J Biomol NMR (2011) 51:253–263
DOI 10.1007/s10858-011-9532-2

ARTICLE

Narrowing the conformational space sampled by two-domain proteins with paramagnetic probes in both domains

Soumyasri Dasgupta · Xiaoyu Hu · Peter H. J. Keizers · Wei-Min Liu · Claudio Luchinat · Malini Nagulapalli · Mark Overhand · Giacomo Parigi · Luca Sgheri · Marcellus Ubbink

Received: 17 May 2011 / Accepted: 14 July 2011 / Published online: 9 August 2011

Abstract Calmodulin is a two-domain protein which in solution can adopt a variety of conformations upon reorientation of its domains. The maximum occurrence (MO) of a set of calmodulin conformations that are representative of the overall conformational space possibly sampled by the protein, has been calculated from the paramagnetism-based restraints. These restraints were measured after inclusion of a lanthanide binding tag in the C-terminal domain to supplement the data obtained by substitution of three paramagnetic lanthanide ions to the calcium ion in the second calcium binding loop of the N-terminal domain. The analysis shows that the availability of paramagnetic restraints arising from metal ions placed on both domains, reduces the MO of the conformations to different extents, thereby helping to identify those conformations that can be mostly sampled by the protein.

Keywords Conformational heterogeneity · Maximum allowed probability · Maximum occurrence · Paramagnetic proteins · Paramagnetic tag · Calmodulin

Introduction

Conformational flexibility is often a crucial feature for proteins to perform their function in solution (Huang and Montelione 2005; Fragai et al. 2006). During biological processes, different kinds of conformational changes may occur, such as side chain rotations, loop motions, interdomain reorientations, intermolecular rearrangements, and random-coil motions in unfolded proteins or protein regions. NMR spectroscopy has long been used for structural and dynamic studies of proteins in solution. In NMR experiments, solution conditions such as temperature, pH and salt concentration can be adjusted to closely mimic the physiological fluid in which the protein performs its function. Most studies on protein dynamics are focused on the analysis of relaxation data (R_1, R_2 and NOE) and provide information on the protein tumbling times and on the presence of local motions (Fischer et al. 1998; Kay 2005) NMR relaxation dispersion experiments can also provide quantitative information on the dynamics of interconverting protein conformation and on lowly populated protein states, when the conformational dynamics occur on the μs–ms time scale (Wang et al. 2007; Eichmüller and Skrynnikov 2007; Vallurupalli et al. 2008; Hass et al. 2010; Korzhnev et al. 2010; Cho et al. 2010). Paramagnetism-based restraints have been shown to monitor the presence of conformational rearrangements among protein domains (Bertini et al. 2004), to detect the presence of minor interconverting conformations (Iwahara and Clore 2006; Tang et al. 2007; Xu et al. 2008), to determine whether

Electronic supplementary material The online version of this article (doi:10.1007/s10858-011-9532-2) contains supplementary material, which is available to authorized users.

S. Dasgupta · X. Hu · C. Luchinat (✉) · M. Nagulapalli · G. Parigi
Magnetic Resonance Center (CERM), University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy
e-mail: luchinat@cerm.unifi.it

P. H. J. Keizers · W.-M. Liu · M. Overhand · M. Ubbink
Leiden Institute of Chemistry, Gorlaeus Laboratories, Leiden University, PO box 9502, 2300 RA Leiden, Netherlands

C. Luchinat · G. Parigi
Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Italy

L. Sgheri
Istituto per le Applicazioni del Calcolo, Sezione di Firenze, CNR, Via Madonna del Piano 10, 50019 Sesto Fiorentino, Italy

regions in the conformational space must be occupied or cannot be occupied by protein complexes (Volkov et al. 2006; Bashir et al. 2010), and to provide information on the maximum occurrence (MO) of any conformation that is sterically allowed (Bertini et al. 2010).

The MO strategy aims to determine the maximum weight that any given conformation can have in any conformational ensemble according to all available experimental data obtained, e.g., through solution NMR or small angle scattering (SAS) measurements (Longinetti et al. 2006; Bertini et al. 2007; Bertini et al. 2010). These measurements in fact provide weighted averages over all the conformations experienced by the system. They cannot be used to recover the actual protein conformational ensemble, but do provide the maximum percent of time that a system can spend in a particular conformation.

Calmodulin (CaM) is a calcium(II) EF-hand protein, which contains two similar globular domains connected by a flexible linker (Barbato et al. 1992). This structural feature makes it easy for the two domains to adopt a variety of different orientations with respect to one another. In order to describe the interdomain conformational variability, lanthanide ions were used as paramagnetic probes, and NMR experiments were performed to obtain pseudocontact shifts (pcs) and residual dipolar couplings (rdc) (Bertini et al. 2004; Bertini et al. 2007; Schmitz et al. 2008; Su and Otting 2010). These data were used as restraints to calculate the conformations with largest MO and to analyze the different MO of the possible protein conformations. The N60D CaM mutant was actually used to selectively substitute the Ca^{2+} ion located in the second binding site of the N-terminal domain with a paramagnetic lanthanide ion (Tb^{3+}, Tm^{3+} or Dy^{3+}) (Bertini et al. 2003). The same analysis was performed to detect the conformational heterogeneity of CaM bound to α-synuclein (Bertini et al. 2007) or to a peptide from the myelin basic protein (MBP, unpublished results from the CERM laboratory). Pcs and rdc restraints were also used to detect slight conformational changes in CaM when bound to peptides representing the interaction sequence of two protein partners, the death-associated protein kinase (DAPk) and the DAPk-related protein 1 (DRP-1), on passing from crystal to solution (Bertini et al. 2009). The results suggest that the two domains are relatively flexible with respect to one another in free CaM and that mobility changes after target peptide binding (Mal et al. 2002; Maximciuc et al. 2006).

Paramagnetism-based restraints provide an advantage for the study of domain dynamics because of the possibility to retrieve the magnetic susceptibility anisotropy tensors from the pcs values collected for nuclei belonging to the domain in which the metal ion is coordinated (the N-terminal domain in this case). The pcs and rdc values collected for the nuclei belonging to the other domain act as reporters of the interdomain conformational variability. In fact, they are the weighted average of the pcs and rdc values corresponding to all sampled conformations, and these values are determined by the magnetic susceptibility anisotropy tensor calculated for the metal bearing domain.

It is expected that the MO value calculated for each conformation decreases towards the actual probability when the number of independent experimental restraints is increased (Bertini et al. 2010). However, if the set of experimental restraints is restricted, some conformations sampled very differently by the protein may result with similar MO. The MO values of less probable conformations are thus expected to decrease more than those of the most probable conformations when new independent restraints are included in the calculation.

In this work we analyze how MO values are affected by the availability of restraints provided by paramagnetic metal ions located in both domains of CaM. Pcs and rdc collected for the N-terminal domain, when the paramagnetic metal is placed in the C-terminal domain of CaM, actually represent independent information and provide a different perspective on the protein conformational variability. Therefore, when the paramagnetic restraints obtained from both the metal positions are used together, the difference in the MO values of highly occurring and lowly occurring conformations is expected to increase. Furthermore, besides providing additional information on the relative position of the domains, the addition of these restraints could also remove some of the possible "ghost" solutions determined by the degeneracy of pcs and rdc restraints (Longinetti et al. 2006; Bertini et al. 2007).

In order to place a metal ion in the C-terminal domain of CaM, the Caged Lanthanide NMR Probe 5 (CLaNP-5) (Keizers et al. 2008) was attached to the H107C/N111C CaM mutant. This tag was chosen because it can bind rigidly to the protein backbone through two cysteine residues (Keizers et al. 2007). Rigid binding (Su et al. 2006; Su et al. 2008; Vlasie et al. 2008; Zhuang et al. 2008; Häussinger et al. 2009) is essential to obtain the correct magnetic susceptibility anisotropy tensor from the pcs of the attached domain and to easily interpret the pcs of the other domain (Bertini et al. 2008).

Materials and methods

Protein preparation

^{15}N labeled N60D CaM was purchased from ProtEra s.r.l. (Florence, Italy, www.proterasrl.com). The NMR samples were prepared in 20 mM MES, 200 mM KCl, pH 6.8. For the attachment of the ClaNP-5 tag, the H107C/N111C mutations were introduced in wild type CaM via site-

directed mutagenesis. ^{15}N labeled His-tagged H107C/N111C CaM was expressed in E. coli BL21(DE3) Gold cells and purified with Ni–NTA column and size exclusion chromatography in the same buffer as the N60D mutant. The His-tag was at the N-terminus of CaM. The entire process of tag attachment was performed under reducing conditions. The Ca^{2+}-CaM mutant was incubated with 5 mM DTT for 30 min to reduce all possible disulfide bridges and ensure that the protein existed in monomeric state. DTT was then washed out in the glove box to ensure the presence of reduced conditions. The protein was diluted to a concentration of 30 μM. Seven equivalents of Ln^{3+}-loaded CLaNP-5 ($Ln^{3+} = Lu^{3+}$, Yb^{3+} and Tm^{3+}) was added to it in the glove box. The mixture was incubated overnight at 4°C for the reaction to reach completion. To separate the tagged monomeric protein from aggregates and free tag present in solution, a purification was performed using a Superdex 200 gel filtration column. Approximately three-fourth of the total protein was found to be monomeric and reacted while the remaining one-fourth formed aggregates. In all the above steps of tagging, the CaM mutant was prepared in 20 mM MES, 200 mM KCl, 20 mM $CaCl_2$, pH 6.8. An excess of calcium was always used in the buffer to avoid exchange of Ca^{2+} from the binding sites with any free Ln^{3+} ions present in the tag solution. ^{1}H-^{15}N HSQC spectra of monomeric Ln^{3+}-CLaNP-5 CaM samples were checked immediately after gel filtration. The samples were monitored for a period of 15 days. No changes in the spectrum were observed with time. Further, non-reducing gels showed that only monomers were present.

NMR Measurements

All NMR experiments were performed at 298 K. ^{15}N labeled N60DCaM (0.4 mM) was titrated to $(Ca_2)_N(Ca_2)_C$-CaM and $(CaLn)_N(Ca_2)_C$-CaM ($Ln^{3+} = Tm^{3+}$, Tb^{3+} and Dy^{3+}) by addition of small amounts of calcium(II) and subsequently lanthanide(III) solutions. The titrations were performed by following the ^{1}H-^{15}N HSQC spectra at 700 MHz as previously reported (Bertini et al. 2003). ^{1}H-^{15}N IPAP HSQC spectra were also acquired to obtain the rdc values.

^{1}H-^{15}N HSQC and IPAP-HSQC spectra of Ln^{3+}-CLaNP-5 Ca_4CaM ($Ln^{3+} = Lu^{3+}$, Yb^{3+} and Tm^{3+}) were acquired at 298 K and 700 MHz. HNCO (Muhandiram and Kay 1994), HNCA (Bax and Ikura 1991), CBCACONH (Muhandiram and Kay 1994) and HNCACB (Kay et al. 1990) experiments at 500 MHz were performed on ^{15}N and ^{13}C labeled H107C/N111C CaM tagged with Ln^{3+}-CLaNP-5 ($Ln^{3+} = Lu^{3+}$ and Yb^{3+}) to obtain the backbone assignment. The backbone resonance signals of Tm^{3+}-CLaNP-5 Ca_4CaM were assigned based on the assigned ^{1}H-^{15}N HSQC spectra of the diamagnetic Lu^{3+} form and of the paramagnetic Yb^{3+} form.

Pcs data were obtained from the difference in ^{1}H chemical shift between corresponding nuclei in the paramagnetic and diamagnetic CaM derivatives. Rdc data were obtained as the difference in the doublet splitting in the indirect ^{15}N dimension in ^{1}H-^{15}N IPAP-HSQC spectra between the paramagnetic form and the diamagnetic form.

Maximum occurrence (MO) calculation of CaM conformations

The pcs values measured for the domain in which the paramagnetic metal is located were used to calculate the magnetic susceptibility anisotropy tensors of the different metals. For the N60D $(CaLn)_N(Ca_2)_C$-CaM samples, the program FANTASIAN (Bertini et al. 2002) was used to determine the anisotropy tensors. For the Ln^{3+}-CLaNP-5 Ca_4CaM samples the programs FANTASIAN and PARAMAGNETICCYANA-2.1 (Banci et al. 1996; Balayssac et al. 2006) were used to determine the anisotropy tensors and the position of the metal ions with respect to the backbone of the C-terminal domain. These tensors were then fixed in all subsequent calculations.

The program for the calculation of the MO of any given conformation (Bertini et al. 2010) was modified to incorporate paramagnetic restraints arising from metal ions located in both protein domains. In this way, pcs and rdc measured for the C-terminal domain when the paramagnetic metal is located in the N-terminal domain could be analyzed together with pcs and rdc measured for the N-terminal domain when the metal is located in the tagged C-terminal domain. More details are reported in the Supplementary Material.

A total of 400 conformations with different inter-domain orientations were obtained through the program RANCH (Bernado et al. 2007; Bertini et al. 2010). The MO value of each conformation was calculated from the paramagnetic restraints (pcs and rdc) obtained for the different lanthanides located in either the N- or the C-terminal domain. Taking each conformation as a starting point, a simulated annealing minimization was performed to generate an ensemble with a maximum of 14 other conformations which, together with the starting conformation, provides the best fit of the experimental data. In the minimization, both the translation and orientation parameters defining the protein conformations and the weights of the latter were allowed to change. This fit was performed by minimizing a target function (TF) defined as the sum of the squared difference between the values obtained from the weighted average of pcs and rdc calculated for all conformations of the ensemble and the corresponding experimental data (see Eq. S2). The weight of each starting conformation was

fixed in the minimization. Several calculations were performed by changing the weight of the starting conformation. The MO value of this conformation was calculated as the weight for which the TF is 10% larger than the minimum value. More details are reported in the Supplementary Material.

Results and discussion

N60D $(CaLn)_N(Ca_2)_C$-CaM

Pcs and rdc for $(CaLn)_N(Ca_2)_C$-CaM (Ln^{3+} = Tm^{3+}, Tb^{3+} and Dy^{3+}) were measured in buffer 20 mM MES, 200 mM KCl and pH 6.8, the same used for the CLaNP-5 tagged CaM samples. Some rdc values measured in these conditions differed outside the estimated error of 1.5 Hz throughout the domain from those previously measured in 400 mM KCl, pH 6.5 (Bertini et al. 2007).

The structure of the CaM domains in solution was fixed to the coordinates deposited in PDB 1J7O and 1J7P (Chou et al. 2001). These structures were chosen because they are refined with an extensive use of rdc derived by external orienting media. The position of the lanthanide ions in the N60D $(CaLn)_N(Ca_2)_C$-CaM samples was fixed to the coordinates of the calcium ion in the second binding loop of the N-terminal domain.

The best fit of the pcs of the N-terminal domain amide protons to the protein structure provided the magnetic susceptibility anisotropy tensors reported in Table 1. They are in good agreement with those previously obtained (Bertini et al. 2004; Bertini et al. 2007). The quality of the fit is good (squared Pearson correlation coefficient $R^2 = 0.95$) as shown in Fig. 1a.

The best fit of the rdc of the amide protons of the C-terminal domain to the protein structure provided the anisotropy tensors $\overline{\Delta\chi}$ reported in Table 1. As previously found, these $\overline{\Delta\chi_{ax}}$ values are sizably smaller than the $\Delta\chi_{ax}$ values obtained from the fit of the pcs of the N-terminal domain nuclei. The good quality of the fits ($R^2 = 0.88$), shown in Fig. 1b, however indicate that the data are in good agreement with the protein structure, which thus moves as a rigid body, so that the obtained tensors are averages of the magnetic susceptibility anisotropy tensors positioned in the N-terminal domain as seen from a nucleus in the C-terminal domain.

CLaNP-5 Tagged CaM

Paramagnetic ions were placed in the C-terminal domain of CaM using the CLaNP-5 tag (Keizers et al. 2008). The mutation H107C/N111C was performed in order to allow the tag to be attached to the protein through sulfur bonds. The residues to be mutated were chosen (i) positioned on

Table 1 Magnetic susceptibility anisotropies and average tensors of the different lanthanides in the second calcium binding loop of the N-terminal domain of CaM and in the ClaNP-5 CaM

	$\Delta\chi_{ax}$ ($10^{-32}m^3$)	$\Delta\chi_{rh}$ ($10^{-32}m^3$)	Euler angles (ϕ, θ, ω) (degrees)		
Metals in the N-terminal domain—from N-terminal domain pcs					
Tb	36 ± 1	−16.5 ± 0.9	160 ± 2[a]	29 ± 3[a]	104 ± 2[a]
Tm	31 ± 3	−9 ± 1	−153 ± 2[a]	−151 ± 6[a]	76 ± 2[a]
Dy	36 ± 1	−13 ± 2	75 ± 5[a]	−41 ± 6[a]	18 ± 1[a]
Metals in the C-terminal domain—from C-terminal domain pcs					
Yb	9.7 ± 0.4	−3 ± 2	−129 ± 2[b]	98 ± 11[b]	76 ± 3[b]
Tm	56 ± 1	−7 ± 3	−131 ± 1[b]	65 ± 11[b]	72 ± 1[b]
	$\overline{\Delta\chi_{ax}}$ ($10^{-32}m^3$)	$\overline{\Delta\chi_{rh}}$ ($10^{-32}m^3$)	Euler angles (ϕ, θ, ω)		
Metals in the N-terminal domain—from C-terminal domain rdc					
Tb	3.0 ± 0.1	−2.7 ± 0.1	72 ± 1[b]	−90 ± 1[b]	63 ± 2[b]
Tm	1.4 ± 0.3	−1.0 ± 0.2	52 ± 6[b]	−63 ± 11[b]	52 ± 6[b]
Dy	2.5 ± 0.1	−1.5 ± 0.1	57 ± 11[b]	−92 ± 11[b]	142 ± 3[b]
Metals in the C-terminal domain—from N-terminal domain rdc					
Tm	2.8 ± 0.3	−2.4 ± 0.3	23 ± 11[a]	34 ± 11[a]	43 ± 3[a]

The Euler angles are in the ZYZ convention. Errors have been calculated with a Monte Carlo approach (standard deviation of 2000 calculated values obtained after removal of 35% of randomly selected data)

[a] With respect to structure 1J7O

[b] With respect to structure 1J7P

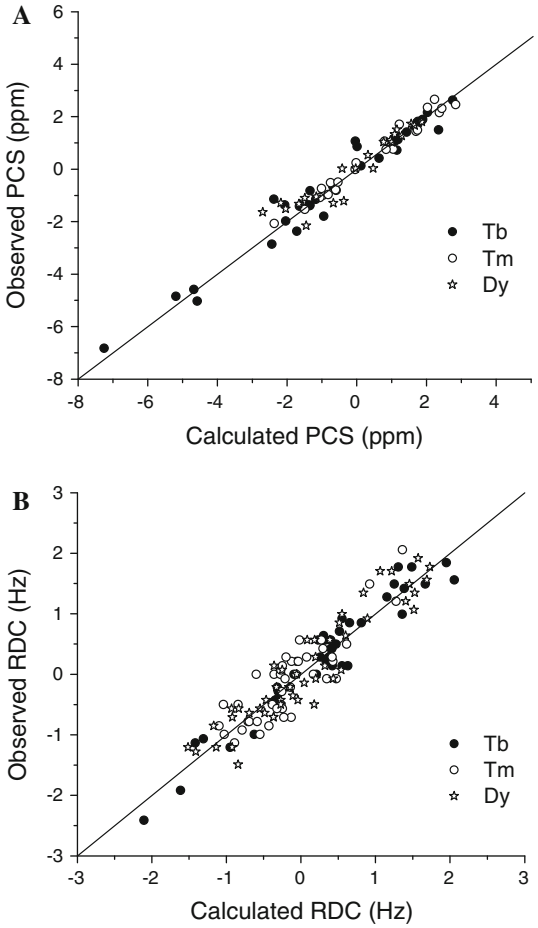

Fig. 1 **a** Observed versus calculated values of pcs of N-terminal domain nuclei for the terbium(III), thulium(III) and dysprosium(III) ions substituted in the second binding site of CaM N-terminal domain. **b** Observed versus calculated values of rdc of C-terminal domain HN for the terbium(III), thulium(III) and dysprosium(III) ions substituted in the second binding site of CaM N-terminal domain

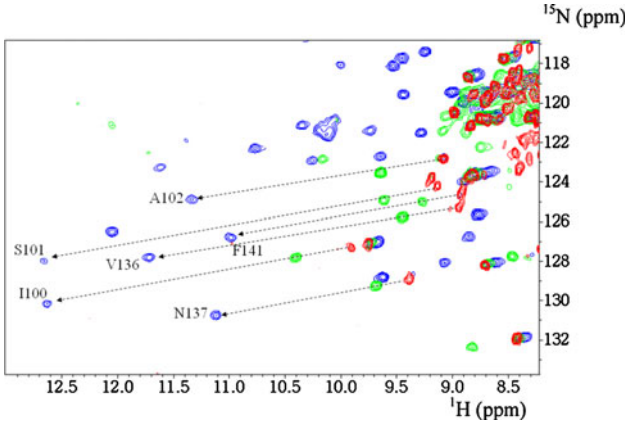

Fig. 2 Superposition of a region of the ^{1}H-^{15}N HSQC spectra of Lu^{3+} (*red*), Yb^{3+} (*green*) and Tm^{3+} (*blue*) ClaNP-5 CaM

one helix (the second of the C-terminal domain) in order to provide rigidity to the CLaNP-5 tag; (ii) so that the cysteine side chains are exposed on the surface of the structure with the C^{β} atoms pointing away and the C^{α} atoms neither closer than 6 Å, nor farther than 10 Å from one another; (iii) in order to attach the tag in a position far enough from the N-terminal domain to avoid steric clashes that may affect the conformational heterogeneity of the protein.

Lu^{3+}-CLaNP-5 was used as the diamagnetic reference. The ^{1}H-^{15}N HSQC spectrum of Lu^{3+}-CLaNP-5 H107C/N111C CaM is similar to that of Ca_4CaM with differences limited to the residues in close proximity to ClaNP-5, indicating that the protein structure is maintained after binding of the tag (see Figure S3). Both Yb^{3+} and Tm^{3+} CLaNP-5 induced positive paramagnetic shifts, which in the Tm^{3+} form are much larger than in the Yb^{3+} form (see Fig. 2), due to the larger magnetic susceptibility anisotropy of Tm^{3+}.

The program PARAMAGNETICCYANA-2.1 was first used to determine the position of the metal ions with respect to the C-terminal domain structure (PDB 1J7P (Chou et al. 2001)) using typical values for the magnetic susceptibility anisotropies and the observed pcs measured in the presence of Tm^{3+} or Yb^{3+} bound to the tag. The magnetic susceptibility anisotropy values were then refined using the program FANTASIAN through the best fit of the pcs to the C-terminal domain structure and the calculated position of the lanthanides. In the calculations, the metal positions of the two lanthanides were constrained to coincide. The two programs were cycled iteratively until convergence of both the metal position and the susceptibility anisotropy tensors was reached.

The observed pcs values fit very well versus the calculated data (Fig. 3a, $R^2 = 0.996$). The resulting axial and rhombic components of the magnetic susceptibility anisotropy tensors as well as the Euler angles providing the orientation of the tensors are reported in Table 1. These results are in agreement with the values reported in (Keizers et al. 2008). The *z*-axes of the Yb^{3+} and Tm^{3+} tensors are parallel, and the *x*- and *y*-axes of the two metals experience a difference in the orientation of only 21° (Fig. 4). The calculations show that the metals are located at about the same distance from the protein backbone as in (Keizers et al. 2008). The positions of Yb^{3+} and Tm^{3+} are in fact similar for the two metals and at distances of 8.3 and 6.5 Å from the C^{α} atoms of residues Cys-107 and Cys-111, respectively. The magnitude of the calculated anisotropies and the correct definition of the lanthanide position indicates that the CLaNP-5 probe binds Ca_4CaM rigidly.

The rigidity of the tag is confirmed by the rdc values measured for the C-terminal domain amide protons. These rdc, when fitted to the domain structure ($R^2 = 0.90$), provide a nice agreement with the values calculated from the best fit tensors. This is clear in particular for the Yb^{3+} sample, the C-terminal domain nuclei of which are less

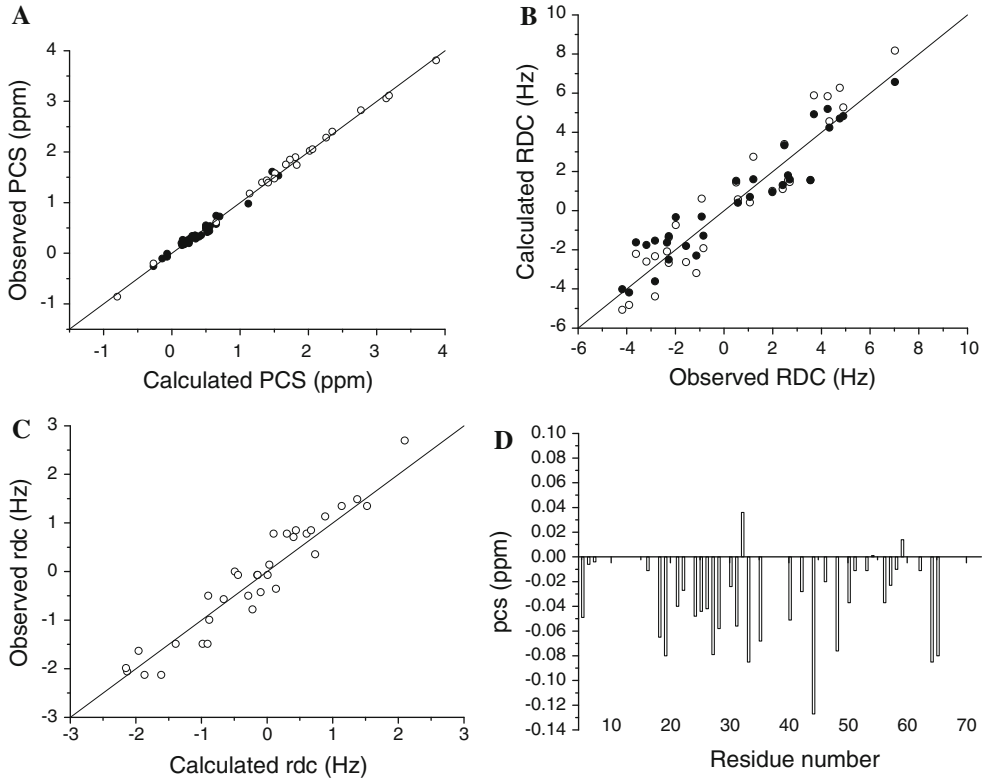

Fig. 3 **a** Observed versus calculated pcs of C-terminal domain nuclei for the Yb^{3+} (*filled circle*) and Tm^{3+} (*circle*) ClaNP-5 Ca_4CaM. **b** Observed versus calculated rdc of C-terminal domain HN for the Yb^{3+}-ClaNP-5 Ca_4CaM. The solid symbols (*filled circle*) indicate the values calculated from the best fit parameters ($\Delta\chi_{ax} = 9.1 \times 10^{-32}$ m^3, $\Delta\chi_{rh} = -2.8 \times 10^{-32}$ m^3), the open symbols (*circle*) indicate the values calculated using the pcs-derived tensor. **c** Observed versus calculated rdc of N-terminal domain HN for the Tm^{3+} ClaNP-5 Ca_4CaM. **d** Observed pcs of the N-terminal domain for Tm^{3+}-ClaNP-5 Ca_4CaM

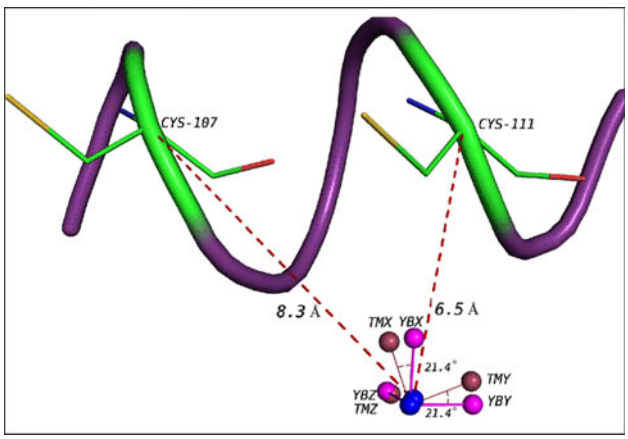

Fig. 4 The lanthanide ions (in *blue*) are placed at a distance of 8.3 and 6.5 Å from the C^{α} atoms of residues Cys-107 and Cys-111, respectively. The orientations of the magnetic susceptibility anisotropy tensors are shown for the Yb^{3+} (*magenta*) and Tm^{3+} (*brown*) metals. The z axes of the anisotropy tensors of the two metals are essentially coinciding; the angle between the x (and y) axes of the anisotropy tensors of the two metals is 21.4°

affected by paramagnetic line broadening due to the smaller susceptibility tensor of Yb^{3+} than that of Tm^{3+}, resulting in a smaller Curie relaxation (Fig. 3b). The best fit anisotropy tensors ($\Delta\chi_{ax} = 9.1 \pm 0.4 \times 10^{-32}$ m^3, $\Delta\chi_{rh} = -2.8 \pm 0.5 \times 10^{-32}$ m^3) are actually very similar to those calculated from the pcs, indicating that no (or very modest) reduction due to motional averaging occurs. Motional averaging is in fact expected to be much more effective, if present, for the rdc-derived tensors than for the pcs-derived tensors (Assfalg et al. 2003; Banci et al. 2004). This is because the different conformations experienced by the system in the presence of modest mobility are characterized by relatively different nitrogen-amide proton orientations (on which the rdc depend) whereas the metal-nuclei vectors (on which the pcs depend) remain similar. The Euler angles ϕ, θ, ω defining the orientation of the rdc-derived tensor are equal to −132°, 104° and 77°, respectively, indicating that the pcs-derived and the rdc-derived tensor axes are within 6° with respect to one another. When the measured rdc are compared with the rdc values calculated from the tensor derived from pcs, a good agreement is indeed observed (Fig. 3b, $R^2 = 0.89$), the differences between calculated and observed data being all within 2 Hz. The rdc-derived tensor for Tm^{3+} is defined by the axial and rhombic anisotropy values of 49.5×10^{-32} m^3 and -18.3×10^{-32} m^3 and by the Euler angles −127°, 72°

and 75°, so that the z axis forms an angle of 5° with respect to the z axis of the pcs-derived tensor, and the x and y axes form angles of 10° with respect to the corresponding axes of the pcs-derived tensor. The rdc-derived $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ values provided above for both the Yb^{3+} and Tm^{3+} tags would be even larger if the presence of an order parameter smaller than 1 is considered.

Due to the much smaller magnetic susceptibility anisotropy of Yb^{3+} with respect to Tm^{3+}, a reliable set of pcs and rdc for the N-terminal nuclei could be observed only for the Tm^{3+} derivative. The rdc, quite reduced with respect to those measured for the C-terminal domain, can be described by a single average tensor. The latter was obtained by fitting the rdc to the N-terminal domain CaM structure. As expected, the calculated average tensor is sizably smaller than that obtained from the C-terminal domain pcs data (Table 1) as a consequence of extensive orientation averaging. The fit of the observed rdc versus the calculated values is shown in Fig. 3c ($R^2 = 0.92$). Pcs of the N-terminal domain ^{N}H nuclei were also collected, and they are reported in Fig. 3d.

MO analysis

Calculations of the MO values for 400 CaM conformations randomly generated as representative of all possible conformations (Bertini et al. 2010) have been performed using the derived magnetic susceptibility anisotropy tensors and the pcs and rdc data observed for the domain without the paramagnetic metal. The MO value of each of the 400 conformations was obtained from the largest weight that the conformation can have when included in any possible ensemble together with 14 other conformations, with different weights, freely chosen among all possible conformations without overlaps between the two domains and with a maximum distance between the last residue in the N-terminal domain (78) and the first residue in the C-terminal domain (81) not exceeding that corresponding to the fully extended conformation. These ensembles were found as the families of structures in best agreement with the experimental data by minimizing the target function (TF), defined as a measure of the disagreement from the experimental data of the average pcs and rdc calculated according to the ensemble itself (see Materials and Methods and Supplementary Material for further details). It was checked that increasing the number of the other conformations above 14 does not decrease the TF; therefore, this number of conformations was chosen for the calculations. During the minimization, the weight of the fixed conformation (one of the 400 randomly generated conformations) was set to different values. The MO of such conformation was defined as the largest weight for which the TF is smaller than a given threshold. The latter was defined 10% larger than the lowest possible TF value (Bertini et al. 2007).

The results obtained for the Tb^{3+}, Tm^{3+} and Dy^{3+} ions positioned in the N-terminal domain provide the map of MO values shown in Fig. 5a. The position of the C-terminal domain of CaM is indicated by an orientation tensor centered in the center of mass of the C-terminal domain, color-coded with respect to the MO of the corresponding conformation from blue (MO lower than 15%) to red (MO greater than 35%). Different orientations of the tensor reflect different orientations of the CaM C-terminal domain

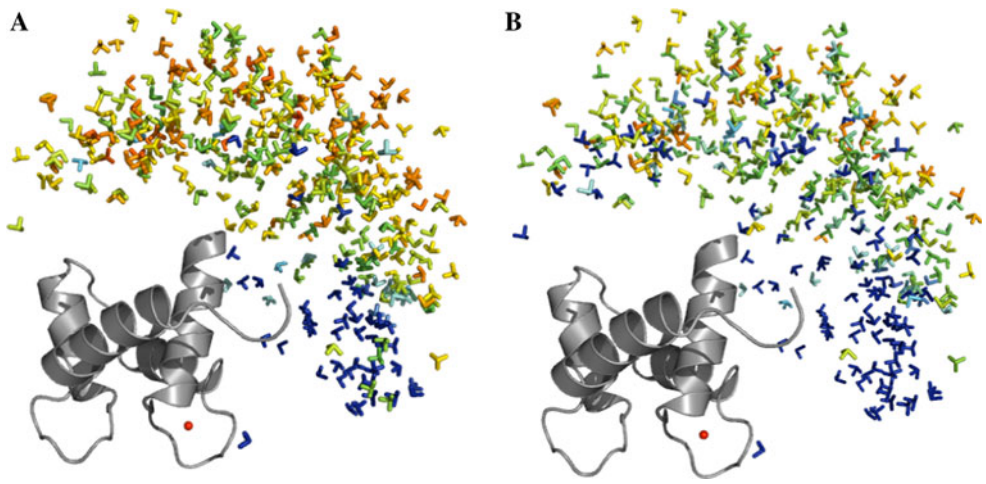

Fig. 5 Orientation tensors positioned at the centre of mass of the C-terminal domain, color-coded with respect to the MO of the corresponding conformation from *blue* (lower than 15%) to *red* (greater than 35%) for 400 structures generated randomly with RANCH. The N-terminal domain is depicted as a cartoon, with the position of the Ln^{3+} ion as a red sphere. Panel **a** shows the results obtained from pcs and rdc arising with metals in the N-terminal domain; panel **b** shows the results obtained when pcs and rdc of Tm^{3+}-ClaNP-5 Ca_4CaM are also included

with respect to the N-terminal domain. The minimum value for the TF was calculated by generating structure ensembles without forcing the ensembles to include any fixed conformation without predefined weight, yielding 0.203, so that a threshold of 0.223 was chosen. The overall distribution of the MO values is indeed relatively similar to that previously calculated for data acquired with a higher salt concentration in solution and with inclusion of SAXS restraints (Fig. 3a of (Bertini et al. 2010)). As already seen, the conformations having the C-terminal domain in the lower right quadrant of the frame have in general low MO, while the conformations with the highest MO are clustered in the central part of the distribution, corresponding to relatively—but not fully—elongated conformations. On average, the MO values are in this case somewhat larger than those previously calculated in (Bertini et al. 2010) due to the absence of the SAXS restraints in the present calculations. This is advantageous for the purpose of this work as it allows us to better estimate the impact of another set of restraints obtained by placing a paramagnetic ion in the other protein domain. The calculations were also performed with inclusion of the SAXS restraints together with the paramagnetic restraints [although SAXS data were acquired in different conditions (400 mM KCl, pH 6.5)] with the aim of showing that differences in the MO values can be still appreciated when the restraints arising by placing the paramagnetic ion in the other protein domain are added (see Supplementary Material, Figure S4).

The set of pcs and rdc data acquired for the N-terminal domain when the Tm^{3+} tag is placed in the C-terminal domain of the CaM mutant was then added to the previous data, and MO calculations were repeated for the same 400 conformations. In this case, the minimum TF was 0.224, so that a threshold of 0.246 was chosen. The number of conformations with an MO of less than 0.1 increased from 19 to 46, and the number of conformations with MO values of less than 0.25 increased from 99 to 213 (Fig. 6). This indicates that pcs and rdc of the Tm-tagged protein are effective in decreasing by different extents the MO of different conformations. Therefore, a larger number of less sampled or even not actually sampled conformations could be identified as well as a smaller number of conformations with largest MO. The number of conformations with MO values in the range from the maximum value to a value 15% smaller than the maximum decreased from 120 (MO range of 27.9–32.8%) to 59 (MO range 27.2–32.0%) upon addition of the restraints obtained for the C-terminal tagged protein. A color-coded map of the MO values for the selected conformations is shown in Fig. 5b. Quite a large reduction in MO is evident for conformations with the C-terminal domain in the upper left side of the frame, although the effect is present throughout the whole conformational space. Figure 7 shows the TF values for all the

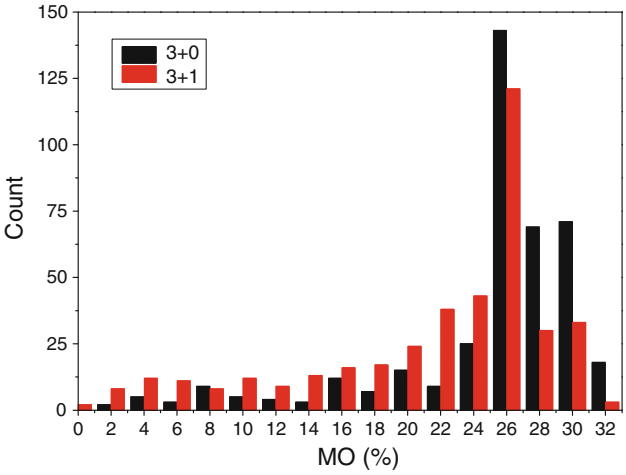

Fig. 6 Number of conformations as a function of the MO values calculated from the pcs and rdc arising with metals in the N-terminal domain (3 + 0 case) and by including pcs and rdc of Tm^{3+}-ClaNP-5 Ca_4CaM (3 + 1 case)

conformations as a function of their weight. The substantial differences in the weight at which the TF value starts increasing reflect the resulting markedly different MO.

The MO values obtained by positioning the C-terminal domain in such a way as to match the conformations observed in the crystal structures of free CaM (Babu et al. 1988; Fallon and Quiocho 2003) and of CaM in complexes (Meador et al. 1992; Meador et al. 1993; Kurokawa et al. 2001; Schumacher et al. 2001; Aoyagi et al. 2003; Maximciuc et al. 2006; Bertini et al. 2009) were also calculated, and the latter conformations are reported in Fig. 8, color-coded according to their MO. The inclusion among the experimental restraints of the pcs and rdc of the Tm-tagged protein actually reduces the MO of many of these conformations sizably. All the regularly closed structures have quite small MO values. The crystal structure corresponding to the extended CaM conformation observed for the free protein (1CLL) has MO of 27%. The crystal structure with largest MO, equal to 31%, is the structure of CaM in its adduct with the SK channel CaM-binding domain (Schumacher et al. 2001) (1G4Y), where CaM binds two Ca^{2+} ions only in the N-terminal domain and the protein adopts a less compact conformation with respect to other CaM-peptide complexes by the extension of the CaM linker region.

It is theoretically predicted that rdc arising from metals located in different domains are not fully independent (Sgheri 2010). Therefore, in terms of independent elements of the anisotropy tensors $\overline{\Delta\chi}$ available from the same number of metals, the total information content is larger when the metals are all located in the same domain. The present calculations show, however, that experimental data from metals located in both domains can improve the ranking of the different conformations according to their

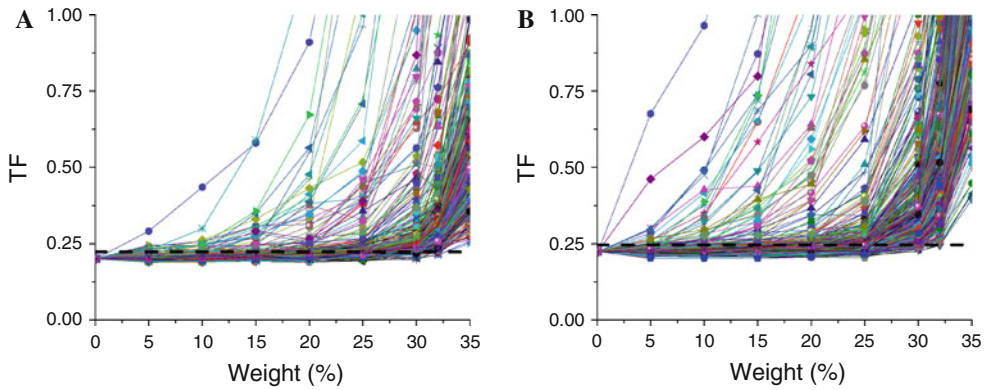

Fig. 7 Target Function describing the best agreement between any conformational ensemble and the experimental data as a function of the weight of the conformation for which the MO is calculated. The *curves* corresponds to the 400 conformations used. The MO is defined by the intersection between the TF curve and the threshold (*dashed line*) chosen to be 10% larger than the smallest TF. Panel **a** shows the results obtained from pcs and rdc arising with metals in the N-terminal domain; panel **b** shows the results obtained when pcs and rdc of Tm^{3+}-ClaNP-5 Ca_4CaM are also included

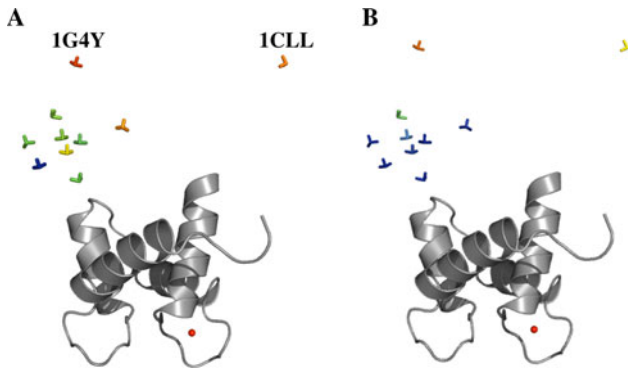

Fig. 8 Orientation tensors centered in the center-of-mass of the C-terminal domain, color-coded with respect to the MO of the corresponding conformation from *blue* (<15%) to *red* (>35%), for crystallographic CaM structures (PDB codes 1CLL, 1PRW, 1CDL, 1CDM, 1G4Y, 1IQ5, 1NIW, 1YR5, 2BCX, 2X0G). Panel **a** shows the results obtained from pcs and rdc arising with metals in the N-terminal domain; panel **b** shows the results obtained when pcs and rdc of Tm^{3+}-ClaNP-5 Ca_4CaM are also included

MO values. The use of pcs in the calculations may actually compensate for the smaller information content of rdc on the conformational distribution of the protein. Furthermore, the redundant information given by the rdc induced by metals placed in different domains produces relations among the average rdc-derived tensors and the pcs-derived magnetic susceptibility anisotropy tensors (Sgheri 2010) that are well suited to assess the consistency of the data. Therefore, the use of pcs and rdc data arising from metals placed in both protein domains of CaM (i) permits an internal check of the quality of the data from the overall agreement of all sets of the experimental rdc with those calculated from the best-fit conformational ensembles and (ii) provides a better discrimination among the different protein conformations depending on the calculated MO values.

Conclusions

We have shown that the simultaneous use of paramagnetism-based restraints arising from paramagnetic metal ions located in different domains of proteins experiencing interdomain mobility is quite informative for the determination of the maximum occurrence of protein conformations. In the case of CaM we have shown that the MO of several conformations is sizably reduced by the addition of pcs and rdc arising from the presence of a single metal ion rigidly attached to the C-terminal domain with respect to the MO values calculated using only data obtained from three metal ions placed in the N-terminal domain of the protein. As a result, the conformations likely experienced by the protein can be more accurately mapped.

Acknowledgments This work has been supported by MIUR-FIRB contracts RBLA032ZM7, RBRN07BMCT and RBIP06LSS2, by the European Commission, contracts Bio-NMR n. 261863, East-NMR n. 228461, SPINE2-COMPLEXES 031220, and We-NMR 261572, and by the Netherlands Organisation for Scientific Research (NWO), grants 700.58.405 (P.H.J.K.) and 700.58.441 (W.M.L. and M.U.).

References

Aoyagi M, Arvai AS, Tainer JA, Getzoff ED (2003) Structural basis for endothelial nitric oxide synthase binding to calmodulin. EMBO J 22:766–775

Assfalg M, Bertini I, Turano P, Mauk AG, Winkler JR, Gray BH (2003) 15 N–1H residual dipolar coupling analysis of native and alkaline-K79A S. cerevisiae cytochrome c. Biophys J 84:3917–3923

Babu YS, Bugg CE, Cook WJ (1988) Structure of calmodulin refined at 2.2 Å resolution. J Mol Biol 204:191–204

Balayssac S, Bertini I, Luchinat C, Parigi G, Piccioli M (2006) ^{13}C direct detected NMR increases the detectability of residual dipolar couplings. J Am Chem Soc 128:15042–15043

Banci L, Bertini I, Bren KL, Cremonini MA, Gray HB, Luchinat C, Turano P (1996) The use of pseudo contact shifts to refine solution structures of paramagnetic metalloproteins: Met80Ala cyano-cytochrome *c* as an example. J Biol Inorg Chem 1:117–126

Banci L, Bertini I, Cavallaro G, Giachetti A, Luchinat C, Parigi G (2004) Paramagnetism-based restraints for Xplor-NIH. J Biomol NMR 28:249–261

Barbato G, Ikura M, Kay LE, Pastor RW, Bax A (1992) Backbone dynamics of calmodulin studied by ^{15}N relaxation using inverse detected two-dimensional NMR spectroscopy; the central helix is flexible. Biochemistry 31:5269–5278

Bashir Q, Volkov AN, Ullmann GM, Ubbink M (2010) Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome *c* and cytochrome *c* peroxidase. J Am Chem Soc 132:241–247

Bax A, Ikura M (1991) An efficient 3D NMR technique for correlating the proton and 15 N backbone amide resonances with the alpha-carbon of the preceding residue. J Biomol NMR 1:99–104

Bernado P, Mylonas E, Petoukhov MV, Blackledge M, Svergun DI (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc 129:5656–5664

Bertini I, Luchinat C, Parigi G (2002) Paramagnetic constraints: an aid for quick solution structure determination of paramagnetic metalloproteins. Concepts Magn Reson 14:259–286

Bertini I, Gelis I, Katsaros N, Luchinat C, Provenzani A (2003) Tuning the affinity for lanthanides of calcium binding proteins. Biochemistry 42:8011–8021

Bertini I, Del Bianco C, Gelis I, Katsaros N, Luchinat C, Parigi G, Peana M, Provenzani A, Zoroddu MA (2004) Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. Proc Natl Acad Sci USA 101:6841–6846

Bertini I, Gupta YK, Luchinat C, Parigi G, Peana M, Sgheri L, Yuan J (2007) Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. J Am Chem Soc 129:12786–12794

Bertini I, Luchinat C, Parigi G, Pierattelli R (2008) Perspectives in NMR of paramagnetic proteins. Dalton Trans 2008:3782–3790

Bertini I, Kursula P, Luchinat C, Parigi G, Vahokoski J, Willmans M, Yuan J (2009) Accurate solution structures of proteins from X-ray data and minimal set of NMR data: calmodulin peptide complexes as examples. J Am Chem Soc 131:5134–5144

Bertini I, Giachetti A, Luchinat C, Parigi G, Petoukhov MV, Pierattelli R, Ravera E, Svergun DI (2010) Conformational space of flexible biological macromolecules from average data. J Am Chem Soc 132:13553–13558

Cho JH, O'Connell N, Raleigh DP, Palmer AG III (2010) Phi-value analysis for ultrafast folding proteins by NMR relaxation dispersion. J Am Chem Soc 132:450–451

Chou JJ, Li S, Klee CB, Bax A (2001) Solution structure of Ca^{2+} calmodulin reveals flexible hand-like properties of its domains. Nature Struct Biol 8:990–997

Eichmüller C, Skrynnikov NR (2007) Observation of μs time-scale protein dynamics in the presence of Ln^{3+} ions: application to the N-terminal domain of cardiac troponin C. J Biomol NMR 37:79–95

Fallon JL, Quiocho FA (2003) A closed compact structure of native Ca^{2+}-calmodulin. Structure 11:1303–1307

Fischer MWF, Zeng L, Majumdar A, Zuiderweg ERP (1998) Characterizing semi local motions in proteins by NMR relaxation studies. Proc Natl Acad Sci USA 95:8016–8019

Fragai M, Luchinat C, Parigi G (2006) "Four-dimensional" protein structures: examples from metalloproteins. Acc Chem Res 39:909–917

Hass MAS, Keizers PHJ, Blok A, Hiruma Y, Ubbink M (2010) Validation of a lanthanide tag for the analysis of protein dynamics by paramagnetic NMR spectroscopy. J Am Chem Soc 132:9952–9953

Häussinger D, Huang J, Grzesiek S (2009) DOTA-M8: an extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy. J Am Chem Soc 131:14761–14767

Huang YJ, Montelione GT (2005) Structural biology: proteins flex to function. Nature 438:36–37

Iwahara J, Clore GM (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. Nature 440:1227–1230

Kay LE (2005) NMR studies of protein structure and dynamics. J Magn Reson 173:193–207

Kay LE, Ikura M, Tschudin R, Bax A (1990) Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. J Magn Reson 89:496–514

Keizers PH, Desreux JF, Overhand M, Ubbink M (2007) Increased paramagnetic effect of a lanthanide protein probe by two-point attachment. J Am Chem Soc 129:9292–9293

Keizers PHJ, Saragliadis A, Hiruma Y, Overhand M, Ubbink M (2008) Design, synthesis, and evaluation of a lanthanide chelating protein probe: CLaNP-5 yields predictable paramagnetic effects independent of environment. J Am Chem Soc 130:14802–14812

Korzhnev DM, Religa TL, Banachewicz W, Fersht AR, Kay LE (2010) A transient and low-populated protein-folding intermediate at atomic resolution. Science 329:1312–1316

Kurokawa H, Osawa M, Kurihara H, Katayama N, Tokumitsu H, Swindells MB, Kainosho M, Ikura M (2001) Target-induced conformational adaptation of calmodulin revealed by the crystal structure of a complex with nematode Ca^{2+}/calmodulin-dependent kinase kinase peptide. J Mol Biol 312:59–68

Longinetti M, Luchinat C, Parigi G, Sgheri L (2006) Efficient determination of the most favored orientations of protein domains from paramagnetic NMR data. Inv Probl 22:1485–1502

Mal TK, Skrynnikov NR, Yap KL, Kay LE, Ikura M (2002) Detecting protein kinase recognition modes of calmodulin by residual dipolar couplings in solution NMR. Biochemistry 41:12899–12906

Maximciuc AA, Putkey JA, Shamoo Y, MacKenzie KR (2006) Complex of calmodulin with a ryanodine receptor target reveals a novel, flexible binding mode. Structure 14:1547–1556

Meador WE, Means AR, Quiocho FA (1992) Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. Science 257:1251–1255

Meador WE, Means AR, Quiocho FA (1993) Modulation of calmodulin plasticity in molecular recognition on the basis of x-ray structures. Science 262:1718–1721

Muhandiram DR, Kay LE (1994) Gradient-enhanced triple resonance three-dimensional NMR experiments with improved sensitivity. J Magn Reson Ser B 103:203–216

Schmitz C, Stanton-Cook MJ, Su XC, Otting G, Huber T (2008) Numbat: an interactive software tool for fitting $\Delta\chi$-tensors to molecular coordinates using pseudo contact shifts. J Biomol NMR 41:179–189

Schumacher MA, Rivard AF, Bächinger HP, Adelman JP (2001) Structure of the gating domain of a Ca^{2+}-activated K^+ channel complexed with Ca^{2+}/calmodulin. Nature 410:1120–1124

Sgheri L (2010) Joining RDC data from flexible protein domains. Inv Probl 26(12):115021

Su XC, Otting G (2010) Paramagnetic labelling of proteins and oligonucleotides for NMR. J Biomol NMR 46:101–112

Su XC, Huber T, Dixon NE, Otting G (2006) Site-specific labelling of proteins with a rigid lanthanide-binding tag. Chem Bio Chem 7:1599–1604

Su XC, Man B, Beeren S, Liang H, Simonsen S, Schmitz C, Huber T, Messerle BA, Otting G (2008) A dipicolinic acid tag for rigid lanthanide tagging of proteins and paramagnetic NMR spectroscopy. J Am Chem Soc 130:10486–10487

Tang C, Schwieters CD, Clore GM (2007) Open-to-close transition in apo maltose-binding protein observed by paramagnetic NMR. Nature 449:1078–1082

Vallurupalli P, Hansen DF, Kay LE (2008) Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. Proc Natl Acad Sci USA 105:11766–11771

Vlasie MD, Fernández-Busnadiego R, Prudêncio M, Ubbink M (2008) Conformation of pseudoazurin in the 152 kDa electron transfer complex with nitrite reductase determined by paramagnetic NMR. J Mol Biol 375:1405–1415

Volkov AN, Worrall JAR, Holtzmann E, Ubbink M (2006) Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. Proc Natl Acad Sci USA 103:18945–18950

Wang X, Srisailam S, Ye AA, Lemak A, Arrowsmith C, Prestegard JH, Tian F (2007) Domain-domain motions in proteins from time-modulated pseudocontact shifts. J Biomol NMR 39:53–61

Xu X, Reinle W, Hannemann F, Konarev PV, Svergun DI, Bernhardt R, Ubbink M (2008) Dynamics in a pure encounter complex of two proteins studied by solution scattering and paramagnetic NMR spectroscopy. J Am Chem Soc 130:6395–6403

Zhuang T, Lee HS, Imperiali B, Prestegard JH (2008) Structure determination of a Galectin-3-carbohydrate complex using paramagnetism-based NMR constraints. Protein Sci 17:1220–1231