Journal of Biomolecular NMR, **24:** 143–148, 2002.
KLUWER/ESCOM

Breaking symmetry in the structure determination of (large) symmetric protein dimers

Vadim Gaponenko, Amanda S. Altieri, Jess Li & R. Andrew Byrd*
Structural Biophysics Laboratory, National Cancer Institute, P.O. Box B, Frederick, MD 21702-1201, U.S.A.

Received 1 July 2002; Accepted 5 September 2002

Key words: asymmetry, protein dimer, protein structure, pseudocontact shifts, residual dipolar coupling

Abstract

We demonstrate a novel methodology to disrupt the symmetry in the NMR spectra of homodimers. A paramagnetic probe is introduced sub-stoichiometrically to create an asymmetric system with the paramagnetic probe residing on only one monomer within the dimer. This creates sufficient magnetic anisotropy for resolution of symmetry-related overlapped resonances and, consequently, detection of pseudocontact shifts and residual dipolar couplings specific to each monomeric component. These pseudocontact shifts can be readily incorporated into existing structure refinement calculations and enable determination of monomer orientation within the dimeric protein. This methodology can be widely used for solution structure determination of symmetric dimers.

Solution structures of symmetric protein dimers are difficult to solve using conventional NMR techniques, because it is often impossible to distinguish between intra- and intermolecular NOEs. The problem is exacerbated in head-to-head dimers. Two general approaches have been proposed to circumvent this problem. The first relies on asymmetric isotope labeling and filtered NOESY experiments (Leupin et al., 1990). The labeling schemes vary from selective deuteration (Arrowsmith et al., 1990; Weiss, 1990) to asymmetric ^{13}C (Folkers et al., 1993; Starich et al., 1996) or asymmetric ^{13}C and ^{15}N labeling (Handel and Domaille, 1996). The second approach is to define a target function describing contributions from intra- and intermolecular NOEs and to computationally obtain the proper orientation of the monomeric subunits within the dimer (Nilges, 1993; O'Donoghue et al., 2000). This methodology has been successfully used for structure calculations of dimers (Clore et al., 1994; Lodi et al., 1994) and higher order symmetric oligomers (O'Donoghue et al., 2000). However, when the size of the protein dimer is such that deuteration is necessary, observation of H^N to side-chain or side-chain to side-chain intermolecular NOEs becomes problematic and H^N to H^N intermolecular NOEs are rare. To overcome this problem we introduce a method to obtain unambiguous, long-range intermolecular restraints, and we illustrate their use for symmetric dimer structure determination.

The strategy is based on introduction of a paramagnetic anisotropic susceptibility (PAS) into the structure of an otherwise symmetric protein dimer. The paramagnetic probe is incorporated by modification of a unique cysteine residue with thiol-reactive EDTA. Binding of subequimolar amounts of a paramagnetic metal creates an asymmetric system with the paramagnetic probe residing on only one monomer. Measurement of pseudocontact shifts (PCS) (La Mar et al., 1978) and residual dipolar couplings (RDC) (Arnesano et al., 2000; Barbieri et al., 2002; Biekofsky et al., 1999; Tolman et al., 1995) for each monomer within the dimer provides long-range distance and orientational restraints for proper positioning of the molecules with respect to each other. The presence of a paramagnetic reference frame provides valuable long-range structural information in metal-binding proteins (Bertini et al., 2001a,b; Gochin, 1998, 2000; Hus et al., 2000; Ubbink et al., 1998). However, introduction of paramagnetic probes into non-metal binding

*To whom correspondence should be addressed. E-mail: rabyrd@ncifcrf.gov

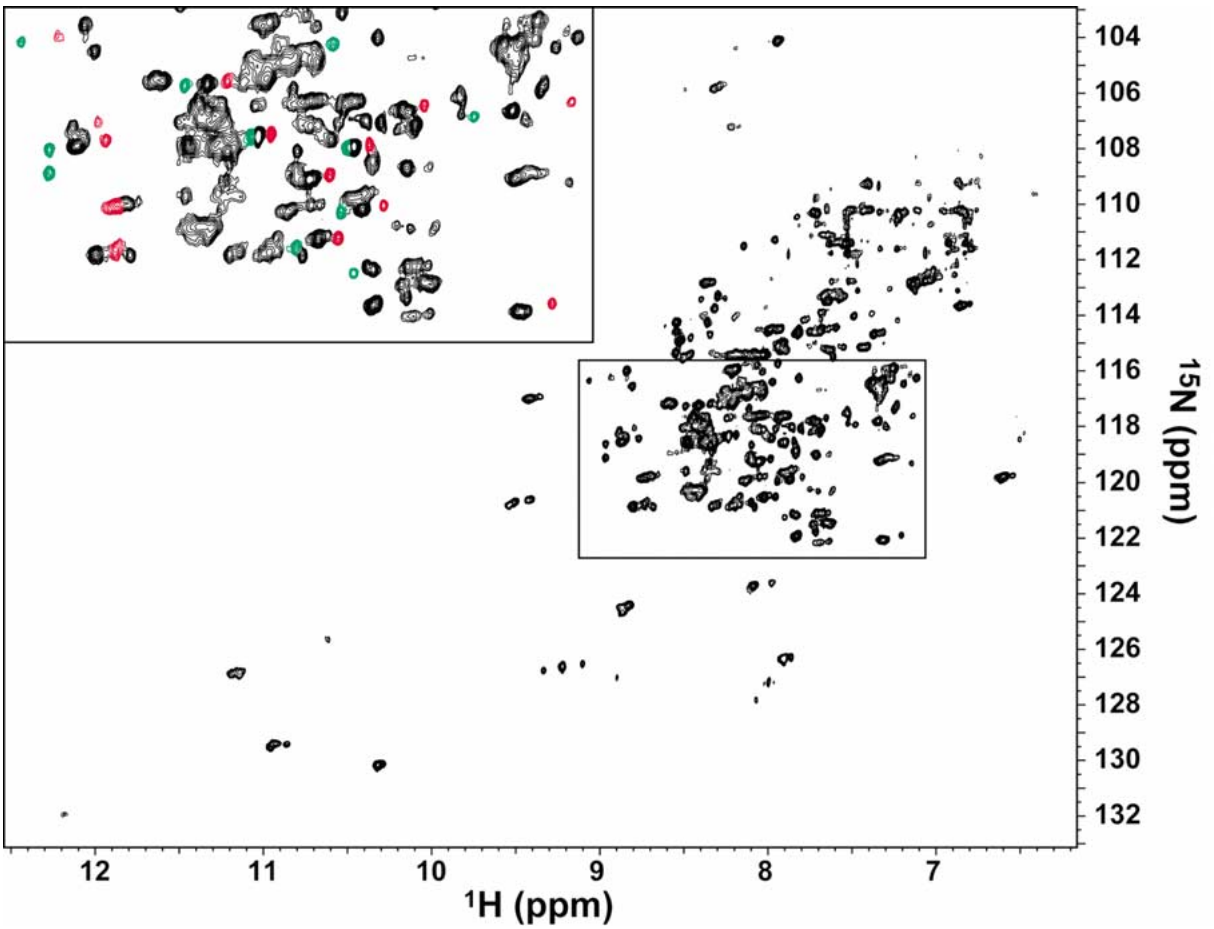

Figure 1. ^{15}N TROSY spectrum collected at 900 MHz on ^{2}H, ^{15}N, ^{13}C enriched ILV labeled sample of STAT4$_{NT}$-EDTA-Co^{2+}. Spectral widths were 16 kHz and 3.2 kHz in the direct and indirect dimensions, respectively, and the number of *t1* increments was 128. The insert in the upper left corner shows an expanded region corresponding to the boxed area in the middle of the spectrum. The peaks belonging to the paramagnetic species are colored red and green to represent the intra- and intermonomer PCSs, respectively.

proteins has not yet been widely used (Donaldson et al., 2001; Feeney et al., 2001; Gaponenko et al., 2000). We show for the first time that it is possible to obtain long-range structural information specific to each monomeric component of a symmetric dimer using paramagnetic probes.

This approach is demonstrated on the N-terminal domain of STAT4 (STAT4$_{NT}$), which forms a 28kDa symmetric dimer (Baden et al., 1998). Backbone ($^{\alpha}$C, $^{\beta}$C, C′, N, H^{N}) and ILV-methyl resonance assignments for STAT4$_{NT}$ have been made previously (S. P. Sarma and R. A. Byrd, unpublished results; Zwahlen et al., 1998). All NMR experiments were performed on Varian INOVA 600, 800 and 900 MHz spectrometers. Data were processed and analyzed using NMRPipe (Delaglio et al., 1995).

Site-directed paramagnetic labeling was achieved by modification of the cysteine residue 107 with S-(2-pyridylthio)-cysteaminyl-EDTA (Toronto Research Chemicals). After an extensive reduction with DTT, the ^{15}N/^{13}C/^{2}H-labeled protein was transferred into 50 mM Tris (pH 8.0), 150 mM NaCl, and 1 mM EDTA using a P10 buffer exchange column (Bio-Rad). S-(2-pyridylthio)-cysteaminyl-EDTA was added to 100 μM protein solution to a final concentration of 1 mM. The conjugation reaction proceeded for three days in the dark at room temperature (Ebright et al., 1992). The reaction products and excess thiol-reactive EDTA were removed by buffer exchange in an Amicon concentrator (Millipore) into 20 mM sodium acetate (pH 5.25), 25 mM NaCl, 90% H_2O/10% D_2O. The completeness of the conjugation reaction was de-

termined using electrospray mass-spectrometry. The protein was concentrated to 0.8 mM in monomer.

To assess possible structural perturbations caused by thiol-reactive EDTA, ^{1}H-^{15}N correlation spectra were collected for native $STAT4_{NT}$ and for the EDTA-modified protein. Few chemical shift differences were observed between the native and modified proteins (data not shown). The largest differences were of the order of 0.1 ppm for H^N in the region surrounding the C107 site. These results are consistent with very minor local structural perturbations introduced by the conjugation reaction.

We have utilized cobalt, bound to the ligated EDTA, as a paramagnetic probe to introduce sufficient magnetic anisotropy allowing observation of resonances belonging to individual monomeric components of the $STAT4_{NT}$ dimer. Co^{2+} was used, instead of other metal ions or lanthanides (Arnesano et al., 2000), in order to minimize broadening and maximize the number of observable resonances. Sub-stoichiometric binding was accomplished by addition of 300 μM $CoCl_2$ to the 0.8 mM (monomer) protein sample. The asymmetry is readily apparent in the TROSY-HSQC spectrum (Figure 1), where the number of signals increases to reveal three species in solution: the Co^{2+}-free, diamagnetic species and two non-equivalent monomeric species. The asymmetry arises due to non-equivalent PCSs to residues within the monomer bearing the paramagnetic Co^{2+} (referred to as the intramonomer) compared to PCSs to the same residue across the interface in the other monomer (referred to as the intermonomer). The sample conditions should generate a ratio of 1:0.6:0.6 for the diamagnetic:intramonomer:intermonomer resonances, in the absence of relaxation effects. Due to paramagnetic relaxation enhancement, the intensities of the PCS resonances are reduced to varying degrees, with typical ratios of 1:0.3:0.3. It should be noted that full saturation with Co^{2+} would result in a single, symmetric species, where the observed PCS at any site would be the sum of effects from two centers (Biekofsky et al., 1999) with a different PAS, thus rendering the shifts less useful for structural restraints, compared to the asymmetric species. Resonances from the fully saturated species were very weak in the spectra in Figure 1 due to relaxation effects. A three-dimensional HNCO experiment was also performed at 900 MHz. The PCSs were measured as differences between the diamagnetic chemical shifts and those representing the paramagnetic species, and chemical shift differences as small as 0.02 ppm could be reliably measured. A total of 78 PCSs were measured for H^N protons in the dimer.

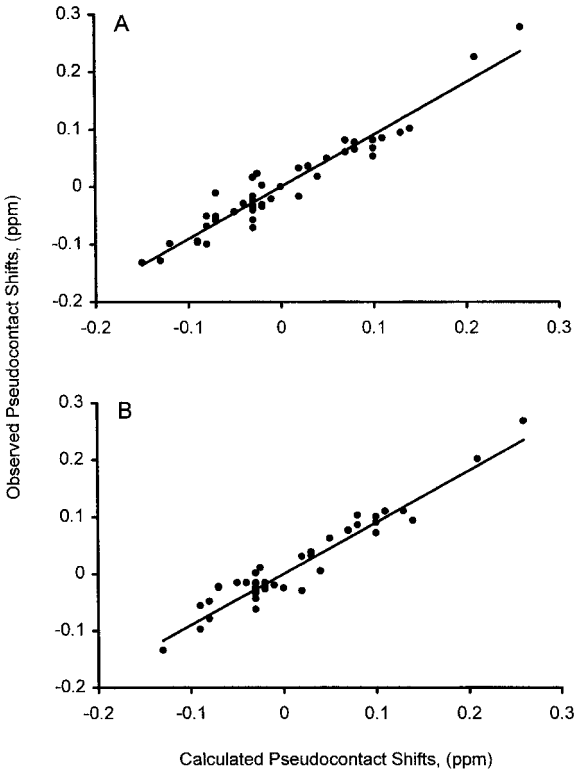

Figure 2. Correlation between observed and calculated intramonomer PCSs for H^N protons in $STAT4_{NT}$-EDTA-Co^{2+}. The calculated PCS values were obtained in a least squares minimization procedure using (A) the NOE-derived structure and (B) the crystal structure. The linear correlation coefficients are 0.93 and 0.92, respectively.

The resolution afforded by the 3D HNCO spectrum permits identification of previously assigned diamagnetic species and assignment of the associated paramagnetically shifted resonances to specific residues. To assign each shifted resonance to a specific monomer, we utilize the fact that the distances from the unpaired electron to H^N, N, and C′ are long and, thus, the difference between the electron-H^N_i and electron-CO_{i-1} vectors is very small. Therefore, to a first approximation, the sign and magnitude of the PCSs for residues i and $i-1$ in the same monomer are nearly equal, which assists the grouping of PCSs to a specific monomer and enables some unambiguous assignments directly. The second step in the assignment is a least squares optimization of the PAS orientation with respect to the coordinates of the $STAT4_{NT}$ derived from either the crystal structure (Vinkemeier et al., 1998) or the NOE-based structure of $STAT4_{NT}$

(S. P. Sarma et al., unpublished results) using the unambiguous assignments from the HNCO spectrum. The initial value of the PAS was assumed to be axial with $\Delta\chi_{ax} = 1.7 \times 10^{-31}$ m^3, similar to the value determined previously for barnase modified with thiol-reactive EDTA-Co^{2+} (Gaponenko, 2000). Since the PAS depends on the nature of the paramagnetic metal probe and the crystal field, this is a reasonable assumption. The third step involves a least squares selection of those PCSs that are predicted to be intramonomer based on the derived PAS orientation. This discrimination does not require a high-resolution structure and can readily be accomplished with an approximate structure based on limited NOEs (Allegrozzi et al., 2000). Observed PCSs were significantly different for each monomer within the symmetric dimer, exhibiting the opposite sign for most of the residues; hence, the sub-stoichiometrically introduced PAS created a truly asymmetric system. Good correlation between observed and calculated intramonomer PCSs for the crystal- and NOE-derived monomer structures are shown in Figure 2. Similar correlations were found for the intermonomer PCSs, which confirms the equivalence of the two monomer structures within the dimer.

Following PCS assignment, the measured shifts were used in a least squares minimization procedure that optimized the orientation of the cobalt induced PAS tensor with respect to the atomic coordinates of the monomer from either the crystal structure or the NOE-based solution structure of $STAT4_{NT}$. The following target function was minimized with respect to the axial component of the PAS:

$$\Delta_{pc} = \sum (\Delta\delta_{pc} - \Delta\hat{\delta}_{pc})^2, \tag{1}$$

where $\Delta\delta_{pc}$ and $\Delta\hat{\delta}_{pc}$ are the observed and calculated PCS values, respectively. We utilized the following equation to calculate $\Delta\delta_{pc}$:

$$\Delta\delta_{pc} = \frac{P_{ax}}{3r^3}\left(3\cos^2\theta - 1\right) + \frac{3P_{rh}}{2}\sin^2\theta\cos(2\varphi), \tag{2}$$

where P_{ax} and P_{rh} are the axial and rhombic components of the PAS, r is the distance between the unpaired electron and the observed nucleus, and θ and ϕ are the polar angles describing the orientation of the electron-nucleus vector in the PAS tensor frame. The optimized value was found to be axially symmetric with $\Delta\chi_{ax} = 1.56 \times 10^{-31}$ m^3, which confirms the validity of the initial assumption (*vide supra*).

The PCS can be incorporated into the structure refinement protocols as orientational and distance restraints, since they follow approximately the same functional form as RDCs (e.g., Equation 2, Bax and Tjandra, 1997). The PCSs were used as unambiguous intermolecular restraints in restrained molecular dynamics and simulated annealing protocols to obtain the orientation of the two monomers in the dimer observed in solution. The resulting dimer structure satisfies the PCS restraints to a high degree (Figure 3). It is also possible to use the absence of PCS as a structural restraint, similar to the case of ^{1}H-^{1}H RDCs (Tjandra et al., 2000). A detailed analysis of the resulting dimer structure will be presented elsewhere.

In addition to the structural information obtained from the PCSs, asymmetric introduction of paramagnetic probes presents a unique opportunity for measurement of RDCs specific to each monomeric component of a symmetric dimer. Paramagnetic alignment has been used extensively for metalloproteins (Arnesano et al., 2000). The degree of alignment is proportional to the square of the applied field; hence, use of the highest applied magnetic field is quite valuable. Residual dipolar couplings for H^N-N bond vectors in the asymmetric $STAT4_{NT}$ were measured in an IPAP HSQC experiment (Ottiger et al., 1998) at 900 MHz. An example of the measured RDCs (E49) is shown in Figure 4, and a total of 46 H^N-N RDCs were measured in the dimer. In this case, the RDC alignment tensor does not lie along the symmetry axis, which is the common occurrence for symmetric dimers in orienting media (Drohat et al., 1999), and emphasizes the significance of our approach.

Observed H^N-N RDCs exhibited magnitudes up to 4.5 Hz, indicating a relatively small degree of alignment even at 21 T. This is consistent with previously reported values of RDCs in other paramagnetic systems aligned with transition metals (Tolman et al., 1995). Optimization of the alignment tensor orientation with respect to the atomic coordinates of $STAT4_{NT}$ using the observed RDCs revealed that the alignment tensor and the PAS tensor are oriented very similarly within the molecular frame, as expected. The deviation was only 9°, which is within the experimental error. The optimized alignment tensor was axially symmetric with $\Delta\chi_{ax} = 1.01 \times 10^{-31}$ m^3, which is 30% smaller than the value deduced from PCSs. A similar ratio of PAS and RDC alignment tensors derived from the same protein molecule has been reported previously (Banci et al., 1998; Bax and Tjandra, 1997; Demene et al., 2000). The dif-

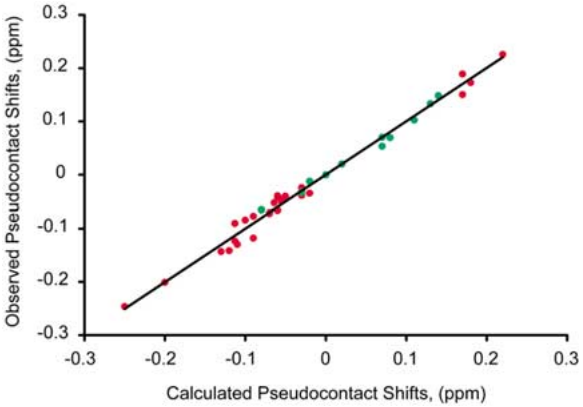

Figure 3. Correlation between observed and calculated PCSs for H^N protons in the $STAT4_{NT}$-EDTA-Co^{2+} dimer. The calculated PCS values were obtained from the dimer structure using XPLOR-NIH. Intramonomer PCSs are shown in red and intermonomer PCSs are shown in green. The linear correlation coefficient is 0.99.

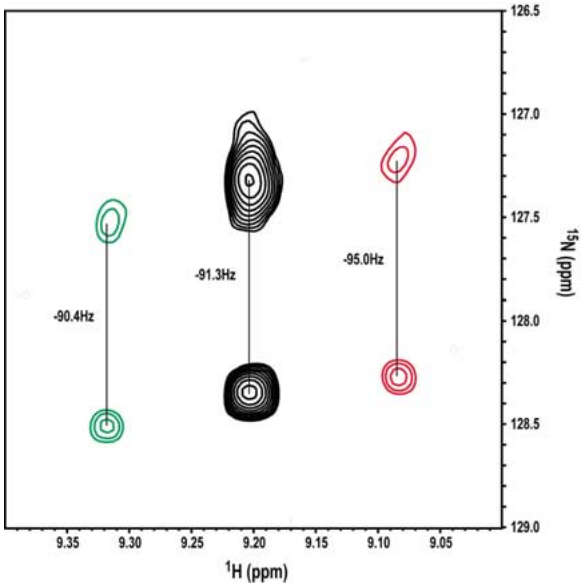

Figure 4. Example of H^N-N RDC measurements in a ^{15}N IPAP HSQC experiment at 900 MHz. The figure shows resonances for Glu 49 of $STAT4_{NT}$-EDTA-Co^{2+} representing the diamagnetic species (black) and intra- (red) and intermonomer (green) pseudocontact shifted species.

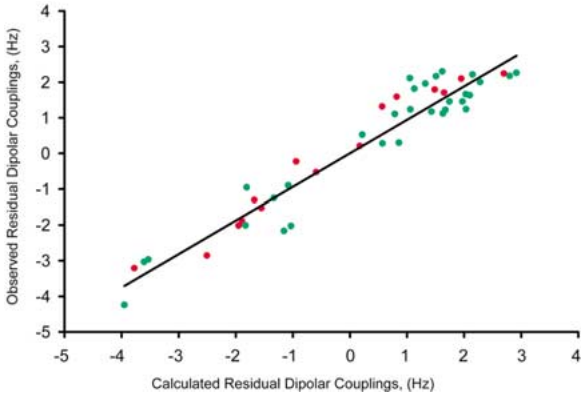

Figure 5. Correlation between observed and calculated RDCs for H^N-N bond vectors in the symmetric dimer of $STAT4_{NT}$-EDTA-Co^{2+}. The calculated RDCs were obtained using XPLOR-NIH structure refinement protocols. Intramonomer RDCs are shown in red and intermonomer RDCs are shown in green. The linear correlation coefficient is 0.93.

ference may be due to two factors: (1) It has been demonstrated that the alignment tensor determined from RDCs is often underestimated by 15–20% (Clore et al., 1998; Tjandra et al., 1997); and (2) the underestimation of the alignment tensor may also have a contribution from dynamic frequency shifts resulting from cross-correlation effects Ghose and Prestegard (1997). Nevertheless, it has been shown that such underestimates have a very minor effect on structure calculations (Clore et al., 1998; Tjandra et al., 1997).

The relative orientation of the two monomers within the dimer may be determined using RDCs alone, independent of the PCSs. The resulting orientation of the two monomers is the same as that obtained based on PCSs. Furthermore, the PCSs and RDCs may be used jointly in structure refinement. As an example, the dimer structure was calculated using both PCS and RDC restraints, and the quality of the orientation is illustrated by the correlation of observed and calculated RDCs (Figure 5), where the data for both monomers is included.

In cases where ligation of a natural cysteine results in structural perturbations or when there is no free cysteine residue, it would be possible to engineer single cysteine mutations into other sites within the protein, as is common in fluorescence spectroscopy. Introduction of additional sites for ligand attachment also provides a direct means of obtaining multiple sets of long-range structural restraints, which improves the accuracy of structures determined using this method and has been done for $STAT4_{NT}$ (unpublished results). This approach is similar to that suggested for RDCs using multiple steric alignments in different liquid crystalline environments (Ramirez and Bax, 1998) or the substitution of different paramagnets within metalloproteins (Arnesano et al., 2000).

In conclusion, we have created an asymmetric system in a naturally symmetric dimer by introducing PAS at a single site within the dimer. This al-

lowed observation of PCSs and RDCs specific to each monomeric component of the dimer. The measured PCSs and RDCs provided valuable distance and orientational restraints to properly position the monomeric subunits with respect to each other within the dimer. The proposed methodology may be generally applicable to solving symmetry problems in dimeric proteins and is of even greater significance for higher molecular weight dimers when perdeuteration is required. A limitation of this approach occurs for proteins with multiple structurally important cysteine residues. In this case, alternative approaches to attachment of the metal binding ligand might be sought.

Acknowledgements

We gratefully acknowledge Varian, Inc., Oxford Instruments, Inc., and Eriks Kupce for access to the 900 MHz spectrometer used in this study. We also acknowledge helpful discussions with Nico Tjandra regarding incorporation of PCS restraints into XPLOR-NIH calculations.

References

Allegrozzi, M., Bertini, I., Janik, M.B.L., Lee, Y.M., Liu, G. and Luchinat, C. (2000) *J. Am. Chem. Soc.*, **122**, 4154–4161.

Arnesano, F., Banci, L., Bertini, I., Van Der Wetering, K., Czisch, M. and Kaptein, R. (2000) *J. Biomol. NMR*, **17**, 295–304.

Arrowsmith, C.H., Pachter, R.B., Altman, R.B., Iyer, S.B. and Jardetzky, O. (1990) *Biochemistry*, **29**, 6332–6341.

Banci, L., Bertini, I., Huber, J.G., Luchinat, C. and Rosato, A. (1998) *J. Am. Chem. Soc.*, **104**, 12903–12909.

Barbieri, R., Bertini, I., Cavallaro, G., Lee, Y.M., Luchinat, C. and Rosato, A. (2002) *J. Am. Chem. Soc.*, **124**, 5581–5587.

Bax, A. and Tjandra, N. (1997) *Nat. Struct. Biol.*, **4**, 254–6.

Bertini, I., Janik, M.B., Lee, Y.M., Luchinat, C. and Rosato, A. (2001a) *J. Am. Chem. Soc.*, **123**, 4181–4188.

Bertini, I., Luchinat, C. and Piccioli, M. (2001b) *Meth. Enzymol.*, **339**, 314–340.

Biekofsky, R.R., Muskett, F.W., Schmidt, J.M., Martin, S.R., Browne, J.P., Bayley, P.M. and Feeney, J. (1999) *FEBS Lett.*, **460**, 519–526.

Clore, G.M., Gronenborn, A.M. and Tjandra, N. (1998) *J. Magn. Reson.*, **131**, 159–162.

Clore, G.M., Omichinski, J.G., Sakaguchi, K., Zambrano, N., Sakamoto, H., Appella, E. and Gronenborn, A.M. (1994) *Science*, **265**, 386–391.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.

Demene, H., Tsan, P., Gans, P. and Marion, D. (2000) *J. Phys. Chem.*, **104**, 2559–2569.

Donaldson, L. W., Skrynnikov, N.R., Choy, W.Y., Muhandiram, D.R., Sarkar, B., Forman-Kay, J.D. and Kay, L.E. (2001) *J. Am. Chem. Soc.*, **123**, 9843–9847.

Drohat, A.C., Tjandra, N., Baldisseri, D.M. and Weber, D.J. (1999) *Protein Sci.*, **8**, 800–809.

Ebright, Y.W., Chen, Y., Pendergrast, P.S. and Ebright, R.H. (1992) *Biochemistry*, **31**, 10664–10670.

Feeney, J., Birdsall, B., Bradbury, A., Biekofsky, R.R. and Bayley, P.M. (2001) *J. Biol. NMR*, **21**, 41–48.

Folkers, P.J.M., Folmer, R.N.H., Konings, R.N.H. and Hilbers, C.W. (1993) *J. Am. Chem. Soc.*, **115**, 3798–3799.

Gaponenko, V. (2000) *The Use of Paramagnetic Systems in Protein Global Fold Determination*, University of Cincinnati, Cincinnati.

Gaponenko, V., Dvoretsky, A., Walsby, C., Hoffman, B.M. and Rosevear, P. R. (2000) *Biochemistry*, **39**, 15217–15224.

Ghose, R. and Prestegard, J.H. (1997) *J. Magn. Reson.*, **128**, 138–143.

Gochin, M. (1998) *J. Biomol. NMR*, **12**, 243–257.

Gochin, M. (2000) *Structure Fold Des.*, **8**, 441–452.

Handel, T.M. and Domaille, P.J. (1996) *Biochemistry*, **35**, 6569–6584.

Hus, J.C., Marion, D. and Blackledge, M. (2000) *J. Mol. Biol.*, **298**, 927–936.

La Mar, G.N., Overkamp, M., Sick, H. and Gersonde, K. (1978) *Biochemistry*, **17**, 325–361.

Leupin, W., Otting, G., Amacker, H. and Wüthrich, K. (1990) *FEBS Lett.*, **263**, 313–316.

Lodi, P.J., Garrett, D.S., Kuszewski, J., Tsang, M.L., Weatherbee, J.A., Leonard, W.J., Gronenborn, A.M. and Clore, G.M. (1994) *Science*, **263**, 1762–1767.

Nilges, M. (1993) *Proteins*, **17**, 297–309.

O'Donoghue, S.I., Chang, X., Abseher, R., Nilges, M. and Led, J.J. (2000) *J. Biomol. NMR*, **16**, 93–108.

Ottiger, M., Delaglio, F. and Bax, A. (1998) *J. Magn. Reson.*, **131**, 373–378.

Ramirez, B.E. and Bax, A. (1998) *J. Am. Chem. Soc.*, **120**, 9106–9107.

Starich, M.R., Sandman, K., Reeve, J.N. and Summers, M.F. (1996) *J. Mol. Biol.*, **255**, 187–203.

Tjandra, N., Omichinski, J.G., Gronenborn, A.M., Clore, G.M. and Bax, A. (1997) *Nat. Struct. Biol.*, **4**, 732–738.

Tolman, J.R., Flanagan, J.M., Kennedy, M.A. and Prestegard, J.H. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 9279–9283.

Ubbink, M., Ejdeback, M., Karlsson, B.G. and Bendall, D.S. (1998) *Structure*, **6**, 323–335.

Vinkemeier, U., Moarefi, I., Darnell, Jr. J.E., and Kuriyan, J. (1998) *Science*, **279**, 1048–1052.

Weiss, M.A. (1990) *J. Magn. Reson.*, **86**, 626–632.

Zwahlen, C., Gardner, K.H., Sarma, S.P., Horita, D.A., Byrd, R.A. and Kay, L.E. (1998) *J. Am. Chem. Soc.*, **120**, 7617–7625.