Journal of Biomolecular NMR **27:** 313–321, 2003.
KLUWER/ESCOM

Smartnotebook: A semi-automated approach to protein sequential NMR resonance assignments

Carolyn M. Slupsky, Robert F. Boyko, Valerie K. Booth & Brian D. Sykes*
Protein Engineering Network Centers of Excellence, 713 Heritage Medical Research Center, University of Alberta, Edmonton, Alberta, T6G 2S2, Canada

Received 19 March 2003; Accepted 10 June 2003

Key words: assignment, automated, chemical shifts, NMR, NMRView, semi-automated

Abstract

Complete and accurate NMR spectral assignment is a prerequisite for high-throughput automated structure determination of biological macromolecules. However, completely automated assignment procedures generally encounter difficulties for all but the most ideal data sets. Sources of these problems include difficulty in resolving correlations in crowded spectral regions, as well as complications arising from dynamics, such as weak or missing peaks, or atoms exhibiting more than one peak due to exchange phenomena. Smartnotebook is a semi-automated assignment software package designed to combine the best features of the automated and manual approaches. The software finds and displays potential connections between residues, while the spectroscopist makes decisions on which connection is correct, allowing rapid and robust assignment. In addition, smartnotebook helps the user fit chains of connected residues to the primary sequence of the protein by comparing the experimentally determined chemical shifts with expected shifts derived from a chemical shift database, while providing bookkeeping throughout the assignment procedure.

Introduction

High throughput structure determination is a requisite for structural genomics and proteomics and, of course, is also desirable in more traditional protein structure determination applications. In determining protein structures using NMR, the critical step of obtaining complete and accurate chemical shift assignments is often the most tedious and time consuming. Thus, the automated assignment of NMR spectra of biomacromolecules is a critical objective for automating structure determinations by NMR.

There are many software packages that address the issue of chemical shift assignment. Automated and semi-automated procedures for the assignment of 2D homonuclear NOESY and COSY ^{1}H NMR spectra include ANSIG (Kraulis, 1989), PROSPECT (van de Ven, 1990), EASY (Eccles et al., 1991), CLAIRE (Kleywegt et al., 1991), CPA (Xu and Sanctuary, 1993), CAMRA (Gronwald et al., 1998), and others (Cieslar et al., 1988; Weber et al., 1989; Eads and Kuntz, 1989; Xu et al., 1994). Most of these programs implement at least a portion of the basic Wüthrich sequential assignment strategy (Wüthruch, 1986) where the spin systems are detected and classified, followed by inter-residue correlation of these spin systems and matching to the primary sequence.

The advent of modern triple resonance three and four-dimensional experiments has resulted in a number of programs for automated assignment of multinuclear NMR data that include ALFA (Bernstein et al., 1993), ANSRS (Kraulis, 1994), Pronto (Kjaer et al., 1994), AUTOASSIGN (Zimmerman et al., 1994, 1997; Szyperski et al., 2002), CONTRAST (Olson and Markley, 1994), GARANT (Bartels et al., 1996), PASTA (Leutner et al., 1998), AURELIA (Görler et al., 1999), RESCUE (Pons and Delsuc, 1999), JIGSAW (Bailey-Kellogg et al., 2000), NOAH/DIAMOND (Oezguen et al., 2002), CANDID

*To whom correspondence should be addressed. E-mail: brian.sykes@ualberta.ca.

(Herrmann et al., 2002), and others (Oschkinat et al., 1991; Hare and Prestegard, 1994; Friedrichs et al., 1994; Meadows et al., 1994; Croft et al., 1997; Lukin et al., 1997; Andrec and Levy, 2002). All have slightly different strategies for obtaining resonance assignment (for reviews see Oschkinat and Croft, 1994; Zimmerman and Montelione, 1995; Moseley and Montelione, 1999).

Regardless of the strategy, automating NMR spectral assignment requires several steps including processing of the spectra, abstraction of the data, detection and correlation of patterns, followed by relating the data to the primary sequence. While data abstraction by means of peak-picking can be done manually, automatically, or a combination of both, no peak-picking procedure can be expected to yield precise peak positions in all practical situations. In addition, overlapping and closely spaced peaks cause ambiguity in the peak tables. Effects such as sample heating, slow sample degradation, and/or small changes in buffer conditions between samples can lead to changes in peak positions. This leads to the necessity of defining larger tolerances within and between experiments in order to find all possible spin system correlations. A major drawback with defining generous error limits is the increased overlap in the peak lists, resulting in the inability to correctly match peak patterns. Moreover, spectral noise can complicate assignment either by obscuring peaks or simulating peaks, a situation that is difficult to address in an automated process. Although these problems may be partially mitigated through the use of several redundant data sets, in addition to the undesirability of having to acquire extra NMR spectra, a fully automated assignment algorithm will still likely contain errors, forcing the spectroscopist to perform extensive error checking. Unless such user error checking is built directly into the automating software, the user has to follow a procedure comparable to a full manual assignment, resulting in meager time-savings. Another automated assignment approach is to find the set of all possible assignments consistent with the data. This can easily be sizeable leaving the user with the task of finding the correct assignment. Thus, although there have been many attempts at automating NMR spectra assignment, so far no single program has become a paradigm.

Our approach has not been to fully automate the assignment procedure, but rather to work instead towards semi-automation, combining the best features of computer based assignment (i.e., rapidity) and manual assignment (i.e., robustness). This program, called smartnotebook, has been implemented in a flexible way, permitting the spectroscopist to use whichever set of NMR spectra that is preferred allowing any complications in the assignment procedure that might arise, whether or not they are anticipated by the issues discussed above, to be appropriately dealt with. In addition, we anticipate smartnotebook will be useful in providing a way to rapidly check assignments produced using fully automated procedures.

The heart of the smartnotebook procedure lies with the interfacing of different programs such as the graphical based NMRView for display and user interaction, with our in-house written peakcon program for producing inter-residue connections. Because peakcon and NMRView are separate from the smartnotebook interface, a user could conceivably replace either. Indeed, any software that will withstand the test of time should be adaptable, allowing users to make modifications that suit their needs. The flexibility of smartnotebook will make it useful for people wanting to use other types of connections programs, to obtain inter-residue correlations, or other types of graphical interfaces.

The connections procedure peakcon is based upon a rules based deterministic algorithm that uses peak lists to find all connectivity information between a reference data set (for example an HSQC or HNCO) and data sets that give rise to intra- and inter-residue connectivity information (for example HNCACB and CBCA(CO)NH). The user decides on the optimal connectivity by inspecting a quality factor and, if required, inspecting the appropriate regions of the spectra, which are displayed by the software at a click of a button. Once a particular connection has been chosen, smartnotebook uses a chemical shift database to fit the connections to the primary sequence to obtain protein backbone ^{15}N, ^{1}HN, ^{13}Cα and side-chain ^{13}Cβ assignment. The chemical shift data base is, by default, generated for the protein sequence based on shifts from the BioMagRes database (Seavey et al., 1991), but may also be provided by the user, which could be very useful in cases where assignments for a homologous protein or proteins are available (Gronwald et al., 1997). Smartnotebook also performs extensive bookkeeping by keeping track of all decisions made by the user. This information includes the assignments that have been made, as well as chains of residues that have not been assigned. All information can be updated if the user decides to 'unassign' a residue or residues. In addition, smartnotebook is able to interact with NMRView to update the chemical shift (ppm.out) and peak-pick (*.xpk) files. These types of files have

been in wide use for several years and software already exists to convert them into files suitable for input into a number of structure calculation programs.

While smartnotebook has been optimized for use with the HNCACB and CBCA(CO)NH spectra, the software has been designed in such a way as to allow users to modify the spectra displayed and the search pattern to look at other types of intra/inter-residue correlations. This suite of programs is written in C and Tk/Tcl (Ousterhout, 1994) and runs within the graphical based program NMRView (Johnson and Blevins, 1994). Because the computational part of the program has been kept as separate as possible from the graphical interface, there is the potential for users to choose whatever graphical display interface they prefer, providing the user can successfully incorporate the smartnotebook tcl script with their graphical interface of choice as we have done with NMRView. In the same manner, if a user had a different way of making connections, the user could create the connections files quite easily and import them into smartnotebook.

Methods

Description of Smartnotebook software

Smartnotebook was designed in a modular fashion that makes use of the program NMRView (Johnson and Blevins, 1994) as a graphical interface. Each module of smartnotebook has a well-defined task that completes fairly rapidly. A powerful feature of the program is that the user can control the patterns used to define connectivities, allowing the software to be used with any set of NMR spectra. The implementation of the software, however, has been done in the context of the HNCACB (Wittekind and Mueller, 1993) and CBCA(CO)NH (Grzesiek and Bax, 1992a) experiments and thus will be described within the framework of these experiments.

Included below are descriptions of the Tcl/Tk algorithms used in this work. The names of the routines are listed, followed by their functional descriptions. Smartnotebook consists of three main modules: (1) makeshift, which is a chemical shift prediction program (2) peakcon, which finds all connectivities in a set of data based on a rule-set and (3) smartnotebook.tcl, which is a tcl program controlling assignment and chain-building.

Makeshift

The makeshift algorithm produces a list (called 'shifts' stored in the smartnotebook.out directory), specific to the protein's sequence, that contains information about the range of backbone ^{1}H, ^{15}N, and ^{13}Cα and side-chain ^{13}Cβ chemical shifts expected for each residue type. The starting point for makeshift is, by default, the BioMagResBank list of average chemical shifts and standard deviations for diamagnetic amino acids. The shifts file can also be acquired from a database provided by the user (for example if the user has used the homology-based program ORB (Gronwald et al., 1997) to predict chemical shifts or the user wishes to use chemical shifts derived from secondary structure predictions). The tolerance for an acceptable fit between the shifts in the database and the experimental shifts does not need to be pre-defined but rather can be adjusted interactively as the assignment proceeds.

peakcon

The peakcon software finds connections between the CBCA(CO)NH experiment and the HNCACB experiment by matching ^{13}Cα and ^{13}Cβ, using reference peaks from either an ^{15}N-^{1}H HSQC (Bodenhausen and Ruben, 1980) or 3D HNCO (Grzesiek and Bax, 1992b) experiment. Input to the program is peak-picked files from the HNCACB, CBCA(CO)NH and ^{15}N-^{1}H HSQC or 3D HNCO experiments.

Peakcon uses a relational database that is constructed from a set of rules files that describe the association between sets of experiments. These rules files are set up specifically for the HNCACB and CBCA(CO)NH experiments but could be modified for use with other experiments (instructions on how to do this are provided with the software distribution). The rule sets provided with the software are dxx, which finds connections between consecutive non-glycine residues, dgx, for glycine followed by non-glycine, dxg, for non-glycine followed by glycine, and dgg, for glycine followed by glycine. Special consideration is given to glycine residues because they lack a Cβ atom. The glycine rules are also useful for cases where Cα and Cβ are overlapped or one of the peaks is obscured for some reason.

Given the rule sets defined above it is logical to assume that there are going to be a lot of potential glycine connections. In particular, con.dgx will contain many more connections than con.dxg due to the fact that only one peak is correlated for con.dgx in the HNCACB(i$-$1) and CBCA(CO)NH(i) versus two peaks for con.dgx. It is more difficult to correlate

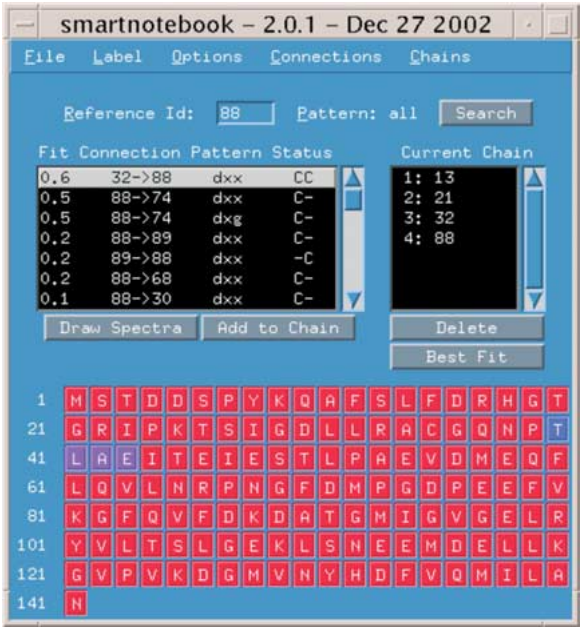

Figure 1. The smartnotebook user interface, implemented with the Tk/Tcl programming language, features a window with the amino acid sequence color-coded for assigned and unassigned residues as well as the current chain. Smartnotebook allows the user to choose and modify connections patterns as well as chains. The user can directly observe each possible connection within the spectral window by clicking on the 'Draw Spectra' button.

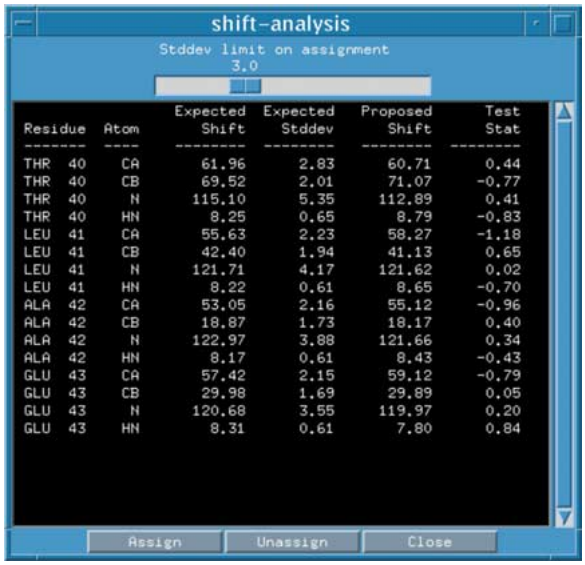

Figure 2. 'Best Fit' window. When the 'Best Fit' button is selected from the smartnotebook window, the user can compare the chemical shifts thought to be associated with the selected sequence with the expected chemical shifts for that sequence. The error limit can be made larger by sliding the standard deviation scale.

two peaks versus one. Similarly, the con.dgg file will contain peaks due to the correlation of a single peak between the HNCACB(i−1) and CBCA(CO)NH(i) spectra. The rules files were set up in this manner to be sure that no connections would be missed. The user has the choice to ignore this data completely when looking at possible chains.

smartnotebook.tcl

This is a Tcl program that interfaces the peak connection algorithm, peakcon, as well as the chemical shift file created with makeshift, with the graphical interface NMRView (Johnson and Blevins, 1994) (Figure 1). The 'smartnotebook' panel of the interface contains a number of buttons and menus for modulating the display, building chains of sequential residues and controlling the assignment bookkeeping.

The display consists of the 'smartnotebook' panel as well as the spectra. A button on the smartnotebook panel allows the regions of the spectra appropriate to the selected connectivity to be drawn in the spectral windows. The 'Label' menu allows the display of peak I.D. numbers to be toggled on or off within the windows of selected spectra.

Central to the smartnotebook procedure is defining connections that identify successive residues followed by their assembly into chains to sequentially connect the residues. The 'Ref I.D.' menu allows the user to specify a particular peak in the ^{15}N HSQC to start with, either by typing in a particular reference peak I.D. number, or by placing the cursor on a peak and choosing 'set from spectra'. The user may also scroll up and down through unassigned peaks. The 'Pattern' menu allows the user to choose various patterns to be queried. The 'Connections' menu allows deleting, undeleting and editing of connections. Deleting a connection simply places an 'x' next to it indicating that the user feels that the connection is not worth considering further. Editing a connection allows the user to change the peak I.D. numbers that make that connection. This feature is useful in crowded regions of the spectrum, or when tolerances are set generously. The 'Folded' menu allows a user to define peaks that have been folded into a particular spectrum so that the correct chemical shift as well as the correct connection pattern may be observed if the folding resulted in negative peaks. Once folded peaks are identified, the connections are re-calculated by peakcon. The 'Chains' menu allows the user to bring up a chain that has been previously started, so that the user may con-

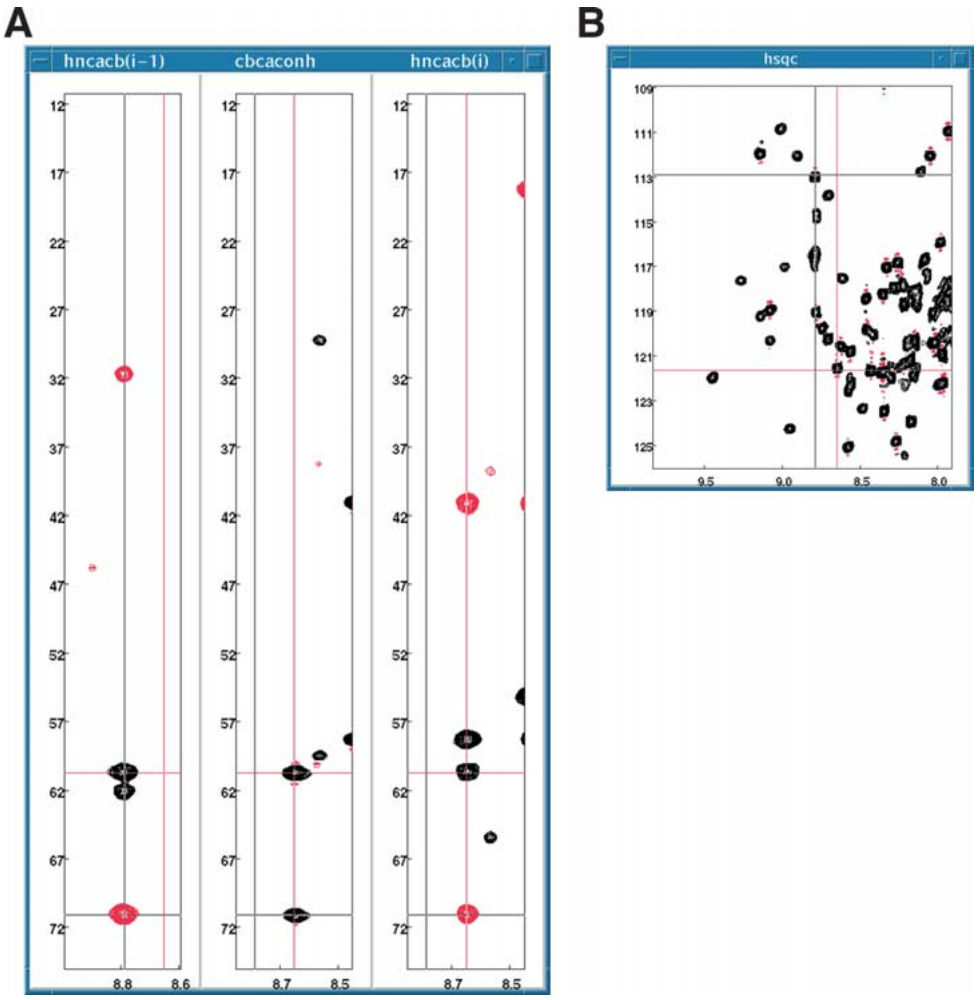

Figure 3. Spectral window layout in smartnotebook. (A) The Strip plots are shown as HNCACB (i-1) – CBCA(CO)NH (i) – HNCACB (i). (B) The single plot is showing a ^{1}H-^{15}N HSQC. The black cursors denote residue 'i−1' ^{1}HN and ^{15}N frequencies on the ^{1}H-^{15}N HSQC. In the HNCACB(i−1) spectrum, the black vertical cursor denotes the (i-1) ^{1}HN frequency. The red cursors represent residue 'i' ^{1}HN and ^{15}N frequencies on the ^{1}H-^{15}N HSQC spectrum. In the HNCACB(i) and CBCA(CO)NH(i) spectra, the red vertical cursor represents the ^{1}HN frequency of residue 'i'. The horizontal cursors on the HNCACB(i) and CBCA(CO)NH(i) spectra represent the correlations to the (i−1) Cα and Cβ.

tinue to add or delete connections. It also allows the user to 'start a new chain'.

The smartnotebook.tcl program also displays the primary sequence of the protein being studied. As peaks are connected together to form chains, the sequence changes color to indicate possible assignments. All matches between the chemical shift of the connecting peaks in the HNCACB – CBCA(CO)NH/ HNCACB spectra and the primary sequence of the protein are colored blue (as a default) on the 'sequence' portion of the smartnotebook panel. The first occurrence of 2 successive amino acids that fit the data the best (for example, the first occurrence of the dipeptide ala – thr) will be displayed as the 'best fit'. The user can then continue adding connections to the dipeptide until there is only one possible fit to the sequence. The 'Best Fit' window shows the comparison between expected chemical shifts for a sequence of amino acids with the actual chemical shifts. If the sequence chosen for the 'best fit' is not what the user feels is the optimal fit, the user has the option of choosing another starting place simply by clicking on a different residue in the sequence. The error limits on the observed chemical shift, in the form of standard deviations, may also be modified by sliding the error bar at the top of the 'Best Fit' sub-window (Figure 2).

Input to Smartnotebook includes a peak-picked reference spectrum (HNCO or HSQC) as well as peak-picked HNCACB and CBCA(CO)NH spectra. The user also needs processed spectra from the HNCACB, CBCA(CO)NH, HSQC (as well as HNCO if using this spectrum for referencing). In addition, the sequence of the protein in NMRView format is needed.

Output from Smartnotebook includes a ppm.out file containing all assignments made; hncacb.xpk, cbcaconh.xpk, hsqc.xpk, and hnco.xpk (if used) containing assignments; a chains file containing a record of all chains made; connectivity files based on the different connectivity types (con.dgg, con.dgx, con.dxg and con.dxx); an edit-connect file that contains the user-editing information concerning connectivity information; fold files containing information about folded peaks in the spectra (cbcaconh.fold, hncacb.fold, hsqc.fold, hnco.fold); as well as a shifts file that contains the expected chemical shift information with the appropriate standard deviations.

Outline of the assignment procedure

By invoking smartnotebook within NMRView, the smartnotebook interface is brought up together with two spectral windows: one with the reference HSQC spectrum, and the other with the HNCACB – CBCA(CO)NH/HNCACB spectra displayed as strips. The left hand HNCACB spectrum displays the i−1 connection whereas the right hand CBCA(CO)NH/HNCACB display the spectra corresponding to 'residue i' (Figure 3). A third window displaying the HNCO reference spectrum may also be displayed if that spectrum is used for referencing. The NMRView message window is used by smartnotebook to display all messages concerning connections, chains, editing, and errors.

The procedure begins with the user deciding upon a peak from the reference spectrum to start with. The user then selects a pattern (dxx, dxg, dgx, dgg, or all patterns) to search for connections to this peak. After the 'search' button is selected, a list of possible connections is displayed in the interface, ranked according to how well the frequencies of the peaks match. Also shown are the peak numbers (from the reference spectrum) of the peaks involved in the connectivity, the pattern used to generate the connectivity and the status of each peak. The status is used to indicate information such as if the peak has already been used to form a chain, or is already assigned. At this point, the user may select a particular connection if the correct one is evident from this list or may choose to inspect the choices by viewing the appropriate regions of the spectra, using the 'draw spectra' button. The correct connection can be indicated by selecting a particular connection on the displayed list and pressing the 'add to chain' button. At this point, the chemical shifts of the chain and the chemical shift database are used to display all possible positions in the sequence that the chain might fit. The tolerance for the fit can be modified from the 'Best Fit' window. As the chain is lengthened the number of possibilities is greatly reduced. At any point, the user can decide to assign the chain to a segment of the sequence using the assign button on the 'best fit' window. If the sequence chosen by the software for the 'best fit' window is not satisfactory, another segment of the sequence can be chosen by clicking on another residue in the displayed sequence. The user can also choose to start a new chain without assigning the first one – it will be saved and can be re-examined later by using the 'chains menu'. Options on the chains menu allow the user to add to chains, delete from chains, assign or unassign chains, until the assignment is complete.

Results and discussion

The Smartnotebook assignment protocol was evaluated by using experimental data obtained on CDC4p (Slupsky et al., 2001) (BMRB assession number 4851) and Pnt (Slupsky et al., 1998) (BMRM assession number 4205). The input data included a 2D ^{1}H-^{15}N HSQC as well as 3D HNCACB and CBCA(CO)NH experiments. Pnt also had a 3D HNCO experiment that was used as a reference spectrum.

Peak connection

The peak connections for both proteins, stored in files within the 'smartnotebook.out' directory, were very large and represented many more connections than what was expected. As previously discussed, the connection files represent connections between two glycine residues (con.dgg), a glycine residue and any other residue (con.dgx), any residue (except glycine) and a glycine residue (con.dgx) and any two residues other than glycine (con.dxx). One reason for the large number of correlations in the connections files has to do with peak overlap and tolerances. When more than one peak in the reference spectrum overlaps, or when two or more peaks are very close together (Figure 4),

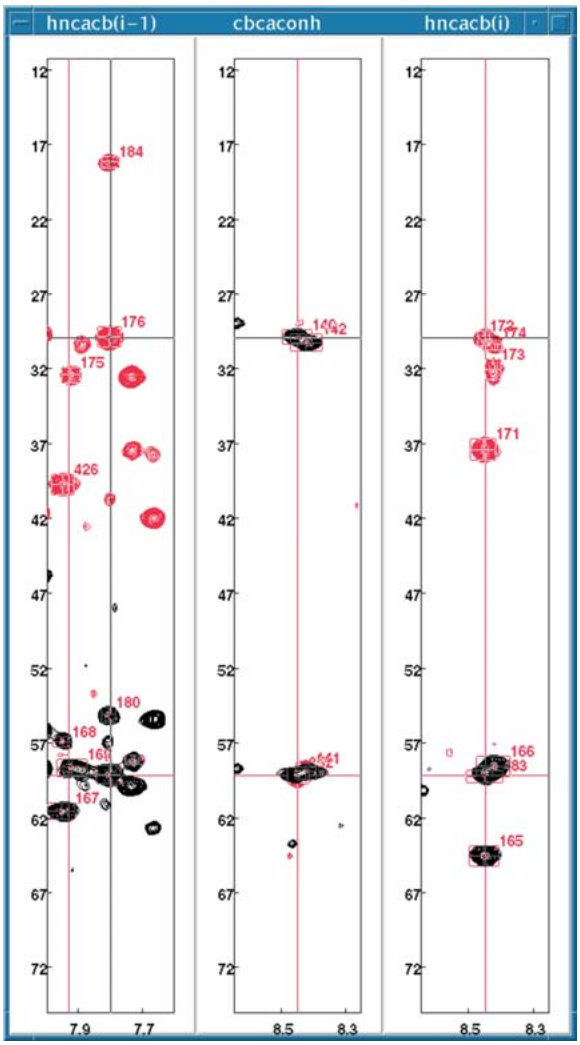

Figure 4. Severe overlap of several peaks in the ^{1}H-^{15}N HSQC spectrum makes it difficult for the program to accurately define associating correlations between the HNCACB and CBCA(CO)NH spectra. Thus, all combinations of peaks will be contained within the connections files and can be specifically selected by the 'edit-connect' function within the 'Connections' menu.

there are many different ways the peaks could correlate. This problem highlights why automated programs generally fail for crowded spectra, in that picking the correct correlation amongst many possibilities is difficult for a program to do. Human judgment is often needed to resolve complicated regions of spectra. Another reason for the large number of correlations is that, upon finding all connections, no filtering is done. This means that for glycine patterns (i.e., dgg, dxg, dgx) there will be an extraordinary number of correlations even if the protein sequence doesn't contain glycine residues. This was done to ensure that the lists of possible correlations were complete and thus no connections were lost due to overlap or spectral artifacts. Complete connection lists ensure that the chosen connections and assignments will be accurate and complete.

When applied to the CDC4p and Pnt datasets, the only difficulty encountered using smartnotebook for assignment involved sequential residues with peaks that were exactly overlapped (N, NH, Cα and Cβ). It was fairly easy for an experienced user to determine that the peaks were overlapped. We chose not to change the program to include peak correlation to itself due to the high number of extra peaks that would need to be queried for each correlation, and the fact that this type of correlation would be difficult to program.

Sequential chemical shift assignment

For Cdc4p, a 141 residue protein, a ^{1}H-^{15}N HSQC spectrum was used as a reference spectrum. In this case, 1388 potential con.dxx connections were found. In addition, 2144 con.dgx, 277 con.dxg, and 771 con.dgg connections were found. Based on these lists of potential residue connections, almost all residues were assigned. Residues not assigned included M1-T3, the N-terminal residues, F86, T91 and G92, residues in the flexible central linker region that were unobservable in the HNCACB and CBCA(CO)NH spectra, and H132, another residue unobservable in the HNCACB and CBCA(CO)NH spectra. Several proline residues broke up continuous chains and included residues 7, 24, 39, 52, 67, 73, 76, and 123. The assignments were based on the identification and assignment of 14 different chains: 1) Asp4 to Ser6, 2) Tyr8 to Ile23, 3) Lys25 to Asn38, 4) Thr40 to Leu51, 5) Ala53 to Arg66, 6) Asn68 to Gly69, 7) Phe70 to Met72, 8) Glu74 to Asp75, 9) Glu77 to Val85, 10) Asp87 to Ala90, 11) Met93 to Gly95, 12) Val96 to Val122, 13) Val124 to Tyr131, and 14) Asp133 to Asn141. All chemical shifts were consistent with previous published chemical shifts (Slupsky et al., 2001). It took approximately 2 days to complete the assignment procedure using Smartnotebook.

For Pnt, a 110 residue protein, the HNCO spectrum was used as a reference spectrum. In this case, 1838 con.dxx connections were found. For the glycine connections, 1393 con.dgx, 139 con.dxg and 353 con.dgg potential connections were obtained. The assignment proceeded in 12 chains. Residues Met1 and Glu2 at the

N-terminus could not be assigned due to rapid amide exchange. Residues Pro7, Pro11, Pro38, Pro41, and Pro90 served to break chains. Residues that could not be assigned included Ser12, Ser13, Lys14, Gly64 and Gly74. These were residues that were not previously assigned using manual methods and were found in loop regions. It took less than one day to perform the assignment procedure using Smartnotebook.

Conclusions

Smartnotebook is a convenient package that aids a user in obtaining sequential assignment. It can be used to attain assignments with speeds approaching fully automated procedures, with all the robustness of traditional manual procedures. Smartnotebook has been set-up with flexibility in mind, so that the user can modify it for use with any set of three-dimensional experiments. The software uses NMRView as the graphical interface, includes extensive bookkeeping and can input and output the commonly used NMRView format peak list and assignment files. The software should be easily combined with single amino acid type labeling as well as specific amino acid identification utilizing the new MUSIC pulse sequences (Schubert et al., 1999, 2001a, b). The program is open source and is freely available from http://www.pence.ualberta.ca/software/smartnotebook. Also available on the website are extensive instructions on how to use and customize smartnotebook, including how to define the rules files for use with additional types of spectra.

Acknowledgements

The authors wish to express their gratitude to Ms Leigh Willard for helpful suggestions and Dr Lawrence McIntosh for allowing the use of data sets from his laboratory. The authors wish to acknowledge funding from PENCE.

References

Andrec, M. and Levy, R.M. (2002) *J. Biomol. NMR*, **23**, 263–270.
Bailey-Kellogg, C., Widge, A., Kelley, J.J., Berardi, M.J., Bushweller, J.H. and Donald, B.R. (2000) *J. Comput. Biol.*, **7**, 537–558.
Bartels, C., Billeter, M., Güntert, P. and Wüthruch, K. (1996) *J. Biomol. NMR*, **7**, 207–213.
Bernstein, R., Cieslar, C., Ross, A., Oschkinat, H., Freund, J. and Holak, T.A. (1993) *J. Biomol. NMR*, **3**, 245–251.
Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.
Cieslar, C., Clore, G.M. and Gronenborn, A.M. (1988) *J. Magn. Reson.*, **80**, 119–127.
Croft, D., Kemmink, J., Neidig, K. and Oschkinat, H. (1997) *J. Biomol. NMR*, **10**, 207–219.
Eads, C.D. and Kuntz, I.D. (1989) *J. Magn. Reson.*, **82**, 467–482.
Eccles, C., Güntert, P., Billeter, M. and Wüthruch, K. (1991) *J. Biomol. NMR*, **1**, 111–130.
Friedrichs, M.S., Mueller, L. and Wittekind, M. (1994) *J. Biomol. NMR*, **4**, 703–726.
Görler, A., Gronwald, W., Neidig, K. and Kalbitzer, H.R. (1999) *J. Magn. Reson.*, **137**, 39–45.
Gronwald, W., Boyko, R.F., Sönnichsen, F.D., Wishart, D.S. and Sykes, B.D. (1997) *J. Biomol. NMR*, **10**, 165–179.
Gronwald, W., Willard, L., Jellard, T., Boyko, R.F., Rajarathnam, K., Wishart, D.S., Sönnichsen, F.D. and Sykes, B.D. (1998) *J. Biomol. NMR*, **12**, 395–405.
Grzesiek, S. and Bax, A. (1992a) *J. Am. Chem. Soc.*, **114**, 6291–6293.
Grzesiek, S. and Bax, A. (1992b) *J. Magn. Reson.*, **96**, 432–440.
Hare, B.J. and Prestegard, J.H. (1994) *J. Biomol. NMR*, **4**, 35–46.
Herrmann, T., Güntert, P. and Wüthruch, K. (2002) *J. Mol. Biol.*, **319**, 209–227.
Johnson, B.A. and Blevins, R.A. (1994) *J. Biomol. NMR*, **4**, 603–614.
Kjaer, M., Andersen, K.V. and Poulsen, F.M. (1994) *Meth. Enzymol.*, **239**, 288–307.
Kleywegt, G.J., Boelens, R., Cox, M., Llinás, M. and Kaptein, R. (1991) *J. Biomol. NMR*, **1**, 23–47.
Kraulis, P.J. (1989) *J. Magn. Reson.*, **84**, 627–633.
Kraulis, P.J. (1994) *J. Mol. Biol.*, **243**, 696–718.
Leutner, M., Gschwind, R.M., Liermann, J., Schwarz, C., Gemmecker, G. and Kessler, H. (1998) *J. Biomol. NMR*, **11**, 31–43.
Lukin, J., Gove, A.P., Talukdar, S.N. and Ho, C. (1997) *J. Biomol. NMR*, **9**, 151–166.
Meadows, R.P., Olejniczak, E.T. and Fesik, S.W. (1994) *J. Biomol. NMR*, **4**, 79–96.
Moseley, H.N.B. and Montelione, G.T. (1999) *Curr. Opin. Struct. Biol.*, **9**, 635–642.
Oezguen, N., Adamian, L., Xu, Y., Rajarathnam, K. and Braun, W. (2002) *J. Biomol. NMR*, **22**, 249–263.
Olson, J.B. and Markley, J.L. (1994) *J. Biomol. NMR*, **4**, 385–410.
Oschkinat, H. and Croft, D. (1994) *Meth. Enzymol.*, **239**, 308–318.
Oschkinat, H., Holak, T.A. and Cieslar, C. (1991) *Biopolymers*, **31**, 699–712.
Ousterhout, J.K. (1994) *Tcl and the Tk Toolkit*, Addison-Wesley Professional Computing Series., Reading, MA.
Pons, J.L. and Delsuc, M.A. (1999) *J. Biomol. NMR*, **15**, 15–26.
Seavey, B.R., Farr, E.A., Westler, W.M. and Markley, J.L. (1991) *J. Biomol. NMR*, **1**, 217–236.
Schubert, M., Oschkinat, H. and Schmieder, P. (2001a) *J. Magn. Reson.*, **148**, 61–72.
Schubert, M., Oschkinat, H. and Schmieder, P. (2001b) *J. Magn. Reson.*, **153**, 186–192.
Schubert, M., Smalla, M., Schmieder, P. and Oschkinat, H. (1999) *J. Magn. Reson.*, **141**, 31–43.
Slupsky, C.M., Desautels, M., Huebert, T., Zhao, R., Hemmingsen, S.M. and McIntosh, L.P. (2001) *J. Biol. Chem.*, **276**, 5943–5951.
Slupsky, C.M., Gentile, L.N., Donaldson, L.W., Mackereth, C.D., Seidel, J.J., Graves, B.J. and McIntosh, L.P. (1998) *Proc. Natl. Acad. Sci.*, **95**, 12129–12134.

Szyperski, T., Yeh, D.C., Sukumaran, D.K., Moseley, H.N.B. and Montelione, G.T. (2002) *Proc. Natl. Acad. Sci.*, **99**, 8009–8014.
van de Ven, F.J.M. (1990) *J. Magn. Reson.*, **86**, 633–644.
Weber, P.L., Malikayil, J.A. and Mueller, L. (1989) *J. Magn. Reson.*, **82**, 419–426.
Wittekind, M. and Mueller, L. (1993) *J. Magn. Reson.*, **B101**, 201–205.
Wüthruch, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, Inc., New York.
Xu, J. and Sanctuary, B.C. (1993) *J. Chem. Inf. Comput. Sci.*, **33**, 490–500.
Xu, J., Straus, S.K., Sanctuary, B.C. and Trimble, L. (1994) *J. Magn. Reson.*, **B103**, 53–58.
Zimmerman, D.E. and Montelione, G.T. (1995) *Curr. Opin. Struct. Biol.*, **5**, 664–673.
Zimmerman, D.E., Kulikowski, C.A., Huang, Y., Feng, W., Tashiro, M., Shimotakahara, S., Chien, C., Powers, R. and Montelione, G.T. (1997) *J. Mol. Biol.*, **269**, 592–610.
Zimmerman, D.E., Kulikowski, C.A., Wang, L., Lyons, B. and Montelione, G.T. (1994) *J. Biomol. NMR*, **4**, 241–256.