

Rapid Measurement of Pseudocontact Shifts in Paramagnetic Proteins by GFT NMR Spectroscopy

Ravi Pratap Barnwal¹, Ashok K. Rout¹, Kandala V.R. Chary*¹ and Hanudatta S. Atreya*²

¹Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai-400005, India

²NMR Research Centre, Indian Institute of Science, Bangalore-560012, India

Abstract: Pseudocontact shifts (PCSs) provide valuable structural and dynamic information in (large) proteins. We propose a methodology based on the principle of G-matrix Fourier transform (GFT) NMR spectroscopy for simultaneous and rapid measurement of a large number of PCSs in proteins with a paramagnetic centre. Four experiments, namely, (3,2)D HNNCO, (3,2)D HNN(CO)CA, (3,2)D HNN(COCA)CB and (3,2)D HNHA taken together facilitate accurate measurement of six PCSs corresponding to ¹H^N, ¹H^α, ¹³C^α, ¹³C^β, ¹³C^γ and ¹⁵N nuclei. In addition, a new algorithm is presented for unambiguous sequence specific resonance assignments of peaks shifted due to PCS. This avoids the need to record multiple 3D correlation experiments. The utility of the proposed experiments is demonstrated with an 8.5 kDa protein, Calbindin. This provides new avenues to a wide range of applications in structure determination/refinement/verification and dynamical studies of proteins in general and paramagnetic proteins in particular.

INTRODUCTION

In recent years, pseudocontact shifts (PCSs) have emerged as one of the most important NMR parameters in calculation, refinement and validation of high-resolution three-dimensional (3D) structures of proteins in solution [1]. These PCSs provide long-range distance information between various NMR active nuclei and the paramagnetic ion as far apart as 45 Å in 3D space [2]. Such information is complimentary to the widely used distance constraints derived from ¹H-¹H NOEs and hence serves to enhance the quality of the resulting structures [3].

PCSs originate due to the presence of a paramagnetic metal centre in a protein possessing anisotropy in its electronic g-factor [4]. A strong anisotropic dipolar interaction between the NMR active nucleus and the unpaired electron gives rise to large changes in chemical shifts of nuclei located within a specific distance of the metal-ion [4]. From the knowledge of the 3D structure of the protein, nature of the metal ion and the assignments available under diamagnetic conditions, the PCS shifted peaks are assigned and corresponding PCSs are measured [5]. These PCSs, in turn, are used for further refinement of the 3D structure [2]. In many instances, spectral overlap due to line-broadening or/and increase in the number of peaks can result in ambiguities for assignments of residues that are pseudocontact shifted. In such cases, a suite of 3D triple-resonance experiments are recorded to aid in their assignment [1, 5b]. While degeneracy in chemical shifts and the spectral overlap can be reduced using 3D NMR experiments, their utility for rapid resonance assignment and structure determination is hampered by long 'minimal' measurement time required to acquire data with high digital resolution [6]. Further, this becomes critical if

the protein under study is unstable or when the individual PCSs have to be measured with several different metal-ions. This necessitates the development of methods to speed up the process of data collection. In this back drop, we propose a novel methodology that uses a suite of NMR experiments based on the principle of G-matrix Fourier transform (GFT) NMR spectroscopy to rapidly collect the data and analyze and quantify individual PCSs of different nuclei. All the proposed experiments can be acquired in a few hours time and facilitate unambiguous and accurate assignment of PCSs.

GFT NMR spectroscopy [6] is based on phase sensitive joint sampling of two or more chemical shifts in a single dimension, thereby providing higher dimensional spectral information rapidly with high precision. In recent years, GFT NMR has found several applications in resonance assignments and structure determination of biomolecules [6]. Here we describe the concerted use of four GFT NMR experiments namely, (3,2)D HNNCO [6c,d], (3,2)D HNN(CO)CA [6c], (3,2)D HNN(COCA)CB [6c] and (3,2)D HNHA [7] facilitating the measurement of PCS of ¹H^N, ¹H^α, ¹³C^α, ¹³C^β, ¹³C^γ and ¹⁵N nuclei. In these experiments, the chemical shifts are jointly sampled for the nuclei shown underlined [6]. Phase sensitive joint sampling of chemical shifts for the two different nuclei is achieved by co-incrementing their respective chemical shift evolution periods wherein one of the shifts is scaled by a factor 'κ' relative to the other [6d,e]. After the G-matrix transformation, it results in two sub-spectra each comprising of peaks at a given linear combination of chemical shifts along the indirect dimension: ω₁: Ω(¹⁵N_i) ± κ*Ω(¹³C^γ_{i-1}) in (3,2)D HNNCO, ω₁: Ω(¹⁵N_i) ± κ*Ω(¹³C^α_{i-1}) in (3,2)D HNN(CO)CA, ω₁: Ω(¹⁵N_i) ± κ*Ω(¹³C^β_{i-1}) in (3,2)D HNN(COCA)CB and ω₁: Ω(¹⁵N_i) ± κ*Ω(¹H^N_i) / ω₁: Ω(¹⁵N_i) ± κ*Ω(¹H^α_i) in (3,2)D HNHA. The scaling factor, κ, allows one to increase the dispersion of peaks or to restrict the chemical shift evolution to avoid losses in sensitivity due to transverse relaxation [6, 7]. A 2D [¹⁵N, ¹H] HSQC provides central peak information (ω₁:Ω(¹⁵N)) for all the spectra [6-7].

*Address correspondence to these authors at the Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai-400005, India; Fax: (+) 91 22 2280 4610; E-mail: chary@tifr.res.in
NMR Research Centre, Indian Institute of Science, Bangalore-560012, India; Fax: (+) 91 80 2360 1550; E-mail: hsatreya@sif.iisc.ernet.in

The experiments are demonstrated with an 8.5 kDa protein called Calbindin that has two Ca^{2+} -binding sites [8]. A paramagnetic lanthanide ion (Yb^{3+}) is introduced in the protein by selective displacement of one of the two metal ions. PCSs are then determined by measuring the difference in chemical shifts observed in the presence and absence of the paramagnetic metal ion.

METHODOLOGY

Sample Preparation and NMR Spectroscopy

Two uniformly ^{13}C and ^{15}N doubly labeled calbindin samples, one with Ca^{2+} -bound to both N- and C-terminal sites ($[\text{Ca}^{2+}]_2\text{Cb}$) and the other Yb^{3+} -substituted in the C-terminal site ($[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$) were purchased from ProtEra SRL (<http://www.protera.it>). Yb^{3+} was chosen as the paramagnetic metal ion due to its favorable PCS/line-broadening ratio compared to other Ln^{3+} [2a, 5]. The powder sample was dissolved in a mixed solvent of 90% $\text{H}_2\text{O}/10\% \text{ } ^2\text{H}_2\text{O}$ (pH 6.0) upto a concentration of ~ 1.5 mM. The radio frequency (r.f.) pulse schemes used for the different experiments are shown in the Supporting Information (Fig. **S1-S3**). All NMR experiments were performed at 25°C on Bruker Avance 500 MHz spectrometer and Varian Inova 600 MHz spectrometer equipped with a cryogenic probe. The scaling factor, κ , was set to 0.5 in all experiments except for (3,2)D HNCO (where $\kappa=2.0$). In the case of (3,2)D $\text{HNN}(\text{CO})\underline{\text{CA}}$ and (3,2)D $\text{HNN}(\text{COCA})\underline{\text{CB}}$ this results in a $t_{1\text{max}}$ of ~ 8 -10 ms for $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ nuclei, respectively, which is necessary to avoid loss of sensitivity due to ^{13}C - ^{13}C couplings [9]. Further, in (3,2)D $\text{HNN}(\text{COCA})\underline{\text{CB}}$, complete transfer of mag-

netization from $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ was implemented to avoid spectral crowding arising from $^{13}\text{C}^\alpha$ PCS peaks overlapping with those of $^{13}\text{C}^\beta$. The total measurement time for the entire set of GFT experiments was 2.5 and 5 hours in case of ($[\text{Ca}^{2+}]_2\text{Cb}$) and ($[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$), respectively (details of acquisition parameters are provided in Table 1). Notably, the corresponding 3D experiments which can provide an *equivalent* spectral resolution would have required several fold measurement times. The spectra on $[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$ were recorded with twice the measurement time compared to the $[\text{Ca}^{2+}]_2\text{Cb}$ to take care of the lower signal-to-noise (S/N) observed in the former case due to line-broadening. The data were pre-processed with the G-matrix using in-house written scripts (using the methodology outlined in refs. 6) and subsequently processed with NMRPipe [10] and analyzed using XEASY [11] and CARA [12]. The information about the ^1H , ^{13}C and ^{15}N resonance assignments thus obtained for $[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$ has been deposited in the BMRB under accession code 15594. The PCS values measured were checked for their accuracy by comparing them with: (i) those obtained using conventional 3D NMR experiments [9] (3D HNCO , 3D $\text{CBCA}(\text{CO})\text{NH}$ and 3D HNHA ; total measurement time ~ 24 hrs) and (ii) those predicted with the program FANTASIAN [13] using the 3D structure of Calbindin (PDB ID: 1ksm.pdb) [3].

Assignment of PCS Peaks to their Respective Non-Shifted Peaks

For each spin system (i.e., amino acid residue), the identification of a pair of peaks along the indirect dimension: $\omega_1: \Omega(^{15}\text{N}_i) \pm \kappa * \Omega(X)$ ($X = ^1\text{H}^N, ^1\text{H}^\alpha, ^{13}\text{C}^\alpha, ^{13}\text{C}^\beta, ^{13}\text{C}^\gamma$ and ^{15}N) in

Table 1. Acquisition Parameters of GFT NMR Experiments

SrNo	Experiment ^a	Sample	Indirect dimension: t_{max} (ms); Complex points; Digital Resolution (Hz/Pt) ^b	Meas. time (hrs)	(S/N) ^c
1	(3,2)D HNCO	$[\text{Ca}^{2+}]_2\text{Cb}$	$\omega_1(^{15}\text{N}; ^{13}\text{C}^\gamma)$: 23.0(^{15}N)/46.0 ($^{13}\text{C}^\gamma$); 74; 12.2	0.3	34.9 \pm 6.7
		$[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$	$\omega_1(^{15}\text{N}; ^{13}\text{C}^\gamma)$: 23.0(^{15}N)/46.0 ($^{13}\text{C}^\gamma$); 74; 12.2	0.6	20.5 \pm 7.1
2	(3,2)D $\text{HNN}(\text{CO})\underline{\text{CA}}$	$[\text{Ca}^{2+}]_2\text{Cb}$	$\omega_1(^{15}\text{N}; ^{13}\text{C}^\alpha)$: 22.0(^{15}N)/11.0 ($^{13}\text{C}^\alpha$); 52; 12.2	0.2	14.3 \pm 2.8
		$[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$	$\omega_1(^{15}\text{N}; ^{13}\text{C}^\alpha)$: 22.0(^{15}N)/11 ($^{13}\text{C}^\alpha$); 52; 12.2	0.4	11.3 \pm 3.0
3	(3,2)D $\text{HNN}(\text{COCA})\underline{\text{CB}}$	$[\text{Ca}^{2+}]_2\text{Cb}$	$\omega_1(^{15}\text{N}; ^{13}\text{C}^\beta)$: 20.0(^{15}N)/10 ($^{13}\text{C}^\beta$); 128; 11.7	0.6	28.3 \pm 5.3 ^d
		$[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$	$\omega_1(^{15}\text{N}; ^{13}\text{C}^\beta)$: 20.0(^{15}N)/11.5 ($^{13}\text{C}^\beta$); 128; 11.7	0.7	13.2 \pm 3.3 ^d
4	(3,2)D HNHA	$[\text{Ca}^{2+}]_2\text{Cb}$	$\omega_1(^{15}\text{N}; ^1\text{H})$: 39.0(^{15}N)/18 (^1H); 180; 8.0	1.5	32.1 \pm 3.2 24.1 \pm 1.1
		$[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$	$\omega_1(^{15}\text{N}; ^1\text{H})$: 39.0(^{15}N)/18 (^1H); 180; 8.0	3.0	28.1 \pm 3.6 15.7 \pm 1.1

^aA 2D [^{15}N , ^1H] HSQC was recorded as central peak spectra for calbindin without Yb (~ 5 minutes) and with Yb (~ 10 minutes).

^bDirect dimension $\omega_2(^1\text{H})/\omega_3(^1\text{H})$: 75; 512; 10.0.

^cS/N is calculated as ratio of the peak height to 2.5*standard deviation of the noise.

^dWith cryogenic probe on 600 MHz.

$[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$ (shown in red; Fig. 1) were assigned to their non-shifted parent peaks corresponding to $[\text{Ca}^{2+}]_2\text{Cb}$ (shown in blue) using the algorithm described above. The generally observed fact that both ^{15}N and $^1\text{H}^{\text{N}}$ nuclei in a given amino acid residue exhibit up-field or down-field PCS [5] served to verify these assignments.

In 2D ^{15}N , ^1H HSQC of $[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$ and the different GFT spectra, peaks were not observed for ~30% of residues. On inspection of the 3D structure, the amide moieties of these residues (18-26, 52-66) were found to be located within a radius of $6.9 \pm 2.7 \text{ \AA}$ from Yb^{3+} (Fig. S6) leading to loss of signal due to extensive line broadening [2a]. A total of 104 $^1\text{H}^{\text{N}}$ and ^{15}N PCS were measured which is similar to the number of PCS measured in the earlier studies [2a].

Fig. (2) shows the four GFT (3,2)D spectra recorded on the two samples. The respective acquisition parameters and sensitivity are given in Table 1. Fig. (3) shows the measured PCS as a function of residue number for different nuclei obtained using the methodology described above. The GFT spectra acquired with $[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$ yield a lower number of PCSs compared to that obtained from 2D ^{15}N , ^1H HSQC (Fig. 2, 3). This is due to the fact that the various GFT experiments employ additional delay periods for polarization transfer to various nuclei ($^{13}\text{C}'$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, $^1\text{H}^{\alpha}$), which have stronger dipolar interaction with the paramagnetic centre and hence have reduced sensitivity for residues close to the metal ion. Fig. (4) depicts the magnitude and the sign of PCS measured for different nuclei mapped onto the high-resolution 3D structure of lanthanide substituted Calbindin [3]. Peaks corresponding to residues 18-26 and 52-66 come close in space to the paramagnetic metal ion Yb^{3+} (Fig. S6) and hence are not observed in many of the spectra due to severe line broadening [2a]. For other residues, the magnitude of PCS decreases with increase in the distance from the metal ion in accordance with theory.

One of the challenging tasks in structural studies of proteins containing paramagnetic metal ions is the sequence specific assignment of PCS shifted peaks [15]. This is typically accomplished using a set of 3D triple resonance experiments that are used for sequential resonance assignment of diamagnetic proteins [16]. In the present study, complete unambiguous assignments were obtained using the novel algorithm described above without recourse to any sequential assignment procedure. This methodology uses the chemical shifts of $^1\text{H}^{\alpha}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, $^{13}\text{C}'$ in conjunction with those of ^{15}N and $^1\text{H}^{\text{N}}$. The underlying assumption of the algorithm is that PCS acts as a perturbation of diamagnetic chemical shift values and therefore a given residue is expected to remain in its vicinity (quantified by r.m.s.d.) in the six-dimensional chemical shift space comprising the six chemical shifts being measured. Thus, in the presence of a paramagnetic metal ion, any spin undergoing PCS can be unambiguously traced back to its source (original) peak in the diamagnetic protein by measuring different types of chemical shifts and using the methodology described above (Eqn. [1]). The utility of the algorithm was tested using the data acquired on Calbindin by successively including various chemical shifts in Eqn. [1]. It is observed (Fig. S5) that using only ^{15}N and ^1H chemical shifts, 8 residues in Calbindin undergoing PCS have lower r.m.s.d. to residues other than their own in the diamagnetic state. Thus, use of information about

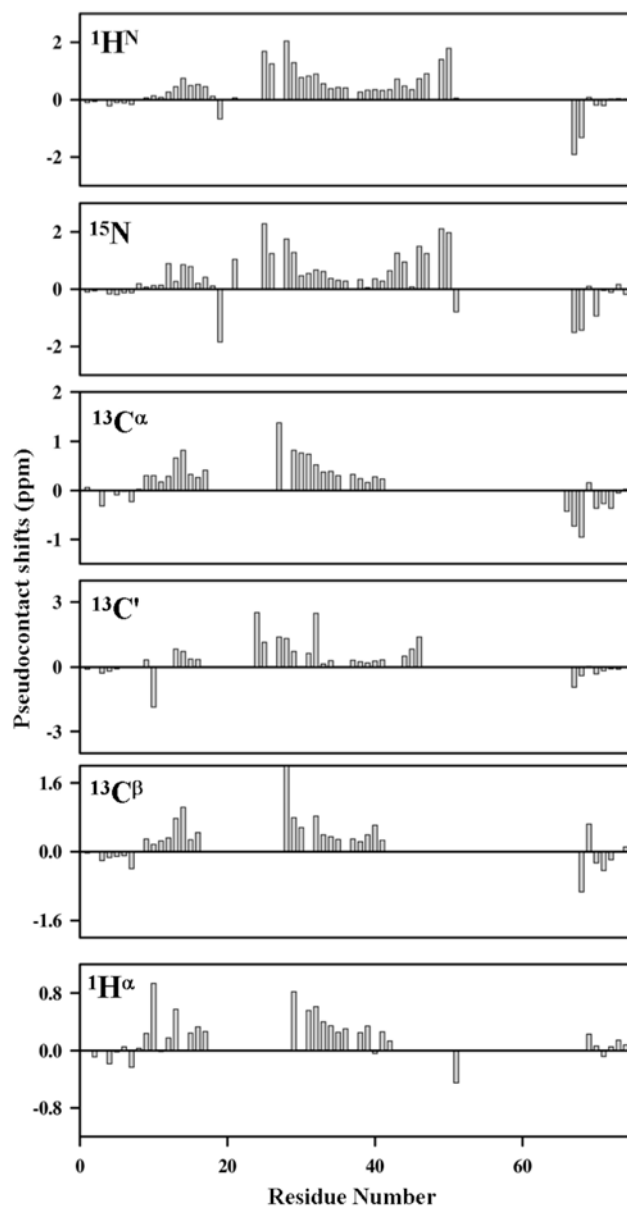


Fig. (3). Plot of the different PCSs measured using (3,2)D GFT spectra as a function of amino acid residues of calbindin. The ^1H , ^{13}C and ^{15}N resonance assignments obtained for $[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$ sample has been deposited in the BMRB under accession code 15594.

the $^1\text{H}^{\text{N}}$ and ^{15}N chemical shifts alone renders their assignment ambiguous. It would then require information from 3D NMR experiments used for sequential resonance assignments to arrive at residue specific assignment. On the other hand, the number of residues that have lower r.m.s.d decreases on inclusion of additional chemical shifts ($^1\text{H}^{\alpha}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, $^{13}\text{C}'$) in Eqn. [1]. This implies that the pair comprising the PCSs and the corresponding diamagnetic counter-part gets uniquely identified by their lowest r.m.s.d. compared to that of all other possible pairs of residues when the different chemical shifts are included. Thus, PCS shifted peaks can be automatically assigned without any recourse to 3D triple resonance correlation experiments [1, 15]. This will hold true even for residues with large PCS because it is unlikely for

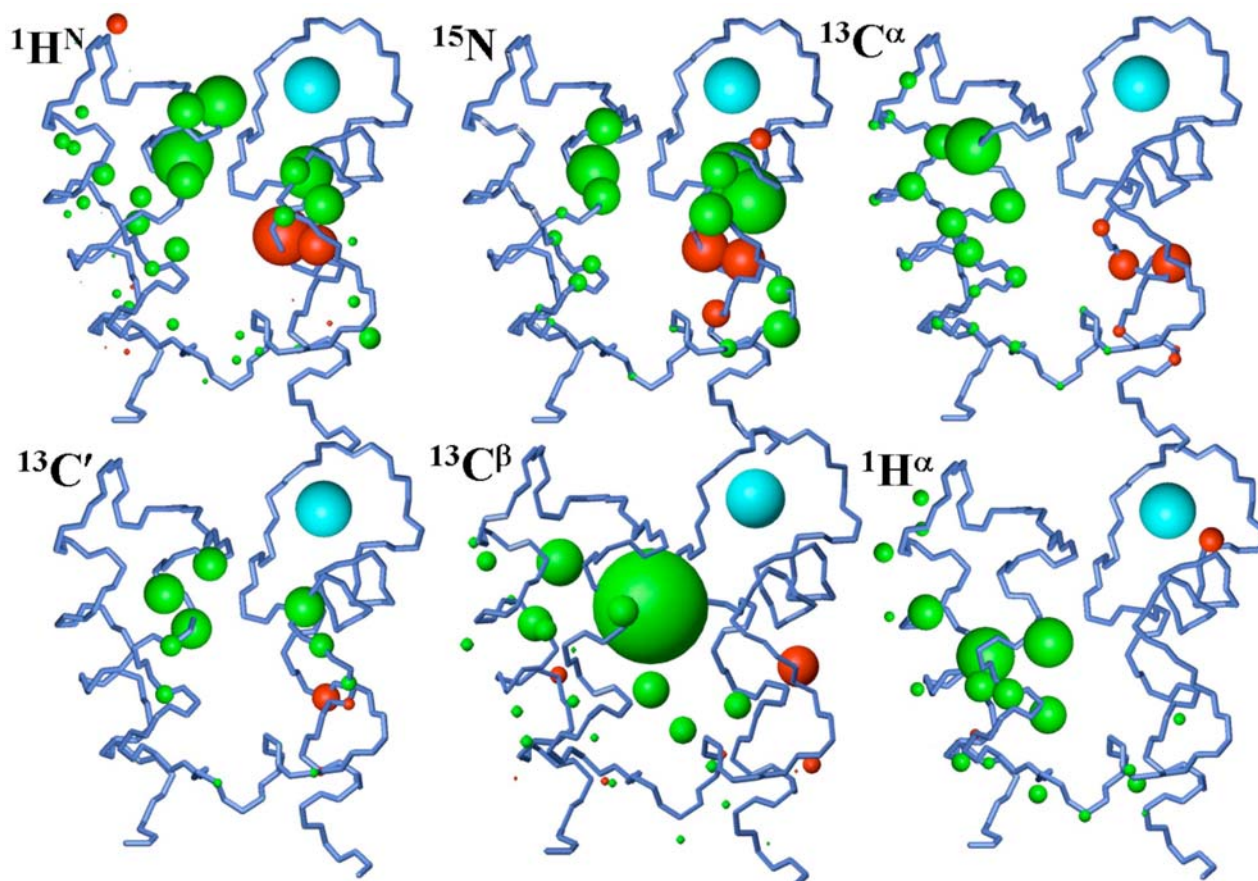


Fig. (4). PCSs of $^1\text{H}^{\text{N}}$, ^{15}N , $^{13}\text{C}^{\alpha}$, $^{13}\text{C}'$, $^{13}\text{C}^{\beta}$ and $^1\text{H}^{\alpha}$ as measured in various GFT-NMR experiments using the $[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$ sample (1ksm.pdb), depicted as spheres using MOLMOL [14]. The radii of the spheres are proportional to the relative magnitude of the observed shifts (shown in Fig. (3)); For each nucleus, the sphere having the largest size denotes the highest PCS for that nucleus. The red and green spheres represent negative and positive shifts, respectively. The cyan colored sphere represents the paramagnetic center (Yb^{3+}). The shifts seen in 10-16 and 30-50 regions (Fig. 3) indicate a higher range of paramagnetic interaction.

two residues with a different set of six chemical shifts (i.e. well separated in 6D chemical shift space) in the diamagnetic protein to overlap in the six-dimensional chemical shift space upon addition of the paramagnetic metal-ion (verified using a chemical shift database of 235 proteins obtained from BioMagResBank (<http://www.bmrb.wisc.edu>); Fig. S5). Notably, this algorithm can be applied to any set of higher dimensional chemical shift correlation data recorded for the assignment of PCS shifted peaks and are not limited to the chosen set of GFT experiments described here. In case of residues in close vicinity of the metal-ion, severe line-broadening leads to absence of peaks. The extent of line-broadening and hence the sensitivity of the experiments will depend on the choice of Ln^{3+} [2a]. Thus for a given residue, if the chemical shifts of specific nuclei are not measured, the r.m.s.d. (Eqn [1]) is calculated using the available chemical shifts for that residue.

The accuracy of the PCS values measured with the GFT spectra was evaluated by comparison with (i) PCS measured using conventional triple resonance 3D NMR experiments (Fig. 5A; for comparison of ^{15}N and $^1\text{H}^{\text{N}}$ PCS, the values given in Ref. 2a was used) and ii) PCS predicted with FANTASIAN [13] (Fig. 5B) using the high-resolution 3D structure of Calbindin (1ksm.pdb; Fig. 4) as the reference. Since systematic errors are negligible for GFT NMR-based measurement of PCS (Fig. 5), one can compare the precision of

measurements for each type of PCS by considering that $\text{r.m.s.d.} = \sqrt{\sigma^2(\text{GFT}) + \sigma^2(\text{conventional})}$, where σ denotes the error/standard deviation of PCS measured. Assuming both GFT and its 3D congener have similar errors in PCS, $\text{r.m.s.d.} = \sigma \cdot \sqrt{2}$. Based on this, a precision (σ) of 0.03-0.2 ppm is obtained for the measurement of different PCS. Measuring the PCS with high precision and accuracy renders them useful for 3D structure refinement protocols employing PCS [17] where the accuracy of the structure depends critically on the accuracy of the calculated alignment tensor.

CONCLUSIONS

We propose a novel methodology based on the principle of GFT NMR spectroscopy for rapid and accurate measurement of six different PCSs in paramagnetic proteins. The concerted use of the six different PCSs helps in accurate prediction of molecular alignment tensor for 3D structure refinement. The measurement times can be reduced further in combination with other fast data collection methods, such as longitudinal ^1H relaxation optimization [6b]. In the case of large proteins, the experiments presented here can be combined with deuteration [18] and TROSY [19] for increased resolution/sensitivity. Taken together, these experiments will have wide range of applications in structure determination and refinement of proteins in general and paramagnetic proteins in particular.

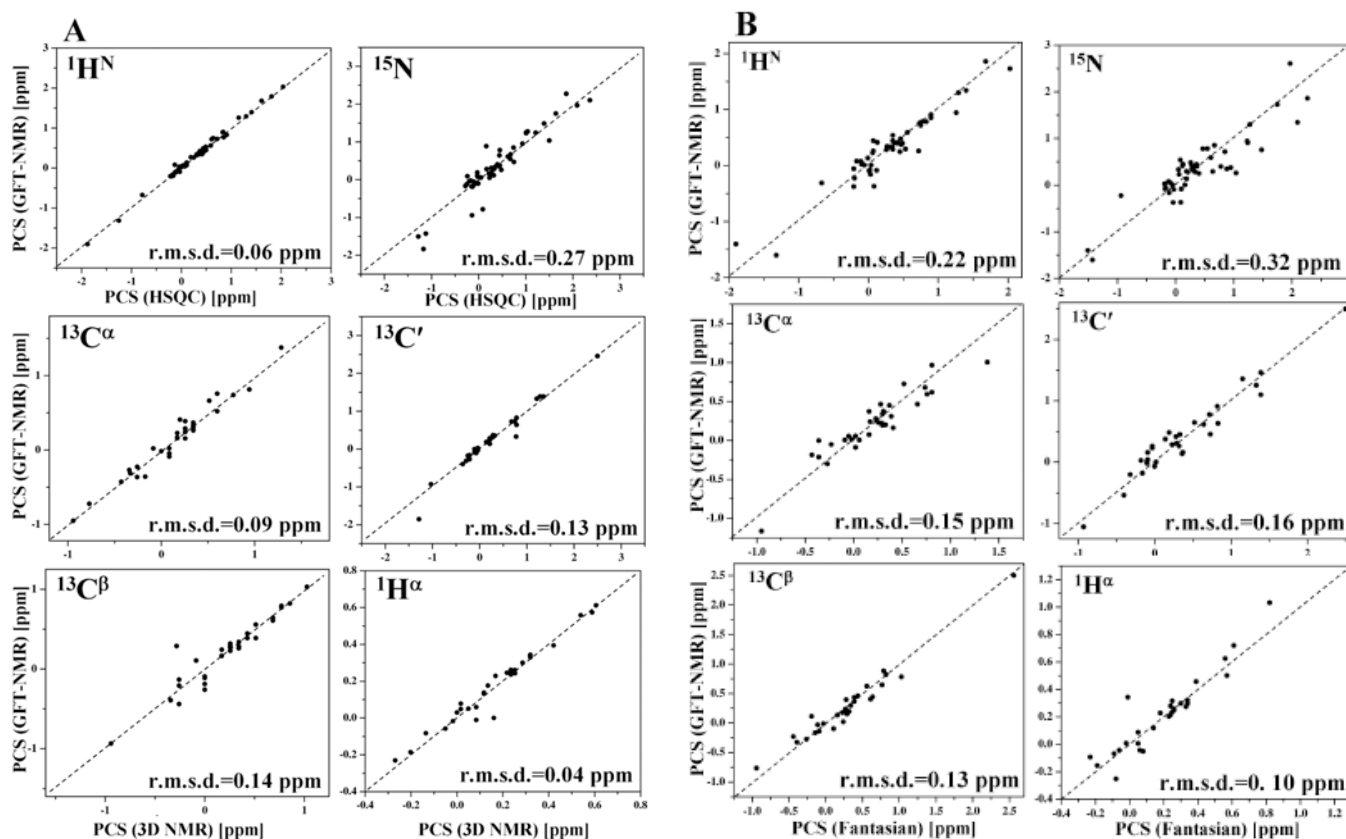


Fig. (5). (A) A comparison of PCS measured in calbindin using GFT NMR spectroscopy with those measured using conventional FT-NMR 2D [^{15}N , $^1\text{H}^{\text{N}}$] HSQC/3D experiments [2a, 16] (3D HNCO for $^{13}\text{C}'$; 3D CBCA(CO)NH for $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$; 3D HNHA for $^1\text{H}^{\alpha}$). For comparison of ^{15}N and $^1\text{H}^{\text{N}}$ PCS, the values given in Ref. 2a was used. (B) A comparison of PCS measured in calbindin using GFT NMR spectroscopy with those predicted using the program FANTASIAN [13]. Least-squares fit of pseudocontact shifts from GFT spectra and HSQC/3D experiments/FantAsian to a linear function reveals that slope and intercepts are, respectively, 1 and 0 within the experimental error. Hence, no systematic errors are detected in the GFT spectra.

ACKNOWLEDGEMENTS

The facilities provided by NMR Research Centre at IISc supported by Department of Science and Technology (DST) and National Facility for High Field NMR at TIFR supported by DST, Department of Biotechnology (DBT), Council of Scientific and Industrial Research (CSIR), and Tata Institute of Fundamental Research, Mumbai, India, are gratefully acknowledged. HSA acknowledges support from Department of Atomic Energy (DAE) BRNS research award. We also acknowledge Dr. Sourajit M. Mustafi for his critical comments during the project.

REFERENCES

- [1] (a) Bertini I, Luchinat C. NMR of paramagnetic molecules in biological systems. Benjamin-Cummings Pub Co; 1986. (b) Bertini I, Turano P, Vila AJ. Nuclear magnetic resonance of paramagnetic metalloproteins. *Chem Rev* 1993; 93: 2833-932. (c) John M, Otting G. Strategies for measurement of Pseudocontact shifts in protein NMR spectroscopy. *Chemphyschem* 2007; 8: 2309-13. (d) John M, Schmitz C, Park AY, Dixon NE, Huber T, Otting G. Sequence-specific and stereospecific assignment of methyl groups using paramagnetic lanthanides. *J Am Chem Soc* 2007; 129: 13749-57.
- [2] (a) Allegrozzi M, Bertini I, Janik MBL, Lee YM, Liu G, Luchinat C. Lanthanide-induced pseudocontact shifts for solution structure refinements of macromolecules in shells up to 40 Å from the metal ion. *J Am Chem Soc* 2000; 122: 4154-61. (b) John M, Pintacuda G, Park AY, Dixon NE, Otting G. Structure determination of protein-ligand complexes by transferred paramagnetic shifts. *J Am Chem Soc* 2006; 128: 12910-6.
- [3] Bertini I, Donaire A, Jimenez B, *et al.* Paramagnetism-based versus classical constraints: An analysis of the solution structure of Ca Ln calbindin D_{9k}. *J Biol NMR* 2001; 21: 85-98.
- [4] Bertini I, Luchinat C, Parigi G. Magnetic susceptibility in paramagnetic NMR. *Prog NMR Spect* 2002; 40: 249-273.
- [5] (a) Atreya HS, Mukherjee, Chary KV, Lee YM, Luchinat C. Structural basis for sequential displacement of Ca(2+) by Yb(3+) in a protozoan EF-hand calcium binding protein. *Protein Sci* 2003; 12: 412-25. (b) Mustafi SM, Mukherjee S, Chary KVR, Bianco CD, Luchinat C. Energetics and mechanism of Ca²⁺ displacement by lanthanides in a calcium binding protein. *Biochemistry* 2004; 43: 9320-31.
- [6] (a) Kim S, Szyperski T. GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. *J Am Chem Soc* 2003; 125: 1385-93. (b) Atreya HS, Szyperski T. G-Matrix Fourier transform NMR spectroscopy for complete protein resonance assignments. *Proc Natl Acad Sci USA* 2004; 101: 9642-9647. (c) Xia Y, Zhu G, Veeraghavan S, Gao X. (3,2)D GFT-NMR experiments for fast data collection from proteins. *J Biol NMR* 2004; 29: 467-476. (d) Atreya HS, Szyperski T. Rapid NMR data collection. *Methods Enzymol* 2005; 394: 78-108. (e) Szyperski T, Atreya HS. Principles and applications of GFT projection NMR spectroscopy. *Magn Reson Chem* 2006; 44: S51-60.
- [7] Barnwal RP, Rout AK, Chary KV, Atreya HS. Rapid measurement of $^3\text{J}(\text{H}^{\text{N}}-\text{H}^{\alpha})$ and $^3\text{J}(\text{N}-\text{H}^{\beta})$ coupling constants in polypeptides. *J Biol NMR* 2007; 39: 259-63.
- [8] Bertini I, Janik MBL, Lee YM, Luchinat C, Rosato A. Magnetic susceptibility tensor anisotropies for a lanthanide ion series in a fixed protein matrix. *J Am Chem Soc* 2001; 123: 4181-88.

- [9] Cavanagh J, Fairbrother WJ, Palmer AG, Skelton NJ, Rance M. Protein NMR spectroscopy: Principles and practice. Academic Press 2 edition. 2006.
- [10] Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biol NMR* 1995; 6: 277-93.
- [11] Bartels C, Xia TH, Billeter M, Guntert P, Wuthrich K. The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J Biol NMR* 1995; 6: 1-10.
- [12] Keller R. The Computer Aided Resonance Assignment Tutorial first edition 2004, ISBN 3-85600-112-3, CANTINA Verlag.
- [13] (a) Banci L, Bertini I, Bren KL, *et al.* The use of pseudocontact shifts to refine solution structures of paramagnetic metalloproteins: Met80Ala cyano-cytochrome *c* as an example. *J Bio Inorg Chem* 1996; 1: 117-26. (b) Banci L, Bertini I, Gori Savellini G, *et al.* The pseudocontact shifts as constraints for energy minimization and molecular dynamic calculations on solution structures of paramagnetic metalloproteins. *Proteins Struct Funct Genet* 1997; 29: 68-76.
- [14] Koradi R, Billeter M, Wuthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 1996; 14: 51-55.
- [15] Bertini I, Lee YM, Luchinat C, Piccioli, Poggi L. Locating the metal ion in calcium binding proteins by using cerium (III) as a probe. *ChemBiochem* 2001; 2: 550-58.
- [16] Sattler M, Schleucher J, Griesinger C. Heteronuclear multidimensional NMR experiments for structure determination of proteins in solution employing pulsed field gradients. *Prog NMR Spect* 1999; 34: 93-158.
- [17] Banci L, Bertini I, Cremonini MA, *et al.* PSEUDODYANA for NMR structure calculation of paramagnetic metalloproteins using torsion angle molecular dynamics. *J Biol NMR* 1998; 12: 553-557.
- [18] Gardner KH, Kay LE. The use of ²H, ¹³C, ¹⁵N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* 1998; 27: 257-406.
- [19] Pervushin K. Impact of Transverse Relaxation Optimized Spectroscopy (TROSY) on NMR as a technique in structure biology. *Q Rev Bio* 2000; 33: 161-197.

Received: December 16, 2007

Revised: February 03, 2008

Accepted: February 04, 2008

SUPPORTING INFORMATION FOR:

Rapid Measurement of Pseudocontact Shifts in Paramagnetic Proteins by GFT NMR Spectroscopy

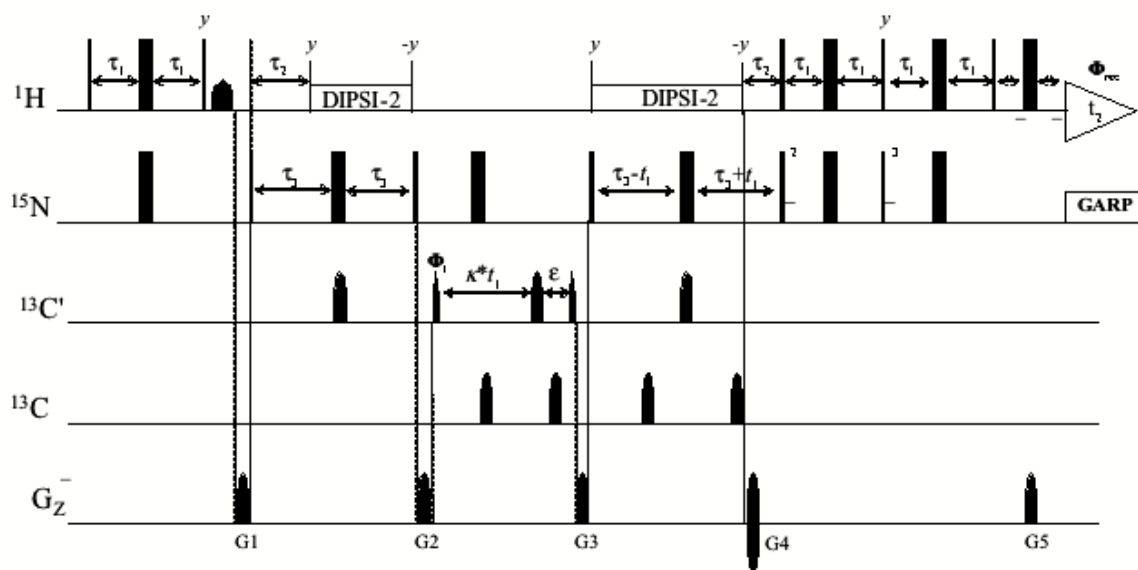
Ravi Pratap Barnwal,[†] Ashok K. Rout,[†] Kandala V. R. Chary*[†] and Hanudatta S. Atreya*[‡][†]Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai-400005, India and [‡]NMR Research Centre, Indian Institute of Science, Bangalore-560012, India

Figure S1. R.f. pulse scheme of GFT (3,2)D HN(NC)O. Rectangular 90° and 180° pulses are indicated by thin and thick vertical bars, respectively, and phases are indicated above the pulses. Where no r.f. phase is marked, the pulse is applied along x . High-power 90° pulse lengths are: $10.0 \mu\text{s}$ for ^1H , $35 \mu\text{s}$ for ^{15}N and $14.8 \mu\text{s}$ for ^{13}C . $\kappa=0.5$ (see text). The ^1H r.f. carrier is placed at the position of the solvent line at 4.7 ppm. The ^{15}N carrier position is set to 121.0 ppm. The $^{13}\text{C}'$ carrier is placed at 176 ppm. A Gaussian shaped ^1H r.f. pulse [1] of length 1.0 ms is used for water flip-back prior to the first r.f. pulse on ^{15}N . DIPSII-2 [2] is used for ^1H decoupling during heteronuclear coherence transfer periods. GARP [3] is employed to decouple ^{15}N (r.f.=1.50 kHz) during acquisition. All pulsed z -field gradients (PFGs) are sine-bell shaped with gradient recovery delay of $200 \mu\text{s}$. The duration and strengths of the PFGs are: G1 (0.5 ms, 30 G/cm); G2 (0.5 ms, 20 G/cm); G3 (1.0 ms, 10 G/cm); G4 (2.0 ms, 40 G/cm); G5 (1.0 ms, 4 G/cm). The delays are: $\tau_1 = 2.3 \text{ ms}$, $\tau_2 = 5.5 \text{ ms}$, $\epsilon=100\mu\text{s}$ and $\tau_3 = 12 \text{ ms}$. Phase cycling: $\phi_1 = x, -x$; $\phi_2 = 2(x), 2(-x)$; $\phi_3 = 2(y), 2(-y)$; $\phi(\text{receiver}) = x, -x, -x, x$. A sensitivity enhancement scheme [4] is employed, *i.e.*, the sign of G2 is inverted in concert with a 180° shift of ϕ_3 . GFT NMR phase-cycle: $\phi_1 = x, y$ yields, in conjunction with quadrature detection in $t_1(^{15}\text{N})$, 2 data sets which are linearly combined employing a G-matrix transformation with the G-matrix [5-6].



Figure S2. R.f. pulse scheme of GFT (3,2)D HNN(CO)CA. Rectangular 90° and 180° pulses are indicated by thin and thick vertical bars, respectively, and phases are indicated above the pulses. Where no r.f. phase is marked, the pulse is applied along x . High-power 90° pulse lengths are: $10.0 \mu\text{s}$ for ^1H , $35 \mu\text{s}$ for ^{15}N and $14.8 \mu\text{s}$ for ^{13}C . $\kappa=0.5$ (see text). The ^1H r.f. carrier is placed at the position of the solvent line at 4.7 ppm. The ^{15}N carrier position is set to 121.0 ppm. The $^{13}\text{C}'$ carrier is placed at 176 ppm before the first pulse on $^{13}\text{C}^\alpha$ and shifted back to 176 ppm after the last 90° r.f. pulse on $^{13}\text{C}^\alpha$. A Gaussian shaped ^1H r.f. pulse [1] of length 1.0 ms is used for water flip-back prior to the first r.f. pulse on ^{15}N . DIPSII-2 [2] is used for ^1H decoupling during heteronuclear transfer periods. GARP [3] is employed to decouple ^{15}N (r.f.=1.50 kHz) during acquisition. All pulsed z-field gradients (PFGs) are sine-bell shaped with gradient recovery delay of $200 \mu\text{s}$. The duration and strengths of the PFGs are: G1 (0.5 ms, 30 G/cm); G2 (2.0 ms, 40 G/cm); G3 (2.0 ms, 4 G/cm). The delays are: $\tau_1 = 2.3 \text{ ms}$, $\tau_2 = 5.5 \text{ ms}$, $\tau_3 = 12.5 \text{ ms}$, $\tau_4 = 4.0 \text{ ms}$. Phase cycling: $\phi_1 = x$; $\phi_2 = 2(x), 2(-x)$; $\phi_3 = 2(y), 2(-y)$; $\phi(\text{receiver}) = x, -x, -x, x$. A sensitivity enhancement scheme [4] is employed, *i.e.*, the sign of G2 is inverted in concert with a 180° shift of ϕ_3 . GFT NMR phase-cycle: $\phi_1 = x, y$ yields, in conjunction with quadrature detection in $t_1(^{15}\text{N})$, 2 data sets which are linearly combined employing a G-matrix transformation with the G-matrix [5-6].

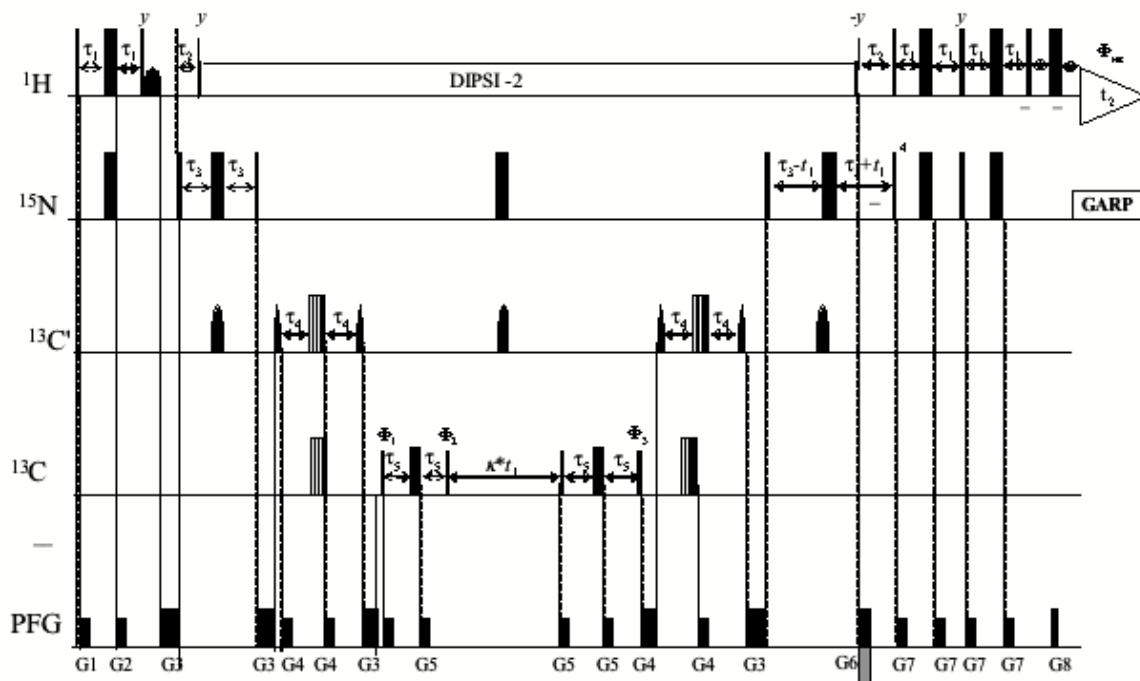


Figure S3. R.f. pulse scheme of GFT (3,2)D HN(COCA)CB. Rectangular 90° and 180° pulses are indicated by thin and thick vertical bars, respectively, and phases are indicated above the pulses. Where no r.f. phase is marked, the pulse is applied along x . High-power 90° pulse lengths are: $10.0 \mu\text{s}$ for ^1H , $35 \mu\text{s}$ for ^{15}N and $14.8 \mu\text{s}$ for ^{13}C . $\kappa=0.5$ (see text). The ^1H r.f. carrier is placed at the position of the solvent line at 4.7 ppm . A Gaussian shaped ^1H r.f. pulse [1] of length 1.0 ms is used for water flip-back prior to the first r.f. pulse on ^{15}N . The ^{15}N carrier position is set to 121.0 ppm . The $^{13}\text{C}'$ carrier is placed at 176 ppm and switched to 56 ppm before the first r.f. pulse on $^{13}\text{C}'$ and switched back to 176 ppm after the last r.f. pulse on $^{13}\text{C}'$. A six-pulse composite sequence [7] is used to simultaneously invert/refocus $^{13}\text{C}^\alpha/^{13}\text{C}'$ magnetization during $^{13}\text{C}^\alpha\text{-}^{13}\text{C}'$ polarization transfer. the 90° and 180° pulse lengths of $^{13}\text{C}^\alpha$ are adjusted (at a ^1H resonance frequency of 600 MHz) to $54 \mu\text{s}$ and $46 \mu\text{s}$, respectively, to minimize perturbation of the $^{13}\text{C}'$ spins. DIPSII-2 [2] is used for ^1H decoupling during heteronuclear transfer periods. GARP [3] is employed to decouple ^{15}N (r.f.= 1.50 kHz) during acquisition. All pulsed z-field gradients (PFGs) are rectangular shaped. The duration and strengths of the PFGs are: G1 (0.5 ms , 10 G/cm); G2 (0.5 ms , 10 G/cm); G3 (1.0 ms , 20 G/cm); G4 (0.5 ms , 10 G/cm); G5 (0.5 ms , 10 G/cm); G6 (1.25 ms , 22 G/cm); G7 (0.5 ms , 10 G/cm); G8 (0.125 ms , 22 G/cm). The delays are: $\tau_1 = 2.3 \text{ ms}$, $\tau_2 = 5.5 \text{ ms}$, $\epsilon=100\mu\text{s}$, $\tau_3 = 12.5 \text{ ms}$, $\tau_4 = 4.0 \text{ ms}$; $\tau_5 = 7.2 \text{ ms}$. Phase cycling: $\phi_1 = x, -x$; $\phi_2 = y, -y$; $\phi_3 = 2(x), 2(-x)$; $\phi(\text{receiver}) = x, -x, -x, x$. A sensitivity enhancement scheme [4] is employed, *i.e.*, the sign of G6 is inverted in concert with a 180° shift of ϕ_4 . GFT NMR phase-cycle: $\phi_1 = x, y$; $\phi_1 = y, -x$ yields, in conjunction with quadrature detection in $t_1(^{15}\text{N})$, 2 data sets which are linearly combined employing a G-matrix transformation with the G-matrix [5-6].

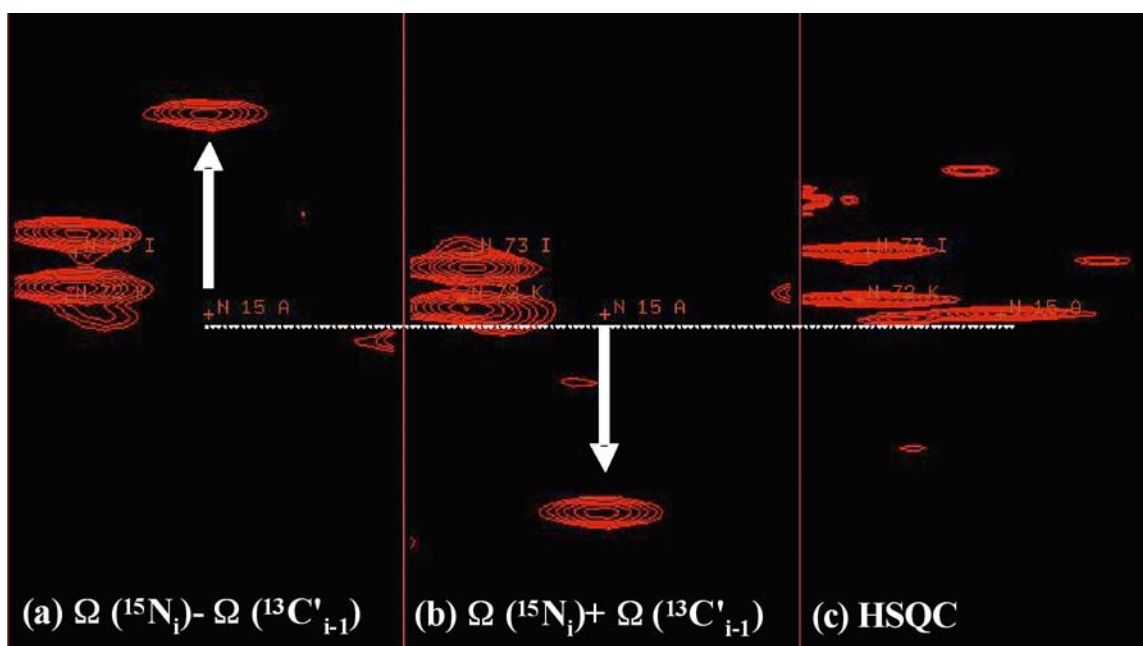
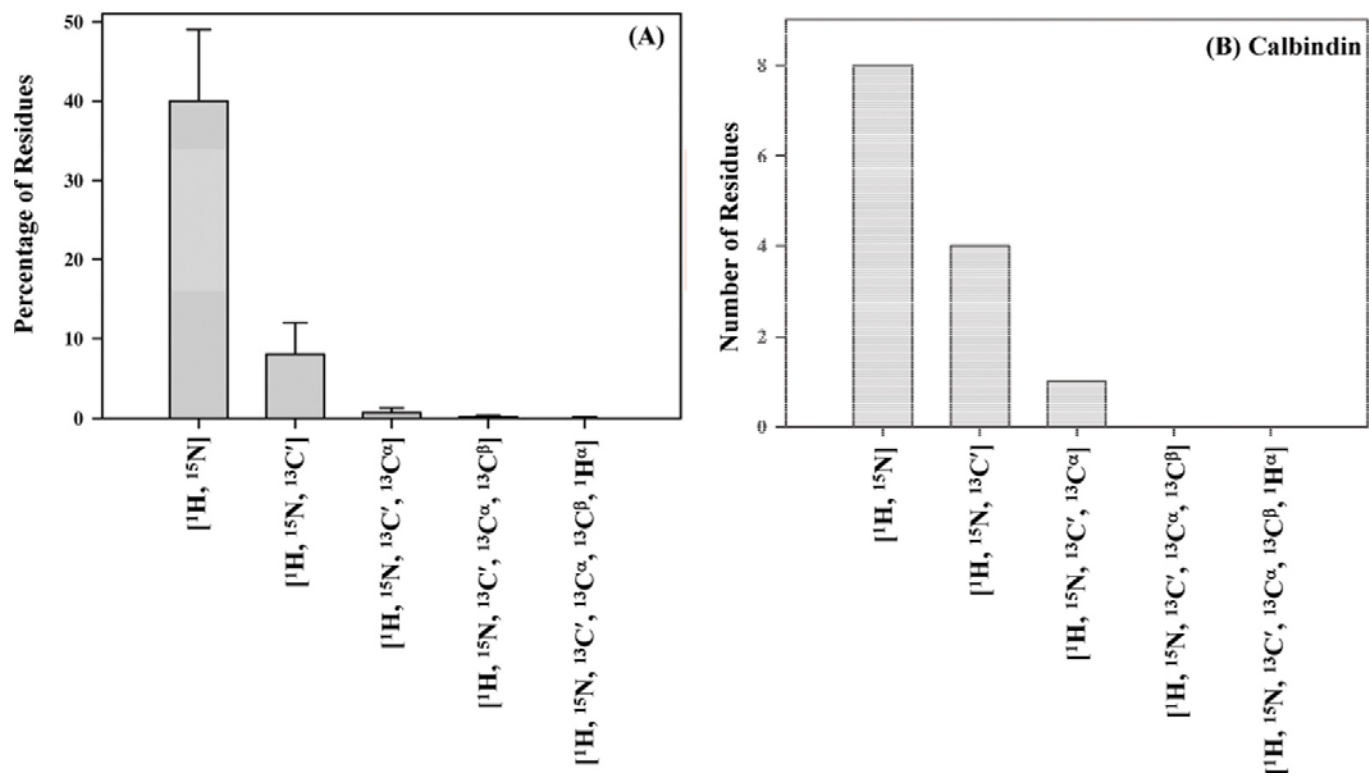


Figure S4: A screen capture using the software XEASY illustrating the identification/assignment of a pair of peaks belonging to an amino acid residue in (3,2)D GFT sub-spectra (shown in (a), (b)) along the indirect dimension: $\omega_1: \Omega(^{15}\text{N}_i) \pm k \cdot \Omega(X)$ ($X = ^1\text{H}^N, ^1\text{H}^\alpha, ^{13}\text{C}^\alpha, ^{13}\text{C}^\beta, ^{13}\text{C}'$ and ^{15}N) using the central peak information ($\omega_1: \Omega(^{15}\text{N})$) obtained from the 2D [$^{15}\text{N}, ^1\text{H}$] HSQC spectrum (shown in (c)). The particular spectrum shown corresponds to (3,2)D GFT HNNCO. The linear combination of the chemical shifts observed is indicated on the top. The symmetrical distribution of the peaks in GFT sub-spectra encoding the two linear combination of chemical shifts about the central peak ($\omega_1: \Omega(^{15}\text{N})$) facilitates the identification/assignment. In the event of two or more amino acids having degenerate $^1\text{H}^N / ^{15}\text{N}$ chemical shifts, the methodology described in Ref. 7d (in main text) is adopted.

**Figure S5:**

(A) Estimation of percentage of residues in proteins having degenerate chemical shifts. The estimation was carried out using a database of 235 proteins with known chemical shift assignments downloaded from the BioMagResBank (<http://www.bmrb.wisc.edu>). The percentage degeneracy was calculated by counting the number of residues that have degenerate chemical shifts with spins belonging to other residues in any protein within a given tolerance limit (which was chosen as 0.3 ppm for all nuclei corresponding to the average PCS observed in Calbindin). Thus, about 30-40% of residues in a given protein can have degenerate $^1\text{H}^{\text{N}}$ and ^{15}N chemical shifts within this tolerance limit. However, the percentage degeneracy decreases as additional chemical shifts ($^1\text{H}^\alpha$, $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, $^{13}\text{C}'$) are included in the calculation. Thus, the chemical shift information of the chosen set of six spins uniquely identifies an amino acid residue. This implies that no two amino acid residues in a given protein will have all the six chemical shifts degenerate.

(B) Number of amino acid residues having non-zero PCSs in Yb^{+3} substituted Calbindin that have a lower r.m.s.d. (calculated using Eqn. [1]; See main text) to the residues other than their own in the diamagnetic state. As is evident, use of information about the $^1\text{H}^{\text{N}}$ and ^{15}N chemical shifts alone renders their assignment ambiguous. It would then require information from 3D NMR experiments used for sequential resonance assignments to arrive at residue specific assignment. On the other hand, this number of residues that have lower r.m.s.d decreases on inclusion of additional chemical shifts ($^1\text{H}^\alpha$, $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, $^{13}\text{C}'$) in Eqn. [1] and thus aid in their rapid assignment. This implies that the pair comprising the PCSs and the corresponding diamagnetic counter-part gets uniquely identified by their lowest r.m.s.d. compared to that of all other possible pairs of residues when 4 or more different chemical shifts are included.

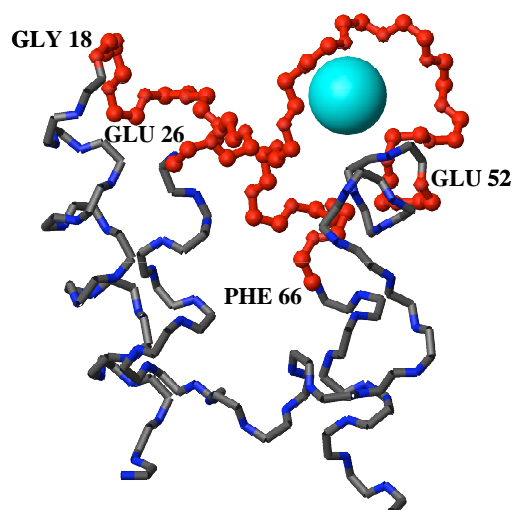


Figure S6: Residues for which PCS were not observed in both conventional 3D NMR experiments and their GFT-NMR counterparts of $[\text{Ca}^{2+}][\text{Yb}^{3+}]\text{Cb}$ sample (1ksm.pdb) are depicted in red color. The cyan color represents the paramagnetic ion (Yb^{3+}). The residues from 18-26 and 52-66 regions show higher broadening in the case of Yb^{3+} . The amide moieties of these residues were found to be located within a radius of $6.9 \pm 2.7 \text{ \AA}$ from Yb^{3+} .

REFERENCES

- [1] Freeman R. Shaped radiofrequency pulses in high resolution NMR. *Prog NMR Spectroscopy* 1998; 32: 59-106.
- [2] Shaka AJ, Lee CJ, Pines A. Iterative schemes for bilinear operators: Application to spin decoupling. *J Magn Reson* 1988; 77: 274-293.
- [3] Shaka AJ, Barker PB, Freeman R. Computer-optimized decoupling scheme for wideband applications and low-level operation. *J Magn Reson* 1985; 64: 547-552.
- [4] Kay LE, Keifer P, Saareinen T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc* 1992; 114: 10663-10665.
- [5] Kim S, Szyperski T. GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. *J Am Chem Soc* 2003; 125: 1385-1393.
- [6] Atreya HS, Szyperski T. Rapid NMR data collection. *Methods Enzymol* 2005; 394: 78-108.
- [7] Shaka AJ. Composite pulses for ultra-broadband spin inversion. *Chem Phys Lett* 1985; 120: 201-205.