Article

Efficient χ-tensor determination and NH assignment of paramagnetic proteins

Christophe Schmitz[a,b], Michael John[c], Ah Young Park[c], Nicholas E. Dixon[c], Gottfried Otting[c], Guido Pintacuda[d] & Thomas Huber[a,*]

[a]*Departments of Mathematics and Biochemistry, University of Queensland, Brisbane, QLD, 4072, Australia;* [b]*INRIA Rocquencourt, 78153, Le Chesnay, France;* [c]*Research School of Chemistry, Australian National University, Canberra, ACT, 0200, Australia;* [d]*Ecole Normale Superieure de Lyon, Laboratoire de Chimie, 46 Allee d'Italie, 69364, Lyon, France*

Abstract

Anisotropic magnetic susceptibility tensors χ of paramagnetic metal ions are manifested in pseudocontact shifts, residual dipolar couplings, and other paramagnetic observables that present valuable long-range information for structure determinations of protein-ligand complexes. A program was developed for automatic determination of the χ-tensor anisotropy parameters and amide resonance assignments in proteins labeled with paramagnetic metal ions. The program requires knowledge of the three-dimensional structure of the protein, the backbone resonance assignments of the diamagnetic protein, and a pair of 2D ^{15}N-HSQC or 3D HNCO spectra recorded with and without paramagnetic metal ion. It allows the determination of reliable χ-tensor anisotropy parameters from 2D spectra of uniformly ^{15}N-labeled proteins of fairly high molecular weight. Examples are shown for the 185-residue N-terminal domain of the subunit ε from *E. coli* DNA polymerase III in complex with the subunit θ and La^{3+} in its diamagnetic and Dy^{3+}, Tb^{3+}, and Er^{3+} in its paramagnetic form.

Abbreviations: θ – subunit θ of *E. coli* polymerase III; ε186 – N-terminal 185 residues of the *E. coli* polymerase III subunit ε; PCS – pseudocontact shift; PRE – paramagnetic relaxation enhancement; RACS – residual anisotropic chemical shifts.

Introduction

Nuclear magnetic resonance (NMR) spectroscopy provides a rich source of structural information in solution. Pseudocontact shifts (PCS) induced by paramagnetic metal ions with an anisotropic magnetic susceptibility tensor χ present particularly valuable data as long-range restraints for the conformational analysis of chemical compounds (Peters et al., 1996; Bertini et al., 2001b, 2002), 3D structure refinements of proteins (Banci et al., 1997, 1998; Bertini et al., 2001a; Gaponenko et al., 2004), and the modeling of protein-protein and DNA-ligand complexes (Ubbink et al., 1998; Gochin, 1998, 2000; Gaponenko et al., 2002; Goodfellow et al., 2002; Diaz-Moreno et al., 2005). Besides giving rise to PCS, the anisotropic part of the χ tensor, Δχ, causes weak alignment of the

paramagnetic molecule in the magnetic field, resulting in residual dipolar couplings (RDC; Tolman et al., 1995; Barbieri et al., 2002; Wöhnert et al., 2003; Ikegami et al., 2004; Pintacuda et al., 2004; Leonov et al., 2005) with concomitant long-range structural information (Tjandra and Bax, 1997; van Dijk et al., 2005; Bertini et al., 2002).

PCS and RDC values are most readily measured by the comparison of paramagnetic and diamagnetic NMR spectra recorded of samples prepared in a paramagnetic and diamagnetic state, respectively. This can be achieved by attachment of a paramagnetic or diamagnetic metal ion. Lanthanide ions are a particularly attractive choice, since their chemical properties are very similar while their paramagnetic properties vary widely (Allegrozzi et al., 2000). Lanthanide binding sites on proteins can be generated by derivatization of cysteine side chains with a metal chelator (Gaponenko et al., 2002, 2004) or by fusion with a lanthanide binding peptide sequence (Ma and Opella, 2000; Wöhnert et al., 2003).

As a drawback, measurement of PCS values requires that both the paramagnetic and diamagnetic NMR spectra are assigned. However, the transfer of assignments from the diamagnetic to the paramagnetic spectrum is straightforward only in special situations, where the chemical shift changes are small (Diaz-Moreno et al., 2005; Ubbink et al., 1998). In general, the assignment of a paramagnetic NMR spectrum is a non-trivial task since, besides the usual problems of accidental overlap between NMR signals, many resonances disappear due to paramagnetic relaxation enhancements. The same problems hinder the sequential assignments of paramagnetic proteins by standard NOESY/TOCSY or triple-resonance techniques (Gochin, 1998; Gaponenko et al., 2002; Goodfellow et al., 2002).

The resonance assignment of paramagnetic spectra is much facilitated if isotope-labeled proteins are available, where pairs of diamagnetic and paramagnetic peaks can quite readily be identified based on the similarity of the PCS values of directly bonded ^{15}N and $^1H^N$ or ^{13}C and ^1H spins. In this case, the resonances in ^{15}N- or ^{13}C-HSQC spectra are displaced along lines that are nearly parallel (Baig et al., 2004; Pintacuda et al., 2004). In addition, the temperature dependence of PCS can be exploited to extrapolate from a given paramagnetic cross-peak to the corresponding cross-peak in the diamagnetic spectrum (Allegrozzi et al., 2000; Bertini et al., 2001a). Assuming that the 3D structure of the pro-

tein and the position of the metal ion are known, five PCS values are sufficient to define an initial set of $\Delta\chi$-tensor parameters (axial and rhombic components and Euler angles) that allows the prediction of PCS values for all nuclear spins of the protein based on its 3D structure. Starting from the assignments of the diamagnetic protein, the assignment of further paramagnetic cross-peaks and refinement of the $\Delta\chi$-tensor parameters can be achieved iteratively. This strategy has been applied to small proteins in an interactive manner and has been successful also without measurements of the temperature dependence of PCS (Babini et al., 2004; Baig et al., 2004).

For larger proteins, the determination of the $\Delta\chi$ tensor from 2D ^{15}N-HSQC spectra becomes non-trivial due to limited spectral resolution which makes it difficult to associate five diamagnetic peaks confidently with their paramagnetic partners. First, the PCS values of directly bonded ^{15}N and ^1H spins are not identical because of non-identical locations with respect to the $\Delta\chi$ tensor. Second, anisotropic chemical shifts are only incompletely averaged due to weak alignment of the paramagnetic protein in the magnetic field (Bertini et al., 2002; Cornilescu and Bax, 2000). This effect modifies the observed paramagnetic shift for spins with significant chemical shift anisotropy (John et al., 2005).

The challenge is illustrated by the complex between the N-terminal 185 residues of the *E. coli* DNA polymerase subunit ε, ε186, and the subunit θ (DeRose et al., 2002, 2003; Hamdan et al., 2002). ε186 contains a site for binding of two divalent metal ions that can be filled by a single lanthanide(III) ion (Frey et al., 1996; Brautigam et al., 1999; Pintacuda et al., 2004). The diamagnetic spectrum of ε 186, recorded of the 1:1 complex with La^{3+}, displays 152 backbone NH cross-peaks, whereas only 81 backbone NH cross-peaks were found in the spectrum recorded with Dy^{3+} (Figure 1). The problem is confounded by the presence of mobile polypeptide segments without structural information and incomplete assignments for the diamagnetic protein.

Here we present the program Echidna (efficient χ-tensor determination and NH assignment) that uses the assigned NMR spectrum and the 3D structure of a diamagnetic protein for simultaneous determination of (i) the χ-tensor anisotropy of its paramagnetic counterpart and (ii) many of the assignments of the paramagnetic NMR signals. The program relies on a single pair of diamagnetic
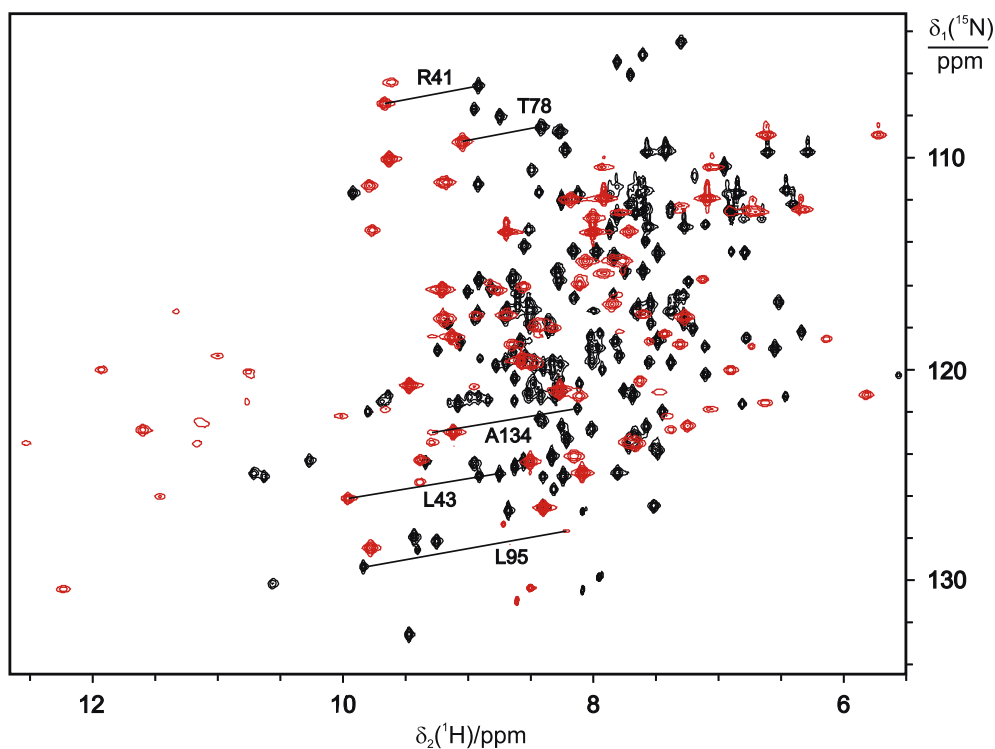
Figure 1. ^{15}N-HSQC spectra recorded with 0.6 mM aqueous solutions of the 30 kDa complex between the ^{15}N/^{2}H-labeled N-terminal domain of the ε subunit, ε186, and the (unlabeled) θ subunit of DNA polymerase III (DeRose et al., 2002, 2003; Hamdan et al., 2002) at 25 °C and pH 7.2. The spectra were recorded using a Bruker AV800 NMR spectrometer. The superposition is shown of spectra recorded with Dy^{3+} and La^{3+} (red and black peaks, respectively). Solid lines connect selected pairs of diamagnetic and paramagnetic cross-peaks.

and paramagnetic HNCO or ^{15}N-HSQC spectra that are among the most sensitive 3D and 2D NMR spectra that can be recorded of ^{15}N/^{13}C or ^{15}N-labeled proteins, respectively. In particular, it is shown to produce reliable $\Delta\chi$-tensor parameters from 2D ^{15}N-HSQC data in situations where these would be difficult to obtain manually. The program is of particular interest for rapid structure determinations of protein-ligand complexes with a paramagnetic metal tag, where the $\Delta\chi$ tensor of a paramagnetic center provides a reference coordinate system for both molecules in the complex.

Methods

Description of the algorithm

The program Echidna requires as input a single pair of diamagnetic and paramagnetic

^{15}N-HSQC or HNCO spectra, the assignment of the diamagnetic spectrum and the protein structure. It provides efficient and fully automatic determination of the χ-tensor anisotropy parameters and assignments of paramagnetic resonances. The iterative process is illustrated in Figure 2.

The program comprises the following steps:

(a) *Selection of diamagnetic peaks.* The ^{1}H NMR resonances of protons close to the paramagnetic center are broadened beyond detection due to paramagnetic relaxation enhancement (PRE). Therefore, Echidna disregards any diamagnetic peak for which the amide proton is located within a sphere of radius r_{cutoff} centered at the metal ion position. Assuming that no cross-peaks can be observed if the PRE effect exceeds 70 Hz, the (user defined) cutoff distance can be estimated from the equation for Curie spin relaxation

$$r_{\text{cutoff}} \approx \sqrt[6]{\frac{1}{5\pi(70\text{Hz})}\left(\frac{\mu_0}{4\pi}\right)^2 B_0^2\gamma_H^2 \frac{(g_e\mu_B)^4 S^2(S+1)^2}{(3k_B T)^2}\left(4\tau_r + \frac{3\tau_r}{1+\omega_H^2\tau_r^2}\right)} \tag{1}$$

where γ_H is the proton magnetogyric ratio, μ_0 the vacuum permeability, B_0 the magnetic field, k_B the Boltzmann constant, T the temperature, μ_B the Bohr magneton, g_e the electron g-factor (g_J for lanthanides), S the total spin quantum number (J for lanthanides), τ_r the rotational correlation time of the molecule, and ω_H the Larmor frequency of the proton.

(b) *Preliminary assignment of an initial subset of signals (diagonal search)*. To a first approximation, $^{13}C'$, ^{15}N and $^1H^N$ spins of a given peptide bond experience comparable paramagnetic shifts (measured in ppm). Therefore, in the presence of a paramagnetic center, the cross-peaks are displaced from their position in the diamagnetic spectrum along lines which are diagonal or nearly diagonal.

This principle is used as the starting point for assigning a first subset of peaks. For each diamagnetic peak, a search for a possible paramagnetic partner is performed in diagonal directions within a user-defined bandwidth w (Figure 3, shaded areas). Only those peak pairings are retained, where a single paramagnetic peak is found for a given diamagnetic starting peak within the bandwidth w and vice versa (case i in Figure 3). No peak pairings are accepted if the paramagnetic peak lies outside the bandwidth (case ii), or more than one paramagnetic peak can be paired with a diamagnetic peak (case iii). Too narrow bandwidths lead to few peak pairings, since paramagnetic and diamagnetic peaks often are not displaced along
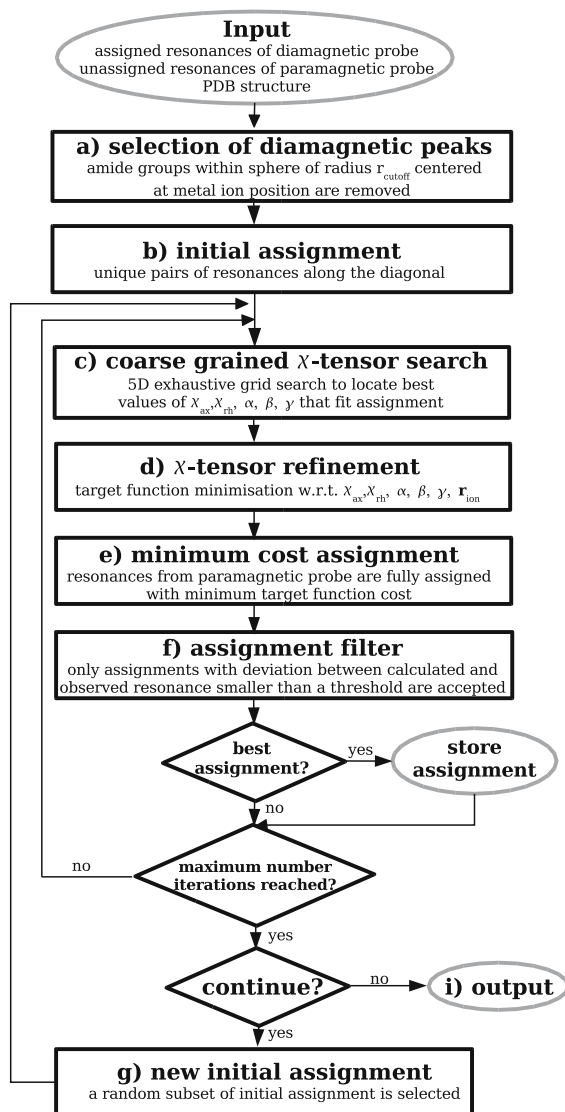

Input
assigned resonances of diamagnetic probe
unassigned resonances of paramagnetic probe
PDB structure

↓

a) selection of diamagnetic peaks
amide groups within sphere of radius r_{cutoff} centered
at metal ion position are removed

↓

b) initial assignment
unique pairs of resonances along the diagonal

↓

c) coarse grained χ-tensor search
5D exhaustive grid search to locate best
values of χ_{ax}, χ_{rh}, α, β, γ that fit assignment

↓

d) χ-tensor refinement
target function minimisation w.r.t. χ_{ax}, χ_{rh}, α, β, γ, $\mathbf{r}_{ion}$

↓

e) minimum cost assignment
resonances from paramagnetic probe are fully assigned
with minimum target function cost

↓

f) assignment filter
only assignments with deviation between calculated and
observed resonance smaller than a threshold are accepted

↓

best assignment? — yes → **store assignment**

no ↓

maximum number iterations reached? — no →

yes ↓

continue? — no → **i) output**

yes ↓

g) new initial assignment
a random subset of initial assignment is selected
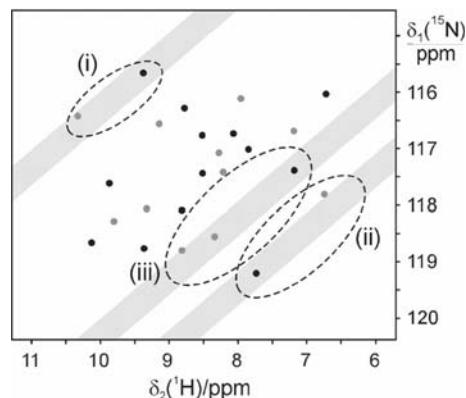
Figure 2. Flow diagram of the Echidna algorithm.

Figure 3. Schematic representation of the superposition of a diamagnetic (black) and a paramagnetic (grey) ^{15}N-HSQC spectrum. Assignments of paramagnetic peaks are made only if single pairs of paramagnetic and diamagnetic peaks are found within a user defined bandwidth (shaded).

strictly diagonal lines, whereas case (iii) becomes dominant for bandwidths that are too wide.

(c) *Coarse grained χ tensor search.* The initial assignments are used to derive a first set of pseudocontact shifts $\Delta\delta_{\exp,n}^{PCS} = \delta_{\mathrm{para},n} - \delta_{\mathrm{dia},n}$, where $\delta_{\mathrm{para},n}$ and $\delta_{\mathrm{dia},n}$ are the chemical shifts measured for the paramagnetic and diamagnetic peak of residue n, respectively. They depend on the anisotropy of the χ tensor and are independent of the isotropic part of the χ tensor. Therefore, the program only needs to determine the parameters of the χ-tensor anisotropy that is represented by the traceless $\Delta\chi$ tensor. The axes of the $\Delta\chi$ tensor also define the axes of the full χ tensor. Starting from the metal position provided with the structure coordinates, a set of at least five $\Delta\delta_{\exp,n}^{PCS}$ values is needed to estimate the χ-tensor anisotropy parameters, including the axial and rhombic components $\Delta\chi_{\mathrm{ax}}$ and $\Delta\chi_{\mathrm{rh}}$, respectively ($\Delta\chi_{\mathrm{ax}} = \chi_z - (\chi_x + \chi_y)/2$ and $\Delta\chi_{\mathrm{rh}} = (\chi_x - \chi_y)$, where χ_x, χ_y and χ_z denote the Cartesian components of the χ tensor) and the three Euler angles α, β, and γ that define its orientation with respect to the molecular frame. The program varies these five tensor parameters in a five-dimensional grid search in order to find the minimum of the cost function Q:

$$Q = \sum_n \sum_k \left(\frac{\Delta\delta_{\exp,k,n}^{PCS} - \Delta\delta_{\mathrm{cal},k,n}^{PCS}}{\Delta_k}\right)^2 = \sum_n q_n \tag{2}$$

where Δ_k is the error tolerance specific for each type of nucleus k (typical values are $\Delta_H = 0.1$ ppm, $\Delta_N = 0.2$ ppm and $\Delta_C = 0.2$ ppm) and

minimization of Q with respect to the five tensor parameters and, optionally, the position of the paramagnetic ion $\mathbf{r}_{\mathrm{ion}}$. In our experience, a discrete grid search in which $\Delta\chi_{\mathrm{ax}}$ and $\Delta\chi_{\mathrm{rh}}$ are varied from 0 to 15,000 ppm Å^3 and 0 to 10,000 ppm Å^3, respectively, in steps of 1,000 ppm Å^3, and α, β, and γ are varied in steps of 5° (corresponding to a total of 32,940 rotations) followed by continuous function minimization presents a good tradeoff between computing time and the ability to locate the global minimum of the cost function. The program internally uses the $\Delta\chi$ parameters in ppm Å^3 units that differ from $\Delta\chi$ in SI units (m^3) by the factor $12\pi \times 10^{-36}$. A single iteration of coarse grained grid search followed by minimization typically takes less than 1 min on a desktop PC.

(e) *Full assignment with the Hungarian method.* Once a set of $\Delta\chi$ values has been determined in step (d), the program predicts $\Delta\delta^{PCS}$ values for all $^1\mathrm{H}^N$, $^{15}\mathrm{N}$, and (if applicable) $^{13}\mathrm{C}'$ nuclei in the protein and generates a complete assignment table according to the algorithm known as the "Hungarian method for minimal cost assignment" (Kuhn, 1955). This algorithm associates pairs of paramagnetic and diamagnetic cross-peaks, finding a global matching that minimizes the general distance measure Q (Eq. (2)) over the entire set of possible assignments (Carpaneto et al., 1980; Hus et al., 2002).

(f) *Assignment filter.* Not all the assignments made in step (e) are used in further iterations. A pair of resonances is only included in the assignment table, if the weighted squared deviation of calculated and experimentally observed pseudo-

$$\Delta\delta_{\mathrm{cal},k,n}^{PCS} = \frac{1}{12\pi r_{k,n}^3}\left[\Delta\chi_{\mathrm{ax}}\left(3\cos^2\theta_{k,n} - 1\right) + \frac{3}{2}\Delta\chi_{\mathrm{rh}}\sin^2\theta_{k,n}\cos 2\varphi_{k,n}\right] \tag{3}$$

is the back-calculated pseudocontact shift. $r_{k,n}$, $\theta_{k,n}$, and $\varphi_{k,n}$ define the position of the nucleus k in residue n with respect to the paramagnetic center and the main axes of its χ tensor.

(d) *Refinement of the $\Delta\chi$-tensor parameters.* The discrete grid search and its computationally highly efficient implementation provides rapid access to approximate values of tensor parameters that minimize the cost function Q (Eq. (2)). Starting from the best $\Delta\chi$ tensor found in the grid search, the tensor parameters are further refined by Powell

contact shift (the individual contribution q_n from a residue n to the cost function Q) is smaller than 2:

$$q_n = \sum_k \left(\frac{\Delta\delta_{\exp,k,n}^{PCS} - \Delta\delta_{\mathrm{cal},k,n}^{PCS}}{\Delta_k}\right)^2 \leq 2 \tag{4}$$

This condition identifies signals contained within an ellipse of axes Δ_H and Δ_N (or, in the case of 3D data, an ellipsoid with the axes Δ_H, Δ_N, and Δ_C) that is centered at the calculated peak positions. The total number of pairs assigned by this criterion is

recorded. Since correct tensor parameters are more likely to result in a large number of assigned paramagnetic peaks, the number of peak pairs found by the algorithm can be used as a quality measure of the χ-tensor parameters determined in step (d).

Steps (c)–(f) are repeated for further refinement of the $\Delta\chi$-tensor parameters. The refinement stops as soon as a user-defined maximum number of iterations is reached (typically 4).

(g) *Consistency check of the results.* Steps (c)–(f) are repeated with randomly selected subsets comprising a user-defined number of peak pairings from the initial assignments made at step (b). These additional cycles reduce the impact of incorrect initial assignments. In our hands, best convergence was achieved when the number of peak pairings in each subset exceeded the minimum number (five); incorrect assignments present in the initial subset often were eliminated in the subsequent refinement steps. From all cycles, the program retains the $\Delta\chi$-tensor parameters resulting in the largest number of peak pairings assigned by the criterion of Eq. (4).

(h) *Calculation of residual anisotropic chemical shifts (RACS).* In calculating the chemical shifts of the paramagnetic peaks, the program includes correction terms that account for incomplete averaging of the anisotropic chemical shifts of $^1H^N$, ^{15}N and ^{13}C' spins due to paramagnetic alignment of the molecule with the external magnetic field:

$$\Delta\delta_{k,n}^{RACS} = \sum_{ij} \delta_k^{CSA,ii} \cos^2 \theta_{k,n}^{ij} A^{ij} \tag{5}$$

where $\delta_k^{CSA,ii}$ and A^{ij} ($i, j = 1, 2, 3$) are the diagonal elements of the chemical shift anisotropy (CSA) tensor $\boldsymbol{\delta}_k^{CSA}$ of nucleus k and the molecular alignment tensor $\mathbf{A}$ in their respective principal axes systems, and $\theta_{k,n}^{ij}$ is the angle between the principal axes i of $\boldsymbol{\delta}_k^{CSA}$ and j of $\mathbf{A}$ for nucleus k in residue n. At a magnetic field strength of 18.8 T the effect can be as large as 0.2 ppm for ^{15}N and ^{13}C' spins (John et al., 2005). Values for $\delta_H^{CSA,ii}$ and $\delta_N^{CSA,ii}$ were taken from Cornilescu and Bax (2000) and assumed to be uniform for all residues, whereas values for $\delta_C^{CSA,ii}$ are calculated from the diamagnetic isotropic ^{13}C' chemical shift according to Markwick and Sattler (2005). Any alignment of the diamagnetic protein is neglected and the alignment tensor of the paramagnetic protein is calculated using (Bertini et al., 2002)

$$\mathbf{A} = \frac{B_0^2}{15\mu_0 k_B T}\Delta\boldsymbol{\chi} \tag{6}$$

The angles $\theta_{k,n}^{ij}$ for the individual peptide groups are recalculated from the coordinate file for every refinement step of the $\Delta\chi$ tensor, and the resulting residual anisotropic chemical shifts $\Delta\delta_{k,n}^{RACS}$ are added to the calculated pseudocontact shifts $\Delta\delta_{k,n}^{PCS}$ to yield the $\Delta\delta_{cal,k,n}^{PCS}$ values of Eq. (4). For improved computational speed, the $\Delta\delta_{k,n}^{RACS}$ correction was not included in the tensor grid search (Eq. (2)).

(i) *Output.* The output of the program provides the best $\Delta\chi$ tensor parameters and writes an assigned peak list of paramagnetic resonances in Sparky format (Goddard and Kneller, 2004) to facilitate manual inspection of the generated assignments. In addition, PCS values, residual dipolar couplings, residual anisotropic chemical shifts, and longitudinal and transverse Curie relaxation rates are reported for all residues for which cross-peak assignments are available for the diamagnetic protein to support further experimental validation of the automatically generated assignment of the paramagnetic NMR data.

Results and discussion

In order to establish reference assignments of ε186 in complex with θ and Dy^{3+}, Er^{3+}, or Tb^{3+}, we recorded HNCO spectra of $^{15}N/^{13}C$-labeled samples. The improved spectral resolution afforded by the 3D NMR spectra allowed the straightforward manual assignment of most of the paramagnetic cross-peaks (see Supporting Information). Using the same 3D data, Echidna closely reproduced the $\Delta\chi$-tensor parameters and nearly all proposed assignments agreed with the manual assignments (Table 1). In the case of near-axially symmetric tensors, as in the case of the ε186/θ/Dy^{3+} complex, significant variations were limited to $\Delta\chi_{rh}$ and the angle γ as expected. Remarkably, Echidna also reproduced the $\Delta\chi$ values using exclusively the ^{15}N and ^1H chemical shifts from the HNCO spectra. This indicates that acceptable tensor parameters can be obtained without expensive ^{13}C-labeling (Allegrozzi et al., 2000; La Mar et al., 2001).

To assign 2D ^{15}N-HSQC data, the program compensates for the shortage of diamagnetic-paramagnetic peak pairs unambiguously identified

Table 1. Δχ-tensor parameters derived from pairs of 3D HNCO and 2D ^{15}N-HSQC spectra by the program Echidna[a]

Protein	Number of Assigned Peaks[b]	Δχ$_{ax}$ (10^{-32} m^3)	Δχ$_{rh}$ (10^{-32} m^3)	α (deg)	β (deg)	γ (deg)
ε186/θ/Dy^{3+} (3D manual)[c]	64	40.5	4.5	27	78	2
ε186/θ/Dy^{3+} (3D Echidna)[c]	46(1)	39.4	4.0	27	79	23
ε186/θ/Dy^{3+} (2D Echidna)[c]	37(2)	40.0	5.8	25	78	14
ε186/θ/Er^{3+} (3D manual)	80	−10.3	−4.0	23	91	42
ε186/θ/Er^{3+} (3D Echidna)	78 (0)	−10.5	−4.3	24	90	41
ε186/θ/Er^{3+} (2D Echidna)	77 (4)	−11.1	−4.0	24	90	41
ε186/θ/Tb^{3+} (3D manual)	66	27.2	5.6	20	84	1
ε186/θ/Tb^{3+} (3D Echidna)	63 (0)	27.8	5.9	20	84	0
ε186/θ/Tb^{3+} (2D Echidna)	56 (2)	27.7	3.7	19	84	3

[a]The Euler angles α, β, and γ refer to the crystal coordinates 1J53 for ε186 (Hamdan et al., 2002).[b]Number of incorrectly assigned peaks in parenthesis.[c] Note an error in Pintacuda et al. (2004), where the Δχ-tensor parameters were reported 9 times too small.

in the first iteration by forming random sub-groups of peak pairs from an initial pick that includes erroneous pairings. The number of peak pairs identified in the first iteration can be controlled by the bandwidth parameter. In the case of ε186, a maximum number of initial peak pairings was obtained using a bandwidth of about 0.1 ppm, but 10–20% of these assignments were incorrect (Figure 4). In contrast, a bandwidth of 0.35 ppm resulted in no incorrect assignments in any of the initial sets, but too few (less than five) peak pairings were made for the ^{15}N-ε186/θ/Dy^{3+} complex.

Convergence strongly depends on the Δχ tensor used in the first iteration; not all starting points converged to acceptable solutions. Independent of the initial pairings made, the solutions assigning the largest number of peaks in the paramagnetic spectrum consistently produced
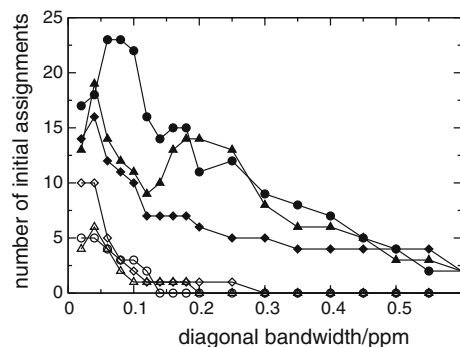

Figure 4. Number of total initial assignments made by the diagonal search module of Echidna as a function of the diagonal bandwidth *w* chosen (closed symbols), and number of incorrect assignments among them (open symbols). The data were calculated using the ^{15}N and ^{1}H chemical shifts observed for ^{15}N/^{13}C-ε186/θ in complex with Dy^{3+} (diamonds), Er^{3+} (circles) and Tb^{3+} (triangles), respectively.

very small percentages of incorrect assignments (Figure 5). Therefore, the total number of assigned paramagnetic cross-peaks can be used as a criterion to identify the best Δχ tensor and the most reliable assignment. Equally important, the Δχ tensor parameters derived from 2D ^{15}N-HSQC data were close to those determined from the 3D HNCO data and insensitive with regard to a few incorrect assignments in the last iteration. The algorithm also proved to be remarkably resilient with respect to uncertainties in the protein coordinates, provided a large number of paramagnetic peaks could be assigned (see Supporting Information).

The algorithm conservatively assigns the paramagnetic NMR spectrum, attempting to minimize the number of incorrect assignments. Therefore, the number of assigned paramagnetic cross-peaks is smaller than the number of observable peaks. A similar number of paramagnetic cross-peaks was assigned when the program was used with or without correction for residual anisotropic chemical shifts (RACS, Eq. (5)). Notably, however, use of the RACS correction accelerated convergence (data not shown).

As a starting point for additional interactive assignments, Echidna also provides complete peak lists of the diamagnetic cross-peaks shifted by the finally determined PCS values. Most important, however, Echidna delivers useful Δχ-tensor parameters in an automatic manner.

A previously reported program, Platypus, achieves the determination of the Δχ tensor together with the assignment of both the paramagnetic and diamagnetic ^{15}N-HSQC spectra, using selectively ^{15}N-labeled protein and knowledge of its 3D
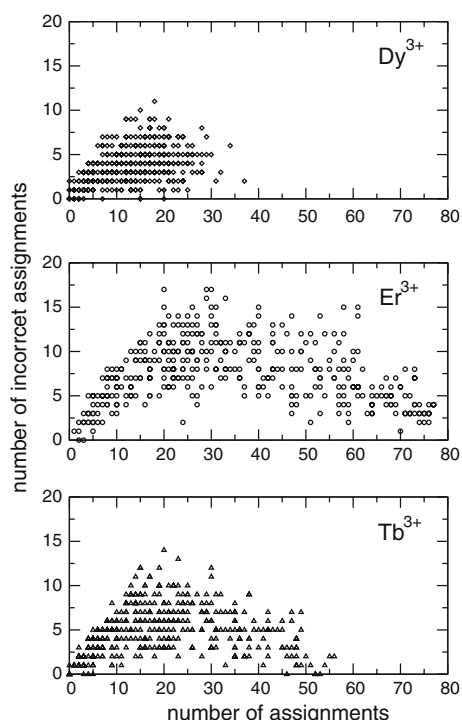
Figure 5. Number of incorrectly versus total number of assigned paramagnetic cross-peaks. Results are shown for Echidna runs starting from different subsets of initially selected paramagnetic-diamagnetic ^{15}N-HSQC peak pairs for complexes of ε186/θ with Dy^{3+}, Er^{3+} and Tb^{3+}, respectively.

structure (Pintacuda et al., 2004). The Platypus algorithm cannot be used, however, to assign ^{15}N-HSQC spectra of uniformly ^{15}N-labeled proteins because of the vast number of assignment possibilities that make the problem computationally unmanageable.

Since the $\Delta\chi$-tensor parameters determined with Echidna are based on many pairings of diamagnetic and paramagnetic cross-peaks, they tend to be more accurate than those determined by Platypus from a protein with a single ^{15}N-labeled amino acid type. Furthermore, uniformly ^{15}N-labeled proteins are more generally available than selectively ^{15}N-labeled proteins. The convenience with which $\Delta\chi$-tensor parameters can be determined with Echidna or Platypus considerably increases the attractiveness of PCS data as long-range restraints for the study of protein-protein and protein-ligand interactions. In addition, the alignment tensor caused by the paramagnetism is closely aligned with the $\Delta\chi$ tensor (Bertini et al., 2002), allowing the straightforward evaluation of residual dipolar couplings measured in ^{15}N-HSQC spectra of samples with and without paramagnetic metal ion.

The Echidna software package is available upon request from the communicating author.

Electronic supplementary material

Supporting Information Available: a compilation of the 10 best Echidna results obtained with 2D and 3D data with Dy^{3+}, Er^{3+}, or Tb^{3+}, respectively; Echidna results obtained for the ε186 structure with structural noise added; backbone assignments of ε186 in complex with θ in the presence of La^{3+}, Dy^{3+}, Er^{3+}, or Tb^{3+}, respectively.

The supplementary material is available in electronic format at http://dx.doi.org/10.1007/s10858-006-9002-4.

Acknowledgements

References

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

Babini, E., Bertini, I., Capozzi, F., Felli, I.C., Lelli, M. and Luchinat, C. (2004) *J. Am. Chem. Soc.*, **126**, 10496–10497.

Baig, I., Bertini, I., Del Bianco, C., Gupta, Y.K., Lee, Y.M., Luchinat, C. and Quattrone, A. (2004) *Biochemistry*, **43**, 5562–5573.

Banci, L., Bertini, I., Gori Savellini, G.., Romagnoli, A., Turano, P., Cremonini, M.A., Luchinat, C. and Gray, H.B. (1997) *Proteins*, **29**, 68–76.

Banci, L., Bertini, I., Cremonini, M.A., Gori-Savellini, G., Luchinat, C., Wüthrich, K. and Güntert, P. (1998) *J. Biomol. NMR*, **12**, 553–557.

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

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

Bertini, I., Luchinat, C. and Parigi, G. (2001b) *Solution NMR of Paramagnetic Molecules: Applications to Metallobiomolecules and Models.* Elsevier, Amsterdam, London.

Bertini, I., Luchinat, C. and Parigi, G. (2002) *Prog. NMR Spectrosc.*, **40**, 249–273.

Brautigam, C.A., Aschheim, K. and Steitz, T.A. (1999) *Chem. Biol.*, **6**, 901–908.

Carpaneto, G. and Toth, P. (1980) *ACM TOMS*, **6**, 104–111.

Cornilescu, G. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 10143–10154.

DeRose, E.F., Li, D., Darden, T., Harvey, S., Perrino, F.W., Schaaper, R.M. and London, R.E. (2002) *Biochemistry*, **41**, 94–110.

DeRose, E.F., Darden, T., Harvey, S., Gabel, S., Perrino, F.W., Schaaper, R.M. and London, R.E. (2003) *Biochemistry*, **42**, 3635–3644.

Diaz-Moreno, I., Diaz-Quintana, A., Dela Rosa, M.A. and Ubbink, M. (2005) *J. Biol. Chem.*, **280**, 18908–18915.

Frey, M.W., Frey, S.T., DeW Horrocks, W. Jr., Kaboord, B. and Benkovic, S.J. (1996) *Chem. Biol.*, **3**, 393–403.

Gaponenko, V., Altieri, A.S., Li, J. and Byrd, R.A. (2002) *J. Biomol. NMR*, **24**, 143–148.

Gaponenko, V., Sarma, S.P., Altieri, A.S., Horita, D.A., Li, J. and Byrd, R.A. (2004) *J. Biomol. NMR*, **28**, 205–212.

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

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

Goddard, T.D. and Kneller, D.G. (2004) *SPARKY 3*. University of California, San Francisco.

Goodfellow, B.J., Nunes, S.G., Rusnak, F., Moura, I., Ascenso, C., Moura, J.J.G., Volkman, B.F. and Markley, J.L. (2002) *Prot. Sci.*, **11**, 2464–2470.

Hamdan, S., Carr, P.D., Brown, S.E., Ollis, D.L. and Dixon, N.E. (2002) *Structure*, **10**, 535–546.

Hus, J.-C., Prompers, J.J. and Brüschweiler, R. (2002) *J. Magn. Reson.*, **157**, 119–123.

Ikegami, T., Verdier, L., Sakhaii, P., Grimme, S., Pescatore, B., Saxena, K., Fiebig, K.M. and Griesinger, C. (2004) *J. Biomol. NMR*, **29**, 339–349.

John, M., Park, A.Y., Pintacuda, G., Dixon, N.E. and Otting, G. (2005) *J. Am. Chem. Soc.*, **127**, 17190–17191.

Kuhn, H.W. (1955) *Naval Res. Logistic Quarterly*, **2**, 83–97.

La Mar, G.N., Kolczak, U., Tran, A.T. and Chien, E.Y. (2001) *J. Am. Chem. Soc.*, **123**, 4266–4274.

Leonov, A., Voigt, B., Rodriguez-Castaneda, F., Sakhaii, P. and Griesinger, C. (2005) *Chemistry*, **11**, 3342–3348.

Ma, C. and Opella, S.J. (2000) *J. Magn. Reson.*, **146**, 381–384.

Markwick, P.R.L. and Sattler, M. (2005) *J. Am. Chem. Soc.*, **126**, 11424–11425.

Peters, J.A., Huskens, J. and Raber, D.J. (1996) *Prog. NMR Spectrosc.*, **28**, 283–350.

Pintacuda, G., Keniry, M.A., Huber, T., Park, A.Y., Dixon, N.E. and Otting, G. (2004) *J. Am. Chem. Soc.*, **126**, 2963–2970.

Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1114.

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

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

Dijk, A.D.van, Fushman, D. and Bonvin, A.M. (2005) *Proteins*, **60**, 367–381.

Wöhnert, J., Franz, K.J., Nitz, M., Imperiali, B. and Schwalbe, H. (2003) *J. Am. Chem. Soc.*, **125**, 13338–13339.