Nuclear overhauser spectroscopy of chiral CHD methylene groups

Rafal Augustyniak¹,²,³ · Jan Stanek⁴ · Henri Colaux¹,²,³ · Geoffrey Bodenhausen¹,²,³,⁵ · Wiktor Koźmiński⁴ · Torsten Herrmann⁶ · Fabien Ferrage¹,²,³

Received: 23 June 2015 / Accepted: 12 November 2015 / Published online: 27 November 2015 © Springer Science+Business Media Dordrecht 2015

Abstract Nuclear magnetic resonance spectroscopy (NMR) can provide a great deal of information about structure and dynamics of biomolecules. The quality of an NMR structure strongly depends on the number of experimental observables and on their accurate conversion into geometric restraints. When distance restraints are derived from nuclear Overhauser effect spectroscopy (NOESY), stereo-specific assignments of prochiral atoms can contribute significantly to the accuracy of NMR structures of proteins and nucleic acids. Here we introduce a series of NOESY-based pulse sequences that can assist in the assignment of chiral CHD methylene protons in random fractionally deuterated proteins. Partial deuteration suppresses spin-diffusion between the two protons of CH₂ groups that normally impedes the distinction of cross-relaxation networks for these two protons in NOESY spectra. Three and four-dimensional spectra allow one to distinguish cross-relaxation pathways involving either of the two methylene protons so that one can obtain stereospecific assignments. In addition, the analysis provides a large number of stereospecific distance restraints. Non-uniform sampling was used to ensure optimal signal resolution in 4D spectra and reduce ambiguities of the assignments. Automatic assignment procedures were modified for efficient and accurate stereospecific assignments during automated structure calculations based on 3D spectra. The protocol was applied to calcium-loaded calbindin D₉k. A large number of stereospecific assignments lead to a significant improvement of the accuracy of the structure.

Keywords NMR spectroscopy · Protein structures · Nuclear Overhauser spectroscopy · Automatic structure calculation

Introduction

Understanding protein function requires precise and accurate information about structure at the atomic level. Along with X-ray crystallography, NMR spectroscopy has become a tool of choice to obtain high-resolution structures of proteins. The assignment of resonances is required to analyze NMR data at an atomic level. A suite of NMR techniques is nowadays routinely used for backbone and side-chain resonance assignments. However, stereospecific assignments of substituents of prochiral centers are a demanding task. This problem arises for the two diastereotopic protons of methylene groups in backbones and side-chains, and for the two diastereotopic methyl groups...
of isopropyl residues in valines and leucines. The advantages of stereospecific assignments have been widely documented (Clore et al. 1990; Clore et al. 1991b; Driscoll et al. 1989; Guntert et al. 1989; Kainosho et al. 2006). The restraints that are most commonly used for structure calculations are derived from NOESY experiments (Kumar et al. 1980; Neuhaus and Williamson 2000). In methylene groups CH pro-R-H pro-S rapid spin-diffusion between the two prochiral protons reduces their utility in the interpretation of NOE effects. Random partial deuteration leads to mixtures of CH pro-R-D pro-S and CD pro-R-H pro-S groups where spin-diffusion is suppressed (stereo-specific deuteration of either H pro-R or H pro-S is also possible (Kainosho et al. 2006; Takeda et al. 2012) but will not be considered here). Such CHD groups may be called chiral methylenes, in analogy to chiral methyl groups (Luthy et al. 1969). Even when deuteration is random, it turns out that stereospecific assignments are possible. When stereospecific assignments are not available, distance restraints involving diastereotopic substituents can use so-called pseudo-atoms, which are fictitious atoms that represent all constraints for both diastereotopic substituents. Despite pseudo-atom correction protocols, this leads to uncertainties in distance restraints and to structures that lack precision (Guntert et al. 1989). On the other hand, the use of restraints based on proper stereospecific assignments can lead to a significant improvement in the accuracy of the structures in both backbones and side-chains (Driscoll et al. 1989). The advantages are especially important for Asp, Asn, Glu and Gln side-chains, since their carboxyl and carbonyl groups are frequently involved in hydrogen bonding and other interactions with ligands or metal ions, but lack NOE restraints.

Although in silico tools can be employed either to circumvent the lack of stereospecific assignments or to retrospectively extract assignments based on preliminary 3D structures, (Folmer et al. 1997; Güntert et al. 1998; Orts et al. 2013; Pristovsek and Franzoni 2006) it is obviously preferable to use experimental assignments of diastereotopic protons or methyl groups. In general, protons can be assigned stereospecifically if one can determine vicinal scalar coupling constants and/or intra-residual nuclear Overhauser effects (NOEs). For C αH 2 pro-R groups, J(H α, H β pro-R) and J(H β, H β pro-S) constants (Clore et al. 1991a; Emerson and Montelione 1992; Lohr et al. 1999; Mueller 1987) can be combined with NOEs between either H α or H β on the one hand and the two β-protons on the other (Wagner et al. 1987). Knowledge of heteronuclear J(13C, H) and J(15N, H) can also contribute to stereospecific assignments (Grzesiek et al. 1992). The combination of J(H α, H β), J(15N, 13C), and J(15C, 13C) coupling constants is especially useful for non-native states and for intrinsically disordered proteins (Hahnke et al. 2010). The protons of C αH 2 groups can also be stereospecifically assigned if the assignment of C αH 2 protons is known and at least some of the four J(H α,H β) couplings can be measured. Note that 3 J(H α,H β) couplings are difficult to measure in large proteins (Cai et al. 1995) when signals overlap, although this problem may be overcome in part by non-uniform sampling (NUS) (Kazimierczuk et al. 2008).

Nuclear Overhauser effects alone can also be used to obtain satisfactory stereospecific assignments. In particular, it is possible to apply rotating-frame Overhauser spectroscopy (ROESY) to partially deuterated samples, bearing in mind that signals arising from spin-diffusion and direct cross-relaxation have opposite signs (Clore et al. 1990). However, rapid T 1(H) relaxation makes it impossible to use long mixing times that would help to obtain informative long-range restraints. Finally, it appears that ‘exact’ nuclear Overhauser effects (eNOE’s) (Vogeli et al. 2009) with short mixing times can yield reliable stereospecific assignments (Orts et al. 2013).

The incorporation of stereoselectively deuterated amino acids into proteins (Gardner and Kay 1998; Lemaster 1990) (possibly via suitable precursors) allows selective substitution of either H pro-R or H pro-S protons by a deuteron. Stereospecific assignment can be achieved by comparing spectra of two samples that are stereoselectively deuterated in a complementary manner. Such sophisticated biosynthetic methods can provide reliable results, as shown for C αH 2 protons in aspartic acid and asparagine (Lemaster 1987) or C αH 2 protons in glycines.(Curley et al. 1994) Prochiral methyl groups in leucine and valine can be assigned by stereoselective 2H and 13C labeling (Atreya and Chary 2001; Kainosho et al. 2006; Neri et al. 1989; Ostler et al. 1993; Plevin et al. 2011).

Obviously, random fractional deuteration is less challenging than stereospecific labeling (Nietlispach et al. 1996). Partial deuteration is sufficient to improve NOESY spectra thanks to the reduction of spin diffusion pathways (Gardner and Kay 1998; Lemaster 1990). This improves the interpretation of NOEs (Lemaster 1987). Generally, it has been shown that partial deuteration levels between 50 and 75 % increase the sensitivity of many NMR experiments because of line-narrowing, despite the reduction of the concentration of protons (Gardner and Kay 1998). A deuteration level of 50 % leads to mixtures of CH 2, CH pro-R-D pro-S, CD pro-R-H pro-S and CD 2 each with overall probabilities of ∼ 25 % (site-to-site variations may be significant, as discussed below). Random partial deuteration also makes it possible to implement isotopic filters to focus on selected isotopomers (Gardner and Kay 1998; Kushlan and Lemaster 1993; Muhandiram et al. 1995). The use of such filters can improve spectral resolution (Valluru-palli et al. 2009).
Here, we introduce two new pulse sequences for nuclear Overhauser effect spectroscopy, CHD–NOESY–H(C)CO and CHD–NOEY–HSQC. Each pulse sequence can be run either in 3D or 4D fashion, depending on the extent of overlap. These methods allow one to isolate the signals of chiral CHD methylene groups in randomly deuterated proteins. Both sequences rely on the use of ‘CHD filters’. The CHD–NOEY–H(C)CO experiment allows one to observe NOEs from neighboring ‘source’ protons to ‘target’ protons in CHD groups that are adjacent to carboxyl or carbonyl groups in the side-chains of aspartic acid, asparagine, glutamic acid and glutamine residues (Asp, Asn, Glu and Gln) as well as in some backbone CβH2CO groups. The CHD–NOEY–CT-HSQC experiment (or its non-constant-time variant) allows the detection of cross-relaxation towards all CHD methylene protons. If high-resolution NMR or X-ray structures are available, our new NOE methods allow one to obtain unambiguous stereospecific assignments.

For the sake of illustration, the stereospecific assignment of a set of methylene protons in calbindin D9k has been performed ‘manually’, based on known NMR (Kordel et al. 1997) and X-ray (Svensson et al. 1992) structures. The stereospecific assignment and structure determination were achieved simultaneously using a modified version of the ATNOS/CANDID algorithms implemented in the UNIO software package (Herrmann et al. 2002a, b). This automated procedure, which can be applied directly to raw NMR spectra, can provide reliable stereospecific assignments and enhances the accuracy of the NMR structure. This is illustrated for calbindin by comparing with a standard UNIO–ATNOS/CANDID procedure that copes with the lack of stereospecific assignments by ‘atom swapping’ during NOE assignment and simulated annealing (Folmer et al. 1997).

Results and discussion

NMR experiments

The pulse sequence for the 3D version of CHD–NOEY–H(C)CO is shown in Fig. 1. The initial 1H frequency labeling and NOESY mixing time τm are followed by an H(C)CO pulses for compensation of Bloch–Siegert effects (Emsley and Bodenhausen 1990b). 1H, 2H and 13C decoupling was performed with DIPSI-2 (Shaka et al. 1985) (ω1/2π = 3.1 kHz), WALTZ-16 (Shaka et al. 1985) (ω1/2π = 1 kHz) and GARP (Shaka et al. 1988) (ω1/2π = 2.7 kHz) respectively. The lengths and peak amplitudes of the gradients in the x, y and z directions were respectively: g0 = 0.5 ms, (38, 38, 38) G/cm; g1 = 0.5 ms, (0, 0, 8) G/cm; g2 = 2.4 ms, (21, 21, 35) G/cm; g3 = 1 ms, (28, 0, 28) G/cm; g4 = 0.5 ms, −49, −49, −49) G/cm. The phase cycling employed was: φ1 = 4{x}, 4{−x}; φ2 = 8{−x}, 8{x}; φ3 = 2{x}, 2{−x}; φ4 = y, −y; φaq = x, −x, −x, x, −x, x, −x, x, −x, x, −x, x, −x. Frequency sign discrimination in the indirect dimension was achieved using States method (States et al. 1982). The delays were: NOESY mixing time τm = 200 ms, τa = 1.85 ms ≈ 141CCH−1, (ΔCH ≈ 135 Hz), τb = 2.03 ms > 141CCH−1, τc = 3 ms < 141CCH−1, δ′ = 500 μs. Note that the adiabatic pulse with inverted sweep on the carbon-13 channel and the inversion pulse on the nitrogen-15 channel during the delay δ′ were omitted in this study. They should be used to refocus the evolution under scalar couplings.

Fig. 1 Pulse sequence for the 3D CHD–NOEY–H(C)CO experiment. Narrow black and wide empty rectangles correspond to 90° and 180° pulses, respectively. Unless otherwise mentioned, all pulses are applied along the x-axis of the rotating frame. The 1H carrier was placed on resonance with the water signal (4.7 ppm), the 15N carrier was chosen at 120 ppm and the 13C carrier frequency was switched between 40 (13C0) and 180 ppm (13C0) as marked by arrows. Black bell-shaped pulses on the proton channel represent 90° water flip-back sinc-shaped pulses with duration of 2 (first two pulses) and 1.48 ms (last two pulses) for WATERGATE (Piotto et al. 1992). Wide rectangles on the 13C channel represent frequency-swept chirp pulses (Bohlen and Bodenhausen 1993) with durations of 500 μs (a single pulse is used for an inversion across the entire 13C spectrum). Narrow black and wide gray bell-shaped pulses on the 13C channel represent Q5 and Q3 pulses (Emsley and Bodenhausen 1990a) with durations of 480 and 340 μs respectively. Highly selective inversions were performed with 1.5 ms REBURP pulses (open bell-shaped arrows) as marked by Q5 and Q3 pulses (Emsley and Bodenhausen 1990a) with durations of 480 and 340 μs respectively. Highly selective inversions were performed with 1.5 ms REBURP pulses (open bell-shaped arrows) as marked by Q5 and Q3 pulses (Emsley and Bodenhausen 1990a) with durations of 480 and 340 μs respectively.
Fig. 2 Pulse sequence for the 4D C,C-edited CHD–HMQC–
NOESY–HSQC. Solid and open bars represent non-selective 90°
and 180° pulses, respectively. All pulses are applied along the x-axis
of the rotating frame unless indicated otherwise. 1H, 13C
decoupling was performed with WALTZ-16 (Shaka et al. 1983),
GARP (Shaka et al. 1983) and WURST-40 (Kupeck and Freeman
1995) respectively. Selective sinc-shaped 180° pulses, with yB1
adjusted for proper inversion of all C' spins without affecting the
C-aliph spins, are represented by open sinc-shaped pulses. 'BS' denotes
Bloch–Siegert compensation pulses (Emsley and Bodenhausen
1990b). Suppression of CH 2 resonances (including CHD, and in
part also CH3 signals) is accomplished by insertion of a pair of τa
delays during 13C evolution (t2) in the HMQC block as discussed in
the text. The presence of these delays allows for the insertion of an
additional pair of gradients g1 without causing any losses. The delays
τ1, τ2 and τ3 are set as for 4D HMOC–NOESY–H(C)/CO experiments.
The delay ε is set to minimize the time required for gradient labelling
by g6. Provided that the duration of the gradient g6 satisfies
t6 ≤ 0.5τ3,max it can be set as follows: ε = τ6g6 (1 − 1/t3,max). The delays
areτ3 = 2τ2 = 3.57 ms; τ1 = 0.5/JCH (JCH ≈ 140 Hz),
ε = 0.65 ms, and the NOESY mixing time τm = 0.2 s. The delays
τa, and τd are used for the best compromise in transfer efficiency for
CH (CHD), CH2 (CH2D) and CH3 groups in the

The information content of the 3D CHD–NOESY–H(C)/CO spectra and its 4D counterpart (see Fig. S2) is thus
limited to a subset of methylene groups. By contrast, the 3D
and 4D CHD–NOESY–CT–HSQC shown on Figs. S1 and 2,
respectively, can yield stereospecific distance constraints for
all methylene groups in a protein. Frequency labeling by the
1H chemical shifts is followed by a NOESY mixing time and
a constant-time HSQC experiment (Vuister and Bax 1992)
which is again modified by inserting a pair of 90° pulses on the
1H channel to retain the signals of all CHD groups and
eliminate the responses of non-deuterated CH2 groups. The
selection of CHD groups is compatible with the sensitivity-
enhanced scheme (Palmer et al. 1991). Composite-pulse
decoupling is applied in intervals where 13C
coherences evolve, except during pulsed field gradients.

Four-dimensional (4D) versions of both H(C)/CO- and
HSQC-based experiments (Figs. S2 and 2) allow one to
exploit the chemical shifts of 13C nuclei that have a scalar

sequence (Yamazaki et al. 1993) adapted from our earlier
work (Paquin et al. 2008). Here, CHD groups are selected by
two 90° pulses applied to the 1H channel after a refocused
1H–13C INEPT sequence. Phase cycling of the second proton
sequence (Yamazaki et al. 1993) adapted from our earlier
work (Paquin et al. 2008). Here, CHD groups are selected by
the 1H, 13C and 15N carrier frequencies are set to 4.77, 42.8 and 117.8 ppm, respectively. The
durations and amplitudes of the gradients, which are all applied along the
z axis, are: g1 = 1 ms, 8.9 G/cm, g2 = 2 ms, 17.7 G/cm,
g3 = 2 ms, 14.2 G/cm, g4 = 0.5 ms, 1.8 G/cm, g5 = 0.5 ms,
23.2 G/cm, g6 = 2 ms, 31.9 G/cm, g7 = 0.5 ms, 3.6 G/cm,
g8 = 1 ms, 5.3 G/cm, g9 = 0.5 ms, ± 32.1 G/cm. A recovery delay
of 1.2 s was used between scans. 11,000 sampling points (t1, t2, t3)
were randomly chosen from a 84 × 66 × 110 grid according to a
Gaussian probability distribution p(t) = exp [−(th/τ2)22σ2] with
σ = 0.5. The maximum evolution times in the indirectly detected
dimensions were 12 (t1), 6 (t2) and 10 ms (t3). The spectral widths
were 7.0 (ο1), 11 (ο2), 11 (ο3) and 12 kHz (ο4). The total duration of
the 4D experiment was 149 h
coupling $J(CH)$ to the ‘source’ protons to minimize signal overlap. Two 4D experiments were recorded using non-uniform sampling in all three indirect dimensions (Kazimierczuk et al. 2010) and were modified in the manner of HMQC (Muller 1979) to minimize signal losses by exploiting favorable relaxation properties of heteronuclear proton-carbon multiple-quantum (MQ) coherences (Carlo-magno et al. 2000; Kumar et al. 2000; Marino et al. 1997; Miclet et al. 2004). The additional delays of $4\tau_a \approx J_{CH}$ for coherence transfer between $^1H$ and $^{13}C$ are combined with semi-constant time evolution of $^1H$ chemical shifts (Stanek et al. 2012). In both 4D experiments, the phase cycle was reduced to four steps in order to record as many sampling points as possible and thus minimize artifacts arising from non-uniform sampling (NUS). This was achieved in part through the use of different CHD filters where antiphase $^{13}C$ coherences in CH$_2$ moieties are suppressed by pulse field gradients, twice in 4D CHD–HMOC–NOESY–H(C)CO (gradients G5 and G7 Fig S2) and once in 4D CHD–HMOC–NOESY–HSQC (gradient G2 Fig. 2). In the latter case, the selection is achieved in the HMOC part, by adding a delay $2\tau_a$ during which MQ coherences in CH$_2$ groups are allowed to evolve under $J_{CH}$. Thus the $^1H$ magnetization of the CHD groups, selected before the mixing time, acts as ‘source’ of cross-relaxation, whereas in all other experiments, the selection is performed on the CH$_2$ ‘target’ groups.

Applications to calbindin D$_{9K}$

All pulse sequences were applied to a calcium-loaded sample of the P43G mutant of calbindin D$_{9K}$ with uniform $^{15}N$ and $^{13}C$ labeling and 33 % random average deuteration (evaluated by mass spectrometry), which is expected for a sample obtained from expression in *Escherichia coli* in a minimal medium with protonated glucose and 50 % D$_2$O (Leiting et al. 1998). The assignment of the calcium-loaded form of calbindin D$_{9K}$ P43G is complete (Oktaviani et al. 2011) albeit without systematic stereospecific assignments. The stereospecific assignments of the H$_{pro-R}$ and H$_{pro-S}$ protons in 37 out of 56 C$_{H2}$ groups had been determined from scalar couplings and $^1H$–$^1H$ cross-relaxation networks (Kordel et al. 1993). We have evaluated the distribution of isotopomers in our sample by comparing signal intensities of CH$_2$, CH$_{pro-R}$D$_{pro-S}$, and CD$_{pro-R}$H$_{pro-S}$ groups in CH$_2$ and CHD-filtered constant-time HSQC spectra. Results are

![Fig. 3 Strip-plots extracted from a 3D CHD-filtered NOESY–H(C)CO spectrum of the P43G mutant of calbindin D$_{9K}$ recorded with the pulse sequence of Fig. 1 showing signals of the C$_{H2}$ groups of (a) Glu35 and (b) Glu65. The chemical shifts in the $\omega_2(^{13}C)$ dimension of the 3D spectra are a 181.0 and b 187.6 ppm. The $^1H$–$^1H$ distances corresponding to NOESY cross-peaks are shown for Glu35 and Glu65 using the first model of the structural ensemble (PDB code 1bl1 g). The prochiral hydrogen atoms of the C$_{H2}$ groups are shown in yellow. Oxygen and nitrogen atoms are shown in red and blue. In (a), grey dashed lines show short distances that have been used to obtain stereospecific assignments for Glu35. In (b), the distances with Glu65 are not compatible, suggesting an improper orientation of Glu65 side-chain.](image-url)
shown as supporting information. Overall, the signal of a proton in the CHD-filtered spectrum is similar or higher than in the CH₂-filtered spectrum. A significant exception occurs for amino acids produced by biosynthetic pathways where the β methylene group stems from the C₆ methylene group of glucose. Thus, CβH₂ sites of serines and aromatic residues show lower populations of CHβ-RDβ-S and CDβ-RHβ-S isotopomers compared to CH₂.

Figure 3 illustrates the dramatically different cross-relaxation effects observed for two pairs of diastereotopic protons in 3D H(C)CO-based experiment. When the three-dimensional structure of a protein is known, a large number of distance restraints can be derived to make stereospecific assignments of prochiral proton signals reliable (Fig. 3a). In a few instances, such as the CβH₂ group of Glu65 (Fig. 3b), the cross-relaxation networks are not compatible with the previously reported structure, indicating inconsistencies in side-chain orientations. In order to compare the cross-relaxation networks derived from CHD-filtered experiments with those obtained in non-filtered experiments, we recorded two interleaved 3D NOESY–H(C)CO experiments with opposite values for the phase φ₂ (see Fig. 1). Proton decoupling was avoided to preserve the signals of CH₂ groups. The sum of these two experiments gave a CHD-filtered spectrum while the difference yielded a CH₂-filtered experiment. Figure 4 shows signals of the Hβ-pro-R and Hβ-pro-S protons of Asn21 in calbindin D₉k. The patterns observed in the CHD-filtered experiment (Fig. 4a) are identical to those obtained with the 3D CHD–NOESY–H(C)CO experiment and signal intensities differ significantly for both prochiral protons. On the other hand, the signals in the CH₂–NOESY–H(C)CO experiment (Fig. 4b) are almost identical, which makes stereospecific assignment unreliable.

The 4D versions of H(C)CO- and HSQC-based experiments make stereospecific assignments of all methylene resonances straightforward. Figure 5 shows two examples taken from a 4D CHD–HMQC–NOESY–HSQC. Since the CHD filtration occurs before the NOESY mixing time, we display only ω₁(13C)/ω₄(1H) planes (the latter dimension is detected directly), which can easily be assigned to distinct diastereotropic protons. These planes exhibit excellent spectral resolution and provide structural information that is in agreement with the crystal structure (PDB code 4icb). Like in Figs. 3 and 4, the efficiency of CHD-filters is confirmed by the differentiation between cross-relaxation rates involving diastereotropic protons.

In conventional protein structure determination, stereospecific assignments are performed in a refinement phase at the end of the procedure, through the analysis of preliminary structures. We have explored the possibility of obtaining stereospecific assignments directly during the process of automated NMR structure determination. Both 3D CHD-filtered spectra are used as input for a modified version of the UNIO–ATNOS/CANDID procedure that iteratively identifies NOESY cross-peaks, makes NOE assignments and structure calculations. The 3D structure of the nth iteration is used to obtain an increasingly reliable and complete interpretation of NOESY signals in the (n + 1)th iteration. Diastereotropic atoms that have not been assigned stereospecifically are systematically swapped between pro-R and pro-S assignments. At the outset of the structure calculation, all prochiral atoms are subjected to this atom-swapping procedure. In subsequent iterations, a preliminary 3D structure is used to calculate a score based on distance restraints for the two possible assignments of each prochiral group. If this score shows good agreement with the experimental restraints for only one of the two configurations, then a new stereospecific assignment is made and used in the next iteration of the NOE assignment and structure calculation.

In order to assess the advantages of stereospecific assignments, two separate UNIO–ATNOS/CANDID calculations were performed. Without stereospecific assignments, we obtained 1130 meaningful distance restraints that led to a pairwise backbone RMSD of 0.687 Å and an average backbone RMSD of 1.583 Å for residues 1–41 and 45–76 with respect to the crystal structure (PDB code 4icb).
With our automated stereospecific assignment routine, UNIO–ATNOS/CANDID yielded 1078 meaningful distance restraints, a pairwise backbone RMSD of 0.695 Å, and an average RMSD of 1.293 Å with respect to the X-ray structure, i.e., a significant improvement of ~0.3 Å. No less than 71 of the 110 diastereotopic side-chain methylene groups were automatically assigned, in good agreement with manual assignments based on 3D CHD–NOESY–H(C)CO spectra (see Table 1). While the number of meaningful distance restraints and the precision of the resulting structure bundles were comparable, a significant increase of the accuracy of atomic coordinates (~0.3 Å) was observed for the average coordinates of the NMR structure bundle calculated with automatically determined stereospecific assignments.

In order to compare the approaches described here, stereospecific assignments of methylene groups adjacent to side-chain carboxyl and carbonyl groups have been made with four different methods, as summarized in Table 1. Assignments were obtained by comparing the analysis of 3D CHD–NOESY–H(C)CO spectra with (1) the NMR structure of PDB code 1b1g, (2) the crystal structure PDB code 4icb, and (3) the new NMR structure described above. Automated stereospecific assignments (4) were also compared to manual assignments. Automated and manual assignments from the new NMR structure were identical, showing that the automated process worked well. The consensus assignment for the methylene protons was defined as the one obtained in at least two of the three structures (1, 2, 3). Wrong stereospecific assignments were obtained for only one methylene group in the crystal structure PDB code 4icb and in the NMR structure PDB code 1b1g, as well as for only two methylene groups in the new NMR structure. In at least one case, i.e., for Asp58, the error stems from the improper description of the interaction with a Ca^{2+} ion, which is always a difficult task in NMR.

Differences in populations of isotopomers (for instance for C^{13}H^{2} sites of serines and aromatic residues) have a minor impact on the resulting global and local NMR structure, as documented by the RMSD values given above and the large number of stereo-specific assignments obtained for these residues (we obtained automatic stereo-specific assignments for 3 of 5 serines and for all 5 phenylalanines). UNIO–ATNOS/CANDID uses an $r^{-6}$ relationship between NOE cross peak volumes and upper distance bounds. The lower bounds are set to the van der Waals radii of the involved atoms. The calibration constant is automatically set so that the structure of the preceding iteration does not violate more than a predetermined percentage of all upper distance bounds. When the cross-peak intensities are scaled down, the upper bound will be too large, resulting in a certain loss of information. The cooperative effect of many NOE distance constraints
usually compensates at least in part for this loss of information. When the cross-peak intensities are scaled up, the corresponding upper distance bound might lead to a consistent constraint violation and hence will be eliminated from the structure calculation. In general, the applied automated structure-based calibration method is sufficiently robust to handle slightly ‘incorrect’ cross peak intensities caused by isotopomer effects or common sources of errors such as spin diffusion. Note that differences in populations of CH\textsubscript{pro-R}D\textsubscript{pro-S} and CD\textsubscript{pro-R}H\textsubscript{pro-S} isotopomers could be used to guide stereospecific assignments, as demonstrated for carbon isotopomers in solid-state NMR of proteins (Castellani et al. 2002).

### Conclusions

We have presented a set of novel NOESY experiments aimed at distinguishing the cross-relaxation patterns of prochiral methylene CHD protons in partially deuterated proteins. They can be carried out in a three- or in four-dimensional manner if resolution needs to be boosted. The robustness of the stereospecific assignments benefits from the differences between cross-relaxation rates involving diastereotopic protons. Stereospecific assignments were also obtained by conventional means, by inspecting hypothetical three-dimensional structures. A large number of consistent stereo-specific assignments can be obtained.
by automated NMR structure determination. The use of stereospecific assignments leads to a remarkable improvement of the atomic coordinates.

**Experimental section**

A 100 µL sample of the calcium-loaded P43G mutant of calbindin D<sub>9k</sub> with uniform 15<sup>N</sup> and 13<sup>C</sup> labeling and 50 % random deuteration was used (protein concentration 4 mM, pH 6). Three-dimensional experiments were performed on a 600 MHz Bruker Avance spectrometer equipped with a TXI room-temperature probe and triple-axis gradients. Four-dimensional experiments were run on a Varian 700 MHz spectrometer equipped with a room-temperature triple resonance probe. The two 4D data sets were processed using a signal separation algorithm (SSA) (Stanek et al. 2012) leading to a reduction of the effective noise to the thermal noise level from 1.6 to 1.0, and from 7.4 to 1.1 in 4D HMQC–NOESY–H(C)CO and CC-edited NOESY, respectively. Noise levels in the 4D spectra were calculated by averaging over 38 and 363 points in the ω<sub>d</sub>(1H) dimension at the coordinates of auto-correlation peaks. The durations of the experiments were: 70 h for the 3D CHD–NOESY–H(C)CO, 111 h for the CHD–NOESY–CT-HSQC, 148 h for the 4D CHD–HMQC–NOESY–H(C)CO, and 149 h for the 4D CHD–HMQC–NOESY–HSQC. Such durations are similar or slightly longer than what is typically used to record conventional NOESY spectra.

The UNIO–ATNOS/CANDID algorithm comprises seven iterations that differ in increasing threshold values for the acceptance of NOE assignments.(Herrmann et al. 2002a, b) In the standard UNIO–ATNOS/CANDID protocol, 80 conformers are calculated, and the 20 conformers with the lowest target function values are selected. These conformers are then used to guide the NOESY analysis of the following iteration. For each prochiral group \( j \) (\( j = 1 \ldots M \)), three different average target functions are calculated for each bundle of \( N \) conformers.

\[
\begin{align*}
T_{\text{swap}}^{j} & = \frac{1}{N} \sum_{i=1}^{N} T_{\text{swap}}^{i,j}, \\
T_{\text{proS}}^{j} & = \frac{1}{N} \sum_{i=1}^{N} T_{\text{proS}}^{i,j}, \\
T_{\text{proR}}^{j} & = \frac{1}{N} \sum_{i=1}^{N} T_{\text{proR}}^{i,j}
\end{align*}
\]

(1)

The first target function is calculated by ‘atom swapping’ of all prochiral groups. The second target function is calculated by assuming the pro-R configuration for the prochiral group \( j \) and atom swapping for all other prochiral groups. The third target function is calculated like the second one but assuming the pro-S configuration for the prochiral group \( j \). A prochiral group \( j \) is definitely assigned to the pro-R configuration if \( j \) exhibits at least 2 distance restraints, provided the following three conditions are simultaneously fulfilled:

(a) \[ 1.1 T_{\text{proR}}^{j} < T_{\text{proS}}^{j} \]

(b) \[ T_{\text{proR}}^{j} < 1.1 T_{\text{swap}} \]

(c) \[ T_{\text{proR}}^{j} + 0.2\AA^2 < T_{\text{proS}}^{j} \]

The first condition assures that the energy difference between the two configurations is at least 10 %. The second condition checks if a consistent stereo-assignment can be achieved for all \( N \) conformers. The last criterion guarantees safe stereo-assignments in cases where the set of distance restraints is in good agreement with the bundle of conformers for both configurations. Note that in early iterations, the 3D structure is usually slightly distorted due to erroneous NOE assignments and distance restraints, so that the average target function value will be higher in early runs than in later iterations. Therefore stereo-specific assignments can lead to an increase of the target function compared to the atom-swapping technique, but this increase should be below 10 % in early iterations (criterion b). In later iterations, the target function will always be close to zero, so that criterion c will gain in importance.

**Acknowledgments** We thank Mikael Akke (Lund University) for a sample of partially deuterated calbindin D<sub>9k</sub>, as well as Dominique Frueh (Johns Hopkins University) and Lewis Kay (University of Toronto) for fruitful suggestions. This research was supported by the Polish budget funds for science in 2013–2014 (Project IP2012 057872 awarded to J.S.) Access to the Research Infrastructure at University of Warsaw was financed by the European Commission’s FP7 (Contract 228461, EAST-NMR). Financial support of the Bio-NMR Project No. 261863 is gratefully acknowledged.

**References**

Atreya HS, Chary KVR (2001) Selective ‘unlabeling’ of amino acids in fractionally C-13 labeled proteins: an approach for stereospecific NMR assignments of CH3 groups in Val and Leu residues. J Biomol NMR 19:267–272

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

Cai ML, Liu JH, Gong YX, Krishnamoorthi R (1995) A practical method for stereospecific assignments of gamma-methylene and delta-methylene hydrogens via estimation of vicinal H–1–H–1 coupling-constants. J Magn Reson, Ser B 107:172–178. doi:10.1006/jmrb.1995.1074

Carlomagno T, Peti W, Griesinger C (2000) A new method for the simultaneous measurement of magnitude and sign of D-(1(CH) and D-(1(HH)) dipolar couplings in methylene groups. J Biomol NMR 17:99–109. doi:10.1023/a:10083469020500
Muller L (1979) Sensitivity enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. J Am Chem Soc 101:4481–4484

Neri D, Szyperski T, Otting G, Senn H, Wuthrich K (1989) Stereospecific nuclear magnetic-resonance assignments of the methyl-groups of valine and leucine in the DNA-binding domain of the 434-repressor by biosynthetically directed fractional C-13 labeling. Biochemistry 28:7510–7516

Neuhaus D, Williamson MP (2000) The nuclear Overhauser effect in structural and conformational analysis, 2nd edn. Wiley, New York

Nietlispach D et al (1996) An approach to the structure determination of larger proteins using triple resonance NMR experiments in conjunction with random fractional deuteration. J Am Chem Soc 118:407–415. doi:10.1021/ja952207b

Oktaviani NA, Otten R, Dijkstra K, Scheek RM, Thulin E, Akke M, Mulder FAA (2011) 100% complete assignment of non-labile (1)H, (13)C, and (15)N signals for calcium-loaded calbindin D(9 k). Biomol NMR Assign 5:79–84. doi:10.1007/s12104-010-9272-3

Orts J, Vögel B, Riek R, Güttert P (2013) Stereospecific assignments in proteins using exact NOEs. J Biomol NMR 52:211–218. doi:10.1007/s10858-013-9780-4

Olst er G et al (1993) Stereospecific assignments of the leucine methyl resonances in the H-1-NMR spectrum of lactobacillus-casei dihydrofolate-reductase. FEBS Lett 318:177–180

Palmer AG, Cavanagh J, Wright PE, Rance M (1991) Sensitivity improvement in proton-detected 2-dimensional heteronuclear correlation NMR-spectroscopy. J Magn Reson 93:151–170. doi:10.1016/0022-2364(91)90036-s

Paquin R, Ferrage F, Mulder FAA, Akke M, Bodenhausen G (2008) Multiple-timescale dynamics of side-chain carboxyl and carbonyl groups in proteins by 13C nuclear spin relaxation. J Am Chem Soc 130:15805

Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. J Biomol NMR 2:661–665. doi:10.1007/BF02192855

Plevin MJ, Hamelin O, Boisbouvier J, Gans P (2011) A simple biosynthetic method for stereospecific resonance assignment of prochiral methyl groups in proteins. J Biomol NMR 49:61–67. doi:10.1007/s10858-010-9463-3

Pristovsek P, Franzoni L (2006) Stereospecific assignments of protein NMR resonances based on the tertiary structure and 2D/3D NOE data. J Comput Chem 27:791–797. doi:10.1002/jcc.20389

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

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

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

Stanek J, Augustyniak R, Kozminski W (2012) Suppression of sampling artefacts in high-resolution four-dimensional NMR spectra using signal separation algorithm. J Magn Reson 214:91–102. doi:10.1016/j.jmr.2011.10.009

States DJ, Haberkorn RA, Ruben DJ (1982) A two-dimensional nuclear Overhauser experiment with pure absorption phase in 4 quadrants. J Magn Reson 48:286–292

Svensson LA, Thulin E, Forsén S (1992) Proline cis-trans isomers in calbindin D9K observed by X-ray crystallography. J Mol Biol 223:601–606

Takeda M, Terauchi T, Kainosho M (2012) Conformational analysis by quantitative NOE measurements of the beta-proton pairs across individual disulfide bonds in proteins. J Biomol NMR 52:127–139. doi:10.1007/s10858-011-9587-0

Vallurupalli P, Hansen DF, Lundstrom P, Kay LE (2009) CPMG relaxation dispersion NMR experiments measuring glycine H-1(alpha) and C-13(alpha) chemical shifts in the ‘invisible’ excited states of proteins. J Biomol NMR 45:45–55. doi:10.1007/s10858-009-9310-6

Vögel B, Segawa TF, Leitz D, Sobol A, Choutko A, Trzesniak D, van Gunsteren W, Riek R (2009) Exact distances and internal dynamics of perdeuterated ubiquitin from NOE builds. J Am Chem Soc 131:17215–17225. doi:10.1021/ja905366h

Vuister GW, Bax A (1992) Resolution enhancement and spectral editing of uniformly C-13-enriched proteins by homonuclear broad-band C-13 decoupling. J Magn Reson 98:428–435. doi:10.1016/0022-2364(92)90144-v

Wagner G, Braun W, Havel TF, Schaumann T, Go N, Wuthrich K (1987) Protein structures in solution by nuclear-magnetic-resonance and distance geometry—the polypeptide fold of the basic pancreatic trypsin-inhibitor determined using two different algorithms, DISGEO and DISMAN. J Mol Biol 196:611–639. doi:10.1016/0022-2836(87)90037-4

Yamazaki T, Yoshida M, Nagayama K (1993) Complete assignments of magnetic resonances of ribonuclease-H from Escherichia-coli by double-resonance and triple-resonance 2D and 3D NMR spectroscopies. Biochemistry 32:5656–5669. doi:10.1021/bi00072a023