DOI: 10.1002/anie.200902561

Protein Structures

Use of Relaxation Enhancements in a Paramagnetic Environment for the Structure Determination of Proteins Using NMR Spectroscopy**

*Tobias Madl, Wolfgang Bermel, and Klaus Zangger**

NMR spectroscopy has developed into one of the principal methods in structural biology and enables not only the determination of the three-dimensional structure but also of the dynamic aspects of biological macromolecules. Methodological developments in the last decade (e.g. transverse relaxation-optimized spectroscopy (TROSY[1]), residual dipolar couplings (RDCs[2,3]), deuteration,[4] stereoarray isotope labeling (SAIL[5]), methyl-TROSY,[6] and direct ^{13}C detection)[7–9] have expanded the size of proteins that can be analyzed by solution-state NMR spectroscopy. However, the number of structures of proteins with molecular masses above 30 kDa solved by NMR spectroscopy is still rather small. Despite the recent advent of alternative structural restraints, the current structure determination approach still relies mainly on the use of a large number of NOEs. With increasing molecular mass, the available number of interproton distances is strongly reduced owing to fractional or complete deuteration, which is carried out to sharpen the resonance lines by dilution of the proton density.[10] To enable structure determination of such systems by solution NMR spectroscopy, the NOE data have to be complemented with restraints from complementary methods, for example restraints from paramagnetic relaxation enhancements (PREs) and pseudo contact shifts (PCSs) caused by covalently attached paramagnetic tags or metals bound to proteins.[11–17] However, the preparation of samples with extrinsic paramagnetic groups could be quite laborious as it involves, for example, introduction of cysteines at several positions and chemical modification. Structural changes associated with removal of cysteine residues by mutagenesis and tag attachment cannot be excluded. Furthermore, residues located close to the paramagnetic tag are extensively broadened.

Herein we present a novel approach towards the structure determination of biological macromolecules by using the paramagnetic effects obtained from an inert and freely soluble paramagnetic agent. This type of compound does not rely on chemical or biochemical modification of the macromolecule and allows for tunable relaxation enhancement simply by variation of the concentration. For our study we used the water-soluble complex gadolinium diethylenetriaminepentaacetic acid bismethylamide, [Gd(dtpa-bma)], which is inert towards proteins and cannot penetrate their interior.[18–20] Thus the solvent surrounding the protein is made paramagnetic. Paramagnetic solvent additives such as gadolinium complexes, stable nitroxyl radicals, and dioxygen have been used to probe protein surfaces and protein–protein interaction sites.[21–23]

The overall relaxation enhancement of a specific nucleus depends on the combined effect of the entire paramagnetic environment. As we show herein, this structural information can be used in an alternative structure determination approach. It yields an adjustable immersion-depth-dependent parameter.

The interaction between an NMR-active nucleus and an inert paramagnetic molecule is described by the "second-sphere interaction model".[19] In this formalism the relaxation enhancement of a single paramagnetic probe is given by a $1/r^6$ dependence, where r is the distance between the paramagnetic center and the observed nucleus. For a planar surface and to a good approximation for large spherical systems, integration over the paramagnetic environment yields a $1/d^3$ dependency of the PRE, where d is the distance from the surface.[20] If only one paramagnetic molecule can get close to the nucleus of interest (close to a deep cleft of a protein, which allows access of only a single [Gd(dtpa-bma)] molecule), the relaxation enhancement decays with $\langle d^{-6}\rangle$. Thus, the PRE versus insertion depth dependence is between $\langle d^{-3}\rangle$ and $\langle d^{-6}\rangle$.

Owing to the a priori unknown structure and shape of the molecule, an exact mathematical description of the relaxation phenomena for a certain nucleus cannot be derived, and we decided to use a conservative model-free approach. For a given PRE the distance between the nucleus and the closest approaching paramagnetic center must be the same as or larger than the relaxation enhancement that would be caused by a single paramagnetic center at a distance $(k/\mathrm{PRE})^{-6}$. Without prior knowledge of the structure, the constant k can be obtained by looking for the maximum PREs measured for non-exchanging (carbon-bound) protons. These protons must be located on the surface. The closest possible distance

[*] Dr. T. Madl,[+] Prof. K. Zangger
Institut für Chemie, Organische und Bioorganische Chemie
Karl-Franzens-Universität Graz
Heinrichstrasse 28, 8010 Graz (Austria)
Fax: (+43) 316-380-9840
E-mail: klaus.zangger@uni-graz.at
Dr. W. Bermel
Bruker BioSpin GmbH
Silberstreifen 4, 76287 Rheinstetten (Germany)
[+] Current address : Munich Center for Integrated Protein Science and Biomolecular NMR, Department Chemie, TU München (Germany)

[**] Funding by the Austrian Science Foundation (FWF) under project number 19902 to K.Z. is gratefully acknowledged. T.M. thanks the Austrian Academy of Sciences (ÖAW) for a DOC scholarship. We would like to thank Michael Sattler and Bernd Simon, TU Munich (Germany) for helpful discussions and Jean-Marie Clement, Institut Pasteur, Paris, France for providing the MBP plasmid.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200902561.

between surface protons and the paramagnetic molecules is 5.9 Å, which is the [Gd(dtpa-bma)] radius of 3.5 Å plus two proton van der Waals radii. Furthermore, it can be safely assumed that residues that experience the highest PREs must be located on the surface of the protein. In other words, the available structural data of PREs are: 1) information about which atoms are located on the surface and thus experience the largest PRE and 2) the minimum distance of the shielded atoms to the closest paramagnetic center.

Standard NMR spectroscopy experiments can be easily adapted for PRE measurements by inclusion of a T_1 saturation- or inversion-recovery block preceding the experiment. Preferably, we use saturation recovery, as the magnetization does not need to be in equilibrium at the start of the saturation (in contrast to an inversion recovery or spin-echo experiment). Therefore, the interscan delay can be reduced significantly. Owing to these faster repetition rates, and because T_1 times are longer than T_2 and thus give a higher accuracy when measuring smaller relaxation enhancements, we used longitudinal relaxation rates. Ubiquitin (8 kDa) and the maltodextrin-binding protein (42 kDa, MBP) were chosen as model systems for the PRE approach, because their structures are well-characterized by X-ray crystallography and NMR spectroscopy.[24–27] While ubiquitin serves as a test case for rapid fold determination, MBP was selected to demonstrate that PREs are particularly advantageous for high-molecular-mass proteins.

We used HSQC (^{13}C–^{1}H and ^{15}N–^{1}H) spectra to measure proton PREs. Upon addition of [Gd(dtpa-bma)] the signals are broadened owing to increased T_2 relaxation. However, the resolution of the spectrum remains good enough to allow the extraction of a high number of PREs (Figure 1).

For the high-molecular-mass MBP, additional information on the location (surface or core) of ^{13}C spins was obtained from ^{13}C direct-detected NMR spectra.[7–9] During the first step of the structure calculation, pseudo molecules were placed around the initial structure to simulate the paramagnetic environment. The initial structure with globular shape was obtained using only NOEs between exchangeable protons (amide protons) and is typically far away from the correct structure. The pseudo-molecule cloud was then connected to the surface by defining distances between the nuclei with the highest PREs and the pseudo molecules. For the other spins with lower relaxation enhancements, the minimal distance between each nucleus and the closest approaching pseudo molecule was defined. Using this set of restraints, a simulated annealing structure calculation was carried out. The resulting structure was then searched for pseudo-molecule–spin separations that were shorter than the defined minimal distances based on the measured relaxation enhancements. For any such violations, additional distance restraints were defined between these close encounters to keep these pseudo molecules in the allowed range for all nuclei for which experimental relaxation enhancements are available. Owing to the large number of such defined minimum distances, no pseudo molecules were trapped in the protein during the calculations. Subsequently, the structure calculation was repeated including this enlarged restraint list. This iterative procedure was carried out until no further violations were observed. In principle, it would be possible to define, at the beginning of the structure calculation, distance restraints between each spin and every pseudo molecule. However, even for small proteins this approach would result in a very large number of restraints. For example, we used 380 pseudo molecules for ubiquitin. Thus for the observed 205 PREs, this method would yield 77900 distance restraints (205×380). Therefore, we start by using only one distance restraint per nucleus and introduce additional ones only if a pseudo molecule gets too close to any nucleus during the structure calculation.

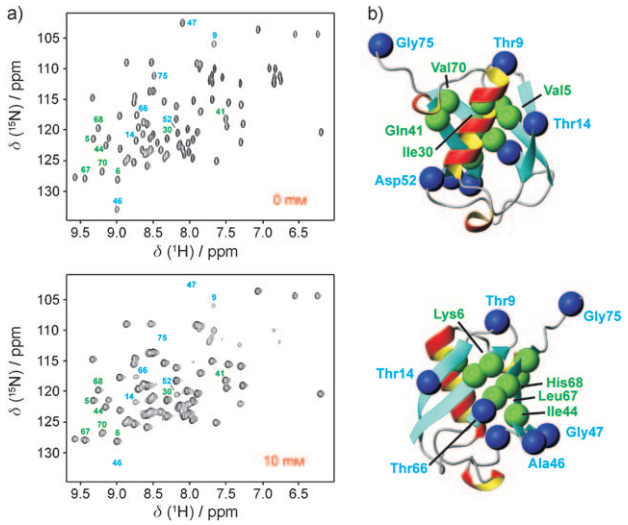

Figure 1. a) ^{15}N–^{1}H HSQC spectra of ubiquitin at concentrations of 0 and 10 mM [Gd(dtpa-bma)]. b) Amino acid NH protons with high PREs are colored blue, residues for which a shielding from the water-soluble paramagnetic agent and therefore a low PRE was observed are represented by green spheres in the ribbon representation of ubiquitin.

For all our model systems we restricted the available NOE data sets to distance restraints between exchangeable protons. The reason is twofold: The strongly limited NOE data set represents 1) a situation typically encountered for perdeuterated proteins and 2) an unbiased set of distance information to demonstrate the impact of PREs on a system with a low number of conventional restraints. According to this algorithm and using only 49 NOEs between exchangeable protons (exclusively backbone HN–HN distances), the PRE solution structure of ubiquitin was determined with a pairwise backbone rmsd of 1.3 Å (Figure 2; rmsd = root mean square deviation). In this regard it is noteworthy that only 7 out of the 49 NOEs comprise long-range distance information (more than four residues apart). The deviation from the X-ray structure[27] is 2.1 Å (backbone rmsd). The impact of PREs on the precision and accuracy of the structure derived from NMR spectroscopy diminishes, as expected, when the number of NOEs increases (see Table 1). Thus, for small proteins PREs are especially useful for the determination of the protein fold, which could then help in the assignment of additional NOEs or as input for determining an X-ray structure by molecular replacement.[28]

To demonstrate that the PRE approach is particularly suitable for large macromolecules in which the NOE assignment procedure is rather challenging, we applied our method

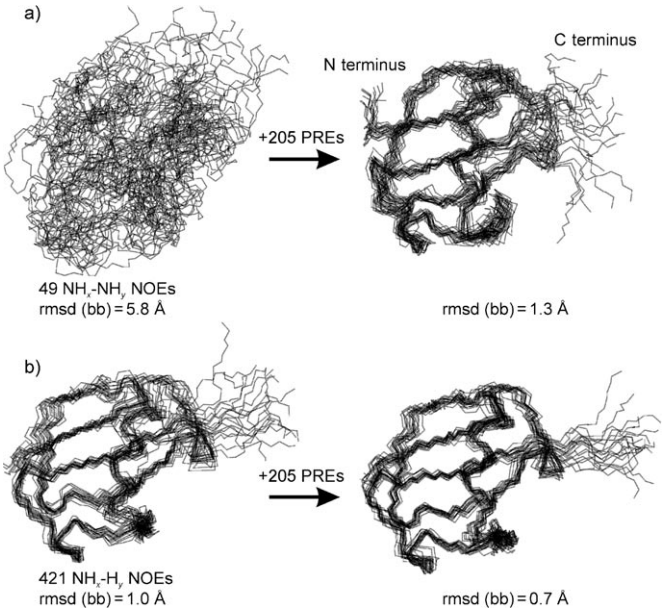

Figure 2. Least-square superposition of the 20 lowest energy structures of ubiquitin showing the impact of PREs on the solution structure using NOEs between a) only exchangeable protons or b) NOEs involving at least one exchangeable proton.

Table 1: Accuracy and precision of NMR structures of ubiquitin and MBP (rmsd [Å]).[a]

	Ubiquitin			
	NOEs		+205 PREs	
NOEs	⟨NMR⟩	NMR_{mean}–X-ray	⟨NMR⟩	NMR_{mean}–X-ray
49 NH–NH	5.80 (6.55)	5.70 (6.65)	1.28 (1.78)	2.05 (2.87)
421 NH–Hx	1.03 (1.74)	2.51 (2.80)	0.74 (1.19)	1.86 (2.29)
all 1318	0.35 (0.74)	0.58 (1.09)	0.37 (0.75)	0.65 (1.19)
	MBP			
	NOEs, RDCs, H-bonds, dihedral angles		+1248 PREs	
NOEs	⟨NMR⟩	NMR_{mean}–X-ray	⟨NMR⟩	NMR_{mean}–X-ray
822 NH–NH	11.12 (11.68)	10.56 (10.89)	1.27 (1.79)	2.56 (2.87)
all 1937	1.48 (1.77)	3.00 (3.51)	1.38 (1.87)	2.87 (3.05)

[a] Accuracy (pairwise rmsd of a bundle of NMR structures) denoted as ⟨NMR⟩ and precision (rmsd of the mean NMR structure versus the X-ray structure) of ubiquitin and MBP. The first number is for the backbone and the one in parenthesis for all heavy atoms.

to the 42 kDa MBP. Only NOEs between amide protons (exclusively backbone HN–HN distances) and RDCs, H-bonds, and dihedral angles were used together with 1248 PRE restraints for ^{1}H and ^{13}C nuclei (Figure 3, Table 1).

Quite encouragingly, the crystal structure of MBP[26] is in better agreement with the PRE structure than with the NOE-based structure from NMR spectroscopy (backbone rmsd of 2.6 Å vs. 3.0 Å). Certain regions in which the PRE and NOE-based structures significantly disagreed were observed (Figure 4). For example a stretch of approximately 10 residues (Ser233 to Asn241) was not well-defined in the structure determined by classical NMR spectroscopy owing to a lack of

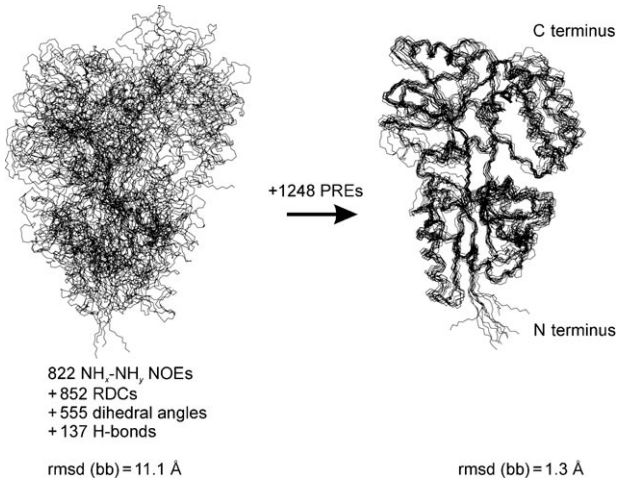

Figure 3. Least-square superposition of the 10 lowest energy structures of MBP showing the impact of PREs on the solution structure using NOEs between exchangeable protons, RDCs, and dihedral angle restraints.

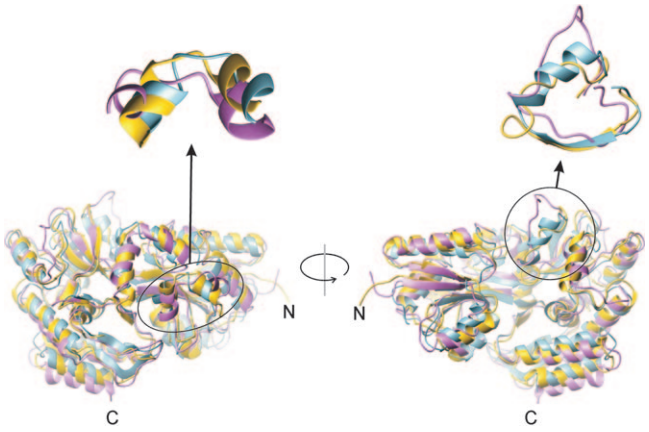

Figure 4. Least-square fitted overlay of the structures of maltodextrin-binding protein determined by X-ray diffraction (blue), NOE-based NMR spectroscopy (pink), and NMR spectroscopy including PREs (yellow). Regions with larger deviations (left: Lys83–Trp94, comprising two α helices; right: Ser233–Asn241 containing an α helix in the X-ray but a loop in the NOE-based NMR structure) are shown expanded.

NOE data.[25] In the crystal structure, these residues are part of an α helix.

Using only PRE restraints for the residues 233 to 241, a helical structure resulted from the refinement procedure (Figure 4). Consequently, the solution structure of this region is better represented by an α helix than by a flexible loop. Other deviations between the structure determined by X-ray diffraction and NOE-based NMR spectroscopy can be seen in the orientation of helices. Owing to the $\langle r^{-6} \rangle$ dependence, PREs are averaged towards smaller distances between the nuclei and the paramagnetic centers. Therefore, an open protein conformation is indicated by high PREs. In contrast, NOEs show a tendency towards compact structures in a conformational equilibrium, as NOEs increase steeply for small interproton distances. Thus, discrepancies between PREs and NOEs can be used to identify conformational

heterogeneity in biological macromolecules. Intrinsically unstructured regions can be identified by their uniformly high PREs.

Owing to their lower gyromagnetic ratio γ, the relaxation enhancements for carbon atoms are much smaller than for protons. This effect has been exploited when paramagnetic centers or tags are used, as switching to lower γ allows the observation of nuclei closer to the paramagnetic center.[15] When a paramagnetic solvent additive is used, as in our approach, the relaxation enhancements can be tuned simply by changing the concentration of [Gd(dtpa-bma)]; thus, there are no intrinsic nucleus-dependent signal cancellations.

In conclusion, we have demonstrated the use of PREs for the rapid structure determination of small to medium-sized proteins by using only a very limited set of NOE data. This method is therefore particularly suitable for high-molecular-mass proteins for which the NOE assignment procedure is rather challenging. The measurement and identification of PREs requires only a resonance assignment without any additional manual analysis of data. PRE structures could be obtained within a few days (ubiquitin) or weeks (MBP). This approach promises significant time savings for the structure determination process and has the potential to contribute to extending the current size limit of NMR spectroscopy.

Experimental Section

[Gd(dtpa-bma)] was purified from the commercial contrast agent Omniscan as described in reference [20]. Proton T_1 relaxation times (acquired in nonselective mode) were obtained from series of saturation-recovery HSQC spectra. ^{13}C T_1 relaxation times were obtained from CaCO and CON experiments[7,8] to which an inversion-recovery block was added at the beginning of the pulse sequence. The PRE structures of ubiquitin and MBP were deposited in the PDB under access numbers 2klg (ubiquitin) and 2klf (MBP).

Received: May 14, 2009
Revised: July 6, 2009
Published online: September 22, 2009

Keywords: maltose-binding protein · NMR spectroscopy · paramagnetic relaxation · protein structures · ubiquitin

[1] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12366.
[2] N. Tjandra, A. Bax, *Science* **1997**, *278*, 1111.
[3] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9279.
[4] K. H. Gardner, L. E. Kay, *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 357.
[5] M. Kainosho, T. Torizawa, Y. Iwashita, T. Terauchi, A. Mei Ono, P. Güntert, *Nature* **2006**, *440*, 52.
[6] V. Tugarinov, P. M. Hwang, J. E. Hollerenshaw, L. E. Kay, *J. Am. Chem. Soc.* **2003**, *125*, 10420.
[7] W. Bermel, I. Bertini, L. Duma, I. C. Felli, L. Emsley, R. Pierattelli, P. R. Vasos, *Angew. Chem.* **2005**, *117*, 3149; *Angew. Chem. Int. Ed.* **2005**, *44*, 3089.
[8] W. Bermel, I. Bertini, I. C. Felli, M. Piccioli, R. Pierattelli, *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 25.
[9] B. H. Oh, W. M. Westler, P. Darba, J. L. Markley, *Science* **1988**, *240*, 908.
[10] K. H. Gardner, X. C. Zhang, K. Gehring, L. E. Kay, *J. Am. Chem. Soc.* **1998**, *120*, 11738.
[11] C. D. Barry, A. C. North, J. A. Glasel, R. J. Williams, A. V. Xavier, *Nature* **1971**, *232*, 236.
[12] J. L. Battiste, G. Wagner, *Biochemistry* **2000**, *39*, 5355.
[13] I. Bertini, *Concepts Magn. Reson.* **2002**, *14*, 259.
[14] I. Bertini, C. Luchinat, G. Parigi, *Solution NMR of Paramagnetic Molecules*, Elsevier, Dordrecht, **2001**.
[15] I. Bertini, C. Luchinat, G. Parigi, R. Pierattelli, *ChemBioChem* **2005**, *6*, 1536.
[16] G. Pintacuda, M. John, X. C. Su, G. Otting, *Acc. Chem. Res.* **2007**, *40*, 206.
[17] J. Wöhnert, K. J. Franz, M. Nitz, B. Imperiali, H. Schwalbe, *J. Am. Chem. Soc.* **2003**, *125*, 13338.
[18] T. Madl, L. Van Melderen, N. Mine, M. Respondek, M. Oberer, W. Keller, L. Khatai, K. Zangger, *J. Mol. Biol.* **2006**, *364*, 170.
[19] G. Pintacuda, G. Otting, *J. Am. Chem. Soc.* **2002**, *124*, 372.
[20] M. Respondek, T. Madl, C. Göbl, R. Golser, K. Zangger, *J. Am. Chem. Soc.* **2007**, *129*, 5228.
[21] A. M. Petros, L. Mueller, K. D. Kopple, *Biochemistry* **1990**, *29*, 10041.
[22] M. Sakakura, S. Noba, P. A. Luchette, I. Shimada, R. S. Prosser, *J. Am. Chem. Soc.* **2005**, *127*, 5826.
[23] A. Bernini, V. Venditti, O. Spiga, N. Niccolai, *Prog. Nucl. Magn. Reson. Spectrosc.* **2009**, *54*, 278.
[24] G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **1998**, *120*, 6836.
[25] G. A. Mueller, W. Y. Choy, D. Yang, J. D. Forman-Kay, R. A. Venters, L. E. Kay, *J. Mol. Biol.* **2000**, *300*, 197.
[26] A. J. Sharff, L. E. Rodseth, F. A. Quiocho, *Biochemistry* **1993**, *32*, 10553.
[27] S. Vijay-Kumar, C. E. Bugg, W. J. Cook, *J. Mol. Biol.* **1987**, *194*, 531.
[28] B. R. Szymczyna, R. E. Taurog, M. J. Young, J. C. Snyder, J. E. Johnson, J. R. Williamson, *Structure* **2009**, *17*, 499.

