Journal of Biomolecular NMR, **16:** 139–146, 2000.
KLUWER/ESCOM

Automation of NMR measurements and data evaluation for systematically screening interactions of small molecules with target proteins

Alfred Ross*, Götz Schlotterbeck, Werner Klaus & Hans Senn
Hoffmann LaRoche AG, Department for Biostructural Research (PRPC-S), Grenzacher Strasse, CH-4070 Basel, Szwitzerland

Received 24 September 1999; Accepted 6 December 1999

Key words: flow injection, protein, screening

Abstract

In this technical note we describe the setup and application of automated sample preparation and usage of flow-through NMR equipment for the characterization of ligand binding on proteins. In addition, we focus on the perspectives of automated analysis of 2D HSQC spectra to identify changes in patterns indicative for ligand binding or changes of sample conditions. In this context we discuss a combination of statistical and non-statistical data analysis.

Abbreviations: BEST, Bruker Efficient Sample Transfer; HSQC, heteronuclear single quantum correlation; PC, principal component; PCA, principal component analysis; 2D, two-dimensional.

Introduction

The idea of using NMR to probe small molecule interactions with proteins has been of practical interest since the early 1960s (Dwek, 1973; Wüthrich, 1976; Otting, 1993). Recently, NMR binding studies have gained new momentum in the search for small ligand molecules binding to target proteins (Shuker et al., 1996; Hajduk et al., 1999; Shapiro and Wareing, 1999). This development is fostered by the systematization of research that is ongoing in many areas of drug discovery (see *'Intelligent drug design'*, Nature, **384** Suppl, 1996). Most of the early NMR binding studies focused on systems with rapid exchange of the ligand molecules between the binding site(s) of the macromolecule and the bulk of the solution containing free ligand(s). The principal manifestations of the interactions in the ligand spectrum are chemical shift changes and line broadening. In these studies, the macromolecular component may actually be present in much lower molar concentrations than the ligand molecule. The protein spectrum is therefore either not visible or can be suppressed by a relaxation filter, allowing an easy evaluation of protein binding effects in the ligand spectrum. This class of simple one-dimensional ^{1}H NMR experiments was soon complemented by investigations of the nuclear Overhauser enhancement resulting from the interactions with the protein (Balaram et al., 1972; Clore and Gronenborn, 1982; Meyer et al., 1997). Recently, the repertoire of homonuclear 1D techniques for the direct measurement of molecular interactions has further been expanded by diffusion edited and NOE-pumping techniques (Hajduk et al., 1997; Lin et al., 1997; Chen and Shapiro, 1998).

In many cases, the conditions to study intermolecular interactions with proteins can greatly be improved by selective isotope labeling of individual components (Wüthrich, 1976). An embodiment of this early idea is represented in a second class of experiments, which relies on ^{13}C and/or ^{15}N labeled protein. Here the ligand–protein interaction is reflected on resonances of the heteronuclear edited protein spectrum. It has been shown that chemical shift perturbation of ^{1}H, ^{13}C and/or ^{15}N resonances is one of the most sensitive methods to monitor specific protein–ligand complex

*To whom correspondence should be addressed. E-mail: alfred.ross@roche.com

formation over a broad affinity range (mM < K_d < nM) (Kallen et al., 1991; Wang et al., 1992; Yu et al., 1992; Chen et al., 1993; Rizo et al, 1994; Feher et al., 1996). In addition, the binding site of the ligand can be localized on the protein surface if the resonances of the protein are assigned. These studies can be complemented by the detection of intermolecular NOEs, which allow a detailed structural analysis of the ligand interaction in the binding cavity of the protein (Otting et al., 1991; Liepinsh and Otting, 1994, 1997; Dötsch et al., 1995).

The mapping of ligand binding surfaces – other than those where substrate and effector molecules bind – has previously been described for crystalline proteins (Fitzpatrick et al., 1993; Mattos and Ringe, 1996). On the basis of such experiments a novel experimental approach to drug design was proposed (Fitzpatrick et al., 1993; Allen et al., 1996). Thereby, the localized binding of different ligand molecules on the protein surface is used to guide the development of a specific lead compound by combining the various functional groups of the ligand molecules. A similar but more systematic approach has recently also been described by Shuker et al. (1996).

The performance of the NMR experiment is a crucial practical issue if the method has to be applied for screening ligand binding with a larger chemical library. Two conditions have to be fulfilled for the NMR experiment to be efficiently applicable for ligand screening: (i) the measurement time for NMR data acquisition must be short (Ross et al., 1997); and (ii) all the steps from sample preparation to data acquisition and evaluation must be automated. In this note we embark on condition (ii) and describe the automation of all the necessary steps, which involve 'just-in-time' sample preparation, transfer of the freshly prepared sample to the magnet, acquisition of the NMR data, back-transfer of the sample, and finally, the analysis of the data. In the selected example, changes in the ^{1}H-^{15}N correlated HSQC spectra (Bax et al., 1990) of the protein after addition of single compounds from a larger library are collectively analyzed to identify specific ligand binding to one or more sites on the target protein.

Our hardware setup, consisting of a Gilson 215 liquid handler (Middleton, WI, U.S.A.) running under XTRAY software (Abimed, Langenfeld, Germany) interfaced to the NMR spectrometer via a capillary to a flow-through probe head (Bruker BEST-system), is schematically shown in Figure 1. A similar setup has recently been described in the context of body fluid NMR with urine samples (Spraul et al., 1997). The automated data analysis of the ^{1}H-^{15}N correlated protein HSQC spectra was done in two steps: First a non-statistical data analysis using the AMIX software (version 2.1.3, Bruker) was used to identify ligands causing spectral changes in the protein spectrum. Second, this information was used in a statistical (Überla, 1968; Henrion and Henrion, 1995) approach to distinguish ligands binding specifically to a binding site from those that show non-specific interaction with the protein or change sample conditions (pH).

Experimental

Compounds

The compounds of the library in use were selected from the ACD database using a molecular weight cutoff filter of 150 amu to monitor a high variety of structural motifs. All ligands were dissolved in DMSO-d_6 at a concentration of 200 mM and stored in amounts of 1 ml in 2 ml 96 deep well plates (referred to as mother-plate). For this purpose the Gilson system was equipped with two standard racks for two well plates each. To prevent aggregation of the ligands in the protein sample, the whole library consisting of 500 ligands was tested for solubility using optical methods (Kansy et al., 1998) at the pH of interest. Based on these results 300 compounds were pre-selected for the NMR based screening. As target protein we used ^{15}N labeled peptide deformylase (Adams, 1968). The protein stock solution with a concentration of 250 μM was stored in a 100 ml bottle fitting in a standard Gilson bottle-rack. The pH of the sample was buffered to 7.5 using 50 mM of sodium phosphate. To prevent aggregation of protein the same buffer was used as system and rinsing liquid for the liquid handling system.

Each compound was defined by its position in the well plates and an identification number. For bookkeeping reasons the data sets taken (see below) were identified by this number.

Sample preparation

The automated sample preparation was integrated as an ICONNMR triggered sub-routine in the Bruker BEST-NMR software. A list of identifiers and well positions for samples of interest was sent automatically via RS232 interface from the NMR spectrometer to the robot at the start of the whole protocol. By this

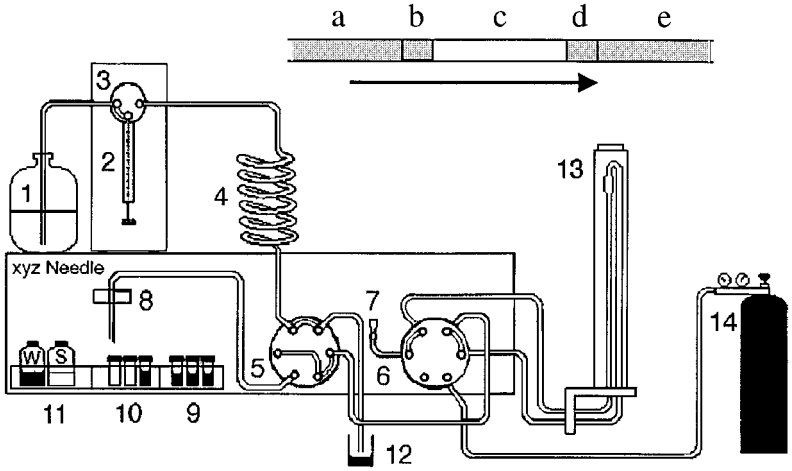

Figure 1. Schematic drawing of the hardware setup: (1) system liquid (protein buffer solution); (2) 5 ml sample dilutor syringe; (3) three-way valve; (4) sample loop; (5) six-way valve for sample loading; (6) six-way valve for sample injection; (7) injection port; (8) 1.5 mm needle; (9) 1 ml deep well plate for sample mixing and deposition (daughter plate); (10) 2 ml 96 deep well plate loaded with 200 mM stock solutions of ligands in DMSO-d_6 (mother plate); (11) W: rinsing liquid (protein buffer solution), S: 0.25 mM protein stock solution; (12) external waste reservoir; (13) ^{15}N self shielded z-gradient NMR flow-through probe with external lock cell; (14) electro-mechanically switched 1.5 bar nitrogen gas to support sample retraction. The insert shows the composition of each sample sandwich: (a) pushing liquid (buffer) – volume optimized for transfer to NMR detection cell; (b) 50 μl heading air gap; (c) 360 μl sample; (d) 50 μl trailing air gap; (e) 800 μl cleaning liquid (buffer).

procedure the samples were selectively picked, independent of their position on the plate. The sample preparation with the Gilson robot started immediately before each NMR measurement to provide freshly mixed samples. The following steps were thereby involved: (a) cleaning of the needle with 1500 μl of rinsing liquid; (b) aspiration of 380 μl of protein stock solution; (c) deposition of this sample in a second 1 ml deep well plate (referred to as daughter-plate) – for simple identification of the sample the protein was put in the same well-position as that of the ligand to be added on the second plate; (d) needle-rinsing step as above; (e) aspiration of 20 μl ligand from the mother-plate; (f) deposition of the ligand aliquot in the corresponding well of the daughter-plate; (g) mixing step: 200 μl of sample in the well was aspired and blown out of the needle two times with the highest flow rate possible (ca. 20 ml/min) – in each well prepared the protein and ligand concentrations were 10 mM and 250 μM respectively, resulting in a ligand to protein concentration ratio of 40:1; (h) transfer step: the freshly mixed sample was taken out of the well to the lowest rest volume possible and injected automatically into the transfer line. The sample was separated from the pushing and heading liquid by an air bubble to prevent dilution. The composition of the resulting 'sandwiched sample' to be transferred via a 500 μm capillary to the spectrometer is shown in the insert of Figure 1. The whole sandwich was pumped to the spectrometer cell with a pump rate of 1 ml/min. The transfer time necessary for pumping the sample from the sample well to the NMR probe was determined before the experiment as follows: the inlet capillaries of a 600 MHz flow-through probe have an approximate dimension of 550 × 0.5 mm. To estimate the correct value for sample transfer the transfer capillary was connected with a junction to a capillary of the same i.d. When the trailing air bubble after the sample was pushed about 500 mm after the junction the sample would thus be positioned correctly in the NMR probe.

After stopping the pump, the NMR data were recorded as described below. The sample-sandwich was then back-transferred to the valve by retraction of the adjusted amount of system liquid supported by a nitrogen pressure of 1.5 bar. This gas-flow was electro-mechanically switched to the outlet of the probe head. Finally, the sample was recovered by the robot to the corresponding well of the deep-well daughter-plate. Now the next sampling cycle was started.

NMR

All spectra were taken on a Bruker DMX 600 MHz spectrometer equipped with a self-shielded z-gradient ^{1}H,^{15}N flow-through probe with a detection volume of 250 μl. This probe was equipped with an internal cell for the lock solvent. As lock solvent we used D_2O.

All data were taken at 300 K. The spectrometer was operating under Bruker XWINNMR version 2.1 mastered by the BEST module of the ICONNMR version 2.0. All spectra were taken by use of the digital filtering and quadrature detection mode as implemented on the spectrometer. Pulses were determined as follows: 8 μs proton 90° pulse, 50 μs ^{15}N pulse, 200 μs ^{15}N pulse for GARP decoupling (Shaka et al., 1985). The sweep width in the proton dimension was set to 16 ppm; that in the ^{15}N indirect dimension was set to 32.1 ppm. The proton carrier was positioned on the water resonance referenced to 4.7 ppm. The nitrogen carrier was set using the internal spectrometer referencing to 118.0 ppm. The ICONNMR module allows, prior to starting the run, to set up a list of wells selected out of two 96 mother-plates each to be handled according to the procedure given above. The NMR protocol consists of: (a) automated locking on D_2O; (b) 1D 2-step z-axis gradient shimming (van Zijl et al., 1994); (c) 1 scan 1D NMR spectrum followed by automated peak picking to determine the water frequency; (d) 1D presaturation spectrum taken with 32 scans at a repetition rate of 1 s acquiring 16K complex data points. Missing resonances in these spectra can easily be used to identify insoluble or aggregating ligands. In this step the inclusion of 1D NMR techniques designed to identify ligand binding (Hajduk et al., 1997; Lin et al., 1997; Meyer et al., 1997; Chen and Shapiro, 1998) is straightforward. (e) 2D ^{1}H-^{15}N correlated HSQC spectrum using the flip-back technique (Grzesiek and Bax, 1993) together with WATERGATE (Sklenar et al., 1993) solvent suppression. For each time increment 32 scans were added up, allowing for a relaxation of 1 s. Sixty-four complex data points were taken along t_1 using the States-TPPI phase-incrementation scheme (Bax et al., 1990). The total acquisition time for each of these 2D data sets was about 1 h. Together with the necessary sample preparation/transportation and all other spectrometer action the repetition rate for each sample was about 70 min.

Prior to Fourier transformation a polynomial baseline correction was applied on the FIDs to reduce the residual water signal (Marion et al., 1989). All 2D correlation spectra were processed using linear prediction (Led and Gesmar, 1991), doubling the number of time domain data points along t_1. Finally the data were zero-filled to a final matrix of 512 × 2K points. In both dimensions we applied π/3 shifted squared cosine windows. A polynomial baseline correction of order 5 was applied along f_2. The processing of all spectra was performed without user interaction in a fully automated setup.

Data analysis

In a first step all spectra were compared to a spectrum of a reference sample prepared by addition of a DMSO-d_6 volume equivalent to that in the ligand samples. This was done fully automated using the spectral comparison tool of the Bruker AMIX software package. In this package the similarity between two spectra is calculated as the normalized scalar product between the results of 2D bucketing calculations (using buckets of decreasing size) performed on the reference and the test spectrum. The calculation of all correlation values is performed without user interaction. The program provides a table ranked according to the calculated correlation coefficients. For further analysis of this result for different underlying patterns of spectral changes, in a second step we performed a statistical analysis (Edlund and Grahn, 1991; Gartland et al., 1991) of all spectra. Using the software AMIX, an automated and not normalized peak picking was performed on the reference spectrum. The result of this procedure was used to define an integration pattern with N integration regions applied to all 300 ligand spectra. The resulting data matrix of size N × 300 was now further analyzed using the software package UNSCRAMBLER version 6.0. In this program, like in most commercially available statistical software packages, an algorithm for principal component analysis (Henrion et al., 1988) is implemented. This method relies on the determination of those axes in a space of dimensionality N (number of integral regions) where the variances of the data (integral values) are maximized. Each data set (integral values for one sample) is projected on these axes. The projected values are named principal components (PC) of the corresponding sample (ligand spectrum). Outliers compared to the 'most probable spectrum' manifest themselves by large PCs along one or more axes. By visual inspection of 2D plots of PC^i versus PC^j one can find combinations showing a clear separation of spectra from the 'average' as determined in the first step of the procedure. The observation of clustering data points in these representations helps to identify different underlying shift patterns. For the 6 PCs which explain 98% of the total variance of the data calculated here, we used the 'center' mode of the algorithm. Due to the large size of the data matrix the application of cross validation testing did not significantly influence the result.

Results and discussion

Experimental

Using the hardware setup described above we measured one reference and 306 single compound samples. We decided for this approach rather than for mixtures, as the interaction of compounds in mixtures has to be studied thoroughly. Otherwise false positive/negative results with respect to protein binding might be obtained. Retrospective analysis of our results showed that about 10% of all tested substances showed an effect on the spectra. For this incidence rate the optimal number of compounds in mixtures would be around 3. We prefer the additional experimental time rather than the time necessary to characterize the physico-chemical properties of all mixtures in use. In addition, we circumvent the task to deconvolute mixtures. A further advantage of not using mixtures is the straightforward use of the 1D proton spectra to select insoluble compounds. We want to emphasize that the hardware setup described here and all of the data analysis shown below is readily applicable to mixtures of substances, without modification.

Due to the failure of the automated lock procedure we were finally left with 306 1D proton and the same number of 2D ^{1}H-^{15}N correlated spectra. As expected, the inspection of the 1D spectra showed that all compounds were soluble at our sample conditions. After the analysis of the 2D data (see below) it became clear that our buffer concentration (50 mM of sodium phosphate) was too low to exclude pH changes of the sample due to basic or acidic groups of substances in our library. To prevent this we recommend to use a buffer concentration above 100 mM for future applications. In addition it would be helpful to characterize the compounds of the library in terms of the pKa's of all titrateable groups. As the manual check of the pH values of hundreds of NMR samples is a very time-consuming task, the addition of a pH sensitive reference substance (e.g. imidazole) would be helpful. The acquired 1D data would serve as NMR based pH detection. An automated measurement of pH in a flow-through detection cell included in the capillary line of our setup would be ideal. To our knowledge, such a system is currently not commercially available. Helpful in the interpretation of the result of the data analysis described below is the inclusion of 2D spectra taken at slightly modified pH (if possible for the protein given) in the data matrix.

As an additional result of the data analysis we found 5 spectra showing typical incidence for protein aggregation. Fortunately, the aggregation in these samples did not lead to a clot occluding the capillary line. Visual inspection of the recovered samples only showed some 'milkyness' of the solvent. To prevent aspiration of seriously aggregated samples to the capillary system, an automated check in the well plates would be needed. Again, to our knowledge, no device to be easily included in the setup described here is currently commercially available.

Data analysis

The result of the first step of the data analysis using AMIX (see Experimental section) is shown in Figure 2. The difficulty to define a clear cutoff to separate interacting from non-interacting compounds is evident. Visual inspection of the five data sets described by the lowest correlation showed incidence for protein aggregation. These sets were excluded from further analysis. Following the recipe described above, we easily identified the combination of PC^2 versus PC^4 as the one which showed the best separation. The result of the PCA is given in Figure 3. It is clear that there are two underlying classes of spectral changes in the data: Class 'A' shows large and small deviations along PC^2 and PC^4, respectively. In Class 'B' the behavior is reversed. To further characterize the spectral changes associated with both classes we superimposed the two most outlying spectra on the reference. The result is shown in Figure 4. The subsets of peaks shifting for ligands of class 'A' is clearly disjunct from those disappearing for class 'B'. For completeness it should be stated that neither PC^1 nor PC^3 showed any clustering which could help to identify distinct spectral changes (data not shown). The origin of the variance associated with these PCs remains unclear. Nevertheless, it is clearly evident that PCA is well suited to separate different binding sites occupied by different structural motifs found in the library. As will be shown, PCA analysis can further help to distinguish real binding effects from experimental artifacts and changes in the sample conditions (pH, aggregation).

To get insight in the physico-chemical origin causing the change of spectral patterns of both classes, the corresponding ligand structures were checked and compared. As most of class 'B' compounds contain an aliphatic amine moiety as functional group, it was suspected that the spectral changes were caused by a small change of sample pH due to the basic properties of these compounds. This was proven by taking two ^{1}H-^{15}N correlated spectra of protein samples with pH values varied by ± 0.5 units. Inclusion of these spectra

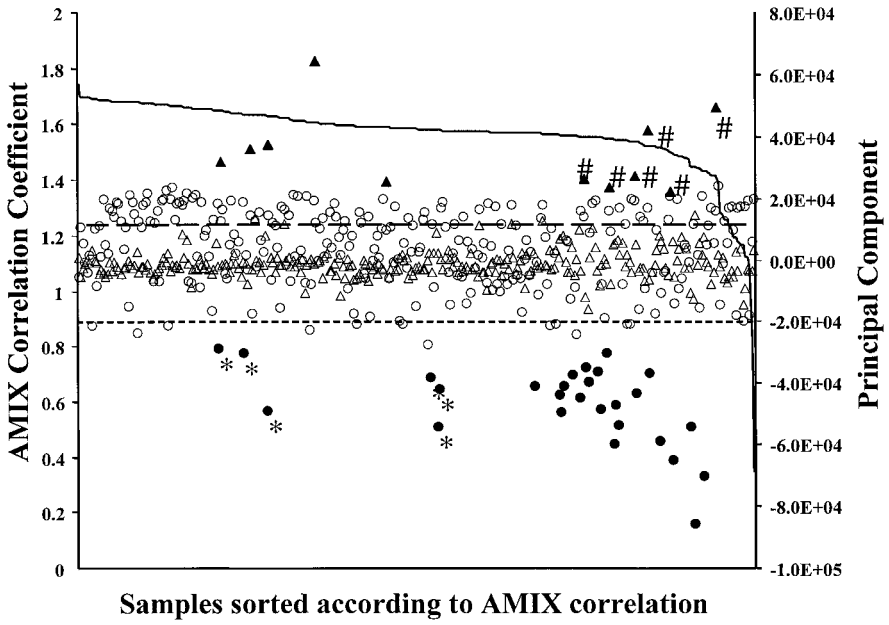

Figure 2. The result of the non-statistical analysis (AMIX). Solid line: sorted correlation coefficient for each sample compared to the reference spectrum. The samples with the lowest coefficients correspond to aggregated protein. For comparison, the values for PC^2 and PC^4 are included as circles and triangles, respectively. Samples regarded as outliers from the main bulk of the PC, as shown in Figure 3, are shown as filled symbols. The lines with the short and long dashes correspond to the cutoff for PC^2 ($-2.0e{-}4$) and PC^4 ($+1.0e4$), respectively. Samples showing a low AMIX correlation that do not contain the common structural motif of binding substances are labeled with '#'. Samples that show a high AMIX correlation containing the binding structural motif are labeled with asterisks.

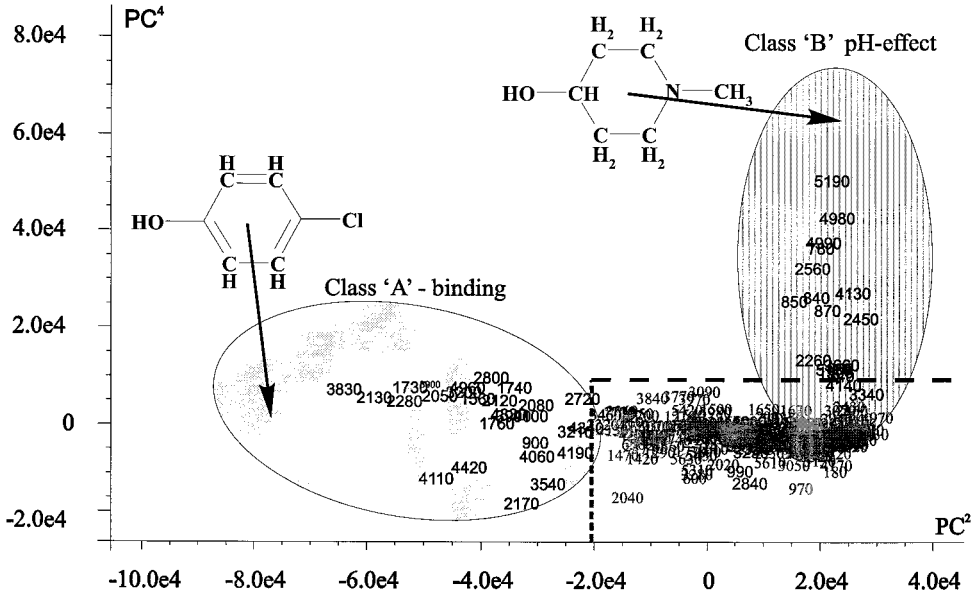

Figure 3. PCA statistical data analysis (UNSCRAMBLER): Shown is the plot of PC^2 versus PC^4. This was selected based on the subset of spectra showing 'lower' correlation in the non-statistical step. Class 'A' ligands are shown to bind in the intended binding pocket, class 'B' ligands inducing a slight pH shift of the sample. With 'A' respectively 'B' we marked data sets whose spectra are shown in Figure 4. The meaning of the dashed lines is the same as given in the legend of Figure 3.

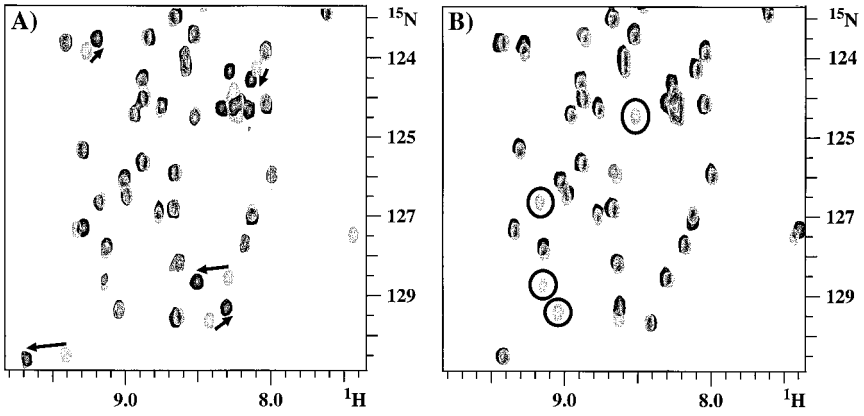

Figure 4. Expansion of ^{1}H-^{15}N correlated spectra. In gray is shown the reference spectrum taken with addition of only DMSO-d_6. In black in (A) binding ligand – arrows indicate shifting signals; in (B) acidic ligand – circles indicate disappearing signals.

in the PC analysis clearly showed the data with the increased pH as members of class 'B'. The data sets with lowered pH did not show up as outliers in the PCA. By doing so the spectral changes due to slight pH changes were included in the data matrix.

The pH-modified samples also clearly indicated that the spectral changes associated with class 'A' are not pH induced but are caused by direct ligand–protein interaction. Inspection of the ligand structures associated with class 'A' allowed identification of a common structural motif responsible for the specific interaction with the protein. Inspection of Figure 2 shows that the data points labeled with asterisks would not have been found by performing only a one-step AMIX data analysis. The same holds true for the pH induced outliers labeled with '#' in Figure 2. Therefore the identification of the common structural motif of class 'A' (which is also found for the substances behind the '$*$'-labeled data points of Figure 2) would have been aggravated. This shows that the similarity-measure as determined using AMIX is of more 'integral' character compared to the information contained in a PCA analysis. It has to be stressed that the AMIX step of the data analysis is nevertheless very helpful in determining the pair of principal components that contains the information of interest.

It has to be emphasized that for the whole analysis the assignment of the protein signals was not necessary. As in our case the assignment of the target-protein spectrum was available, we could show that protein resonances influenced by class 'B' members were mostly Glu or Asp residues. In contrast, class 'A' ligands change the resonances of amino acids in spatial proximity to the active site of the protein.

Finally we determined the binding constants of selected class 'A' ligands in order to achieve a ranking with respect to their affinity to the protein and to provide structural starting points (needles) to medical chemists for further chemical optimization. This was done using ligand titrations in NMR and analytical ultra centrifugation (Eason, 1989).

Conclusions

In summary, we have presented a roboterized sample preparation and handling setup which is, together with the fully automated data acquisition and analysis tool, suited for any application necessitating the preparation of fresh protein compound mixtures, followed by acquisition and evaluation of data to characterize a large number of ligand–protein interactions. Further improvement regarding the sensitivity of probeheads will provide the basis to require less protein or reduce the data acquisition time (Hajduk et al., 1999). The scheme outlined above is easily adapted to any NMR measuring technique. Limiting for this setup is that protein aggregation in the samples has to be safely excluded, as otherwise irreversible occlusion of the flow-through probe might occur. To improve on this problem we currently have a roboterized system under construction being able to perform all above defined tasks by filling freshly prepared samples in individual disposeable NMR sample tubes and transferring those automatically from the lab bench to the magnet. We believe that both setups will be regarded as complementary in the future. The flow-through approach described here has the advantage of speed,

convenience and elegance, whereas the scope of applicability is broader for the single-tube setup under development. This is given by two considerations: (i) In our application the protein was extremely well behaved – for other systems under investigation this might not be the case. For such target proteins irreversible occlusion of the system might occur. (ii) The necessity to perform a solubility pre-screening narrows the range of structural motifs that can be studied safely. The pre-screening excludes all substances not soluble in water, but 'soluble in the protein'. These substances would be interesting ligands per se.

The data analysis described is equally applicable for all hardware setups.

Acknowledgements

We have to thank M. Heller for carefully reading the manuscript and M. Tesche (Abimed GmbH, Germany) for providing us with a developmental version of the XTRAY software.

References

Adams, J.M. (1968) *J. Mol. Biol*., **33**, 571–589.

Allen, K.N., Bellamacina, C.R., Ding, X., Jeffery, C.J., Mattos, C., Petsko, G.A. and Ringe, D. (1996) *J. Phys. Chem.*, **100**, 2605–2611.

Balaram, P., Bothner-By, A.A. and Breslow, E. (1972) *J. Am. Chem. Soc.*, **94**, 4017–4018.

Bax, A., Ikura, M., Kay, L.E., Torchia, D. and Tschudin, R. (1990) *J. Magn. Reson.*, **86**, 304–318.

Chen, A. and Shapiro, M.J. (1998) *J. Am. Chem. Soc.*, **120**, 10258–10259.

Chen, Y., Reizer, J., Saier, H., Fairbrother, W.J. and Wright, P. (1993) *Biochemistry*, **32**, 32–37.

Clore, G.M. and Gronenborn, A.M. (1982) *J. Magn. Reson.*, **48**, 402–417.

Dötsch, V., Wider, G., Siegal, G. and Wüthrich, K. (1995) *FEBS Lett.*, **366**, 6–10.

Dwek, R.A. (1973) *Nuclear Magnetic Resonance in Biochemistry. Application to Enzyme Systems*, Clarendon Press, Oxford.

Eason, R. (1989) In *Centrifugation, a Practical Approach* (Ed., D. Rickwood), IRL Press, New York, NY.

Edlund, U. and Grahn, H. (1991) *J. Pharm. Biomed. Anal.*, **9**, 655–658.

Feher, V.A., Baldwin, E.P. and Dahlquist, F.W. (1996) *Nat. Struct. Biol.*, **3**, 516–521.

Fitzpatrick, P.A., Steinmetz, A.C., Ringe, D. and Klibanov, A.M. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 8653–8657.

Gartland, K.P., Beddell, C.R., Lindon, J.C. and Nicholson, J.K. (1991) *Mol. Pharmacol.*, **39**, 629–642.

Grzesiek, S. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 12593–12594.

Hajduk, P.J., Gerfin, T., Boehlen, J.M., Häberli, M., Marek, D. and Fesik, S.W. (1999) *J. Med. Chem.*, **42**, 2315–2317.

Hajduk, P.J., Olejniczak, E.T. and Fesik, S.W. (1997) *J. Am. Chem. Soc.*, **119**, 12257–12261.

Henrion, G., Henrion, A. and Henrion, R. (1988) *Beispiele zur Datenanalyse mit BASIC Programmen*, Deutscher Verlag der Wissenschaften, Berlin.

Henrion, R. and Henrion, G. (1995) *Multivariate Datenanalyse*, Springer, Berlin.

Kallen, J., Spitzfaden, C., Zurini, M.G., Wider, G., Widmer, H., Wüthrich, K. and Walkinshaw, M.D. (1991) *Nature*, **353**, 276–279.

Kansy, M., Kratzat, K., Parrilla, P., Senner, F. and Wagner, B. (1998) In *Proceedings of the 12th European Symposium on Quantitative Structure-Activity Relationships*, Plenum Press, Copenhagen, in press.

Led, J.J. and Gesmar, H. (1991) *J. Biomol. NMR*, **1**, 237–246.

Liepinsh, E. and Otting, G. (1994) *J. Am. Chem. Soc.*, **116**, 9670–9674.

Liepinsh, E. and Otting, G. (1997) *Nat. Biotechnol.*, **15**, 264–268.

Lin, M., Shapiro, M.J. and Wareing, J.R. (1997) *J. Am. Chem. Soc.*, **119**, 5249–5250.

Marion, D., Ikura, M., Tschudin, R. and Bax, A. (1989) *J. Magn. Reson.*, **85**, 393–399.

Mattos, C. and Ringe, D. (1996) *Nat. Biotechnol.*, **14**, 595–599.

Meyer, B., Thomas, W. and Peters, T. (1997) *Eur. J. Biochem.*, **246**, 705–709.

Otting, G. (1993) *Curr. Opin. Struct. Biol.*, **3**, 760–768.

Otting, G., Liepinsh, E. and Wüthrich, K. (1991) *Science*, **254**, 974–980.

Rizo, J., Liu, Z.P. and Gierasch, L.M. (1994) *J. Biomol. NMR*, **4**, 741–760.

Ross, A., Salzmann, A. and Senn, H. (1997) *J. Biomol. NMR*, **10**, 389–396.

Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.

Shapiro, M.J. and Wareing, J.R. (1999) *Curr. Opin. Drug Discov.*, **2**, 396–400.

Shuker, S.B., Hajduk, P.J., Meadows, R.P. and Fesik, S.W. (1996) *Science*, **274**, 1531–1534.

Sklenar, V., Piotto, M., Leppik, R. and Saudek, V. (1993) *J. Magn. Reson.*, **A102**, 241–245.

Spraul, M., Hofmann, M., Ackermann, M., Nicholls, A.W., Damment, S.J., Haselden, J.N., Shockor, J.P., Nicholson, J.K. and Lindon, J.C. (1997) *Anal. Commun.*, **34**, 339–341.

Überla, K. (1968) *Faktorenanalyse*, Springer, Berlin.

van Zijl, P.C.M., Sukumar, S., Johnson, M.O., Webb, P. and Hurd, R.E. (1994) *J. Magn. Reson.*, **A111**, 203–207.

Wang, J., Hinck, A.P., Loh, S.N., LeMaster, D.M. and Markley, J.L. (1992) *Biochemistry*, **31**, 921–936.

Wüthrich, K. (1976) *NMR in Biological Research: Peptides and Proteins*, Elsevier, New York, NY.

Yu, H., Rosen, M.K., Shin, T.B., Seidel-Dugan, C., Brugge, J.S. and Schreiber, S.L. (1992) *Science*, **258**, 1665–1668.