-1$, i amino acid type combinations.

We have successfully applied the extended combinatorial-labeling based assignment protocol to a protein tumbling with a rotational correlation time of 8.2 ns. The size of proteins for which the approach is usable is clearly limited by the sensitivity of the newly included experiments, HN(CACON) and (H)N(COCON)H, with the upper limit of the correlation estimated to be roughly 10 ns. Similar to other backbone assignment methods, deuteration at non-exchangeable sites can shift this limit to larger molecular sizes. However, it should be noted that the performance of the HN(CACON) and in particular the (H)N(COCON)H experiment depends to a significant degree on carbonyl relaxation which is marginally affected by deuteration. For the same reason, application of these experiments at higher static fields may be counter-

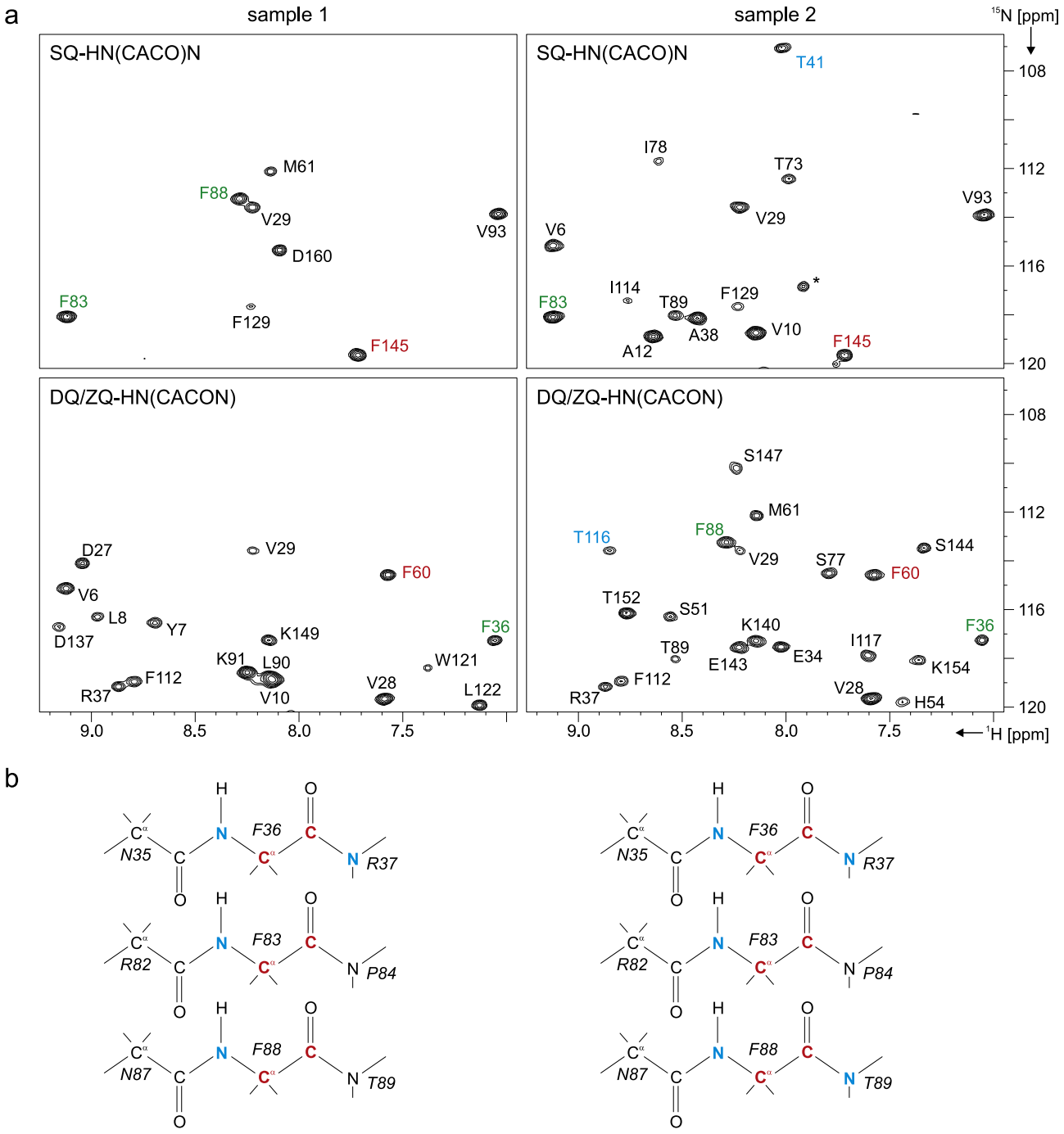

Fig. 7. Unambiguous assignment of residues in non-unique $i-1/i$ pairs in CypD with the help of the HN(CACON) experiment. (a) SQ and DQ/ZQ subspectra of samples 1 and 2 in the labeling scheme of Table 2. Assignments for residues discussed in the text are indicated with colored labels. For Val29 and Thr89, which are sequentially followed by non-^{15}N-labeled amino acids and therefore only expected in SQ subspectra, weak additional correlations are observed in DQ/ZQ subspectra arising from the magnetization transfer pathway via $^{2}J(^{15}N,^{13}C^{\alpha})$ couplings involving $^{13}C/^{15}N$-labeled residues in position $i-1$. The cross peak labeled with an asterisk is due to a degradation product. (b) Labeling patterns of Phe36, Phe83, and Phe88 and the relevant parts of their adjacent residues in samples 1 (left) and 2 (right). ^{15}N and ^{13}C labeling is indicated in blue and red, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

productive. The incorporation of deuterated amino acids using cell-free expression to allow assignment of large systems with the help of triple-selective combinatorial labeling has been shown to be feasible [42], it is however compromised by the limited commercial availability of deuterated and site-specifically labeled amino acids.

Since exclusively 2D spectra were employed in the current implementation, resolution is a critical issue, too. While ^{13}C chemical shift degeneracy is irrelevant here, any overlap in the ^{1}H-^{15}N correlation map will lead to ambiguities. Although alleviated to some extent owing to amino-acid specific labeling, a well-dispersed ^{15}N HSQC is prerequisite for the success of combinatorial labeling when aiming at sequence-specific assignments. Therefore, the method proposed here is preferentially applied to folded proteins, ideally with a high content of β-sheet secondary structure or a mixture of both α-helical and β-sheet secondary structure elements. It is not expected to be particularly useful for the assignment of large IDPs.

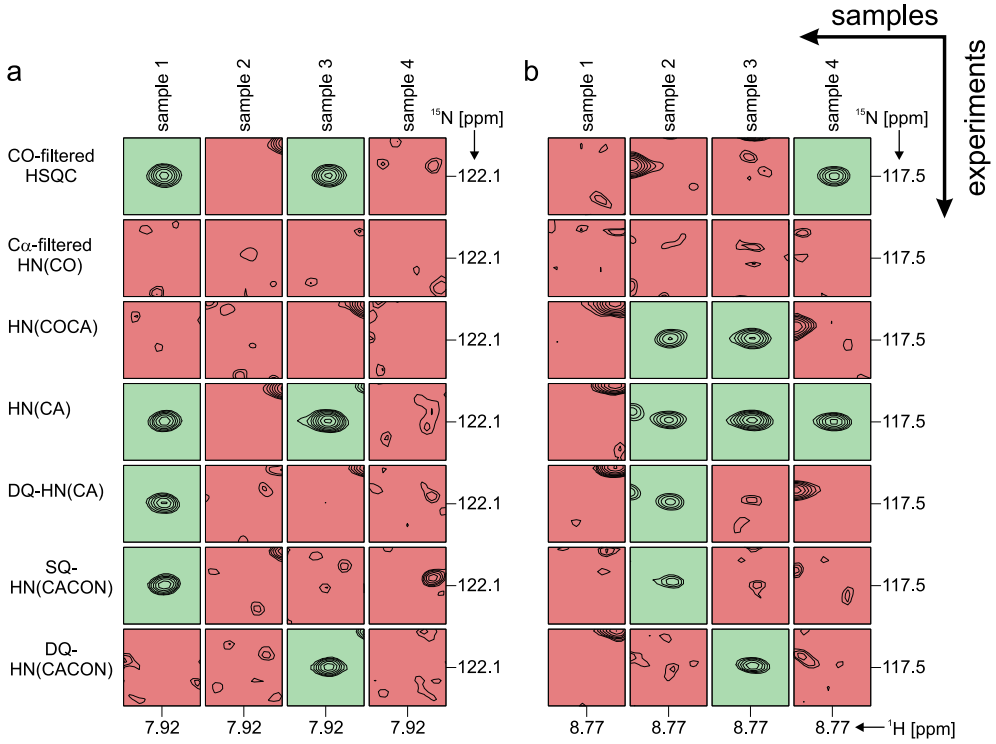

Fig. 8. Array of 2D HN(CX) spectra of the four selectively labeled samples of CypD at the peak positions of (a) Leu39 and (b) Ile114. Boxes are centered at the corresponding amide chemical shifts and have a width of 0.2 ppm along the ^{1}H dimension and 2 ppm along the ^{15}N dimension. Spectra expansions are shown with a green background if a signal is expected in the center and with a red background if it should not contain a signal based on the current labeling scheme. In the latter case, the lowest plot level is chosen to be within the noise to demonstrate the absence of a cross peak.

The assembly of triple-selective combinatorial labeling schemes needs the amino acid sequence of the target protein to be taken into account. Considerations relevant for the identification of individual $i-1$, i dipeptides were given in previous publications [35,40,41] and shall not be reiterated here. As a general guideline it should be added that a maximum of sequence-specific assignments is scored when abundant amino acid types are ^{15}N and/or fully ^{13}C/^{15}N labeled in more samples than less abundant ones. This contrasts the strategy followed for amino-acid type identification with the help of a combinatorial ^{15}N-labeling scheme, where the number of cross peaks in ^{15}N-HSQC spectra was sought to be minimized in order to optimize resolution [32]. The additional identification of sequential neighbors however necessitates a proportional increase of the fraction of samples in which amino acids are detectable, i.e. carry a ^{15}N label [31]. Further differentiation based on the labeling state of residue $i+1$ with the help of the HN(CACON) experiment requires the presence of fully ^{13}C/^{15}N labeled residues i in at least one of the samples, and only one of the sequentially following amino acid types of two degenerate $i-1$, i pairs being ^{15}N labeled in the same sample. In the case of higher degeneracies the amino acid type of position i must be fully ^{13}C/^{15}N labeled in several samples, while the respective amino acid types of positions $i+1$ can be ^{15}N labeled in these samples in a combinatorial manner. In order to employ the (H)N(COCON)H experiment for lifting remaining ambiguities the amino acid type in position $i-1$ needs to be fully ^{13}C/^{15}N or, alternatively, dual-selectively ^{15}N and 1-^{13}C labeled. When the differentiation is based on the presence/absence of cross peaks the backbone carbonyl group of exactly one of the amino acid types in position $i-2$ must be ^{13}C labeled and that of the other(s) non-labeled, or vice versa when there are more than two possible assignments. The latter restriction does not apply if the residue in position $i-1$ is already assigned. In this case residue i can be assigned from sequential linkages in an (H)N(COCON)H spectrum whenever it carries a ^{15}N label itself, residue $i-1$ carries both ^{15}N and 1-^{13}C labels and residue $i-2$ is 1-^{13}C labeled. Since the (H)N(COCON)H experiment is bi-directional the same result can be obtained via correlations with an already assigned residue $i+1$ in a $^{13}CO_{i-1}$-$^{15}N_i$-$^{13/12}C^{\alpha}_i$-$^{13}CO_i$-$^{15}CO_{i-1}$-$^{15}N_i$-$^{13/12}C^{\alpha}_i$-$^{13}CO_i$-$^{15}N_{i+1}$ amino acid stretch.

Optimized labeling schemes can be conveniently found using suitable computer algorithms as demonstrated for dual-selective (^{15}N and 1-^{13}C) combinatorial labeling [34,36,37,87] and, very recently, suggested for a larger variety of labeling types and corresponding NMR coding systems [88]. While a labeling scheme that allowed the complete assignment of the p63 PAD peptide could be easily devised "manually", it is likely that the percentage of assigned NH resonances achieved for the protein CypD would be further increased with the help of software that takes into account all relevant isotopomers (Table S3) and the full set of experiments employed here.

4. Conclusions

The objective of the presented work was to probe the suitability of combinatorial selective labeling for obtaining backbone amide assignments without additional information from 3D (or higher dimensional) triple resonance spectra of uniformly labeled samples. Full assignments were achieved for the 31 residues detectable in a ^{15}N-HSQC of the unstructured p63 PAD peptide using three selectively labeled samples, while 88% of the 152 detectable amide signals of the folded protein CypD could be assigned using four samples. This compares favorably with reported yields of combinatorial labeling schemes, e.g. 8% [40], 14% [41], 37% [35] (using three

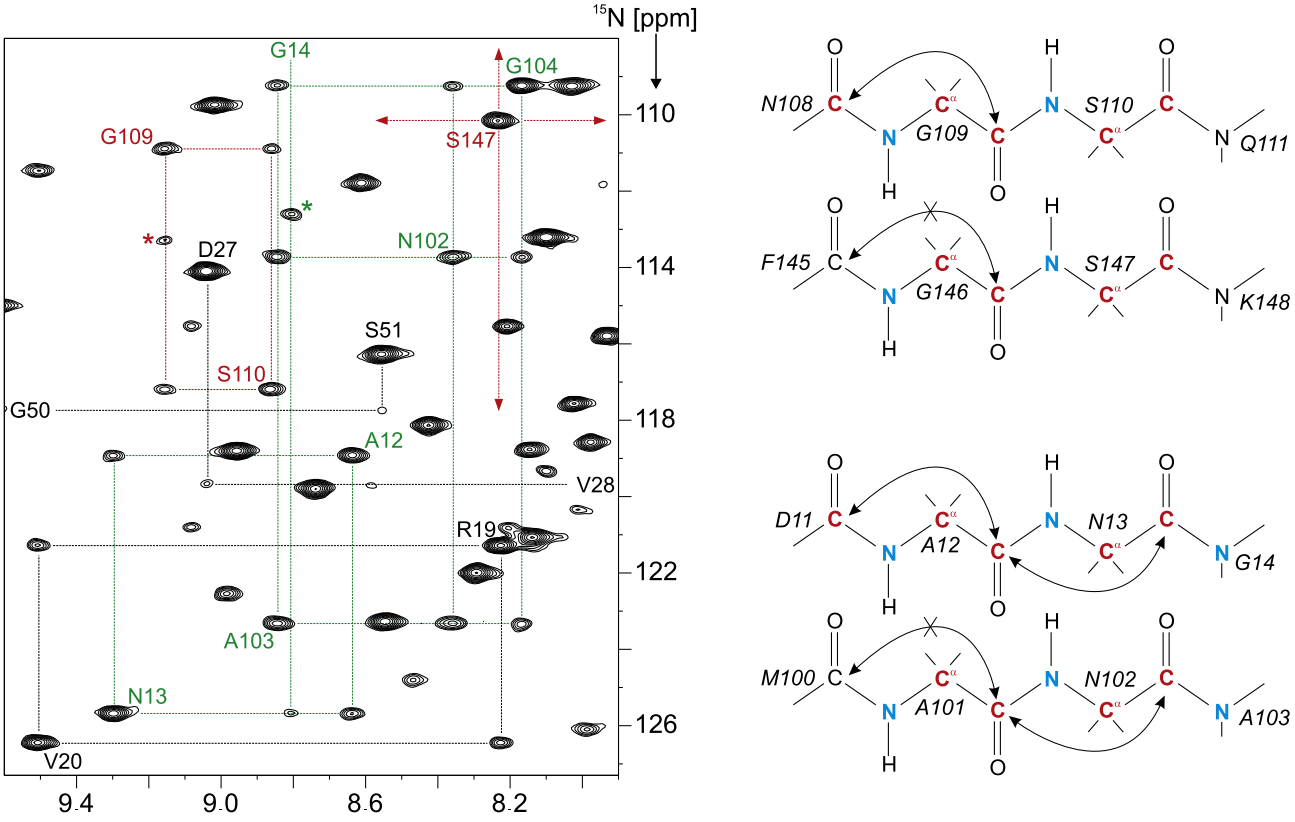

Fig. 9. Application of the (H)N(COCON)H experiment to sample 4 of CypD. The spectrum was recorded at 600 MHz using a MOCCA-XY16 mixing period of 280.5 ms. Assignments are provided for $^{15}N_i$-$^{1}H_i$ auto-correlation peaks that are connected to $^{15}N_i$-$^{1}H_{i\pm1}$ and/or $^{15}N_{i\pm1}$-$^{1}H_i$ cross peaks by dashed lines. Auto-correlation peaks of Gly14, Val28 and Gly50 are located outside the plotted region. The dashed arrows pointing away from the auto-correlation peak of Ser147 indicate the absence of cross peaks along the F_1 and F_2 coordinates of this residue. Cross peaks labeled with green and red asterisks are Asn13 $^{15}N^{\delta}$ - Gly14 $^{1}H^{N}$ and Asn108 $^{15}N^{\delta}$ - Gly109 $^{1}H^{N}$ side-chain-to-backbone correlations, respectively, arising from a transfer via intraresidual $^{3}J_{C\gamma\acute{c}}$ couplings in asparagine residues. For residues whose assignment is discussed in the text (i = Asn13, Asn102, highlighted in green, and i = Ser110, Ser147, highlighted in red) labeling patterns of $i-2$, $i-1$, i, $i+1$ polypeptide stretches are visualized at the right, where ^{15}N and ^{13}C nuclei are drawn in blue and red, respectively. Double-headed arrows delineate a magnetization transfer via $^{3}J_{\acute{c}\acute{c}}$ giving rise to $^{15}N_i$-$^{1}H_{i\pm1}$ and $^{15}N_{i\pm1}$-$^{1}H_i$ cross peaks. When crossed out, the transfer cannot take place due to the lack of a ^{13}CO label. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

samples in each study), ca. 25% (five samples) [31], 27% (six samples) [37] and, representing a theoretical value, 28% (nine samples) [87].

Aggregate measurement time requirements for the three samples of p63 PAD were approximately 7 h. Considering that HN(CACON) spectra of samples 1 and 2 and (H)N(COCON)H spectra of samples 1 and 3 were not needed for the actual assignment and only recorded for demonstration, this time further reduces to 4 h. Clearly, this is competitive with conceivable higher-dimensional methods that address the problem of low ^{13}C chemical shift dispersion. In the case of CypD the advantage is debatable. Obtaining assignments via four-sample combinatorial labeling required a total of 110 h of spectrometer time at 500 and 600 MHz (Table S2). For comparison, a complete backbone assignment has been obtained from a set of five non-uniformly sampled standard 3D triple resonance spectra recorded on a single uniformly $^{13}C/^{15}N$ labeled sample within 302 h on a 500 MHz spectrometer [35]. It is explicitly not claimed here that our approach is generally superior to existing assignment methods. Also, it is recognized that exclusively $^{1}H^{N}$ and ^{15}N chemical shifts assignments are provided while structural investigations will additionally require assignments for other nuclei, that are already partially afforded by $\geq$3D based assignment procedures. Nevertheless, the protocol presented here augments the toolbox of selective labeling to facilitate studies that rely on backbone amide signals, especially in cases of extensive ^{13}C chemical shift degeneracy and limited sensitivity. Based on the results obtained in this study it may be envisioned that for medium sized proteins the corresponding assignments can be established entirely by combinatorial selective labeling.

Compliance with ethical standards

Funding: Deutsche Forschungsgemeinschaft Grant DO545/18.

Conflict of interest: The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by the state of Hesse (Center for Biomolecular Magnetic Resonance), the Deutsche Forschungsgemeinschaft (DO545/18) and the Cluster of Excellence Frankfurt (Macromolecular Complexes). We would like to thank Daniel Schwarz and Matthias Frech (Merck KGaA, 64293 Darmstadt, Germany) for providing reagents and scientific input on the project.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmr.2019.03.010.

References

[1] R.H. Griffey, A.G. Redfield, R.E. Loomis, F.W. Dahlquist, Nuclear magnetic resonance observation and dynamics of specific amide protons in T4 lysozyme, Biochemistry 24 (1985) 817–822.

[2] D.M. LeMaster, F.M. Richards, ^{1}H–^{15}N heteronuclear NMR studies of *Escherichia coli* thioredoxin in samples isotopically labeled by residue type, Biochemistry 24 (1985) 7263–7268.

[3] H. Senn, A. Eugster, G. Otting, F. Suter, K. Wüthrich, 15N-labeled P22 c2 repressor for nuclear magnetic resonance studies of protein-DNA interactions, Eur. Biophys. J. 14 (1987) 301–306.

[4] D.C. Muchmore, L.P. McIntosh, C.B. Russell, D.E. Anderson, F.W. Dahlquist, Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance, Methods Enzymol. 177 (1989) 44–73.

[5] D.W. Hoffman, L.D. Spicer, Isotopic labeling of specific amino-acid types as an aid to nmr-spectrum assignment of the methionine repressor protein, Techn. Protein Chemis. Ii (1991) 409–416.

[6] J. Fiaux, E.B. Bertelsen, A.L. Horwich, K. Wüthrich, Uniform and residue-specific ^{15}N-labeling of proteins on a highly deuterated background, J. Biomol. NMR 29 (2004) 289–297.

[7] M. Kainosho, T. Tsuji, Assignment of the three methionyl carbonyl carbon resonances in Streptomyces subtilisin inhibitor by a carbon-13 and nitrogen-15 double-labeling technique. A new strategy for structural studies of proteins in solution, Biochemistry 21 (1982) 6273–6279.

[8] R.H. Griffey, A.G. Redfield, L.P. Mcintosh, T.G. Oas, F.W. Dahlquist, Assignment of proton amide resonances of T4 lysozyme by ^{13}C and ^{15}N multiple isotopic labeling, J. Am. Chem. Soc. 108 (1986) 6816–6817.

[9] M. Ikura, M. Krinks, D.A. Torchia, A. Bax, An efficient NMR approach for obtaining sequence-specific resonance assignments of larger proteins based on multiple isotopic labeling, FEBS Lett. 266 (1990) 155–158.

[10] M. Ikura, D. Marion, L.E. Kay, H. Shih, M. Krinks, C.B. Klee, A. Bax, Heteronuclear 3D NMR and isotopic labeling of calmodulin. Towards the complete assignment of the 1H NMR spectrum, Biochem. Pharmacol. 40 (1990) 153–160.

[11] H. Takahashi, A. Odaka, S. Kawaminami, C. Matsunaga, K. Kato, I. Shimada, Y. Arata, Multinuclear NMR study of the structure of the Fv fragment of anti-dansyl mouse IgG2a antibody, Biochemistry 30 (1991) 6611–6619.

[12] S. Tate, Y. Kikumoto, S. Ichikawa, M. Kaneko, Y. Masui, T. Kamogashira, M. Ouchi, S. Takahashi, F. Inagaki, Stable isotope aided nuclear magnetic resonance study to investigate the receptor-binding site of human interleukin 1β, Biochemistry 31 (1992) 2435–2442.

[13] W.M. Westler, B.J. Stockman, J.L. Markley, Correlation of C-13 and N-15 chemical-shifts in selectively and uniformly labeled proteins by heteronuclear two-dimensional nmr-spectroscopy, J. Am. Chem. Soc. 110 (1988) 6256–6258.

[14] G.S. Rule, N. Tjandra, V. Simplaceanu, C. Ho, Assignment strategies for ^{15}N–^{1}H correlated spectra of large proteins in solution, J. Magn. Reson. 102 (1993) 126–128.

[15] T. Yabuki, T. Kigawa, N. Dohmae, K. Takio, T. Terada, Y. Ito, E.D. Laue, J.A. Cooper, M. Kainosho, S. Yokoyama, Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis, J. Biomol. NMR 11 (1998) 295–306.

[16] J. Weigelt, M. van Dongen, J. Uppenberg, J. Schultz, M. Wikström, Site-selective screening by NMR spectroscopy with labeled amino acid pairs, J. Am. Chem. Soc. 124 (2002) 2446–2447.

[17] L. Guignard, K. Ozawa, S.E. Pursglove, G. Otting, N.E. Dixon, NMR analysis of in vitro-synthesized proteins without purification: a high-throughput approach, FEBS Lett. 524 (2002) 159–162.

[18] J. Shi, J.G. Pelton, H.S. Cho, D.E. Wemmer, Protein signal assignments using specific labeling and cell-free synthesis, J. Biomol. NMR 28 (2004) 235–247.

[19] N. Vajpai, A. Strauss, G. Fendrich, S.W. Cowan-Jacob, P.W. Manley, S. Grzesiek, W. Jahnke, Solution conformations and dynamics of ABL kinase-inhibitor complexes determined by NMR substantiate the different binding modes of imatinib/nilotinib and dasatinib, J. Biol. Chem. 283 (2008) 18292–18302.

[20] A.D. Gossert, A. Hinniger, S. Gutmann, W. Jahnke, A. Strauss, C. Fernández, A simple protocol for amino acid type selective isotope labeling in insect cells with improved yields and high reproducibility, J. Biomol. NMR 51 (2011) 449–456.

[21] E. Michel, L. Skrisovska, K. Wüthrich, F.H. Allain, Amino acid-selective segmental isotope labeling of multidomain proteins for structural biology, ChemBioChem 14 (2013) 457–466.

[22] K. Ozawa, M.J. Headlam, P.M. Schaeffer, B.R. Henderson, N.E. Dixon, G. Otting, Optimization of an *Escherichia coli* system for cell-free synthesis of selectively ^{15}N-labelled proteins for rapid analysis by NMR spectroscopy, Eur. J. Biochem. 271 (2004) 4084–4093.

[23] K. Ozawa, P.S. Wu, N.E. Dixon, G. Otting, ^{15}N-Labelled proteins by cell-free protein synthesis. Strategies for high-throughput NMR studies of proteins and protein-ligand complexes, FEBS Journal 273 (2006) 4154–4159.

[24] C. Klammt, F. Löhr, B. Schäfer, W. Haase, V. Dötsch, H. Rüterjans, C. Glaubitz, F. Bernhard, High level cell-free expression and specific labeling of integral membrane proteins, Eur. J. Biochem. 271 (2004) 568–580.

[25] D. Staunton, R. Schlinkert, G. Zanetti, S.A. Colebrook, I.D. Campbell, Cell-free expression and selective isotope labelling in protein NMR, Magn. Reson. Chem. 44 (2006) S2–S9.

[26] S. Sobhanifar, S. Reckel, F. Junge, D. Schwarz, L. Kai, M. Karbyshev, F. Löhr, F. Bernhard, V. Dötsch, Cell-free expression and stable isotope labelling strategies for membrane proteins, J. Biomol. NMR 46 (2010) 33–43.

[27] G. Jaipuria, B. Krishnarjuna, S. Mondal, A. Dubey, H.S. Atreya, Amino acid selective labeling and unlabeling for protein resonance assignments, Adv. Exp. Med. Biol. 992 (2012) 95–118.

[28] I. Maslennikov, S. Choe, Advances in NMR structures of integral membrane proteins, Curr. Opin. Struct. Biol. 23 (2013) 555–562.

[29] B. Hoffmann, F. Löhr, A. Laguerre, F. Bernhard, V. Dötsch, Protein labeling strategies for liquid-state NMR spectroscopy using cell-free synthesis, Prog. Nucl. Magn. Reson. Spectrosc. 105 (2018) 1–22.

[30] D. Shortle, Assignment of amino acid type in ^{1}H–^{15}N correlation spectra by labeling with ^{14}N-amino acids, J. Magn. Reson. 105 (1994) 88–90.

[31] M.J. Parker, M. Aulton-Jones, A.M. Hounslow, C.J. Craven, A combinatorial selective labeling method for the assignment of backbone amide NMR resonances, J. Am. Chem. Soc. 126 (2004) 5020–5021.

[32] P.S. Wu, K. Ozawa, S. Jergic, X.C. Su, N.E. Dixon, G. Otting, Amino-acid type identification in ^{15}N-HSQC spectra by combinatorial selective ^{15}N-labelling, J. Biomol. NMR 34 (2006) 13–21.

[33] C.J. Craven, M. Al-Owais, M.J. Parker, A systematic analysis of backbone amide assignments achieved via combinatorial selective labelling of amino acids, J. Biomol. NMR 38 (2007) 151–159.

[34] T. Kasai, S. Koshiba, J. Yokoyama, T. Kigawa, Stable isotope labeling strategy based on coding theory, J. Biomol. NMR 63 (2015) 213–221.

[35] C. Hein, F. Löhr, D. Schwarz, V. Dötsch, Acceleration of protein backbone NMR assignment by combinatorial labeling: Application to a small molecule binding study, Biopolymers 107 (2017).

[36] N. Trbovic, C. Klammt, A. Koglin, F. Löhr, F. Bernhard, V. Dötsch, Efficient strategy for the rapid backbone assignment of membrane proteins, J. Am. Chem. Soc. 127 (2005) 13504–13505.

[37] I. Maslennikov, C. Klammt, E. Hwang, G. Kefala, M. Okamura, L. Esquivies, K. Mörs, C. Glaubitz, W. Kwiatkowski, Y.H. Jeon, S. Choe, Membrane domain structures of three classes of histidine kinase receptors by cell-free expression and rapid NMR analysis, Proc. Nat. Acad. Sci. USA 107 (2010) 10902–10907.

[38] M. Etzkorn, T. Raschle, F. Hagn, V. Gelev, A.J. Rice, T. Walz, G. Wagner, Cell-free expressed bacteriorhodopsin in different soluble membrane mimetics: biophysical properties and NMR accessibility, Structure 21 (2013) 394–401.

[39] J.A. Butterwick, R. MacKinnon, Solution structure and phospholipid interactions of the isolated voltage-sensor domain from KvAP, J. Mol. Biol. 403 (2010) 591–606.

[40] F. Löhr, S. Reckel, M. Karbyshev, P.J. Connolly, N. Abdul-Manan, F. Bernhard, J. M. Moore, V. Dötsch, Combinatorial triple-selective labeling as a tool to assist membrane protein backbone resonance assignment, J. Biomol. NMR 52 (2012) 197–210.

[41] F. Löhr, F. Tumulka, C. Bock, R. Abele, V. Dötsch, An extended combinatorial 15N, 13Cα, and 13C' labeling approach to protein backbone resonance assignment, J. Biomol. NMR 62 (2015) 263–279.

[42] A. Laguerre, F. Löhr, E. Henrich, B. Hoffmann, N. Abdul-Manan, P.J. Connolly, E. Perozo, J.M. Moore, F. Bernhard, V. Dötsch, From nanodiscs to isotropic bicelles: A procedure for solution nuclear magnetic resonance studies of detergent-sensitive integral membrane proteins, Structure 24 (2016) 1830–1841.

[43] A. Dubey, R.V. Kadumuri, G. Jaipuria, R. Vadrevu, H.S. Atreya, Rapid NMR assignments of proteins by using optimized combinatorial selective unlabeling, ChemBioChem: Eur. J. Chem. Biol. 17 (2016) 334–340.

[44] D. Schwarz, F. Junge, F. Durst, N. Frölich, B. Schneider, S. Reckel, S. Sobhanifar, V. Dötsch, F. Bernhard, Preparative scale expression of membrane proteins in *Escherichia coli*-based continuous exchange cell-free systems, Nat. Protoc. 2 (2007) 2945–2957.

[45] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, Attenuated T2 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. Nat. Acad. Sci. USA 94 (1997) 12366–12371.

[46] M. Salzmann, K. Pervushin, G. Wider, H. Senn, K. Wüthrich, TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins, Proc. Nat. Acad. Sci. USA 95 (1998) 13585–13590.

[47] M. Salzmann, G. Wider, K. Pervushin, H. Senn, K. Wüthrich, TROSY-type triple-resonance experiments for sequential NMR assignments of large proteins, J. Am. Chem. Soc. 121 (1999) 844–848.

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

[49] M. Czisch, R. Boelens, Sensitivity enhancement in the TROSY experiment, J. Magn. Reson. 134 (1998) 158–160.

[50] K.V. Pervushin, G. Wider, K. Wüthrich, Single Transition-to-single Transition Polarization Transfer (ST2-PT) in [^{15}N,^{1}H]-TROSY, J. Biomol. NMR 12 (1998) 345–348.

[51] J. Weigelt, Single scan, sensitivity- and gradient-enhanced TROSY for multidimensional NMR experiments, J. Am. Chem. Soc. 120 (1998) 10778–10779.

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

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

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

[55] A. Favier, B. Brutscher, Recovering lost magnetization: polarization enhancement in biomolecular NMR, J. Biomol. NMR 49 (2011) 9–15.
[56] F. Löhr, A. Laguerre, C. Bock, S. Reckel, P.J. Connolly, N. Abdul-Manan, F. Tumulka, R. Abele, J.M. Moore, V. Dötsch, Time-shared experiments for efficient assignment of triple-selectively labeled proteins, J. Magn. Reson. 248 (2014) 81–95.
[57] 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.
[58] S. Grzesiek, A. Bax, A three-dimensional NMR experiment with improved sensitivity for carbonyl-carbonyl J correlation in proteins, J. Biomol. NMR 9 (1997) 207–211.
[59] Y. Yoshimura, N.V. Kulminskaya, F.A. Mulder, Easy and unambiguous sequential assignments of intrinsically disordered proteins by correlating the backbone 15N or 13C' chemical shifts of multiple contiguous residues in highly resolved 3D spectra, J. Biomol. NMR 61 (2015) 109–121.
[60] S. Zerko, P. Byrski, P. Wlodarczyk-Pruszynski, M. Gorka, K. Ledolter, E. Masliah, R. Konrat, W. Kozminski, Five and four dimensional experiments for robust backbone resonance assignment of large intrinsically disordered proteins: application to Tau3x protein, J. Biomol. NMR 65 (2016) 193–203.
[61] F. Kramer, W. Peti, C. Griesinger, S.J. Glaser, Optimized homonuclear Carr-Purcell-type dipolar mixing sequences, J. Magn. Reson. 149 (2001) 58–66.
[62] J. Furrer, F. Kramer, J.P. Marino, S.J. Glaser, B. Luy, Homonuclear Hartmann-Hahn transfer with reduced relaxation losses by use of the MOCCA-XY16 multiple pulse sequence, J. Magn. Reson. 166 (2004) 39–46.
[63] D. Lee, C. Hilty, G. Wider, K. Wüthrich, Effective rotational correlation times of proteins from NMR relaxation interference, J. Magn. Reson. 178 (2006) 72–76.
[64] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, J. Biomol. NMR 6 (1995) 135–140.
[65] M. Ikura, L.E. Kay, A. Bax, A novel approach for sequential assignment of ^{1}H, ^{13}C, and ^{15}N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin, Biochemistry 29 (1990) 4659–4667.
[66] R.T. Clubb, V. Thanabal, G. Wagner, A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $^{1}H^{N}$, ^{1}5N, and ^{1}3C' chemical shifts in $^{1}5N-^{1}3C$-labelled proteins, J. Magn. Reson. 97 (1992) 213–217.
[67] M. Wittekind, L. Mueller, HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins, J. Magn. Reson. 101 (1993) 201–205.
[68] M. Tuppi, S. Kehrloesser, D.W. Coutandin, V. Rossi, L.M. Luh, A. Strubel, K. Hotte, M. Hoffmeister, B. Schäfer, T. De Oliveira, F. Greten, E.H.K. Stelzer, S. Knapp, M. De Felici, C. Behrends, F.G. Klinger, V. Dötsch, Oocyte DNA damage quality control requires consecutive interplay of CHK2 and CK1 to activate p63, Nat. Struct. Mol. Biol. 25 (2018) 261–269.
[69] R. Weisemann, H. Rüterjans, W. Bermel, 3D triple-resonance NMR techniques for the sequential assignment of NH and 15N resonances in 15N- and 13C-labelled proteins, J. Biomol. NMR 3 (1993) 113–120.
[70] S. Grzesiek, J. Anglister, H. Ren, A. Bax, 13C Line narrowing by 2H decoupling in 2H/13C/15N-enriched proteins - application to triple-resonance 4D *J*-connectivity of sequential amides, J. Am. Chem. Soc. 115 (1993) 4369–4370.
[71] S.C. Panchal, N.S. Bhavesh, R.V. Hosur, Improved 3D triple resonance experiments, HNN and HN(C)N, for HN and 15N sequential correlations in (13C, 15N) labeled proteins: application to unfolded proteins, J. Biomol. NMR 20 (2001) 135–147.
[72] C. Wiedemann, N. Goradia, S. Hafner, C. Herbst, M. Görlach, O. Ohlenschläger, R. Ramachandran, HN-NCA heteronuclear TOCSY-NH experiment for $^{1}H^{N}$ and ^{15}N sequential correlations in (^{13}C, ^{15}N) labelled intrinsically disordered proteins, J. Biomol. NMR 63 (2015) 201–212.
[73] S. Mantylahti, M. Hellman, P. Permi, 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 (2011) 99–109.
[74] X. Yao, S. Becker, M. Zweckstetter, A six-dimensional alpha proton detection-based APSY experiment for backbone assignment of intrinsically disordered proteins, J. Biomol. NMR 60 (2014) 231–240.
[75] V. Motackova, J. Novacek, A. Zawadzka-Kazimierczuk, K. Kazimierczuk, L. Zidek, H. Sanderova, L. Krasny, W. Kozminski, V. Sklenar, Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments, J. Biomol. NMR 48 (2010) 169–177.
[76] J. Novacek, A. Zawadzka-Kazimierczuk, V. Papouskova, L. Zidek, H. Sanderova, L. Krasny, W. Kozminski, V. Sklenar, 5D 13C-detected experiments for backbone assignment of unstructured proteins with a very low signal dispersion, J. Biomol. NMR 50 (2011) 1–11.
[77] A. Zawadzka-Kazimierczuk, K. Kazimierczuk, W. Kozminski, A set of 4D NMR experiments of enhanced resolution for easy resonance assignment in proteins, J. Magn. Reson. 202 (2010) 109–116.
[78] W. Bermel, I. Bertini, I.C. Felli, L. Gonnelli, W. Kozminski, A. Piai, R. Pierattelli, J. Stanek, Speeding up sequence specific assignment of IDPs, J. Biomol. NMR 53 (2012) 293–301.
[79] A. Zawadzka-Kazimierczuk, W. Kozminski, H. Sanderova, L. Krasny, High dimensional and high resolution pulse sequences for backbone resonance assignment of intrinsically disordered proteins, J. Biomol. NMR 52 (2012) 329–337.
[80] M. Nowakowski, S. Saxena, J. Stanek, S. Zerko, W. Kozminski, Applications of high dimensionality experiments to biomolecular NMR, Prog. Nucl. Magn. Reson. Spectrosc. 90–91 (2015) 49–73.
[81] A. Piai, T. Hosek, L. Gonnelli, A. Zawadzka-Kazimierczuk, W. Kozminski, B. Brutscher, W. Bermel, R. Pierattelli, I.C. Felli, "CON-CON" assignment strategy for highly flexible intrinsically disordered proteins, J. Biomol. NMR 60 (2014) 209–218.
[82] S. Zerko, W. Kozminski, Six- and seven-dimensional experiments by combination of sparse random sampling and projection spectroscopy dedicated for backbone resonance assignment of intrinsically disordered proteins, J. Biomol. NMR 63 (2015) 283–290.
[83] S. Hiller, C. Wasmer, G. Wider, K. Wüthrich, Sequence-specific resonance assignment of soluble nonglobular proteins by 7D APSY-NMR spectroscopy, J. Am. Chem. Soc. 129 (2007) 10823–10828.
[84] N. Vajpai, A. Strauss, G. Fendrich, S.W. Cowan-Jacob, P.W. Manley, W. Jahnke, S. Grzesiek, Backbone NMR resonance assignment of the Abelson kinase domain in complex with imatinib, Biomol NMR Assign 2 (2008) 41–42.
[85] J.S. Hu, A. Bax, Measurement of three-bond $^{13}C-^{13}C$ *J* couplings between carbonyl and carbonyl/carboxyl carbons in isotopically enriched proteins, J. Am. Chem. Soc. 118 (1996) 8170–8171.
[86] J.M. Schmidt, M. Blümel, F. Löhr, H. Rüterjans, Self-consistent ^{3}J coupling analysis for the joint calibration of Karplus coefficients and evaluation of torsion angles, J. Biomol. NMR 14 (1999) 1–12.
[87] F. Hefke, A. Bagaria, S. Reckel, S.J. Ullrich, V. Dotsch, C. Glaubitz, P. Guntert, Optimization of amino acid type-specific 13C and 15N labeling for the backbone assignment of membrane proteins by solution- and solid-state NMR with the UPLABEL algorithm, J. Biomol. NMR 49 (2011) 75–84.
[88] M. Myshkin, M. Dubinnyi, D. Kulbatskii, E. Lyukmanova, Z. Shenkarev, Design of optimal selective labeling schemes for fast NMR assignment of proteins, XXVIIIth International Conference on Magnetic Resonance in Biological Systems, Dublin, Ireland, 2018, pp. Poster P11.