Lanthanide-Induced Pseudocontact Shifts for Solution Structure Refinements of Macromolecules in Shells up to 40 Å from the Metal Ion

Marco Allegrozzi,[†] Ivano Bertini,*[,†] Matthias B. L. Janik,[†] Yong-Min Lee,[†] Gaohua Liu,[†] and Claudio Luchinat[‡]

Contribution from the Department of Chemistry and the Department of Soil Science and Plant Nutrition, University of Florence, Florence, Italy

Received October 14, 1999. Revised Manuscript Received February 2, 2000

Abstract: A number of pseudocontact shifts (PCS) in monolanthanide-substituted Calbindin D_{9k} (Ca_2Cb hereafter), a protein of 75 amino acids, were measured for Ce(III), Yb(III), and Dy(III). The assignment of the shifts was obtained through the conventional assignment procedures for the Ce(III) derivative (CaCeCb), since the line broadening is not severe, whereas in the case of Dy(III) and Yb(III) the assignment was obtained by analyzing the temperature dependence of the $^1H-^{15}N$ HSQC shifts of the lanthanide derivatives and comparing the results with the $^1H-^{15}N$ HSQC spectrum of Ca_2Cb or CaCeCb. The NOE-based solution structures of Ca_2Cb or CaCeCb were then refined with PCS. Since the three lanthanides span a wide range of magnetic anisotropies, the refinement was effective in shells from the metal of ∼5−15 Å for Ce(III), ∼9−25 Å for Yb(III), and ∼13−40 Å for Dy(III), as useful PCS were observed in these shells. The root-mean-square deviation of 30 conformers from the average for CaCeCb was 0.74 and 1.10 Å for the backbone and all heavy atoms, respectively, obtained from 1539 NOEs, 39 ^3J values, and 6 T_1 values. With 589 pseudocontact shifts for Ce(III) (out of which 280 were larger than 0.1 ppm), 92 PCS for Yb(III), and 74 for Dy(III) the RMSD decreased to 0.54 and 0.95 Å for Ce(III), 0.60 and 0.98 Å for Yb(III), and 0.66 and 1.04 Å for Dy(III) for the backbone and all heavy atoms, respectively. While for Ce(III) resolution improvements are mainly found for the metal binding site itself, Yb(III) and Dy(III) can further constrain regions far away from the metal. These results show that constructing a lanthanide binding site may be a general and convenient tool to "enlighten" shells at variable distances from the metal itself, and may be used for various purposes including the investigation of biomolecular complexes.

Introduction

Lanthanides are well-known to induce pseudocontact shifts (PCS) as they are magnetically anisotropic.[1−3] The PCS values, δ^{pc}, depend on the paramagnetic susceptibility tensor which can be written as:[4,5]

$$\delta_i^{pc} = \frac{1}{12\pi r_i^3}\left[\Delta\chi_{ax}(3n_i^2 - 1) + \frac{3}{2}\Delta\chi_{rh}(l_i^2 - m_i^2)\right] \quad (1)$$

where $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic anisotropy parameters of the magnetic susceptibility tensor of the metal, r_i is the distance between atom i and the metal ion, and l_i, m_i, and n_i are the direction cosines of the position vector of atom i ($\boldsymbol{r}_i$) with respect to the principal axes of the magnetic susceptibility tensor centered on the metal. This compact form of the equation for PCS is the one used in our minimization algorithms,[5−7] and is equivalent to the more standard form reported in textbooks.[3,8,9] On the basis of eq 1, attempts are known in the literature which aim at structurally characterizing lanthanide-containing molecules.[10−13] Protocols recently have been proposed to use δ^{pc} to produce a solution structure together with other structural constraints[3,6,14,15] such as NOE's, ^3J couplings, and residual dipolar couplings (rdc)[16] and to refine the structure through molecular dynamics simulations.[7] In this way the solution

* Address correspondence to this author.
† Department of Chemistry.
‡ Department of Soil Science and Plant Nutrition.

(1) Bleaney, B. *J. Magn. Reson.* **1972**, *8*, 91−100.
(2) Kurland, R. J.; McGarvey, B. R. *J. Magn. Reson.* **1970**, *2*, 286−301.
(3) Bertini, I.; Luchinat, C. *NMR of paramagnetic substances*; Coord. Chem. Rev. 150; Elsevier: Amsterdam, 1996; pp 1−300.
(4) McConnell, H. M.; Robertson, R. E. *J. Chem. Phys.* **1958**, *29*, 1361−1365.
(5) Banci, L.; Bertini, I.; Bren, K. L.; Cremonini, M. A.; Gray, H. B.; Luchinat, C.; Turano, P. *J. Biol. Inorg. Chem.* **1996**, *1*, 117−126.
(6) Banci, L.; Bertini, I.; Cremonini, M. A.; Gori Savellini, G.; Luchinat, C.; Wüthrich, K.; Güntert, P. *J. Biomol. NMR* **1998**, *12*, 553−557.
(7) Banci, L.; Bertini, I.; Gori Savellini, G.; Romagnoli, A.; Turano, P.; Cremonini, M. A.; Luchinat, C.; Gray, H. B. *Proteins Struct. Funct. Genet.* **1997**, *29*, 68−76.
(8) La Mar, G. N.; Horrocks, W. DeW, Jr.; Holm, R. H. *NMR of Paramagnetic Macromolecules*; Academic Press: New York, 1973.
(9) Bertini, I.; Luchinat, C. *NMR of paramagnetic molecules in biological systems*; Benjamin/Cummings: Menlo Park, CA, 1986.
(10) Barry, C. D.; Glasel, J. A.; Williams, R. J. P.; Xavier, A. V. *J. Mol. Biol.* **1974**, *84*, 471−490.
(11) Barry, C. D.; Martin, D. R.; Williams, R. J. P.; Xavier, A. V. *J. Mol. Biol.* **1974**, *84*, 491−502.
(12) Reilley, C. N.; Wood, B. W.; Desreux, J. F. *Anal. Chem.* **1975**, *47*, 2110−2116.
(13) Gochin, M.; Roder, H. *Protein Sci.* **1995**, *4*, 296−305.
(14) Gochin, M. *J. Am. Chem. Soc.* **1997**, *119*, 3377−3378.
(15) Turner, D. L.; Brennan, L.; Chamberlin, S. G.; Louro, R. O.; Xavier, A. V. *Eur. Biophys. J.* **1998**, *27*, 367−375.
(16) Banci, L.; Bertini, I.; Huber, J. G.; Luchinat, C.; Rosato, A. *J. Am. Chem. Soc.* **1998**, *120*, 12903−12909.

10.1021/ja993691b CCC: $19.00
Published on Web 04/18/2000

structure is fully consistent with the δ^{pc}. Care has been taken in neglecting nuclei which could experience contact shifts together with δ^{pc}, and in minimizing possible ligand-centered effects.[17] δ^{pc} allow the metal ion to be located within the protein frame without any assumptions. These protocols have been applied so far for low-spin Fe^{3+} hemes[18−22] and Ce^{3+}-containing proteins.[23]

We would like to explore the possibility of using lanthanides with larger magnetic anisotropy, such as Yb^{3+} and Dy^{3+}, for instance as substitutes in calcium-binding proteins. Such metal ions have $S = {}^1/_2$ and ${}^5/_2$ and $J = {}^{15}/_2$ and ${}^7/_2$, respectively, and broaden the proton signals of nuclei close to them beyond detection limits, so that no structural constraints could be obtained. However, it is a resonable assumption that a calcium-binding protein has the same molecular structure upon replacement of Ca^{2+} by lanthanides. If minor changes occur, they occur in the coordination sphere. Therefore, we propose here to use the δ^{pc} of some paramagnetic lanthanides to refine shells at variable distances from the metal ion for a protein whose structure is obtained from NOEs. We are going to present a simple procedure to structurally refine shells ∼5−15 Å from the metal for Ce^{3+}, ∼9−25 Å from the metal for Yb^{3+}, and ∼13−40 Å from the metal for Dy^{3+}. The procedure can be applied to any macromolecule for which a lanthanide binding site is created either chemically or through molecular engineering, or both. One long-term goal is that of using PCS to enlighten the interface of interacting biomolecules where NOEs may be few. The method is tested on the 75 amino acid Ca^{2+} binding protein Calbindin D_{9k}. The protein contains two EF-hand-type Ca^{2+} binding sites, having different binding constants.[24] While the Ca^{2+} ion of site II (the archetypal EF C terminal site)[25] could be easily exchanged with other metal ions, site I remains unchanged.[26,27] The systems investigated are CaCeCb, CaYbCb, and CaDyCb.

Materials and Methods

Sample Preparation. Protein expression[28] and purification[29] of both the Ca^{2+} and the apo form of Bovine Calbindin D_{9k} were performed as reported. A Pro43→Met43 (P43M) mutant representing the minor A form was prepared avoiding any conformational heterogeneity due to *cis*−*trans* isomerization as found for the wild-type protein.[30,31] The expression system was a generous gift of Prof. S. Forsén. Uniformly ^{15}N-labeled P43M was obtained from M9 minimal medium containing $^{15}NH_4Cl$ as the sole nitrogen source. NMR samples were prepared by dissolving the lyophilized protein in 550 μL of either 100% D_2O or 90% H_2O/10% D_2O to finally obtain 2.0 mM protein solutions. The pH was adjusted to 6.0 by means of 0.1 M NaOH or 0.1 M HCl. Lanthanide-containing Calbindin D_{9k} samples were obtained by titrating either the Ca^{2+} or the apo form of P43M with 0.1 M solutions of analytical grade $CeCl_3$ (prepared by dissolving Ce_2O_3 in concentrated HCl) and $YbCl_3$ and $DyCl_3$ (Fluka). Titrations were followed by 1D ^1H NMR and by 2D $^1H-^{15}N$ HSQC spectroscopy for ^{15}N-labeled samples. The samples were kept at 4 °C between measurements.

NMR Spectroscopy. NMR spectra were acquired on Bruker AVANCE 800, AVANCE 600, and DRX 500 spectrometers operating at 800.13, 600.13, and 500.13 MHz, respectively. The residual water signal was suppressed by either presaturation during both the relaxation delay and mixing time or by gradient tailored excitation (WATERGATE).[32] The spectra were calibrated at different temperatures according to the empirical relationship[33] $\delta_{HOD} = (-0.012T + 5.11)$ ppm, with T being the temperature (in °C). TPPI NOESY,[34,35] TOCSY[36] and DQF COSY[37] spectra were recorded at 300 K with mixing times of 60−120 ms for the NOESY and a spin lock time of 60 ms for the TOCSY spectra, respectively. The spectral width was set to 14.0 ppm and the recycle delay to 800 ms. Spectra consisted of 4K data points in the f_2 and 512 increments in the f_1 dimension with 64 transients each. $^1H-^{15}N$ HSQC[38] spectra were recorded for all three lanthanide derivatives at 280, 300, and 310 K using a spectral width of 14 and 32 ppm in the ^1H and ^{15}N dimension, respectively. 256 increments each with 1024 complex data points and 32 transients were collected. 2D spectra for the detection of connectivities between paramagnetically shifted resonances of the Ce^{3+} derivative of P43M were collected using a spectral window of 45.0 ppm, with acquisition times of the order of 150 ms and short relaxation delays to allow for the accumulation of a high number of scans. A 30 ms TOCSY and a 40 ms NOESY spectrum with a total of 384 increments in f_1 were also recorded. Raw data were multiplied in both dimensions by a squared cosine window function and Fourier transformed to obtain a final matrix of 1024 × 1024, or of 512 × 512 real data points for HSQC spectra. A polynomial baseline correction was applied in the f_2 dimension. 1D NOE difference spectra on pseudocontact-shifted resonances were performed as previous reported.[39,40] Nonselective T_1 relaxation times were obtained by the inversion recovery method[41] and an exponential three-parameter fit of the peak intensities versus the delay time. All NMR data were processed with the Bruker XWINNMR software packages. The program XEASY (ETH Zürich) was used for analysis of the 2D spectra.[42]

Structure Calculation. The 60 ms NOESY spectrum cross-peak intensities were integrated using the elliptical integration routine implemented in XEASY and converted into interatomic upper distance limits by the program CALIBA.[43] The program DYANA[44] employing torsion angle dynamics (TAD) combined with a simulated annealing

(17) Banci, L.; Bertini, I.; Luchinat, C.; Pierattelli, R.; Shokhirev, N. V.; Walker, F. A. *J. Am. Chem. Soc.* **1998**, *120*, 8472−8479.
(18) Banci, L.; Bertini, I.; Bren, K. L.; Gray, H. B.; Sompornpisut, P.; Turano, P. *Biochemistry* **1997**, *36*, 8992−9001.
(19) Banci, L.; Bertini, I.; Gray, H. B.; Luchinat, C.; Reddig, T.; Rosato, A.; Turano, P. *Biochemistry* **1997**, *36*, 9867−9877.
(20) Arnesano, F.; Banci, L.; Bertini, I.; Felli, I. C. *Biochemistry* **1998**, *37*, 173−184.
(21) Arnesano, F.; Banci, L.; Bertini, I.; Felli, I. C.; Koulougliotis, D. *Eur. J. Biochem.* **1999**, *260*, 347−354.
(22) Assfalg, M.; Banci, L.; Bertini, I.; Bruschi, M.; Turano, P. *Eur. J. Biochem.* **1998**, *256*, 261−270.
(23) Bentrop, D.; Bertini, I.; Cremonini, M. A.; Forsén, S.; Luchinat, C.; Malmendal, A. *Biochemistry* **1997**, *36*, 11605−11618.
(24) Linse, S.; Brodin, P.; Drakenberg, T.; Thulin, E.; Sellers, P.; Elmden, K.; Grundstrom, T.; Forsén, S. *Biochemistry* **1987**, *26*, 6723−6735.
(25) Kretsinger, R. H. *CRC Crit. Rev. Biochem.* **1980**, *8*, 119−174.
(26) Vogel, H. J.; Drakenberg, T.; Forsén, S.; O'Neil, J. D.; Hofmann, T. *Biochemistry* **1985**, *24*, 3870−3876.
(27) Akke, M.; Forsén, S.; Chazin, W. J. *J. Mol. Biol.* **1991**, *220*, 173−189.
(28) Brodin, P.; Grundstrom, T.; Hofmann, T.; Drakenberg, T.; Thulin, E.; Forsén, S. *Biochemistry* **1986**, *25*, 5371−5377.
(29) Johansson, C.; Brodin, P.; Grundstrom, T.; Thulin, E.; Forsén, S.; Drakenberg, T. *Eur. J. Biochem.* **1990**, *187*, 455−460.
(30) Malmendal, A.; Carlström, G.; Hambraeus, C.; Drakenberg, T.; Forsén, S.; Akke, M. *Biochemistry* **1998**, *37*, 2586−2595.
(31) Chazin, W. J.; Kördel, J.; Drakenberg, T.; Thulin, E.; Brodin, P.; Grundstrom, T.; Forsén, S. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 2195−2198.
(32) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661−666.
(33) Bertini, I.; Ciurli, S.; Dikiy, A.; Luchinat, C. *J. Am. Chem. Soc.* **1993**, *115*, 12020−12028.
(34) Macura, S.; Wüthrich, K.; Ernst, R. R. *J. Magn. Reson.* **1982**, *47*, 351−357.
(35) Marion, D.; Wüthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967−974.
(36) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355−360.
(37) Derome, A.; Williamson, M. *J. Magn. Reson.* **1990**, *88*, 177−185.
(38) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185−188.
(39) Johnson, R. D.; Ramaprasad, S.; La Mar, G. N. *J. Am. Chem. Soc.* **1983**, *105*, 7205−7206.
(40) Banci, L.; Bertini, I.; Luchinat, C.; Piccioli, M.; Scozzafava, A.; Turano, P. *Inorg. Chem.* **1989**, *28*, 4650−4656.
(41) Vold, R. L.; Waugh, J. S.; Klein, M. P.; Phelps, D. E. *J. Chem. Phys.* **1968**, *48*, 3831−3832.
(42) Eccles, C.; Güntert, P.; Billeter, M.; Wüthrich, K. *J. Biomol. NMR* **1991**, *1*, 111−130.
(43) Güntert, P.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1991**, *217*, 517−530.
(44) Güntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, *273*, 283−298.

algorithm was used to calculate a family of 400 structures starting from randomly generated conformers in 10 000 annealing steps. The quality of structures calculated by DYANA can be assessed by a properly defined function (target function) proportional to the squared deviations of the calculated constraints from the experimental ones, plus the squared van der Waals contact violations. This family was subsequently minimized with respect to the pseudocontact shift constraints obtained for the individual lanthanide ions by a conjugate gradient minimization of the PSEUDYANA[6] module of the program. With preliminary DYANA structures available, the scaling factors for the volume-to-distance conversion for each class of distance constraints were evaluated by plotting volumes of peaks arising from pairs of protons at a fixed distance. As the structure emerged from successive runs of DYANA calculations, selected interatomic distances were taken from these structures and additional calibrations were determined. According to the above protocol, several cycles of the structure calculations were carried out to recalibrate the NOE distance constraints. At this stage, new NOESY cross-peaks were assigned by back-calculating NOEs using the program CORMA.[45] A single isotropic correlation time of 4.1 ns was used.[46] Stereospecific assignments of diastereotopic protons and methyl groups were obtained using the program GLOMSA.[43] Additional proton−proton and proton−metal upper distance limits were determined from NOE difference spectra and nonselective T_1 values of the pseudocontact-shifted signals, as previously reported.[3] No constraints were introduced for the first metal binding loop, while loose lower and upper distance limits of 1.15 and 4.50 Å were imposed for the metal−donor distances in the lanthanide binding site according to the X-ray structure.[47] Pseudocontact shifts were obtained by subtracting the chemical shifts of the diamagnetic (Ca^{2+}) form of Calbindin D_{9k}[48] for the individual metal derivative and subsequently used as structural constraints together with the NOE, T_1, and ^3J value-derived constraints throughout the minimization. Magnetic susceptibility tensors were obtained from preliminary DYANA structures using the program FANTASIA[5,6] and recalculated iteratively during the structure refinement until convergence, which was achieved within six cycles for CaCeCb, eight cycles for CaYbCb, and six cycles for CaDyCb. A fixed tolerance of 0.2 ppm for NH and 0.1 ppm for nonexchangeable protons, plus a proportional tolerance of 10% for the HSQC derived PCS and 30% for the other PCS, was used for the pseudocontact shift constraints. The 30 structures with the lowest target functions were included in the final families and root-mean-square deviations from the mean for the backbone (bb) and all heavy atoms (all) (which include carbon, nitrogen, oxygen, and sulfur atoms from the side chains) were calculated. The programs MOLMOL[49] and PROCHECK[50] were subsequently used for secondary structure analysis.

Results

Solution Structure Determination of CaCeCb. Both assignment[48] and solution structure[51] for the diamagnetic (Ca^{2+}) form of Bovine Calbindin D_{9k} are available in the literature. In the present study this assignment was extended to the Ca^{2+}−Ce^{3+} loaded protein by use of scalar and dipolar spectroscopies. 95.8% of the protons and 73.4% of the nitrogens were assigned, which compare well with the 95.6% and 100% of the diamagnetic Ca_2Bp analysis.[52] The ^{1}H NMR spectrum of CaCeCb (see Supporting Information) displays several hyperfine shifted signals both upfield and downfield in a range from 30 to −15 ppm. By use of 1D NOE measurements these far-shifted resonances were all assigned to either NH or nonexchangeable protons of the residues 54 to 61, confirming that the Ce^{3+} ion is exclusively bound to the archetypal EF C-terminal site (site II) of the protein. The full assignment, together with the ^3J values and the δ^{pc} used, is available as Supporting Information. A total of 2150 NOESY cross-peaks were integrated, from which 1539 were found to be meaningful. By comparing the spectra of the Ca_2Cb and CaCeCb derivatives 589 δ^{pc} were determined, including 60 ^{15}N and 60 ^{1}H δ^{pc} derived from HSQC spectra (Figure 1A) (see later) out of which 280 were larger than 0.1 ppm. An additional 42 distance constraints were calculated from the 1D-NOE and T_1 measurements. The constraints used and their distribution along the sequence are summarized in Figure 2A−C. The final structure consists of 30 conformers with RMSD values from the mean of 0.52 ± 0.11 and 0.90 ± 0.08 Å for the backbone and all heavy atoms, respectively. These figures should be compared with the values of 0.98 and 1.29 Å for the backbone and all heavy atoms reported for the Ca_2Cb structure.[51]

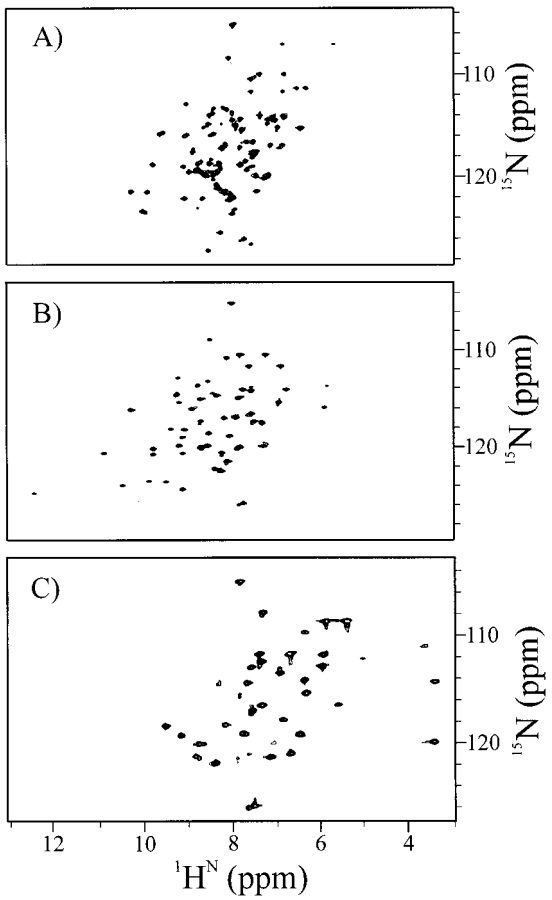

Figure 1. 500 MHz, 300 K ^{1}H−^{15}N HSQC spectra of CaCeCb (A), CaYbCb (B) and CaDyCb (C) at pH 6.0. The total number of observable cross-peaks, out of a total of 72 expected cross-peaks, is 65 in spectrum A, 57 in spectrum B, and 43 in spectrum C. The larger number of downfield shifted peaks in spectrum B and upfield shifted peaks in spectrum C also can be appreciated.

Refinement by Using Other Lanthanides. Two more lanthanide-containing Calbindin D_{9k} derivatives were prepared by titrating the protein with either a 0.1 M Yb^{3+} or Dy^{3+} solution. The 1D spectra of the three lanthanide derivatives are given in the Supporting Information (Figure S1) for comparison.

(45) Borgias, B.; Thomas, P. D.; James, T. L. *Complete Relaxation Matrix Analysis (CORMA)*; University of California, San Francisco, 1989.
(46) Akke, M.; Skelton, N. J.; Kördel, J.; Palmer, A. G., III; Chazin, W. J. *Biochemistry* **1993**, *32*, 9832−9844.
(47) Szebenyi, D. M. E.; Moffat, K. J. *J. Biol. Chem.* **1986**, *261*, 8761−8777.
(48) Kördel, J.; Forsén, S.; Chazin, W. J. *Biochemistry* **1989**, *28*, 7065−7074.
(49) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graphics* **1996**, *14*, 51−55.
(50) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.
(51) Kördel, J.; Skelton, N. J.; Akke, M.; Chazin, W. J. *J. Mol. Biol.* **1993**, *231*, 711−734.
(52) Skelton, N. J.; Akke, M.; Kördel, J.; Thulin, E.; Forsén, S.; Chazin, W. J. *FEBS Lett.* **1992**, *303*, 136−140.

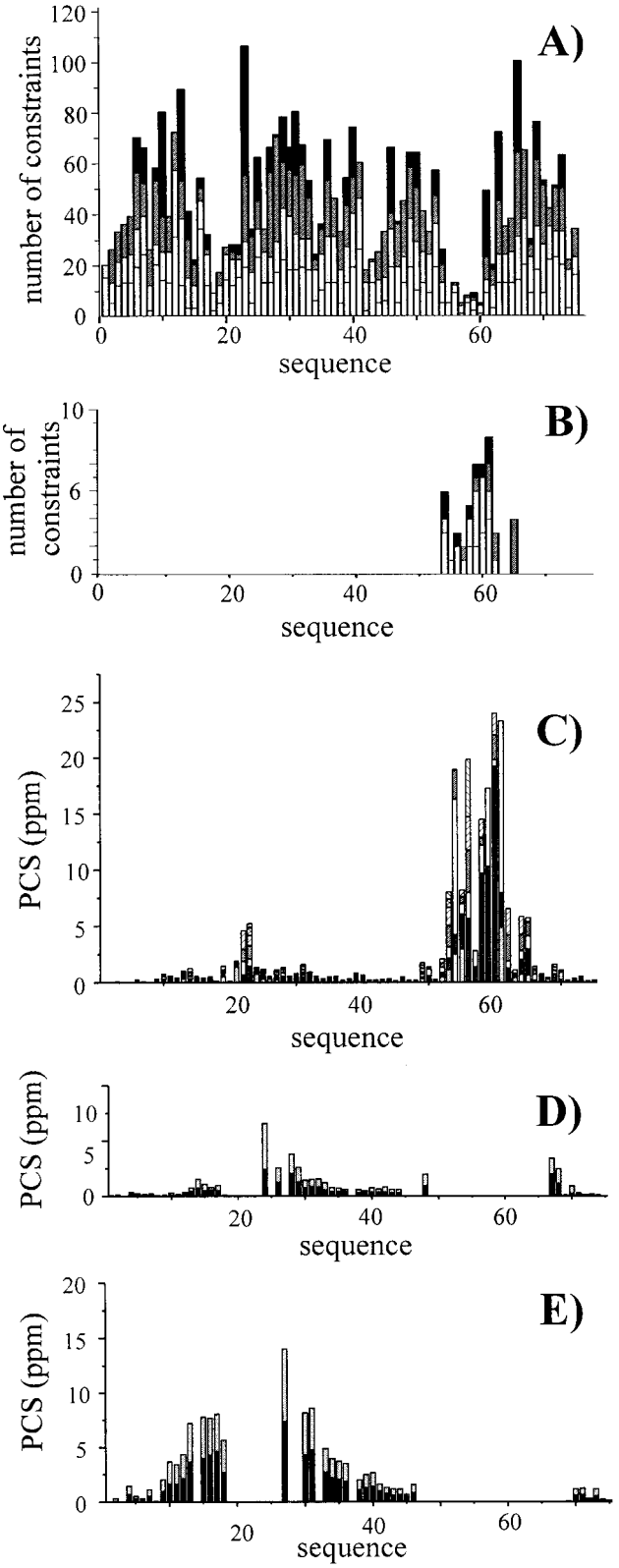

Figure 2. Number of 2D NOE (A), 1D NOE (B), and PCS per residue of CaCeCb (C), CaYbCb (D), and CaDyCb (E) used in the structure refinement. In panels A and B the shades of gray from white to black indicate intraresidue, sequencial, mid-range, and long-range NOEs, respectively. In panels C−E, each section of the bar indicates a different PCS for that residue, and has a length proportional to the absolute value of the PCS. For panels D and E only HSQC-derived constraints (two per residue) were used (see text).

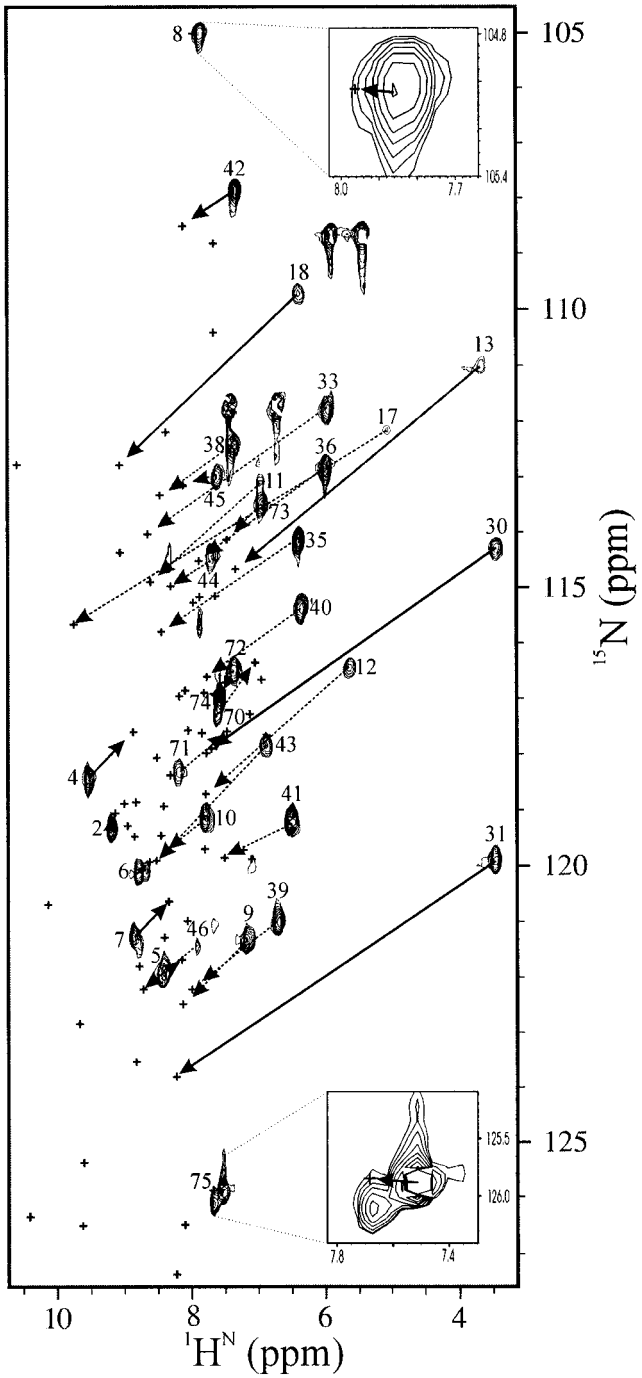

Figure 3. ^{1}H−^{15}N HSQC spectrum of CaDyCb superimposed to the position of the peaks (+) in the Ca_2Cb derivative. The arrows indicate the correspondence with the Ca_2Cb peaks. The solid arrows show the assignments that could be initially made by analysis of the temperature dependence.

A 1:1 (Ln^{3+}:Ca^{2+}) ratio was easily achieved by monitoring the ^{1}H−^{15}N HSQC cross-peaks of ^{15}N-labeled samples, as cross-peaks of the diamagnetic protein disappeared, while new signals of the paramagnetic form were observed. For Dy^{3+} most cross-peaks were shifted upfield, as for Ce^{3+}, while for Yb^{3+} most were shifted downfield (Figure 1B,C). This reversal behavior for Yb^{3+} matches theoretical predictions.[1] Instead of proceeding through the whole assignment of these two derivatives using all kinds of available data, such as analyzing 3D NMR spectra, we propose here a shortcut mainly based on ^{1}H−^{15}N HSQC measurements and on the temperature dependence of the pseudocontact-shifted resonances.

Dy^{3+} has a very large magnetic moment and magnetic anisotropy. Only 43 HSQC cross-peaks are observed (Figure

Table 1. Constraints Used in the Calculations and Final Target Functions Together with the Root-Mean-Square Deviations from the Mean for Lanthanide-Substituted Calbindin D_{9k} with the Results Obtained for CaCeCb without PCS Shown for Comparison

	general constraints	PCS	target function[a] $(Å)^2$	RMSD bb (Å)[a,b]	RMSD all (Å)[a,b]
CaCeCb	1539 NOEs, 6 T_1, 39 ^3J	-	0.10	0.74 ± 0.13	1.10 ± 0.10
CaCeCb	1539 NOEs, 6 T_1, 39 ^3J	589	0.40	0.52 ± 0.11	0.90 ± 0.08
CaYbCb	1539 NOEs, 6 T_1, 39 ^3J	114	0.28	0.61 ± 0.13	0.96 ± 0.16
CaDyCb	1539 NOEs, 6 T_1, 39 ^3J	86	0.10	0.65 ± 0.11	1.00 ± 0.11

[a] For 30 structures. [b] Residues 4 to 74.

1C), most of which are sizably shifted with respect to the Ca^{2+} position. However, pseudocontact shifts are strongly temperature dependent, and different shifts can be observed by recording HSQC spectra at different temperatures. With increasing temperature the pseudocontact shifts decrease, as expected on theoretical grounds (Curie like behavior).[3] From the direction and extent of the movement of the cross-peaks in the $^1H-^{15}N$ plane, it was easy in several cases to visually extrapolate to a plausible peak in the diamagnetic Ca^{2+} spectrum (Figure 3). Based on some unambiguously assigned peaks an initial magnetic anisotropy tensor could be calculated through eq 1 and the program FANTASIA.[5,6] With this tensor, predictions could be made for the other more shifted peaks, which could then be identified in the spectrum and included in a new FANTASIA calculation. With this iterative protocol very quickly a full assignment of all visible HSQC peaks was obtained. In principle, from the tensor and the known structure of Ca_2Cb the ^1H signals shifted far away could also be assigned. This was not done here, because the whole assignment would need to be checked, through either a NOESY spectrum with short mixing time or 1D NOE experiments, which were not performed.

The assignment of the HSQC spectrum of the CaCeCb derivative (Figure 1A) was easy to achieve by visual inspection only, as many of the HSQC cross-peaks were almost coincident with those observed for the Ca^{2+} sample. The not perfectly coincident peaks could then be used to measure small pseudocontact shifts (for both ^1H and ^{15}N) and to generate an initial magnetic anisotropy tensor with the program FANTASIA. For the CaYbCb sample (Figure 1B) a mixed approach was used, by starting with some almost coincident HSQC cross-peaks and some assignments obtained through the temperature dependence.

Chemical shifts of all HSQC resonances at the three temperatures are included in the Supporting Information. Only a few exceptions to the expected Curie behavior were found in the case of Dy^{3+} and Yb^{3+}, namely for residues lysine 71 and glutamine 75. Their anomalous temperature dependence might introduce uncertainty in the correct PCS values to use as constraints. Therefore, the corresponding shifts were not used. In total, 114 δ^{pc} for Yb^{3+} and 86 δ^{pc} for Dy^{3+} were obtained by this protocol (Figure 2D,E), which were subsequently included in individual structure calculations together with all structural constraints such as NOEs, ^3J, and T_1 values already used for the Ce^{3+} calculation. For both metals, families of 30 structures were calculated which where then used to optimize the tensor parameters $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ for a new calculation run. Finally, conformers with RMSD values from the average of 0.61 ± 0.13 and 0.96 ± 0.16 Å for Yb^{3+} and of 0.65 ± 0.11 and 1.00 ± 0.11 Å for Dy^{3+} for the backbone and all heavy atoms, respectively, were obtained. All target functions are on the order of 1 or smaller, with no consistent violations (Table 1). The structure consists of four helices and three loops, of which the first and the third constitute the calcium- and lanthanide-binding site.

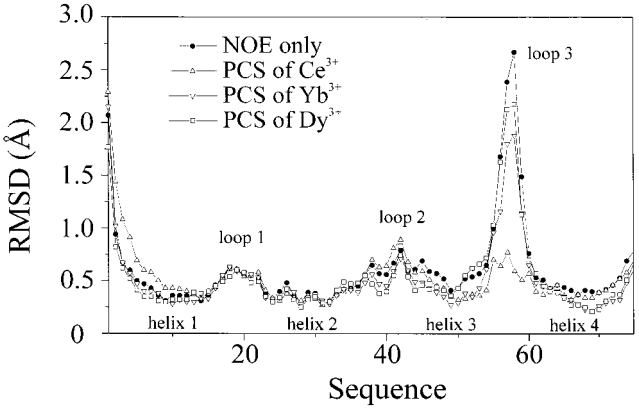

Figure 4. RMSD per residue for CaCeCb without PCS (●) and with PCS from CaCeCb (△), CaYbCb (▽), or CaDyCb (□).

Discussion

Mean structures of all three metal families were calculated and compared to one another. RMSD values of these structures to their mean are 0.56 ± 0.15 and 0.66 ± 0.13 Å for the backbone and all heavy atoms respectively, which are close or below the average RMSD within each family. This means that the sets of pseudocontact shifts for Ce^{3+}, Yb^{3+}, and Dy^{3+} are consistent with each other, and all of them can be used interchangeably to calculate the same final structure.

The RMSD per residue for the three structures calculated using PCS for Ce^{3+}, Yb^{3+}, and Dy^{3+} are shown in Figure 4 in comparison with those obtained in the absence of PCS. The corresponding families are shown in Figure 5A−D. The most stricking improvement is seen in the second metal binding loop, which shows a RMSD peak of above 2.7 Å in the Ce^{3+} structure without PCS, which drops to 0.6 in the presence of PCS. Interestingly, Yb^{3+} and Dy^{3+} show again a peak in this loop, due to the lack of HSQC peaks from the lanthanide binding site. However, Yb^{3+} causes a distinct reduction of the RMSD in helix 3 as well as in the second loop and in helix 4. Moreover, the Ca^{2+} binding site is also improved. Dy^{3+} shows the most important improvements in the regions farthest from the metal, i.e., the N and C terminal residues and in the second loop. It appears that the potential of Dy^{3+} is not fully exploited in this relatively small protein, as measurable PCS would probably be observed at even larger distances from the metal (see later). Figure 5B−D also shows the rather good definition of the position of the metal itself (RMSD to the average of 0.37 Å for Ce^{3+}, 0.48 Å for Yb^{3+}, and 0.44 Å for Dy^{3+}). This is particulary true for Ce^{3+}, as already noted in the case of a Ce^{3+}-substituted fragment of Calmodulin.[23] It should also be mentioned that the mean position of Dy^{3+} is somewhat off with respect to that of Ce^{3+} and Yb^{3+}. This is likely not to reflect a real distortion of the protein. Test calculations performed imposing to Dy^{3+} the same coordinates as those of Ce^{3+} indeed show no appreciable worsening of the agreement between experimental and calculated PCS from Dy^{3+}. It has to be realized that the Dy^{3+}

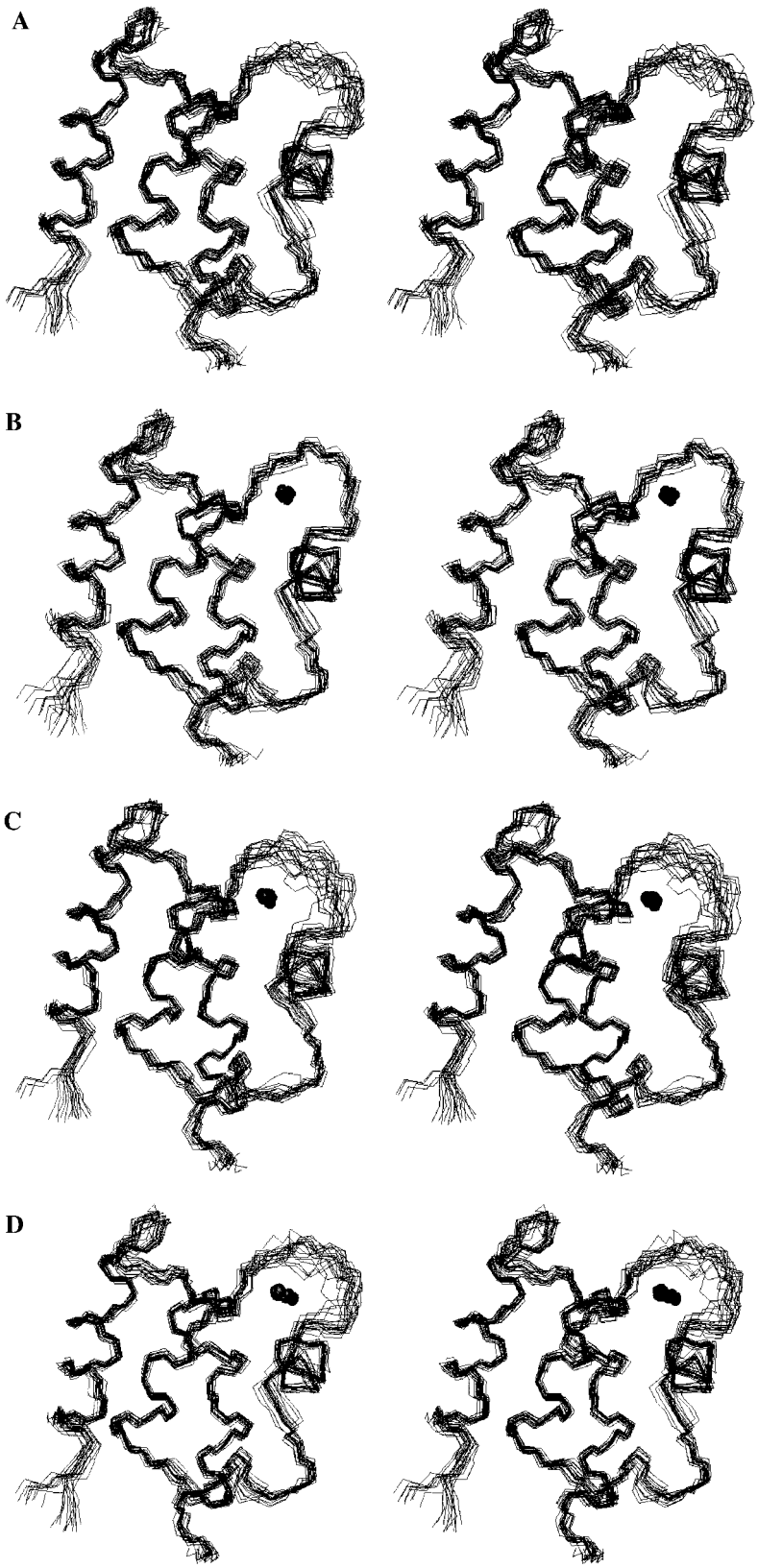

Figure 5. Stereoviews of the families of 30 structures for CaCeCb without PCS (A) and with PCS from CaCeCb (B), CaYbCb (C), or CaDyCb (D).

coordinates are dictated by PCS all arising from distant regions of the protein and, moreover, all from the same side. Again, a larger protein showing a more spherical distrubution of PCS from Dy^{3+} would probably not show this problem.

A good agreement with the X-ray structure of Ca_2Cb[47] is found for the metal-coordinating oxygen atoms at least for the CaCeCb family. Even with loose upper distance limits of 4.5 Å imposed for the lanthanide binding loop, the atoms are still allowed to move and arrange themself, driven by the experimental constraints. However, close metal-to-oxygen distances in the range of 1.7 to 3.0 Å are found for the side chain oxygens of GLU 54, ASN 56, ASP 58, and GLU 65, as well as for the

Table 2. Comparison of the Final Magnetic Susceptibility Tensor Parameters, $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$, for the Three Lanthanide Complexes

	$\Delta\chi_{ax}$ $(10^{-32}\ m^3)^a$	$\Delta\chi_{rh}$ $(10^{-32}\ m^3)^a$		x^b	y^b	z^b
CaCeCb	1.7	0.3	*x*	0	90	90
			y	90	0	90
			z	90	90	0
CaYbCb	8.2	5.8	*x*	14.9	75.4	87.0
			y	84.0	77.8	13.6
			z	76.7	19.1	76.8
CaDyCb	34.9	21.2	*x*	25.7	69.6	76.1
			y	72.5	21.7	77.6
			z	72.5	83.0	18.8

[a] The $\Delta\chi$ values and axes orientations are defined in such a way that the relationships $\chi_z > \chi_x > \chi_y$ and $\Delta\chi_{ax} > \Delta\chi_{rh} > 0$ always hold. [b] Angles in degrees between the *x*, *y*, and *z* axes of the χ tensors for the three metals and a reference system (taken as coincident with the principal axes of the Ce(III) tensor).

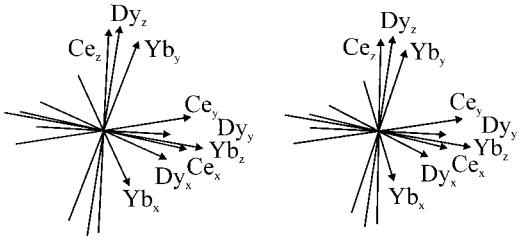

Figure 6. Stereoview of the orientation of the magnetic susceptibility tensors for the three lanthanide derivatives. The two *z* axes for CaCeCb and CaDyCb are close, and close to the *y* axis of CaYbCb.

backbone oxygen of GLU 60, throughout the family, which are indeed the coordinating atoms found in the Ca_2Cb protein.

The parameters characterizing the magnetic susceptibility tensors are summarized in Table 2. Figure 6 shows a comparison of the three tensors. They have similar overall orientation with respect to the protein frame, as is predicted by the localization of the metal ion within the coordination sphere. However, opposite signs are found for the principal axes directions of the tensors, as expected from theory.[1] In particular, the axes of largest magnetic susceptibility for Ce^{3+} and Dy^{3+} (*z*-axes) are very close to one another, and close to the axis of smallest magnetic susceptibility (*y*-axis) for Yb^{3+}. This observation is relevant in the sense that this relationship among magnetic susceptibility tensors for various lanthanides is expected to hold only if the geometry of the coordination polyhedron is maintained, thus confirming the overall validity of the whole approach described here. A more thorough comparative analysis of the magnetic properties of a series of lanthanide-substituted proteins will be the subject of another paper.

A final consideration is in order on the relative magnitudes of the $\Delta\chi_{ax}$ values from Table 2, which qualitatively reflect the overall shifting ability of these three metals. It appears that Yb^{3+} is almost 4 times more effective than Ce^{3+}, and Dy^{3+} about 5 times more effective than Yb^{3+} (20 times more than Ce^{3+}). Indeed, Ce^{3+} and Dy^{3+} are expected to represent the weakest and the strongest paramagnetic lanthanide ions.[1] These considerations, coupled with the line broadening capabilities that are theoretically expected[53−56] and experimentally verified here, are summarized in Figure 7. The figure shows the expected attenuation of proton line widths and PCS with increasing distance from the metal. The solid lines mark the useful ranges

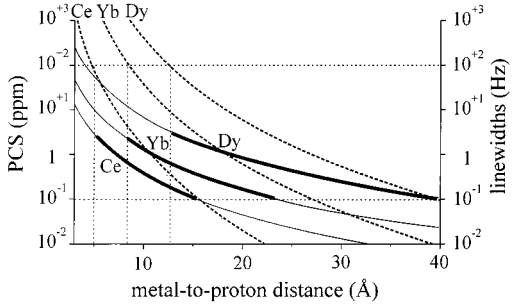

Figure 7. Useful ranges of PCS (solid lines) for Ce(III)-, Yb(III)-, and Dy(III)-containing proteins. Dashed lines indicate observed line widths for the three metals. The lower distance limit is estimated from the observability of HSQC peaks in the present system, and the upper distance limit from the observed PCS being smaller than 0.1 ppm in absolute value. The lower distance limits can be further lowered by including shifts measured from 1D ^{1}H NMR spectra.

for both, as judged from the present data. By taking the lowest metal-to-proton distance for which HSQC cross-peaks could be detected for each derivative and 0.1 ppm as the threshold for the reliability of a measured PCS, we obtain useful ranges of distances covered by PCS for Ce^{3+} of 5−15 Å, for Yb^{3+} of 9−25 Å, and for Dy^{3+} of 13−40 Å. Of course, the range of detectability of ^{1}H line widths in a 1D spectrum is much larger, thereby the lower distance limits will be further lowered by direct detection and assignment of ^{1}H PCS in 1D spectra, as is indeed done here for the Ce^{3+} derivative.

Concluding Remarks

This research has shown that the use of different lanthanides can provide structural constraints in shells at variable distance (up to ∼40 Å) and of different thickness (from 10 to 25 Å). Despite this being known in principle, we provide here an assessment of the improvement in resolution by using PCS and the PSEUDYANA protocol.[6] A simplified strategy has been proposed for the assignment of backbone NH signals affected by paramagnetism, which can easily be expanded, when needed, even to ^{13}C and ^{1}H nuclei of side chains. A next step is that of investigating biomolecular complexes or quaternary structures with one lanthanide binding site. The lanthanide should lower the difficulties in defining the structure of such systems in cases where NOEs at the interface are few due to the large proton−proton distances, whereas PCS can locate the nuclei of interest in the molecular axis frame. An example of the use of PCS from cobalt(II) to structurally characterize a drug−DNA octanucleotide duplex interaction has recently appeared.[57,58] The PCS method thus appears to be quite general and applicable to larger molecules, of course with the same limitations that would apply to the diamagnetic analogues (more peak overlap, line broadening, etc.). An increase in paramagnetic line broadening with molecular weight is also expected from the Curie mechanism, but for larger molecules the percentage of protein protons affected by paramagnetic broadening becomes less and less relevant.

Acknowledgment. Thanks are expressed to Prof. Sture Forsén for providing us with the expression system for Calbindin D_{9k}. Financial support of the EU through contract No. BIO4-98-0156 is gratefully acknowledged. This work also has been supported partly by MURST (40%, 1997) and CNR (contract Nos. 97.01133.49 and 98.01789.CT03).

(53) Alsaadi, B. M.; Rossotti, F. J. C.; Williams, R. J. P. *J. Chem. Soc., Dalton Trans.* **1980**, 2147−2150.
(54) Burns, P. D.; La Mar, G. N. *J. Magn. Reson.* **1982**, *46*, 61−68.
(55) Gueron, M. *J. Magn. Reson.* **1975**, *19*, 58−66.
(56) Vega, A. J.; Fiat, D. *Mol. Phys.* **1976**, *31*, 347−362.
(57) Gueron, M. *J. Biomol. NMR* **1998**, *12*, 243−257.
(58) Tu, K.; Gochin, M. *J. Am. Chem. Soc.* **1999**, *121*, 9276−9285.

Supporting Information Available: 1D spectra of the three lanthanide compounds, tables with proton chemical shifts of the CaCeCb (300 K), and nitrogen and proton chemical shifts of CaCeCb, CaYbCb, and CaDyCb obtained from HSQC spectra at 280, 300, and 310 K and additional tables including PCS, 1D-NOE, and T_1 derived contraints as well as ^3J values for CaCeCb (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA993691B