Graphical interpretation of Boolean operators for protein NMR assignments

Dries Verdegem · Klaas Dijkstra · Xavier Hanoulle · Guy Lippens

Received: 31 March 2008 / Accepted: 9 June 2008 / Published online: 2 September 2008
© Springer Science+Business Media B.V. 2008

Abstract We have developed a graphics based algorithm for semi-automated protein NMR assignments. Using the basic sequential triple resonance assignment strategy, the method is inspired by the Boolean operators as it applies “AND”-, “OR”- and “NOT”-like operations on planes pulled out of the classical three-dimensional spectra to obtain its functionality. The method’s strength lies in the continuous graphical presentation of the spectra, allowing both a semi-automatic peaklist construction and sequential assignment. We demonstrate here its general use for the case of a folded protein with a well-dispersed spectrum, but equally for a natively unfolded protein where spectral resolution is minimal.

Keywords Computer-aided sequential assignment · Graphical semi-automatic protein assignment method · Boolean operators in NMR · Assignment of structured proteins · Assignment of unfolded proteins

Introduction

this strategy run into trouble when (i) overlap of the amide resonances increases due to the size or the unstructured nature of the protein, or (ii) spectral incompleteness due to intermediate line broadening or other phenomena. Both operations of peak list construction and frequency matching will suffer under those conditions, leading the operator back to the physical spectra, where one will manually try to complete the data.

We present here an assignment strategy that simultaneously allows both peak list construction and frequency matching in a semi-automatic manner, while remaining close to the initial spectra. It is based on a graphical interpretation of the Boolean AND operator, i.e. a point-by-point multiplication of 2D spectra. The main advantage is that the operator can walk graphically through the protein sequence, while maintaining a quality evaluation of the experimental data that lead to a decision on a sequential assignment. Although point-by-point operations (addition, subtraction, multiplication or division) are very commonly performed on FID’s, they can also be done on frequency domain spectra and have even been introduced yet in the field of NMR spectra assignments (Masse et al. 2006). However, in our method, these operations play a more prominent role. Demonstrating the principles first on the well-folded Cyclophilin B protein, we extend its application towards a fragment of the natively unfolded Tau protein (Tau F3, amino acids 208–324), where extreme spectral overlap leads to strong degeneracies in the resonance frequencies.

**Theory and methods**

**The assignment principle**

Starting from the root $^1$H,$^{15}$N HSQC spectrum and clicking on a certain appearing peak, our program readily extracts the corresponding $^1$H,$^{13}$C planes from the CBCA(CO)NH and HNCO spectra (Fig. 1). On the basis of these two $^1$H,$^{13}$C spectra, the operator defines with the mouse the carbon frequencies corresponding to the $(i-1)$ residue. These are automatically stored in a peak list (without...
assignment at this moment), and the corresponding $^1$H,$^{15}$N planes are extracted from the HNCACB and HN(CA)CO spectra. Rather than displaying the three corresponding planes together on screen and determine by eye the coordinates where there is simultaneous intensity, we impose this criterion by a point-by-point multiplication of the planes. This corresponds to a graphical interpretation of the Boolean AND operator, that requires simultaneous intensity in the spectra to obtain a resulting spectrum with a detectable intensity (Fig. 2). Applied to the three $^1$H,$^{15}$N spectra extracted at the carbon frequencies of the $(i - 1)$ residue, the point-by-point multiplication therefore defines a novel $^1$H,$^{15}$N HSQC spectrum that contains intensity at the position of the $(i - 1)$ residue.

Once the $(i - 1)$ residue position has been found, it can be used as the starting point of another run of the algorithm. The repeated execution of the routine allows for an assigning walk through the spectrum towards the N-terminus of the protein.

In order to obtain product planes with constant highest peak intensities, the final product plane is initially normalized by dividing it by its maximum value (or minimum value if an odd number of negative peaks was involved in the multiplication) and is afterwards multiplied by a constant factor (e.g. 1e10) to finish with a spectrum with “natural” intensities (i.e. with peaks of more or less the same magnitude as the ones in real spectra).

Boolean operators in NMR

The assignment method is based on a graphical interpretation of the Boolean AND operator, that can be implemented as the point-by-point multiplication of spectral matrices (Fig. 2). Likewise, the OR operator would correspond with point-by-point summation. The NOT operation applied to a spectrum does not result in a new spectrum as such, but rather a 0/1 filled matrix of the same size as the original spectrum. Whether a certain element of this matrix is zero or one is determined by a chosen threshold (see light grey plane in the 2D-case of Fig. 2). If the intensity at a certain point in the original spectrum exceeds this threshold, the corresponding value in the NOT-matrix is set to zero. In the other case, the NOT-matrix value is set to one. This results in a “spectrum” that display holes at the places where the original spectrum contained peaks. Both OR and NOT operations on spectral planes will prove useful further on when trying to assign proteins with unfavorable amino acid sequences.

Input spectra

In its most basic form, the described algorithm uses the HNCACB, CBCA(CO)NH, HN(CA)CO, HNCO and of course HSQC spectra as input. When these five spectra are applied as depicted in Fig. 1, an assigning “walk” towards the N-terminus of the protein is made. It is however interesting to note that a simple exchange of sequential and intra-residue spectra in the algorithm results in the opposite functionality that allows a “walk” in the opposite direction, towards the C-terminus.

Here, the assignment of Cyclophilin B and Tau F3 spectra will be discussed.

The NMR measurements of both protein samples were performed on a Bruker Avance 600 MHz equipped with a cryogenic triple resonance probe head by using standard Bruker pulse programs. The CypB (185aa, 20.4 kDa) sample contained 600 μM CypB in an aqueous buffer with 50 mM Na$_2$HPO$_4$/NaH$_2$PO$_4$, 60 mM NaCl, 1 mM EDTA, 1 mM DTT, pH 6.35 at 293 K. The Tau F3 (124aa, 13.3 kDa) sample contained 250 μM of protein in a 25 mM Tris-D11, 25 mM NaCl, 2.5 mM EDTA, 2.5 mM DTT aqueous buffer (pH 6.8, 293 K). The acquisition

![Fig. 2 The Boolean operators applied to 2D spectra presented as topographic maps](image)
parameters for the CypB spectra were: 2048 (\(^1\)H) and 256 (\(^{15}\)N) complex points and 32 scans per increment for the HSQC (exp time: 2 h 41 min), 1024 (\(^1\)H), 68 (\(^{15}\)N) and 128 (\(^{13}\)C) complex points and 16 scans per increment for the HNCACB (exp time: 1 day 21 h 23 min), 1024 (\(^1\)H), 104 (\(^{15}\)N) and 142 (\(^{13}\)C) complex points and 8 scans per increment for the CBCA(CO)NH (exp time: 1 day 15 h 20 min) and 1024 (\(^1\)H), 104 (\(^{15}\)N) and 128 (\(^{13}\)C) complex points and 8 scans per increment for the HN(CA)CO and HNCO (exp times: 1 day 10 h 58 min and 1 day 10 h 25 min). For the Tau F3 sample, the acquisition parameters were: 2048 (\(^1\)H) and 256 (\(^{15}\)N) complex points and 64 scans per increment for the HSQC (exp time: 5 h 25 min), 2048 (\(^1\)H), 96 (\(^{15}\)N) and 232 (\(^{13}\)C) complex points and 8 scans per increment for the HNCACB (exp time: 2 days 13 h 36 min), 2048 (\(^1\)H), 96 (\(^{15}\)N) and 132 (\(^{13}\)C) complex points and 8 scans per increment for the CBCA(CO)NH (exp time: 1 day 11 h 42 min), 2048 (\(^1\)H), 92 (\(^{15}\)N) and 96 (\(^{13}\)C) complex points and 8 scans per increment for the HN(CA)CO and HNCO (exp times: 1 d 10 h 58 min and 1 d 10 h 25 min) and 2048 (\(^1\)H), 86 (\(^{15}\)N) and 96 (\(^{13}\)C) complex points and 8 scans per increment for the HNN (exp time: 23 h 26 min).

It is important to notice that the different spectra required for an assignment of this kind should all be recorded under the same sample conditions. Any technique based on the point-by-point multiplication of spectrum slices originating from different spectra is obviously quite sensitive to small differences in chemical shift across those spectra, which might arise if nonidentical parameters are used.

Results and discussion

Graphical walk through the triple resonance spectra

To demonstrate the procedure on a real-life example, we start from the cross peak at 7.55, 121.72 ppm in the CypB HSQC spectrum, that we previously assigned to Lys 149 (Hanouille et al. 2007). Extracting the \(^1\)H,\(^{15}\)N planes from the HN(CA)CO, HNCACB (C\(_\alpha\) and C\(_\beta\)) at the (i − 1) \(^{13}\)CO, \(^{13}\)C\(_\alpha\) and \(^{13}\)C\(_\beta\) carbon frequencies as defined by the HNCO and HN(CO)CACB lines at the Lys 149 position yields the three planes shown in Fig. 3. The product plane (4) (in green) comes into being as a result of their point-by-point multiplication. This plane superposed on the Cyclophilin B HSQC indisputably points out the position of the root signal of Arg 148.

When lowering the threshold, we do see other amide resonances of lower intensity, indicating that due to the limited resolution in the carbon dimension residues can have some residual intensity that matches the three required frequencies. When a given (\(^1\)H,\(^{15}\)N) correlation peak represents two or more residues, the operator is faced with the same problem as the number based algorithms. However, as in other semi-automated assignment programs, our method, inherent to its principle, constantly shows the relevant spectrum slices on screen. The obvious advantage is that the raw data with all the information about subtle frequency differences and/or peak forms are still available. Two real situations where only working with raw data helps to exclude ambiguity in the assignment of the CypB protein are considered here.

Figure 4 shows the plane pulled out from the cyclophilin B CBCA(CO)NH spectrum after clicking the Val 12 residue signal. This plane can then, according to the product plane (AND) algorithm, be used to select the C\(_\alpha\) and C\(_\beta\) (i − 1) signals. The Val 12 (\(^1\)H,\(^{15}\)N) correlation peak appears in a more crowded region of the HSQC. Four \(^{13}\)C peaks can be distinguished in the \(^1\)H,\(^{13}\)C plane, but visual inspection readily allows to pair the peaks at 61.5 and 69.0 ppm. The two other signals at 41.0 and 54.0 ppm have a proton frequency that differs by 0.006 ppm from the previous pair, and would therefore probably be assigned to the same peak by automated assignment routines that commonly apply a proton uncertainty of 0.05 ppm.

A second example illustrating the advantage of having ready access to the raw data is found when trying to assign the amide peak that corresponds to Gly 31. The superposition of the HSQC and the product plane obtained after clicking Leu 32 is shown in Fig. 5. We are faced here with the extreme, but possible situation in which the authentic (i − 1) signal is not the most intense one in the product plane. To establish and overcome this problem however, a simple feedback strategy, that exploits once again the usefulness of being able to graphically present the slight chemical shift differences, is sufficient.

This feedback functionality graphically compares the set of peaks involved, as in Fig. 6 for the Leu 32 case, and reveals clearly that the Gly 138 C\(_\alpha\) chemical shift is shifted slightly downfield compared to the Leu 32 C\(_\alpha\) (i − 1) shift.

At any point, the spectroscopist can decide not to include one of the three plane subject to the multiplication (bottom Fig. 1). If for example a certain residue has a weak C\(_\beta\) peak in the HNCACB spectrum, one can exclude the corresponding C\(_\beta\)-plane from the (i − 1) product plane calculation (and thus treat the residue as if it were a glycine). Although this practice will in theory lead to a less selective product plane, it can in some cases avoid the situation where the product plane exhibits a too low signal/noise to be useful.

Using our graphical walk, that is in this case only interrupted when one encounters a proline residue as these do not appear in a HSQC spectrum, we were able to repeat the full assignment of CypB in a minimal time (less than a
day), and comparison with our previous assignment based on peak lists showed perfect agreement.

Extending to natively unstructured proteins

Natively unfolded proteins represent a different challenge to assignment programs. With a reduced amide proton chemical shift range (often inferior to 1 ppm), overlap becomes very severe, leading to many branching points for the automatic matching algorithms. Therefore, manual intervention of the operator becomes even more important than in the case of folded proteins such as CypB, as it allows to alleviate possible ambiguities on the basis of subtle peak position or shape differences.

In this unstructured protein category, the algorithm was powerful enough to determine almost all the proline bordered amino acid stretches of the Tau F3 (amino acids 208–324) protein fragment. We were able to assign all residues except for the S237–S238 pair and the two GGG triplets starting at G271 and G302. The pair and triplets occur in heavily overlapped regions and are moreover preceded by a proline residue preventing the upstream graphical walk.

The performance of the product plane algorithm can however, for natively unfolded proteins, be improved when combined with the information included within the triple resonance HNN spectrum, that can be recorded with a decent sensitivity for these protein due to their narrow line widths. When a certain HSQC root is chosen with the mouse, a corresponding \(^1\text{H},^{15}\text{N}\) plane can also be pulled out of this latter 3D spectrum, in which one can find the \(^1\text{H},^{13}\text{C}\) carbon frequencies yields the three planes (1), (2) and (3). The resulting product plane (in green) superposed on the original HSQC-spectrum is presented in (4). In it, the Arg 148 position can be clearly identified. All clicking required to obtain the first three planes can be done fairly precisely because of an available zoom function. This also enables one to backup precise backbone \(^1\text{H},^{13}\text{C},^{15}\text{N}\), and side-chain \(^1\text{H},^{13}\text{C}\beta\) assignments to an output file.

**Fig. 3** When clicking the Lys 149 \(^1\text{H},^{15}\text{N}\) signal in the CypB HSQC, automatic extraction of the corresponding \(^1\text{H},^{13}\text{C}\) planes of the HNCO and HN(CO)CACB allow the manual definition of the \((i-1)\) CO, \(C_x\) and \(C_\beta\) frequencies. Extracting the \(^1\text{H},^{15}\text{N}\) planes from the HN(CA)CO, HNCACB (\(C_x\) and \(C_\beta\)) at these \((i-1)\) \(^1\text{C}\), \(^1\text{C}_x\) and \(^1\text{C}_\beta\) carbon frequencies yields the three planes (1), (2) and (3).
feedback strategy of graphically comparing the C$_a$ and C$_\beta$ shifts involved shows that there is a perfect match with none of those three. The HNN info on the other hand, allows one to pinpoint the largest product plane signal as the genuine Ser 305 amide resonance. This situation arose because of the almost complete overlap of the Ser 262, Ser 293 and Ser 305 HSQC signals that all have a glycine residue in the (i - 1) position. This caused the individual HNCACB signals to have merged to three new averaged glycine C$_a$, serine C$_a$ and serine C$_\beta$ signals at different chemical shifts.

When the HNN was added to list of input spectra we succeeded in completely assigning the Tau F3 spectra, including the eight earlier mentioned difficult cases. For
unstructured proteins, the HNN information also makes the assignment a lot faster, since it efficiently prevents the need for a signal position based feedback in the assigning walk.

Generating starting points

The above described procedure leads to the ready assignment of stretches of connected resonances. Based on the residue-type specific carbon chemical shifts, that unambiguously define residues such as Gly, Ala, Ser and Thr, those stretches can in most cases be mapped in a straightforward way onto the protein sequence. For the case of a folded protein such as CypB in our example, this information is ample, and a full assignment can be easily obtained. For the Tau fragment, however, the rapid obtention of suitable starting points helps in the procedure, and avoids problems with repeating stretches in the protein sequence. Generally, the existence of suitable starting points in the assignment procedure gives additional confidence in the method, and leads to a more rapid assignment.

The graphical interpretation of the Boolean operators as defined above can equally be used in a similar way by including the OR operator to allow for some spectral degeneracy. A first manner is to define a given residue type by the requirement that both the $C_a$ and $C_b$ frequencies fall within a certain range of the random coil chemical shift values for this residue type. This requirement can be obtained graphically in two steps (Fig. 8): first, the HNCACB $^1H,^{15}N$ planes with the $^{13}C$ chemical shift values within the defined range of the random coil values are summed, leading to a $C_a$- and $C_b$-defining plane. Formally, this sum procedure is equivalent to the Boolean OR operator. In a second stage, we multiply both resulting sum-planes to obtain a novel $^1H,^{15}N$ plane that contains intensity only for those resonances where the $C_a$ and $C_b$ requirement is fulfilled. This procedure is akin to the MUSIC pulse sequences, where one combines carbon selective pulses and multiple quantum filtering to obtain residue-type specific subspectra (Schubert et al. 1999, 2001a,b). However, the present method does not require novel experiments, as it is a post-processing method based only on the existing HNCACB experiment, and thereby does not suffer from the relaxation losses that inevitably accompany the longer pulse lengths required for selectivity. This is a distinct advantage for larger proteins, but also for unfolded proteins where the selectivity of the post-music procedure can easily be fine-tuned on the basis of the same experiment, without requiring the recording of novel experiments.

Following this procedure starting from the HNCACB spectrum, one obtains $(i, i + 1)$ subspectra, as the given residue (i) will also be seen from the $(i + 1)$ amide resonance because of the (weaker) N(i)-C$_a$(i-1) coupling constant. The same principle can equally be applied to the CBCA(CO)NH spectrum, and thereby leads to a subspectrum of only those residues that have the required residue type as their downstream neighbor $(i + 1)$. Applying the third Boolean operator described in Fig. 2, both the HNCACB and CBCA(CO)NH can thus be used to generate...
unique, and the Leu (i + 1) signals are present in the spectrum. In order to generate the pure Asp (i) HSQC (2), the Cα and Cβ windows were 54.2 ± 0.5 and 41.1 ± 0.5 ppm, respectively. The cutoff threshold for the NOT operation (Fig. 2) was put at 0.001% of the most intense CBCA(CO)NH derived product plane peak. The Asp type-specific subspectrum contains all four Asp residues in the Tau F3 sequence. Finally, a Boolean AND operation between (1) and (2) results in a spectrum that only contains intensity at the position of the Asp residue in the unique (L)D283 dipeptide (3). The spectrum manipulations that lead to (3) are done in a few seconds and thus this procedure provides a very fast generation of starting points. Again, as was the case in the assignment method, all planes are normalized before multiplication and multiplied by a constant factor (1e10) after, as to maintain constant intensity. The scales are values in ppm.

However, we found that for unfolded proteins such as Tau, where the defined windows can be kept reasonably small (e.g. 1 ppm) because of the smaller Cα and Cβ chemical shift spreads, this signal/noise reduction is disturbing in only a minor number of cases. Thus, the rapid determination of pivotal points, whereby we can even allow for some ambiguous assignments, greatly enhances not only the initial stages of the assignment procedure. As it provides for suitable anchoring points, it facilitates to connect the sequential stretches to the protein sequence. The complete assignment of Tau F3, using our complete package of assignment tools, was done in 1 day time.

Discussion

We have shown here a graphical implementation of the traditional assignment procedure based on connecting complementary triple resonance experiments. The main advantage of the procedure is that the operator remains very close to the experimental spectra at every moment, without relying on peak lists. Whereas the latter allow a rapid computer-assisted assignment in favorable cases, spectral overlap or differential quality of the data in different zones of the spectra can introduce errors that inevitably will lead to problems requiring manual intervention. The product planes as defined in this work represent the Boolean AND operator in its most simple fashion: point-by-point multiplication guarantees that the only remaining intensity comes from planes that both had only a minor number of cases. Thus, the rapid determination of pivotal points, whereby we can even allow for some ambiguous assignments, greatly enhances not only the initial stages of the assignment procedure. As it provides for suitable anchoring points, it facilitates to connect the sequential stretches to the protein sequence. The complete assignment of Tau F3, using our complete package of assignment tools, was done in 1 day time.

Discussion

We have shown here a graphical implementation of the traditional assignment procedure based on connecting complementary triple resonance experiments. The main advantage of the procedure is that the operator remains very close to the experimental spectra at every moment, without relying on peak lists. Whereas the latter allow a rapid computer-assisted assignment in favorable cases, spectral overlap or differential quality of the data in different zones of the spectra can introduce errors that inevitably will lead to problems requiring manual intervention. The product planes as defined in this work represent the Boolean AND operator in its most simple fashion: point-by-point multiplication guarantees that the only remaining intensity comes from planes that both had only a minor number of cases. Thus, the rapid determination of pivotal points, whereby we can even allow for some ambiguous assignments, greatly enhances not only the initial stages of the assignment procedure. As it provides for suitable anchoring points, it facilitates to connect the sequential stretches to the protein sequence. The complete assignment of Tau F3, using our complete package of assignment tools, was done in 1 day time.

Discussion

We have shown here a graphical implementation of the traditional assignment procedure based on connecting complementary triple resonance experiments. The main advantage of the procedure is that the operator remains very close to the experimental spectra at every moment, without relying on peak lists. Whereas the latter allow a rapid computer-assisted assignment in favorable cases, spectral overlap or differential quality of the data in different zones of the spectra can introduce errors that inevitably will lead to problems requiring manual intervention. The product planes as defined in this work represent the Boolean AND operator in its most simple fashion: point-by-point multiplication guarantees that the only remaining intensity comes from planes that both had only a minor number of cases. Thus, the rapid determination of pivotal points, whereby we can even allow for some ambiguous assignments, greatly enhances not only the initial stages of the assignment procedure. As it provides for suitable anchoring points, it facilitates to connect the sequential stretches to the protein sequence. The complete assignment of Tau F3, using our complete package of assignment tools, was done in 1 day time.

Discussion

We have shown here a graphical implementation of the traditional assignment procedure based on connecting complementary triple resonance experiments. The main advantage of the procedure is that the operator remains very close to the experimental spectra at every moment, without relying on peak lists. Whereas the latter allow a rapid computer-assisted assignment in favorable cases, spectral overlap or differential quality of the data in different zones of the spectra can introduce errors that inevitably will lead to problems requiring manual intervention. The product planes as defined in this work represent the Boolean AND operator in its most simple fashion: point-by-point multiplication guarantees that the only remaining intensity comes from planes that both had only a minor number of cases. Thus, the rapid determination of pivotal points, whereby we can even allow for some ambiguous assignments, greatly enhances not only the initial stages of the assignment procedure. As it provides for suitable anchoring points, it facilitates to connect the sequential stretches to the protein sequence. The complete assignment of Tau F3, using our complete package of assignment tools, was done in 1 day time.

Discussion

We have shown here a graphical implementation of the traditional assignment procedure based on connecting complementary triple resonance experiments. The main advantage of the procedure is that the operator remains very close to the experimental spectra at every moment, without relying on peak lists. Whereas the latter allow a rapid computer-assisted assignment in favorable cases, spectral overlap or differential quality of the data in different zones of the spectra can introduce errors that inevitably will lead to problems requiring manual intervention. The product planes as defined in this work represent the Boolean AND operator in its most simple fashion: point-by-point multiplication guarantees that the only remaining intensity comes from planes that both had only a minor number of cases. Thus, the rapid determination of pivotal points, whereby we can even allow for some ambiguous assignments, greatly enhances not only the initial stages of the assignment procedure. As it provides for suitable anchoring points, it facilitates to connect the sequential stretches to the protein sequence. The complete assignment of Tau F3, using our complete package of assignment tools, was done in 1 day time.
sharp lines, the assignment becomes as trivial as for a folded protein. An extension to the Boolean OR operator allows to define amino-acid specific subspectra based on the original HNCA/CB and CBCA(CO)NH spectrum. When compared to the experimental MUSIC pulse sequences, this procedure does not suffer from additional relaxation losses due to the selective and hence longer carbon pulses, does not require novel experiments and can readily build in differential $^{13}$C selectivity, but can evidently not reproduce the multiple quantum filtering as was done in the some MUSIC sequences. These residue selective subspectra constitute the input of a very straightforward starting point generation method. We showed that this latter procedure is particularly suitable for unfolded proteins, where the random coil $^{13}$C chemical shifts by definition provide an excellent center point for the chemical shift range to be considered. We are currently exploring how the procedure can be combined with quantum-mechanical or semi-empirical chemical shift calculations in order to provide a rapid assignment of the HSQC spectra of proteins with a known 3D structure.

All methods described in this paper were developed using python scripts with the NMR python library functionality (http://linuxnmr02.chem.rug.nl/~dijkstra/NMRpy/). However, we have also implemented them in the CcpNmr software suite (Vranken et al. 2005) as an extension to Analysis2.0 software suite (Vranken et al. 2005) as an extension to ever, we have also implemented them in the CcpNmr software suite. The 600 MHz facility used in this study was funded by the Region Nord—Pas de Calais (France), the CNRS and the Institut Pasteur de Lille. Part of this work was funded by a grant of the Agence National de la Recherche (ANR 05 BLAN 0320-0; Tau:Tubulin). D.V. received a predoctoral grant of the French Minist`ere de la Recherche.

Acknowledgments We thank Dr. I. Landrieu for sample preparation, Dr. J.-M. Wieruszewski for collecting the NMR spectra and Dr. T. Stevens and W. Boucher of the University of Cambridge, Department of Biochemistry for implementing our protein NMR assignment tools in the CcpNmr software suite. The 600 MHz facility used in this study was funded by the Région Nord—Pas de Calais (France), the CNRS and the Institut Pasteur de Lille. Part of this work was funded by a grant of the Agence National de la Recherche (ANR 05 BLAN 0320-0; Tau:Tubulin). D.V. received a predoctoral grant of the French Minist`ere de la Recherche.

References

Andrec M, Levy RM (2002) Protein sequential resonance assignments by combinatorial enumeration using $^{13}$C$_{a}$ chemical shifts and their (i,i – 1) sequential connectivities. J Biomol NMR 23: 263–270


Herrmann T, Güntert P, Wüthrich K (2002a) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Biomol NMR 31:209–227

Herrmann T, Güntert P, Wüthrich K (2002b) Protein NMR structure determination with automated NOE-identification in the NOEASY spectra using the new software ATNOS. J Biomol NMR 24:171–189


