J|A|C|S
ARTICLES
Published on Web 02/13/2004

Fast Structure-Based Assignment of ^{15}N HSQC Spectra of Selectively ^{15}N-Labeled Paramagnetic Proteins

Guido Pintacuda,[†,‡] Max A. Keniry,[†] Thomas Huber,[§] Ah Young Park,[†] Nicholas E. Dixon,[†] and Gottfried Otting*[,†]

Contribution from the Australian National University, Research School of Chemistry, Canberra, ACT 0200, Australia. Karolinska Institute, Department of Medical Biochemistry and Biophysics, Tomtebodavägen 6, S-171 77 Stockholm, Sweden, and the University of Queensland, Department of Mathematics, Brisbane QLD 4072, Australia

Received October 30, 2003; E-mail: gottfried.otting@anu.edu.au

Abstract: A novel strategy for fast NMR resonance assignment of ^{15}N HSQC spectra of proteins is presented. It requires the structure coordinates of the protein, a paramagnetic center, and one or more residue-selectively ^{15}N-labeled samples. Comparison of sensitive undecoupled ^{15}N HSQC spectra recorded of paramagnetic and diamagnetic samples yields data for every cross-peak on pseudocontact shift, paramagnetic relaxation enhancement, cross-correlation between Curie-spin and dipole−dipole relaxation, and residual dipolar coupling. Comparison of these four different paramagnetic quantities with predictions from the three-dimensional structure simultaneously yields the resonance assignment and the anisotropy of the susceptibility tensor of the paramagnetic center. The method is demonstrated with the 30 kDa complex between the N-terminal domain of the ϵ subunit and the θ subunit of *Escherichia coli* DNA polymerase III. The program PLATYPUS was developed to perform the assignment, provide a measure of reliability of the assignment, and determine the susceptibility tensor anisotropy.

Introduction

NMR spectroscopy is commonly used in pharmaceutical industry to screen libraries of chemical compounds for ligands binding to protein targets. Chemical shift changes observed in ^{15}N HSQC spectra of ^{15}N-labeled proteins provide evidence of even weak binding interactions. Provided that the sequence-specific resonance assignment of the ^{15}N HSQC cross-peaks is available, they also yield information about the binding site.[1] Resonance assignments of proteins with molecular weights exceeding about 15 kDa are usually obtained by analysis of ^{15}N/^{13}C double- or ^{15}N/^{13}C/^{2}H triple-labeled protein samples with the help of 3D and 4D NMR experiments.[2] Availability of a crystal structure of the protein is of little help in this process. In modern pharmaceutical research, however, the crystal structure of the protein target is frequently known prior to NMR resonance assignments. Here, we present a novel strategy for rapid sequence-specific assignment of ^{15}N HSQC cross-peaks of selectively ^{15}N-labeled proteins which explicitly relies on prior knowledge of the three-dimensional structure of the protein. The method requires one or several selectively ^{15}N-labeled protein samples and availability of a binding site for a paramagnetic metal ion, but no ^{13}C labeling and no NMR data other than a minimum number of highly sensitive ^{15}N HSQC spectra. Besides resonance assignments, the anisotropy parameters of the susceptibility tensor of the metal ion are obtained, providing a basis for the measurement of important intermolecular restraints to define the orientation of binding ligands.[3]

The assignment strategy was established with the 30 kDa complex between the N-terminal domain of the ϵ subunit, ϵ186, and the θ subunit of *Escherichia coli* DNA polymerase III.[4,5] A 1.8 Å crystal structure of ϵ186 is available.[6] The structure seems to be maintained in solution and in complex with the θ subunit.[7] The active site contains two Mn^{2+} ions which can be replaced by a single lanthanide ion, as in the related domain of DNA polymerase I.[8]

[†] Australian National University.
[‡] Karolinska Institute.
[§] University of Qeensland.

(1) (a) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531−1534. (b) Medek, A.; Hajduk, P. J.; Mack, J.; Fesik, S. W. *J. Am. Chem. Soc.* **2000**, *122*, 1241−1242. (c) Pellecchia, M.; Senn, D. S.; Wüthrich, K. *Nat. Rev. Drug Discovery* **2002**, *1*, 211−219.
(2) Sattler, M.; Schleucher, J.; Griesinger, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93−158.
(3) (a) Ubbink, M.; Ejdebäck, M.; Karlsson, B. G.; Bendall, D. S. *Structure* **1998**, *6*, 323−335. (b) Gochin, M. *J. Biomol. NMR* **1998**, *12*, 243−257. (c) Tu, K.; Gochin, M. *J. Am. Chem. Soc.* **1999**, *121*, 9276−9285. (d) Gochin, M. *Structure* **2000**, *8*, 441−452. (e) Nocek, J. M.; Huang, K.; Hoffman, B. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2538−2543. (f) Goodfellow, B. J.; Nunes, S. G.; Rusnak, F.; Moura, I.; Ascenso, C.; Moura, J. J. G.; Volkman, B. F.; Markley, J. L. *Protein Sci.* **2002**, *11*, 2464−2470. (g) Gaponenko, V.; Altieri, A. S.; Li, J.; Byrd, R. A. *J. Biol. NMR* **2002**, *24*, 43−148.
(4) DeRose, E. F.; Li, D.; Darden, T.; Harvey, S.; Perrino, F. W.; Schaaper, R. M.; London, R. E. *Biochemistry* **2002**, *41*, 94−110.
(5) Keniry, M. A.; Berthon, H. A.; Yang, J. Y.; Miles, C. S.; Dixon, N. E. *Protein Sci.* **2000**, *9*, 721−733.
(6) Hamdan, S.; Carr, P. D.; Brown, S. E.; Ollis, D. L.; Dixon, N. E. *Structure* **2002**, *10*, 535−546.
(7) DeRose, E. F.; Darden, T.; Harvey, S.; Gabel, S.; Perrino, F. W.; Schaaper, R. M.; London, R. E. *Biochemistry* **2003**, *42*, 3635−3644.
(8) (a) Frey, M. W.; Frey, S. T.; DeW Horrocks, W., Jr.; Kaboord, B.; Benkovic, S. J. *Chem. Biol.* **1996**, *3*, 393−403. (b) Brautigam, C. A.; Aschheim, K.; Steitz, T. A. *Chem. Biol.* **1999**, *6*, 901−908.

Experimental Section

Protein Samples. The N-terminal domain of the ϵ subunit of *E. coli* polymerase III containing the first 186 residues (ϵ186) was expressed and purified as described,[6] using a minimal medium supplemented with ^{15}N-labeled amino acids for the expression of selectively ^{15}N-labeled samples.[9]

Three samples were prepared with ^{15}N-Leu, ^{15}N-Phe, and ^{15}N-Val labeling, respectively. Complexes of all samples with unlabeled θ subunit were isolated by chromatography.[10] Complexes with lanthanide ions were prepared at 0.1 mM protein concentrations in a buffer of 20 mM Tris, 100 mM NaCl, and 0.1 mM dithiothreitol at pH 7.0 (NMR buffer) by the addition of 1.2 equiv of $LnCl_3$, followed by six washes with NMR buffer using an ultrafiltration device. The same protein samples were used for measurements with different lanthanides, after regeneration of the apo-protein complex by washing with NMR buffer in the presence of 400 mM NaCl. The NMR measurements were performed with 0.5 mM protein concentration in NMR buffer.

NMR Measurements and Data Evaluation. All NMR data were recorded at 25 °C on a Varian Inova 600 NMR spectrometer equipped with a PENTA probe (^{1}H, ^{13}C, ^{15}N, ^{31}P, ^{2}H) and operating at a ^{1}H NMR frequency of 600 MHz. ^{15}N HSQC spectra were recorded with a pulse sequence without the refocusing INEPT period and without decoupling during the acquisition time to measure the one-bond ^{1}H$-^{15}$N coupling constant (Supporting Information). This scheme has the additional advantage of minimizing loss of magnetization due to paramagnetic relaxation. Loss of magnetization was further minimized by shortening the total INEPT period to 2.2 ms for the paramagnetic samples. The following parameters were used: $t_{1\max} = 51$ ms, $t_{2\max} = 146$ ms, 23 ppm sweep width in the ^{1}H dimension, total recording time 5 and 16 h for diamagnetic and paramagnetic samples, respectively. The spectra were processed with the program Prosa,[11] using a cosine window function in both dimensions and two-fold zerofilling before Fourier transformation.

The antiphase line shape of each doublet observed in the ^{1}H dimension was fitted by Lorentzian line shapes, using the equation:

$$L(\nu) = \frac{A}{1 + 4\pi^2\left(\frac{\nu - \nu_0 - J/2}{\lambda - \eta}\right)^2} - \frac{A\cdot\frac{\lambda - \eta}{\lambda + \eta}}{1 + 4\pi^2\left(\frac{\nu - \nu_0 + J/2}{\lambda + \eta}\right)^2} \qquad (1)$$

where $L(\nu)$ is the peak intensity as a function of the frequency ν, A the amplitude of the low-field doublet component, ν_0 the frequency of the center of the doublet, J the scalar coupling constant, and λ and η the auto- and cross-correlated relaxation rates, respectively. The least-squares fits were performed with the program KaleidaGraph (Synergy Software, Reading, U.S.A.).

Two versions of the code for the assignment were developed in C and in Mathematica (Wolfram Research). The C program, with the name PLATYPUS (*p*aramagnetic *l*abeling for *a*ssignment with *t*errific *y*ields of *p*roteins *u*sing their *s*tructures), took less than one minute on a single processor 2 GHz PC for a typical run with floating assignments. Error estimates were obtained by separate routines in the programs, fitting assignments to structures generated by random addition of structural noise. PLATYPUS is available upon request.

Results

Paramagnetic Parameters. The presence of a fast-relaxing paramagnetic ion in the protein changes the appearance of the ^{15}N HSQC spectrum in several ways which are manifested as changes in chemical shifts, line widths, and one-bond coupling constants:

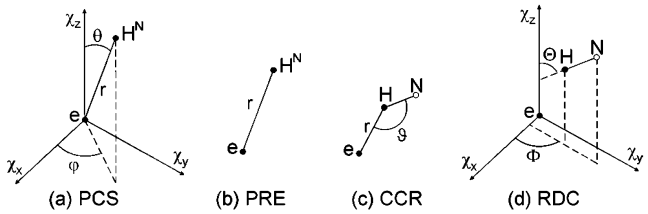

Figure 1. Schematic representations of the geometric dependence of the four paramagnetic effects used in the present work for amide resonance assignment. (a) ^{1}H^N pseudocontact shift (PCS), (b) paramagnetic relaxation enhancement (PRE), (c) cross-correlation between the Curie-spin and ^{1}H$-^{15}$N dipole−dipole relaxation (CCR), and (d) ^{1}H$-^{15}$N residual dipolar coupling (RDC).

(i) pseudocontact shifts (PCS) which change the observed chemical shifts due to interaction with the anisotropic susceptibility tensor χ of the paramagnetic center:[12,13]

$$\delta^{\mathrm{PCS}} = \frac{1}{12\pi r^3}\left[\Delta\chi_{\mathrm{ax}}(3\cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{\mathrm{rh}}\sin^2\theta\cos 2\varphi\right] \qquad (2)$$

(ii) paramagnetic relaxation enhancement (PRE) which is predominantly due to the Curie contribution:[14]

$$\lambda^{\mathrm{PRE}} = \frac{k}{r^6}\left(4\tau_{\mathrm{r}} + \frac{3\tau_{\mathrm{r}}}{1 + \omega_{\mathrm{H}}^2\tau_{\mathrm{r}}^2}\right) \qquad (3)$$

(iii) cross-correlated relaxation (CCR) between the Curie-spin and ^{1}H$-^{15}$N dipole−dipole relaxation:[15]

$$\eta^{\mathrm{CCR}} = \kappa\frac{3\cos^2\vartheta - 1}{r^3}\left(4\tau_{\mathrm{r}} + \frac{3\tau_{\mathrm{r}}}{1 + \omega_{\mathrm{H}}^2\tau_{\mathrm{r}}^2}\right) \qquad (4)$$

(iv) residual dipolar coupling (RDC) due to alignment of the molecule with respect to the magnetic field:[13,16]

$$\Delta\nu^{\mathrm{RDC}} = K\left[\Delta\chi_{\mathrm{ax}}(3\cos^2\Theta - 1) + \frac{3}{2}\Delta\chi_{\mathrm{rh}}\sin^2\Theta\cos 2\Phi\right] \qquad (5)$$

where (r, θ, φ), $(\vartheta, \Theta, \Phi)$ define the position of the amide proton and the N−H bond vector with respect to the metal center and the main axes of its χ tensor (Figure 1), $\Delta\chi_{\mathrm{ax}} = \chi_z - (\chi_x + \chi_y)/2$, $\Delta\chi_{\mathrm{rh}} = (\chi_x - \chi_y)$, ω_{H} is the Larmor frequency of the proton, τ_{r} is the rotational correlation time, and the proportionality factors are:[12−16]

(9) Senn, H.; Eugster, A.; Otting, G.; Suter, F.; Wüthrich, K. *Eur. Biophys. J.* **1987**, *14*, 301−306.
(10) Hamdan, S.; Bulloch, E. M.; Thompson, P. R.; Beck, J. L.; Yang, J. Y.; Crowther, J. A.; Lilley, P. E.; Carr, P. D.; Ollis, D. L.; Brown, S. E.; Dixon, N. E. *Biochemistry* **2002**, *41*, 5266−5275.
(11) Güntert, P.; Dötsch, V.; Wider, G.; Wüthrich, K. *J. Biomol. NMR* **1992**, *2*, 619−629.
(12) Bertini, I.; Luchinat, C. *Coord. Chem. Rev.* **1996**, *150*, 1−256.
(13) Bertini, I.; Luchinat, C.; Parigi, G. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *40*, 249−273.
(14) (a) Guéron, M. *J. Magn. Reson.* **1975**, *19*, 58−66. (b) Vega, A. J.; Fiat, D. *Mol. Phys.* **1976**, *31*, 347−351. The *direct* dipolar relaxation mechanism (Solomon, I.; Bloembergen, N. *J. Chem. Phys.* **1956**, *25*, 261−266) can be neglected for a fast-relaxing paramagnetic ion in a slowly tumbling molecule.
(15) (a) Ghose, R.; Prestegard, J. H. *J. Magn. Reson.* **1997**, *128*, 138−143. (b) Boisbouvier, J.; Gans, P.; Blackledge, M.; Brutscher, B.; Marion, D. *J. Am. Chem. Soc.* **1999**, *121*, 7700−7701. (c) Madhu, P. K.; Grandori, R.; Hohenthanner, K.; Mandal, P. K.; Müller, N. *J. Biomol. NMR* **2001**, *20*, 31−37. (d) Pintacuda, G.; Hohenthanner, K.; Otting, G.; Müller, N. *J. Biomol. NMR* **2003**, *27*, 115−132.
(16) (a) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9279−9283. (b) Barbieri, R.; Bertini, I.; Lee, Y.-M.; Luchinat, C.; Velders, A. H. *J. Biomol. NMR* **2002**, *14*, 365−368.

$$k = \frac{1}{5}\left(\frac{\mu_0}{4\pi}\right)^2 B_0^2\gamma_H^2(g_J\mu_B)^4\, J^2(J+1)^2/(3k_BT)^2 \quad (6)$$

$$\kappa = \frac{1}{30}\left(\frac{\mu_0}{4\pi}\right)^2 B_0\gamma_H^2\gamma_N\hbar(g_J\mu_B)^2\, J(J+1)/(r_{NH}^3k_BT) \quad (7)$$

$$K = -\frac{1}{120\pi^2} B_0^2\gamma_H\gamma_N\hbar S^2/(r_{NH}^3k_BT) \quad (8)$$

where γ_H and γ_N are the proton and nitrogen magnetogyric ratios, r_{NH} the N−H distance (1.03 Å), μ_0 the vacuum permeability, B_0 the magnetic field, $\hbar$ Planck's constant divided by 2π, k_B the Boltzmann constant, μ_B the Bohr magneton, T the temperature, g_J the g-factor, J the total spin moment, and S the order parameter of the molecular alignment. In general, any magnetic anisotropy of the diamagnetic molecule is much smaller than the anisotropy of the paramagnetic metal ion and can be neglected in the present context. Therefore, the axes of the χ tensor can be assumed to coincide with those of the alignment tensor and the anisotropy parameters are the same except for a scaling factor which includes S^2.[13] In the present work, S^2 was set to 0.9 for all amide protons.

All four paramagnetic parameters are readily measured by comparison of undecoupled ^{15}N HSQC spectra recorded in the presence and absence of a paramagnetic metal ion (Figure 2). Line shape fitting of the antiphase doublets observed in the ^{1}H dimension yields the line width and frequency positions of each doublet component, and hence PCS, PRE, CCR, and RDC values (Figure 2, insert). Conversely, all four paramagnetic parameters can be accurately predicted from the three-dimensional structure of the protein once the electronic χ tensor anisotropy and the correlation time τ_r are known. Therefore, the resonance assignment can be obtained from a comparison of the experimental with predicted paramagnetic parameters.

Experimental Measurements. Spectra were recorded for three samples of ϵ186, selectively labeled with ^{15}N-Leu, ^{15}N-Val, and ^{15}N-Phe, respectively, and complexed with the unlabeled θ subunit. Figure 2 shows a superposition of the ^{15}N HSQC spectra of the ^{15}N-Leu labeled sample, recorded in the presence and absence of Dy^{3+}. ϵ186 contains 17 leucine residues, but only eight ^{15}N HSQC cross-peaks could be detected for the ^{15}N-Leu labeled sample in the presence of Dy^{3+} ($J = {}^{15}/_2$, $g_J = {}^4/_3$),[12] since resonances of the $^1H^N$ spins within 14 Å of the metal ion were broadened beyond detection. Similarly, only 6 out of 12 and 6 out of 10 cross-peaks could be observed for the ^{15}N-Val- and ^{15}N-Phe-labeled samples, respectively, in the presence of Dy^{3+}. The PCS effect for each cross-peak is similar in the ^{1}H and ^{15}N dimension, since the ^{1}H and ^{15}N spins of each amide group are located at similar positions with respect to the paramagnetic center. Therefore, the cross-peaks are shifted along nearly parallel diagonal lines, allowing the identification of pairs of diamagnetic and paramagnetic cross-peaks.[17] In the case of the ^{15}N-Leu labeled complex with Dy^{3+}, only the diamagnetic partners of the two most strongly shifted peaks (Leu57, Leu176) could not be identified unambiguously in this way (Figure 2), resulting in corresponding ambiguities in their PCS values and other paramagnetic parameters. These were resolved as discussed below.

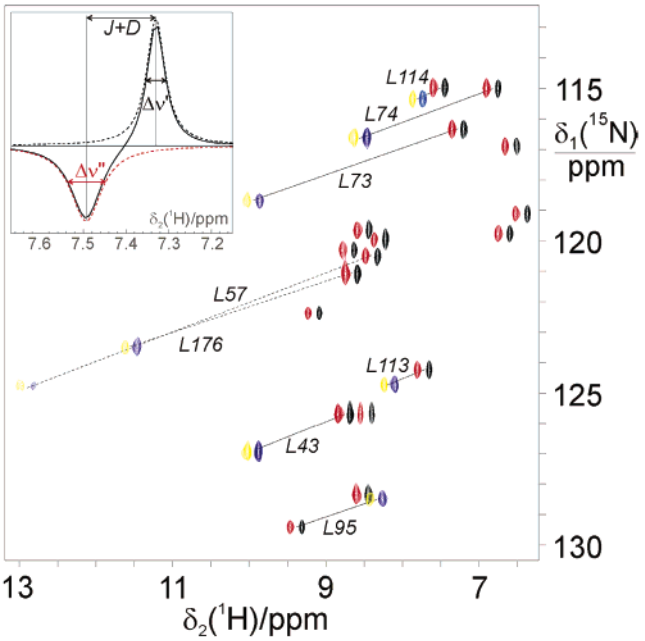

Figure 2. Antiphase doublets of ^{15}N HSQC spectra recorded with 0.5 mM aqueous solutions of ^{15}N-Leu-labeled ϵ186 in complex with unlabeled θ at 25 °C and pH 7.0. Shown is the superposition of two spectra recorded with Dy^{3+} (yellow and blue peaks) and in the absence of a paramagnetic metal ion (red and black peaks). Straight lines connect pairs of diamagnetic and paramagnetic cross-peaks. Dashed lines identify the two most strongly shifted paramagnetic cross-peaks for which the diamagnetic partners could not be identified by simple visual inspection. A cross section through the doublet at $\delta_1 = 115$ ppm/$\delta_2 = 7.4$ ppm is shown in the insert together with the individual fitted components (dashed lines). PCS values are obtained from the chemical shifts, PRE results in overall line broadening, CCR effects result in differential line broadening between the two doublet components of the ^{1}H NMR signal, and RDCs are manifested by the altered separation of the doublet components.

For accurate measurement of PCS values, diamagnetic samples of the $\epsilon186-\theta$ complexes were prepared with La^{3+} in the metal binding site. With few exceptions, however, the ^{15}N HSQC spectra recorded without any lanthanide ion were closely similar to those of the La^{3+} complexes and significant differences in chemical shifts were observed only in the ^{15}N dimension.

Algorithm for Resonance Assignment. Provided that the assignment is known for at least $n \geq 5$ pairs of paramagnetic and corresponding diamagnetic cross-peaks, the orientation and the anisotropy values of the χ tensor can be readily determined from eqs 2 and 5 on the basis of the three-dimensional protein structure, either by singular value decomposition[18] or by multidimensional least-squares minimization of the difference between experimental and back-calculated PCS and RDC values.[13] Similarly, the rotational correlation time τ_r can be fitted from the set of experimental PRE and CCR data. Once these constants have been determined, the paramagnetic effects for all cross-peaks can be back-calculated from the structure.

The correct assignment is assumed to be the one with minimal difference between experimental and predicted paramagnetic data:

(17) (a) Allegrozzi, M.; Bertini, I.; Janik, M. B. L.; Lee, Y.-M.; Liu, G.; Luchinat, C. *J. Am. Chem. Soc.* **2000**, *122*, 4154−4161. (b) Bertini, I.; Janik, M. B. L.; Lee, Y.-M.; Luchinat, C.; Rosato, A. *J. Am. Chem. Soc.* **2001**, *123*, 4181−4188.

(18) Losonczi, J. A.; Andrec, M.; Fisher, M. W. F.; Prestegard, J. H. *J. Magn. Reson.* **1999**, *138*, 334−342.

$$Q = \sum_p k_p \sum_n (d_{p,n}^{exp} - d_{p,n}^{cal})^2 \quad (9)$$

with summation over the n cross-peaks observed for the paramagnetic protein and the $p = 4$ different types of paramagnetic data d_p described above, each of them weighted by a constant k_p. In calculating Q, differences below a predefined uncertainty threshold were set to zero to account for errors in the experimental data.

An exhaustive analysis of all $n!$ possible assignments becomes intractable for large numbers of cross-peaks.[19,20] Consequently, we developed an algorithm which finds the best assignment with a limited number of steps independent of the size of the experimental data set. In our approach, for an initial set of χ tensor anisotropy and orientation as well as the correlation time τ_r the assignment is first optimized according to an algorithm known as the "Hungarian method for minimal cost assignment".[21] χ tensor anisotropy and orientation, as well as the correlation time τ_r, are subsequently calculated by gradient minimization of the residual Q (eq 9), while the assignment is continuously optimized at each step of the minimization. In this way, fitting of the paramagnetic parameters (χ tensor and τ_r) and assignment are achieved simultanously.

To improve the convergence to the global minimum of Q for this local optimization strategy, minimizations are started at trial points from a five-dimensional grid search, where the grid points represent different magnitudes and orientations of the main components of the χ tensor with respect to the molecular frame. m grid points are generated by systematic variation of the tensor components, where $m \ll n!$.

In our experience, the cost function hypersurface is sufficiently smooth that initialization of axial and rhombic components to any values between 10^{-32} m^3 and 10^{-31} m^3, and 10^{-33} m^3 and $5 \cdot 10^{-32}$ m^3, respectively, and systematic rotation of the main components of the χ tensor in 20° steps around the molecular frame axes is sufficient to locate the global minimum in Q. For example, in the case of Dy^{3+}, $m = 2916$ ($= 18 \times 18 \times 9$) was used, initializing the search to axial and rhombic components of 10^{-32} m^3 and 10^{-32} m^3, respectively. The algorithm does not test all $n!$ conceivable resonance assignments but only optimizes the assignments initially suggested for the m grid points, followed by a minimization path with minimum Q values.

For the paramagnetic cross-peaks with ambiguous diamagnetic partner, the apparent paramagnetic parameters were determined for all possible pairings of diamagnetic and paramagnetic cross-peaks, and all different combinations were tested during the fitting procedure. For those cross-peaks, minimization of the residual Q provided not only the resonance assignment of the paramagnetic cross-peaks, but also the identification of their diamagnetic partners.

Error Estimates. The resonance assignments obtained for the eight paramagnetic cross-peaks measured for the ^{15}N-Leu-labeled complex remained unchanged when additional data measured for ^{15}N-Val- and ^{15}N-Phe-labeled $\epsilon 186-\theta$ complexes with Dy^{3+} were included in the fit. The resonance assignments based on these three labeled samples also remained unchanged when random variations of up to 10% were added to the experimental data.[18]

An important point for the optimization is the choice of the constants k_p which determine the relative contribution of the different paramagnetic effects to the target function Q. To find weighting constants such that each of the four types of paramagnetic data contributed with comparable weight to the final target function Q, four different runs of initial optimizations were performed, where each type of paramagnetic data was used individually without any of the other data. In this way, a minimum value Q_P was obtained for each type of data. The assignments made by these optimizations were mostly incorrect and disregarded. The constants k_p were subsequently chosen so that each residual Q_p had a comparable value. Numerical values of the constants k_p employed in the present work are included with Table 1. The threshold values, below which differences between experimental and predicted data no longer contributed to Q, were defined by the experimental inaccuracies.

Table 1 shows that the experimental and predicted data agreed most closely for the PCS data, reflecting the high accuracy of the experimental data, while the other paramagnetic restraints primarily served to resolve assignment ambiguities for residues with similar PCS values. This result was expected, since PCS data are less sensitive with respect to local reorientation of the N−H bonds than CCR and RDC data, while PRE data depend on the reproducibility of the sample properties and magnetic field homogeneity between measurements with different samples. Such experimental uncertainties could explain why the experimental PRE values were generally found to be much larger than the back-calculated data. In any case, the definition of the residual Q (eq 9) ensures that the relative magnitude of the PRE values matters more than their absolute values. The agreement between experimental and back-calculated data was much better for the other paramagnetic parameters (Table 1). In principle, RDC data would be expected to yield the same tensor anisotropy axes as the PCS data, but RDC values are more difficult to measure accurately and can be compromised by nonuniform, residue dependent order parameters S^2 (eq 8).

Clearly, differences between the crystal structure and the solution structure of $\epsilon 186$ in the complex with θ would present a potential source of wrong assignments. A good criterion for the reliability of the individual assignments is thus presented by their resilience toward randomly introduced structural changes.[22] This was achieved by a Monte Carlo routine which generated 1000 new structures from the crystal structure by simultaneously reorienting the N−H bond vectors, translating the N−H groups, and translating the Ln^{3+} ion in a random manner. Using Gaussian distributions of the added structural noise with standard deviations of $\sigma^{HN} = 15°$ and $\sigma^r = \sigma^{Ln} = 0.3$ Å, respectively, the assignments of most cross-peaks were reproduced for more than 99% of the 1000 randomized structures (Table 2). Table 2 further shows that availability of more than one selectively ^{15}N-labeled sample improved the reliability of the resonance assignments, with only two assignments (Val 38 and Val 82) reproduced in as few as 75% of the randomized structures.

Besides providing a criterion for the reliability of individual assignments, addition of structural noise also provides a means

(19) Forsberg, J. H.; Delaney, R. M.; Zhao, Q.; Harakas, G.; Chandran, R. *Inorg. Chem.* **1995**, *34*, 3705−3715.
(20) Zweckstetter, M. *J. Biomol. NMR* **2003**, *27*, 41−56.
(21) Kuhn, H. W. *Nav. Res. Logist. Quart.* **1955**, *2*, 83−97.
(22) Zweckstetter, M.; Bax, A. *J. Biomol. NMR* **2002**, *23*, 127−137.

Table 1. Experimental and Back-Calculated Paramagnetic Effects PCS (δ^{PCS}), RDC (D^{RDC}), PRE (Autocorrelated Relaxation λ^{PRE}), and CCR (Cross-Correlated Relaxation η^{CCR}) for the Three ^{15}N-Labeled Samples of ϵ186 in Complex with Unlabeled θ[a,b]

	PCS (ppm)				RDC (Hz)				PRE (s^{-1})				CCR (s^{-1})			
	δ_{La}	δ_{Dy}	δ^{PCS}_{exp}	δ^{PCS}_{cal}	J_{La}	$(J+D)_{Dy}$	D^{RDC}_{exp}	D^{RDC}_{cal}	λ_{La}	λ_{Dy}	λ^{PRE}_{exp}	λ^{PRE}_{cal}	η_{La}	η_{Dy}	η^{CCR}_{exp}	η^{CCR}_{cal}
L43	8.87	10.04	1.17	1.19	−92.5	−87.6	4.9	2.0	69.1	109.3	40.2	22.1	13.2	21.4	8.2	19.7
L57	8.77	11.64	2.87	2.87	−94.2	−85.8	8.4	4.9	83.6	162.1	78.6	22.1	9.4	10.0	0.6	−12.4
L73	7.38	10.03	2.65	2.71	−91.6	−95.7	−4.1	−5.7	91.7	158.3	66.6	57.7	8.8	33.9	25.1	32.5
L74	6.92	8.63	1.71	1.81	−92.1	−100.5	−8.4	−9.1	74.8	137.0	62.2	28.0	4.4	29.5	25.1	22.9
L95	9.53	8.42	−1.11	−1.13	−91.8	−96.0	−4.2	−4.1	98.0	151.4	53.4	46.6	15.0	20.7	5.7	0.9
L113	7.84	8.25	0.41	0.40	−94.1	−82.8	11.3	6.9	84.2	111.2	27.0	12.8	5.6	−8.2	−13.8	−29.7
L114	7.62	7.88	0.26	0.26	−92.2	−77.0	15.2	17.0	75.4	108.7	33.2	8.7	6.3	5.0	−1.3	−1.8
L176	8.51	13.00	4.49	4.48	−93.5	−88.8	4.7	2.5	77.9	223.1	145.1	81.6	6.9	−3.2	−10.1	−30.7
F72	7.04	9.39	2.35	2.16	−90.7	−100.8	−10.1	−10.0	84.8	197.9	113.2	50.1	5.7	6.3	0.6	11.4
F79	9.81	11.19	1.38	1.36	−93.2	−94.5	−1.3	−0.4	118.1	206.7	88.4	36.7	17.0	45.9	28.9	26.5
F86	8.91	7.68	−1.23	−1.18	−93.7	−77.9	15.8	21.0	70.4	159.6	89.2	26.2	10.1	15.1	5.0	19.2
F111	9.04	9.05	−0.01	−0.01	−94.5	−88.5	6.0	7.8	69.7	177.2	107.4	32.6	9.4	−15.1	−24.5	−7.8
F124	7.48	6.64	−0.84	−0.78	−90.5	−93.2	−2.7	−4.2	82.9	175.9	93.0	9.3	5.7	−11.9	−17.6	−15.6
F138	8.65	9.32	0.67	0.66	−90.6	−77.2	13.4	15.0	74.1	123.2	49.0	14.5	10.7	3.2	−7.5	−9.6
V38	9.69	9.75	0.06	0.04	−93.7	−87.0	6.7	−1.9	103.7	185.4	81.6	43.8	17.6	49.0	31.4	29.3
V39	9.03	9.51	0.48	0.44	−93.8	−96.0	−2.2	−6.9	104.9	125.7	20.8	15.8	30.1	25.1	−5.0	17.4
V58	10.86	15.54	4.68	4.69	−93.6	−100.2	−6.6	−9.1	110.0	194.2	84.2	64.2	19.5	39.6	20.1	22.9
V82	7.52	7.41	−0.09	−0.10	−90.1	−97.6	−7.5	−6.3	77.3	129.4	52.2	23.9	6.9	31.4	24.5	12.4
V127	8.77	7.81	−0.96	−0.88	−92.5	−99.6	−7.1	−10.0	75.4	113.1	37.6	20.4	10.7	−1.2	−11.9	−19.2
V133	7.69	7.92	0.23	0.25	−93.3	−91.2	2.1	3.0	118.1	135.1	16.8	55.8	15.1	−6.3	−21.4	−27.5

[a] For each of the four paramagnetic properties, four columns report the experimentally measured values of the La^{3+} and Dy^{3+} complexes, the net paramagnetic effect and the value calculated from the crystal structure after optimization of resonance assignment, χ tensor and τ_r parameters, using the data from all three samples. The assignment obtained is reported on the left. [b] Uncertainty thresholds of, respectively, ±0.025 ppm, ±2 Hz, ±38 s^{-1} and ±12 s^{-1} were used for the four classes of measurements. The values of the weighting factors used in eq 9 were: $k_{PCS} = 1.0$ ppm^{-2}, $k_{RDC} = 0.01$ Hz^{-2}, $k_{PRE} = 0.00025$ s^2, $k_{CCR} = 0.00057$ s^2.

Table 2. Robustness of the Resonance Assignment against Structural Noise, when Using Only Data from the ^{15}N-Leu-Labeled Sample (column 1), after Addition of the Data from the ^{15}N-Phe-Labeled Sample (column 2), and Using the Combined Data from the ^{15}N-Leu-, ^{15}N-Phe-, and ^{15}N-Val-Labeled Samples (column 3)[a]

	1 aa	2 aa	3 aa
L43	0.89	0.99	0.99
L57	0.98	0.99	0.99
L73	0.94	0.99	0.99
L74	0.87	0.99	0.99
L95	1.00	1.00	1.00
L113	0.97	0.98	0.99
L114	0.97	0.98	0.99
L176	0.99	1.00	1.00
F72		0.99	0.99
F79		0.97	0.97
F86		1.00	1.00
F111		0.99	0.99
F124		0.99	0.99
F138		0.97	0.98
V38			0.75
V39			0.94
V58			1.00
V82			0.75
V127			1.00
V133			0.89

[a] The frequency of the assignment for each residue is reported, as obtained in 1000 calculations performed after addition of structural noise to the protein structure coordinates and angles, as well as the lanthanide coordinates, using Gaussian distributions with standard deviations $\sigma^r = \sigma^{Ln} = 0.3$ Å and $\sigma^{HN} = 15°$.

Table 3. χ Tensor Anisotropy Parameters and Uncertainty of χ Tensor Orientation Obtained from One, Two, and Three Selectively Labeled Samples of ϵ186−θ Complexed with Dy^{3+}

	n_{peaks}[a]	$\Delta\chi_{ax}$[b] (10^{-32} m^3)	$\Delta\chi_{rh}$[b] (10^{-32} m^3)	τ_r[b] (ns)	$\Delta\Omega_z$[c] (deg)	$\Delta\Omega$[d] (deg)
^{15}N-Leu	8	4.6 ± 0.2	1.2 ± 0.6	11.1 ± 2.5	1.8	25.3
^{15}N-Leu,Phe	14	4.5 ± 0.1	0.6 ± 0.1	16.4 ± 3.3	0.7	11.2
^{15}N-Leu,Phe,Val	20	4.5 ± 0.1	0.5 ± 0.1	17.2 ± 2.6	0.6	7.3

[a] Total number of paramagnetic cross-peaks assigned and used for the fit of the χ tensor anisotropy. [b] Error ranges determined by running the assignment algorithm 1000 times after addition of structural noise to the crystal structure coordinates ($\sigma^r = \sigma^{Ln} = 0.3$ Å and $\sigma^{HN} = 15°$). The rotational correlation times were determined from a fit of only CCR data, as PRE values showed large discrepancies between experimental and back-calculated data (Table 1). [c] rmsd of the *z*-axis of the χ tensor relative to the mean orientation after addition of structural noise. [d] Sum of the rmsd values for the three χ tensor axes relative to the respective mean orientations.

of assessing the precision of the fitted χ tensor anisotropy (Table 3). As expected, availability of more than one selectively ^{15}N-labeled sample enhanced the reproducibility of the fitted χ tensor anisotropy. In particular, availability of the ^{15}N-Val-labeled sample reduced the uncertainty in the orientation of the χ tensor axes (Table 3), although it contributed little to the reliability of the assignments of the ^{15}N-Leu and ^{15}N-Phe cross-peaks (Table 2). A graphical view of the orientation of the χ tensor axes shows that the main axis of the tensor is well defined even when larger structural uncertainties are assumed, while the *x*- and *y*-components are more variable (Figure 3).

An independent test of the reliability of the assignment procedure was performed on published data of CN−myoglobin.[15d,16a] This complex contains low-spin Fe^{3+} with a highly anisotropic χ tensor resulting in large PCS and appreciable RDC values. All the correlation signals that are present in the ^{15}N HSQC spectrum of diamagnetic CO−myoglobin can be observed in paramagnetic CN−myoglobin. Using PCS and RDC values from the amide resonances of Ala, Asp, Gln, and Glu, respectively, the assignment algorithm provided correct assignments in all cases, even when only single-residue types were considered. PRE and CCR data were disregarded for most of the resonances as PRE values in CN−myoglobin are small, and CCR data are compromised by second-order effects due to the contribution of magnetic anisotropy to the cross-correlation process.[15d] The protein contains 17 Ala, 14 Glu, 7 Asp, and 5 Gln residues. It is remarkable that data from as few as five residues, as in the case of Gln, could yield correct assignments.

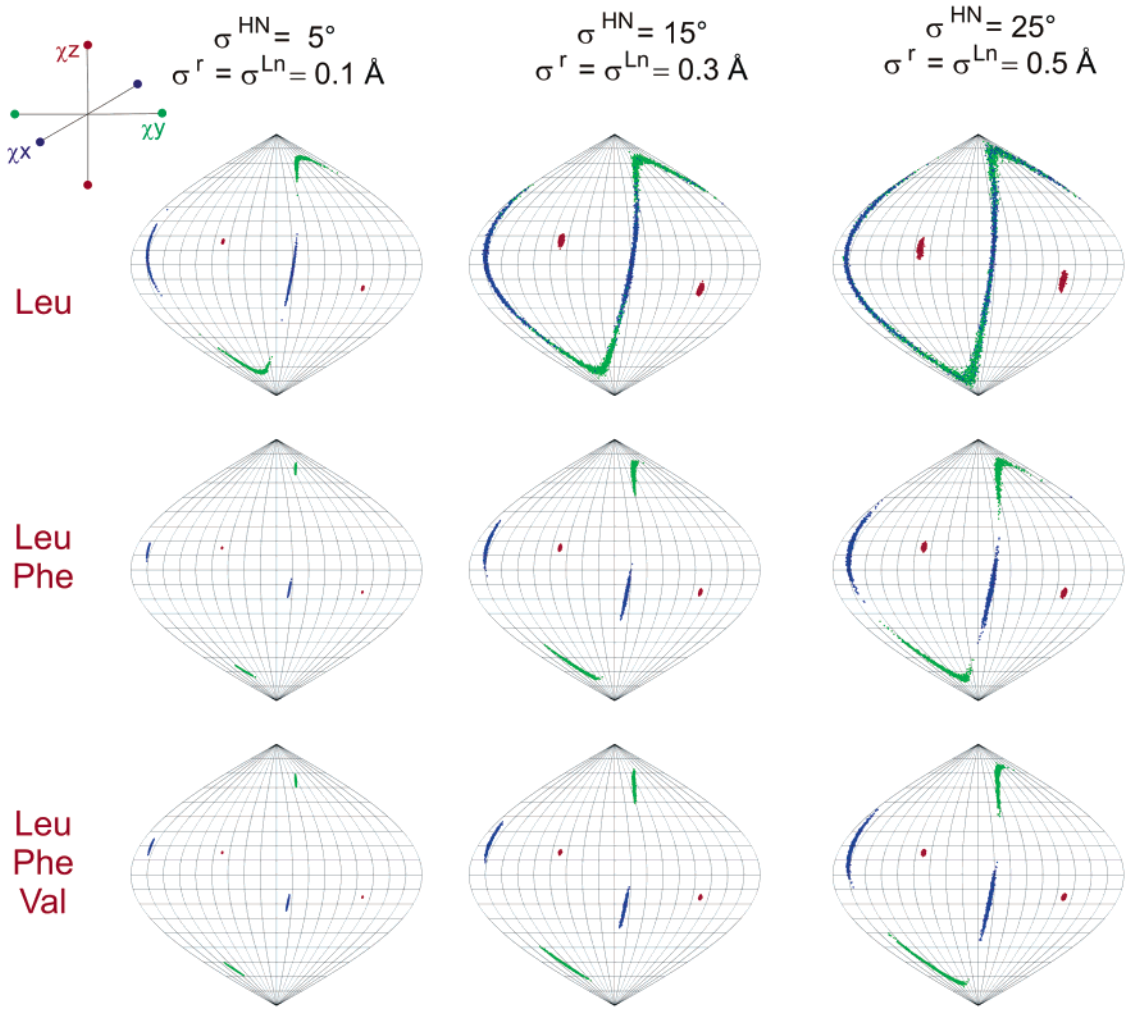

Figure 3. Orientation of the principal axes of the electronic susceptibility tensor χ of Dy^{3+} in the $\epsilon 186-\theta$ complex obtained after fitting PCS and RDC values of the ^{15}N-Leu labeled complex alone (top panel), the ^{15}N-Leu and the ^{15}N-Phe labeled complexes combined (middle row), and the ^{15}N-Leu, ^{15}N-Phe, and ^{15}N-Val labeled complexes combined (bottom panel). The orientations of the tensor axes are visualized with respect to the reference system of the crystal coordinates (PDB code 1J53) using Saunson−Flaumsteed projections[18] (*z*-axis: red; *y*-axis: green; *x*-axis: blue). Increasing levels of structural noise ($\sigma^r = \sigma^{Ln}$, and σ^{HN} as described in the main text) added to the crystal structure illustrate the sensitivity of the tensor parameters with respect to structural deviations between crystal and solution structures. The data in this figure were calculated with the assignments fixed to those of Table 1 to accelerate the computations. The plots were generated with Mathematica.

For comparison, resonance assignment of $\epsilon 186-\theta$ was also performed using exclusively PCS and RDC data. While omission of the PRE and CCR data significantly decreased the robustness of the assignments, with frequencies of less than 0.4 for some of the residue assignments, the preferred assignment was still the same as for the full data set (Supporting Information).

Completeness of Assignments. Using Dy^{3+} as the paramagnetic metal ion precludes the assignment of those ^{15}N HSQC cross-peaks which are broadened beyond detection. Additional cross-peaks can, however, be assigned by the use of different lanthanides. For example, the complex with Ce^{3+} has a χ tensor with 10-fold smaller anisotropy and less paramagnetic relaxation enhancement than the complex with Dy^{3+}, due to the smaller magnetic moment ($J = {}^5/_2$, $g_J = {}^3/_4$).[12,17] All 17 ^{15}N HSQC cross-peaks could thus be observed in the ^{15}N-Leu labeled $\epsilon 186-\theta$ complex in the presence of Ce^{3+}. The eight cross-peaks with large PCS values in the presence of Dy^{3+} were barely shifted in the presence of Ce^{3+}, while the cross-peaks which are strongly shifted by Ce^{3+} are from amides close to the metal center which did not yield observable cross-peaks in the complex with Dy^{3+} (Supporting Information). The Ce^{3+} complex thus resulted in the unambiguous assignment of six additional cross-peaks of the ^{15}N-Leu labeled sample, by use of the algorithm described above.

Discussion

This assignment strategy yields resonance assignments with selectively labeled samples much faster than conventional methods, as only two-dimensional ^{15}N HSQC spectra need to be recorded and evaluated. The resonance assignments established in the present work agree with previous resonance assignments of the $\epsilon 186-\theta$ complex, as far as they have been published.[7] The earlier resonance assignments of $\epsilon 186$ relied on expensive ^{15}N/^{13}C/^{2}H-labeled samples in addition to selectively labeled samples, higher sample concentrations, and higher temperatures.[7] The sensitivity of the method is illustrated by the fact that adequate ^{15}N HSQC spectra could be recorded overnight with 0.5 mM samples of the 30 kDa complex, using a NMR spectrometer equipped with a conventional probehead operating at room temperature.

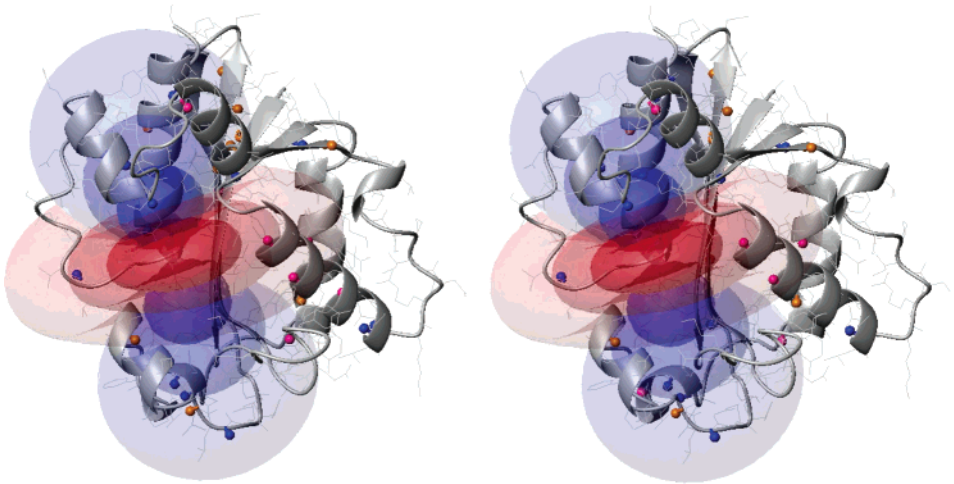

Figure 4. Stereoscopic view of experimentally determined iso-surfaces corresponding to pseudo-contact shifts of ±1, ±0.5, and ±0.2 ppm, respectively, superimposed onto a ribbon representation of the structure of ϵ186 in complex with Dy^{3+}. Positive and negative PCS values are indicated by blue and red colors, respectively. The N−H groups of Leu (blue), Phe (pink), and Val (orange) residues are shown in a ball-and-stick representation. The figure was generated with Molmol.[30]

Selectively ^{15}N-labeled samples are easier to prepare than deuterated samples and are particularly inexpensive to make if the protein can be expressed in a cell-free expression system.[23] While most proteins do not contain a natural lanthanide binding site, different methods are available to attach lanthanide ions site-specifically to diamagnetic proteins.[3g,17,24−26] Lanthanide ions are ideally suited for the present application, as they have similar coordination properties but very different electron spins and χ tensor anisotropies.[17] Furthermore, the large isotropic susceptibilities of lanthanide ions prevent the observation of magnetic anisotropy effects close to the metal ion, so that CCR effects can be measured without interference from second-order effects.[15d] Use of different lanthanides thus provides a straightforward route to resonance assignments of amide groups located at different distances from the paramagnetic center. For lanthanide ions with small paramagnetic effects such as Ce^{3+}, however, the RDC, PRE, and CCR effects are small and may be difficult to measure compared to PCS values. The analysis of the χ tensor in calbindin D_{9k} complexed to different lanthanides showed that the directions (but not the magnitudes) of the principal axes of the χ tensor anisotropies are largely conserved between different lanthanides.[17b] Information on the χ tensor anisotropy axes of a strongly paramagnetic lanthanide can thus be used to support the resonance assignment with weakly paramagnetic lanthanides.

Notably, the present strategy tends to provide more reliable assignments of the paramagnetic spectrum than of the diamagnetic spectrum, where spectral overlap would, in general, be more pronounced. For example, the paramagnetic cross-peaks of Leu57 and Leu176 were shifted so far in the complex with Dy^{3+} (Figure 2) that a choice of several diamagnetic cross-peaks yielded small residuals Q (eq 9). In this situation, the assignment of the diamagnetic cross-peaks can be improved by use of different lanthanides which reduce the magnitude of the pseudocontact shifts.

Our assignment strategy relies on the availability of the coordinates of the three-dimensional protein structure and is, in this sense, related to a recently proposed protocol where many residual dipolar couplings were measured and compared to predictions based on the protein structure to obtain resonance assignments.[20] Compared to the measurement of small residual dipolar couplings, the determination of paramagnetic data from ^{15}N HSQC cross-peaks is much simpler. Furthermore, pseudocontact shifts are less affected by structural variations than residual dipolar couplings. Paramagnetic data thus afford resonance assignments with remarkable reliability even when available data are restricted to a single selectively ^{15}N-labeled sample with a single paramagnetic metal ion. A criterion for the reliability of the assignment of individual cross-peaks is provided by a straightforward Monte Carlo procedure which checks the robustness of the assignment against the addition of structural noise.

In principle, independent verification of the assignments could be obtained by further experiments which rely on the analysis of ^{15}N HSQC spectra of ^{15}N-labeled proteins, including measurements of the relaxation enhancement observed upon addition of an external relaxation probe[27] or of the effects of dipole−dipole cross-correlation on multiplet fine-structure.[28] These experiments, however, tend to be more limited in their information content.

Paramagnetic relaxation enhancements of ligand resonances by proteins tagged with nitroxide radicals have been shown to present a highly sensitive tool to screen for ligands in libraries of chemical compounds.[29] In principle, proteins tagged with Gd^{3+} could be used in the same way. For the purpose of resonance assignments of the protein, however, lanthanide ions generating significant pseudocontact shifts are more suitable.

(23) (a) Kigawa, T.; Yabuki, T.; Yoshida, Y.; Tsutsui, M.; Ito, Y.; Shibata, T.; Yokoyama, S. *FEBS Lett.* **1999**, *442*, 15−19. (b) Guignard, L.; Pursglove, S.; Ozawa, K.; Dixon, N. E.; Otting, G. *FEBS Lett.* **2002**, *524*, 159−162.
(24) Bentrop, D.; Bertini, I.; Cremonini, I.; Forsén, S.; Luchinat, C.; Malmendal, A. *Biochemistry* **1997**, *36*, 11605−11618.
(25) (a) Dvoretsky, A.; Gaponenko, V.; Rosevear, P. R. *FEBS Lett.* **2002**, *528*, 189−192. (b) Griesinger, C. *Book of Abstracts*, 10th Chianti Workshop on Magnetic Resonance, San Miniato (Pisa, Italy), May 25−30, 2003.
(26) (a) Ma, C.; Opella, S. J. *J. Magn. Reson.* **2000**, *146*, 381−384; b) Wöhnert, J.; Franz, K. J.; Nitz, M.; Imperiali, B.; Schwalbe, H. *J. Am. Chem. Soc.* **2003**, *125*, 13338−13339.
(27) Pintacuda, G.; Otting, G. *J. Am. Chem. Soc.* **2002**, *124*, 372−373.
(28) Crowley, P.; Ubbink, M.; Otting, G. *J. Am. Chem. Soc.* **2000**, *122*, 2968−2969.
(29) Jahnke, W.; Rüdisser, S.; Zurini, M. *J. Am. Chem. Soc.* **2001**, *123*, 3149−3150.
(30) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.

Even if only a few ^{15}N HSQC cross-peaks are assigned by using only a single, residue-selectively ^{15}N-labeled protein sample and only a single paramagnetic metal ion, these assignments may already provide sufficient information for the approximate identification of the binding site of a small unlabeled ligand. The anisotropy of the electronic χ tensor determined simultaneously (Figure 4) provides, in addition, long-range restraints for the structural definition of intermolecular interactions,[3] in particular as pseudocontact shifts can be observed for distances up to 40 Å from the metal binding site.[17a]

When determination of the χ tensor anisotropy parameters for studies of intermolecular interactions is the prime objective, the number of selectively labeled samples which must be prepared can be kept to a minimum by a strategic choice of labeled amino acids. The first sample should contain at least five ^{15}N-labeled amides sufficiently far from the metal ion to result in observable cross-peaks. The choice of ^{15}N-labeled amino acids for subsequent sample preparations can then be based on their potential to improve the reliability of the fit for the initially determined χ tensor anisotropy parameters. This choice depends on the number and position of the residues with respect to the χ tensor, whereas the amino acid type is of no importance for the fit.

In conclusion, the assignment strategy described here promises to contribute decisively to modern drug discovery, for the identification of ligand binding sites as well as for the structural characterization of ligand binding.

Acknowledgment. We thank Professor Lorenzo Di Bari for valuable discussions. G.P. thanks the EU for a postdoctoral fellowship within the Research Training Network on Cross-Correlation (HPRN-CT-2000-00092). G.O. thanks the Australian Research Council for a Federation Fellowship. Financial support by the Australian Research Council and access to the NMR facility at the Australian National University is gratefully acknowledged.

Supporting Information Available: ^{15}N HSQC pulse sequence and ^{15}N HSQC spectra of ^{15}N-Val- and ^{15}N-Phe-labeled $\epsilon 186-\theta$ in complex with Dy^{3+} and of the ^{15}N-Leu-$\epsilon 186-\theta$ complex with Ce^{3+}; assignments obtained with the help of PCS and RDC data only; pseudo-code of the algorithm. This material is available free of charge via the Internet at http://pubs.acs.org.

JA039339M