ORIGINAL PAPER

Solution structure and dynamics of S100A5 in the apo and Ca^{2+}-bound states

Ivano Bertini · Soumyasri Das Gupta · Xiaoyu Hu ·
Tilemachos Karavelas · Claudio Luchinat ·
Giacomo Parigi · Jing Yuan

Abstract S100A5 is a calcium binding protein of the S100 family, with one canonical and one S100-specific EF-hand motif per subunit. Although its function is still unknown, it has recently been reported to be one of the S100 proteins able to interact with the receptor for advanced glycation end products. The homodimeric solution structures of S100A5 in both the apo and the calcium(II)-loaded forms have been obtained, and show a conformational rearrangement upon calcium binding. This rearrangement involves, in particular, the hinge loop connecting the N-terminal and the C-terminal EF-hand domains, the reorientation of helix III with respect to helix IV, as common to several S100 proteins, and the elongation of helix IV. The details of the structural changes are important because they must be related to the different functions, still largely unknown, of the different members of the S100 family. For the first time for a full-length S100 protein, relaxation measurements were performed on both the apo and the calcium-bound forms. A quite large mobility was observed in the hinge loop, which is not quenched in the calcium form. The structural differences resulting upon calcium binding change the global shape and the distribution of hydrophobic and charged residues of the S100A5 homodimer in a modest but significantly different manner with respect to the closest homologues S100A4 and S100A6.

I. Bertini (✉) · S. Das Gupta · X. Hu · T. Karavelas ·
C. Luchinat · G. Parigi · J. Yuan
Magnetic Resonance Center (CERM),
University of Florence,
Via Luigi Sacconi 6,
50019 Sesto Fiorentino, Italy
e-mail: ivanobertini@cerm.unifi.it; bertini@cerm.unifi.it

I. Bertini
Department of Chemistry,
University of Florence,
Via della Lastruccia 3,
50019 Sesto Fiorentino, Italy

Present Address:
T. Karavelas
Department of Chemistry,
University of Ioannina,
45110 Ioannina, Greece

C. Luchinat · G. Parigi
Department of Agricultural Biotechnology,
University of Florence,
via Maragliano 75-77,
50144 Florence, Italy

Introduction

S100 proteins have been found to be implicated in a Ca^{2+}-dependent (and, in some cases, Zn^{2+}- or Cu^{2+}-dependent) regulation of a variety of intracellular and extracellular activities, and several biological targets have been identified for the different proteins [1]. A large variability in the sequence is observed within the protein family, which is responsible for the modulation of the shape and the nature of the binding surface. This modulation is needed to bind different targets, although the overall fold of most members of the family is very similar. The structures available for

S100–target peptide adducts actually display a remarkable lack of uniformity in the orientation of the target [2]. Furthermore, individual S100 proteins can bind different targets in different ways [3].

Several S100 protein structures are available, solved either in solution or in the solid state. Together with S100A4 and S100A6, S100A5, the function of which is still unknown [4], belongs to a well-defined subset of the S100 family showing high homology [5]. The structures of S100A4 and S100A6 are available [6, 7], whereas that of S100A5 is not. We report here the solution structures of S100A5 in both the apo and the calcium(II)-bound states, and compare these structures with those of its closest homologues and of other S100 proteins.

The expression level of S100A5 was demonstrated by immunohistochemical analyses to be restricted to a few specific cells [8], i.e., in the olfactory bulb, in the brainstem, and in the spinal trigeminal tract, analogously to S100A3 and differently from other S100 proteins such as S100A6 and S100B. Marked modifications of the levels of expression of different S100 proteins (including S100A5 and S100A6) occur in connection with the progression of astrocytic tumor malignancy [9]. It was found that totally resected WHO grade I meningiomas with high levels of S100A5 either did not recur or recurred later than those with low S100A5 levels [9]. In this respect, the recently found interaction with the receptor for advanced glycation end products, well known to be involved in tumor outgrowth, may acquire further biological relevance [10].

Most S100 proteins are encoded by genes located in the same chromosome 1q21, with the exception of genes encoding S100B (located on chromosome 21q22), calbindin D_{9k} (also called S100G, located on chromosome Xp22) and S100P (located on chromosome 4p16) [11]. Interestingly, S100A1, S100A3, S1004, S100A5, S100A6, S100A8, S100A9, S100A12, and S100A13 genes are all mapped within a short distance. It is widely appreciated that colocalization of genes may imply coexpression of the proteins [12, 13], and in the case of S100 proteins this probably correlates with the observation of functional heterodimers and possibly with concerted functions [13].

All S100 proteins are constituted by two EF-hand motifs, highly conserved helix–loop–helix structural domains that can each bind a calcium(II) ion. Canonical EF-hand proteins have calcium binding loops constituted by 12 residues; S100 proteins are a subgroup where the N-terminal EF-hand loop is constituted by 14 residues [1, 4, 14]. The N-terminal EF-hand comprises helix I, the S100-specific calcium binding loop I, and helix II, which is separated by a flexible linker, called "hinge loop," from the C-terminal EF-hand, which comprises helix III, calcium binding loop II, and helix IV.

Calcium(II) binding is an important mechanism in cells because calcium(II) is toxic at elevated levels to cellular metabolism, and therefore its influx and efflux in the cytosol must be controlled and kept at submicromolar resting levels [15]. Furthermore, calcium(II) ions play a central role in cell signaling. Calcium(II) binding to EF-hand proteins in fact induces in most cases conformational changes that correlate with binding of target proteins/enzymes involved in a wide variety of cellular processes. The helices in the EF-hand motifs can have an almost antiparallel arrangement, called "closed conformation," or an almost orthogonal arrangement, called "open conformation," depending on the presence of bound calcium [16–22]. The latter conformation exposes large hydrophobic clefts on the protein surface, which acts as a binding region for a variety of targets.

Proteins undergoing changes in the orientation of the helices of each EF-hand motif upon calcium(II) binding are generally functionally related to activation of target proteins, whereas proteins not undergoing conformational changes have the function of calcium buffer and transport [23]. In most cases, S100 proteins undergo smaller structural changes upon calcium(II) binding in the N-terminal domain and larger changes in the C-terminal domain [2], although not as large as those observed for the EF-hand protein calmodulin [24].

All S100 structures determined to date (with the exception of that of calbindin D_{9k}) show that these proteins exist as homodimers, heterodimers, or tetramers [14]. Most of the S100 proteins are homodimers. The dimer interface consists of helices I (I′) and IV (IV′) of each subunit arranged in a X-type four-helix bundle, in both the apo and the calcium-bound states [14]. Calcium binding results in minor alterations of the backbone conformation of calcium binding loop I but causes helix III to reorient and form a more open structure with respect to the apo state. As a result, the hydrophobic residues of helices III and IV in the C-terminal EF-hand are more exposed, thus facilitating the interaction with target proteins. The solution structures obtained here for the apo and calcium(II)-bound forms of S100A5 show that the same behavior applies to this protein, and provide the details of the exposed surface and charge distribution responsible for its possible interactions. Despite the high homology, S100A4, S100A5, and S100A6 show modest but significant differences in the pattern of hydrophobic/hydrophilic/charged residues exposed upon calcium binding. The present data thus provide a further example of the diversity of the exposed protein surface, which is likely to be reflected in a diversity in target proteins.

Materials and methods

Sample preparation

Untagged human *S100A5* was cloned into the *Nde*I and *Bam*HI sites of the expression vector pET21a. The recombinant plasmid was transformed into *Escherichia coli* BL21 Gold cells. For the production of ^{15}N-labeled or ^{13}C- and ^{15}N-labeled S100A5, cultures were grown in minimal medium using ^{13}C-glucose and/or ^{15}N-ammonium sulfate as the sole carbon and nitrogen source, respectively. Cells were grown at 310 K to an optical density at 600 nm of 0.6 and growth was induced with 1 mM isopropyl β-D-thiogalactopyranoside. After induction, the temperature was reduced to 298 K and the culture was grown overnight. Cells were harvested by centrifugation at 15,000*g* for 15 min and resuspended in lysis buffer [20 mM tris(hydroxymethyl)aminomethane (Tris)–HCl pH 7.5, 1 mM EDTA, 1 mM phenylmethylsulfonyl fluoride, 2 mM dithiothreitol (DTT)]. Cell lysis was performed by sonicating with eight bursts of 30 s each. The suspension was ultracentrifuged at 200,000*g* for 30 min. The cleared lysate was precipitated by slowly adding ammonium sulfate to 30% and centrifuging at 15,000*g* for 20 min. The supernatant was brought to 2 mM CaCl$_2$, applied to a phenyl Sepharose column equilibrated with 20 mM Tris–HCl, pH 7.5, 2 mM CaCl$_2$. The unbound proteins were washed out from the column with the same buffer. S100A5 was then eluted with 20 mM Tris–HCl, pH 7.5, 5 mM EDTA. The eluate was concentrated and applied to a Superdex 75 column equilibrated with 30 mM 2-morpholinoethanesulfonic acid, pH 6.5, 100 mM NaCl, 5 mM DTT. The fractions containing S100A5 were pooled and washed with excess EDTA to remove all metal ions. The yield of S100A5 was 20 mg L^{-1} of culture.

NMR spectroscopy and structure determination

All NMR experiments for assignments were performed at 298 K using a Bruker 500 MHz spectrometer equipped with a cryoprobe. Apo and Ca$_2$-S100A5 samples (0.4 mM) were ^{13}C- and ^{15}N-labeled, in 30 mM 2-morpholinoethanesulfonic acid, 100 mM NaCl, and 5 mM DTT buffer (pH 6.5), containing 10% D$_2$O. Sequential assignments of the backbone resonance were achieved via HNCO, HNCA, CBCA(CO)NH, and HNCACB spectra. Side-chain assignments were performed through 3D (h)CCH total correlation spectroscopy, HBHA(CBACACO)HN together with ^{13}C nuclear Overhauser effect spectroscopy (NOESY) heteronuclear single quantum coherence (HSQC), and ^{15}N-NOESY-HSQC experiments. Proton–proton distance restraints were derived from the analysis of 2D NOESY, ^{15}N-NOESY-HSQC, and ^{13}C-NOESY-HSQC spectra

acquired using a Bruker 900 MHz spectrometer equipped with a cryoprobe. The spectra were processed using TOPSPIN and analyzed with CARA [25]. The secondary structure elements were predicted from the chemical shift index and the backbone dihedral angles were obtained from TALOS [26], accordingly. The structures were calculated using the program CYANA-2.0 [27]. The two subunits in the dimeric structure were linked together through a chain of dummy atoms with zero van der Waals radii. The calcium(II) ions were included in the calculation of the calcium-loaded form by adding new residues in the amino acid sequence. Four chains of dummy atoms with zero van der Waals radii, that can freely penetrate into the protein, each of them ending with one atom with a radius of 1.8 Å, which mimics the calcium ion, were included for this purpose. Protein ligand atoms were linked to the metal ion through upper distance limits of 3 Å, according to the structure of S100A13.

The best 30 structures out of the calculated 350 structures of the CYANA family were then subjected to restrained energy minimization with AMBER 10 [28]. Nuclear Overhauser effect (NOE) and torsion angle restraints were applied with force constants of 50 kcal mol^{-1} Å^{-2} and 32 kcal mol^{-1} rad^{-2}, respectively. The program PROCHECK-NMR [29] was used to evaluate the quality of the structures.

Relaxation measurements

^{15}N-R_1, R_2, and steady-state heteronuclear ^{1}H–^{15}N NOEs were measured with a 700 MHz spectrometer using standard pulse sequences [30, 31], at 298 K. The longitudinal (R_1) and transverse (R_2) relaxation rates were determined by fitting the cross-peak intensities as a function of the delay to a single-exponential decay through the standard routines of the Sparky software program [32]. The heteronuclear NOE values were obtained from the ratio of the peak height for ^{1}H-saturated and unsaturated spectra. The heteronuclear NOE values and their errors were estimated by calculating the mean ratio and the standard error from the available data sets. R_1, R_2, and NOE values were obtained for 67 and 71 out of the 92 assigned backbone ^{N}H resonances for the apo and the calcium forms, respectively. Estimates of the reorientation time were then calculated with the model-free approach [33]. Theoretical predictions of ^{N}H, R_1, and R_2 values for apo-S100A5 and Ca$_2$-S100A5 were calculated by using the HYDRONMR software program [34].

Metal binding detection

The binding of apo-S100A5 to Ca^{2+} was monitored by following the changes in the chemical shifts of the protein

NMR peaks in the ^{1}H–^{15}N-HSQC spectra upon titration of the apoprotein with calcium ions. The chemical shift perturbation between the free and bound states was obtained for each residue by calculating the composite chemical shifts according to Eq. 1:

$$\delta(\text{HN}) = \sqrt{\frac{\Delta\delta_\text{H}^2 + (\Delta\delta_\text{N}/5)^2}{2}}, \tag{1}$$

where $\Delta\delta_\text{H}$ and $\Delta\delta_\text{N}$ are the differences in chemical shifts between the bound and free states of the amide protons and of the nitrogen atoms, respectively.

Accession numbers

Atomic coordinates, structural restraints, and resonance assignments of apo-S100A5 and Ca$_2$-S100A5 have been deposited in the Protein Data Bank (codes 2KAX and 2KAY) and BioMagResBank (codes 16033 and 16034).

Results

Resonance assignment

The ^{1}H–^{15}N-HSQC of S100A5 in both the apo and the calcium forms show well-dispersed resonances, as expected for a regularly folded protein. The backbone resonance signals were assigned from residue Glu-2 to residue Tyr-83 and from residue Phe-87 to residue Lys-92 in the apo form, and from residue Glu-2 to residue Lys-92 with the exception of Cys-43 in the calcium form. The corresponding assignments are deposited in BioMagResBank together with the ^{1}H–^{13}C–^{15}N assignments of the side chain resonances. The types of NMR spectra used for the assignments are described in "Materials and methods".

Ca^{2+} titration of apo-S100A5

The binding of calcium(II) to apo-S100A5 was monitored by following the changes in the ^{1}H–^{15}N-HSQC NMR spectrum of ^{15}N-labeled apo-S100A5 upon addition of increasing amounts of calcium(II). New peaks appeared in the spectrum during the titration corresponding to the calcium(II)-bound S100A5 form. The intensity of the new peaks increased on increasing the Ca^{2+} to apo-S100A5 ratio. When a 2:1 ratio (with respect to the protein subunit concentration) was reached, the original peaks, corresponding to the apo form, disappeared. Such behavior is indicative of a slow exchange regime, i.e., the exchange rate between the metal-free and the metal-bound forms is much smaller than the chemical shift difference between the two forms. Figure 1 shows the chemical shift changes
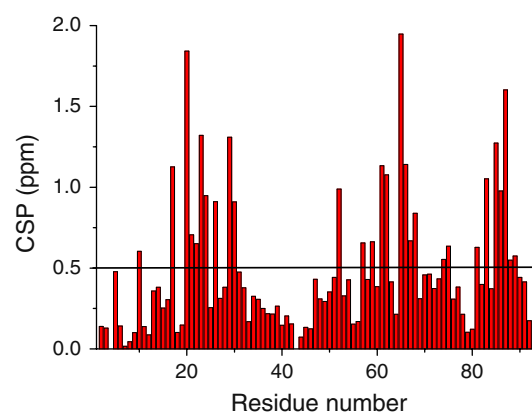

Fig. 1 Composite chemical shift perturbation (*CSP*), as defined in Eq. 1, of apo-S100A5 upon calcium(II) binding. The *horizontal line* indicates the average value

on passing from the apo to the calcium forms of S100A5. The residues undergoing the largest changes in chemical shifts are located in the Ca^{2+} binding loops of the two EF-hand motifs, as expected, and also in the C-terminus. However, significant differences occur throughout the protein, thus indicating that a significant conformational change occurs on passing from the apo form to the calcium form.

The findings of the Ca^{2+} titration experiments are consistent with previous measurements, which provided dissociation constants for the binding of the first and second calcium(II) ions in the submillimolar and submicromolar range, respectively, and a strong positive cooperativity [8]. As already pointed out [8], the affinity of calcium(II) for S100A5 is among the highest in the whole S100 family.

Relaxation measurements

The relaxation parameters for apo and calcium-loaded S100A5 are shown in Fig. 2. Such measurements indicate that the protein is dimeric in both forms. The reorientation times corresponding to the observed relaxation rates were in fact calculated to be 12.6 ± 1.0 and 13.5 ± 1.8 ns for the apo and calcium-loaded forms, respectively, in agreement with the molecular weight and the reorientation times observed for other S100 homodimers [35–39].

In both apo-S100A5 and Ca$_2$-S100A5, the relaxation rate measurements show large mobility on a time scale shorter than the reorientation time (R_1 increases, R_2 decreases, the NOE decreases) in the hinge loop and for the last residues at the C-terminus, thus indicating that such regions may be largely unstructured. Occurrence of motion is also detected form some other residues of the calcium binding loops (21, 26, 27, 61–63 in the apo form; 25, 27, 30 in the calcium form).
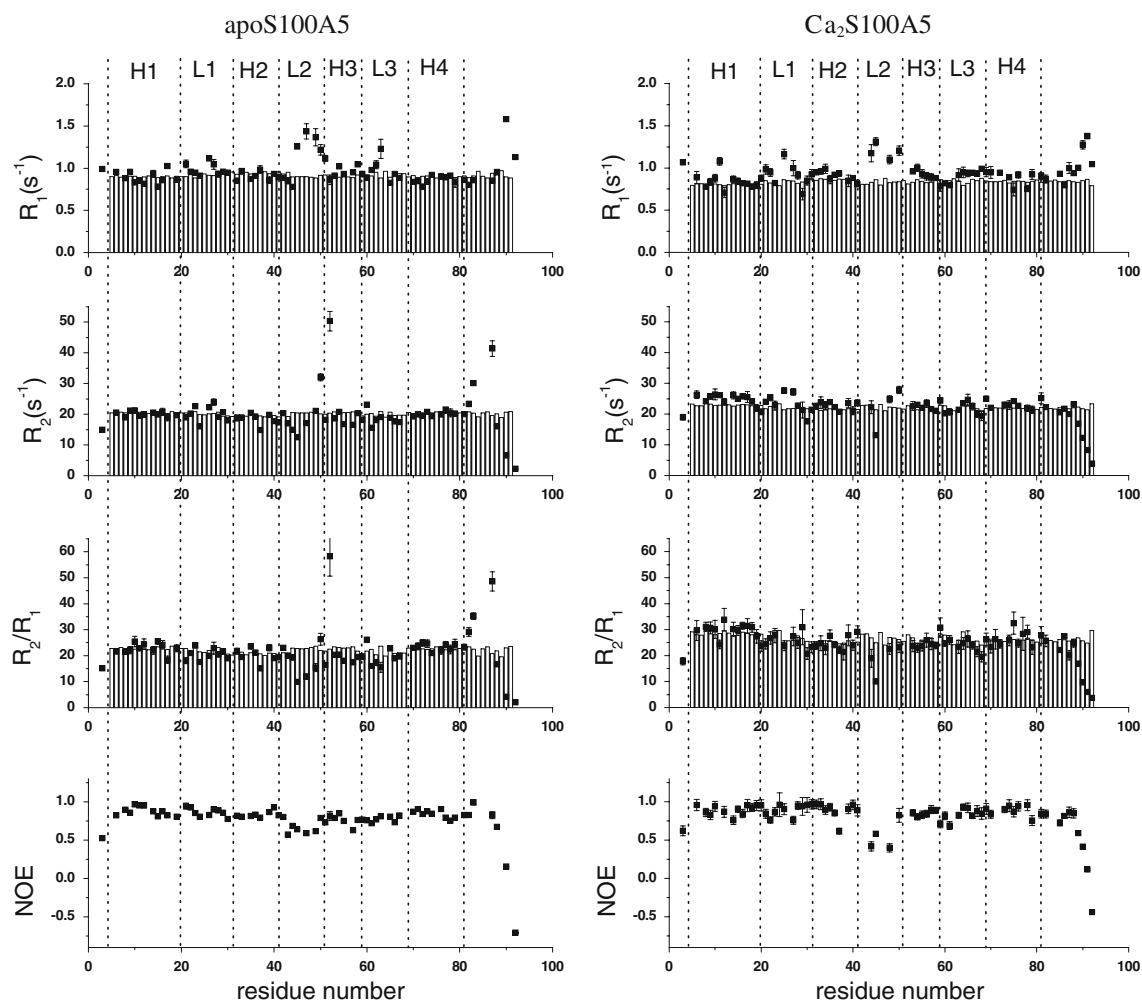
Fig. 2 Sequential plot of the experimental relaxation parameters of apo and Ca^{2+}-S100A5. The values calculated by HYDRONMR are shown as *bars*. *NOE* nuclear Overhauser effect

In the apoprotein, motion on a slower time scale (microseconds to milliseconds) was observed for some residues localized at the beginning of helix III (Asp-50 and Ile-52), at the end of helix IV (Tyr-83), and at the C-terminus (Phe-87), as indicated by the significantly larger R_2 value (for Ile-52, in particular, R_2 is 52 s^{-1} compared with an average value of about 20 s^{-1} observed for helix residues). This mobility may originate from backbone amide conformational exchange and/or side-chain rotation. Since the time scale for the conformational exchange (microseconds to milliseconds) is sizably shorter than the mixing time (100 ms) of the NOESY spectra, the conformational reorientations of these residues may cause $^1H–^1H$-NOEs to be observed between nuclei of the side chains of such residues and residues located in quite different positions (see later). Upon calcium binding, residues in the slow motion regime are not observed any more. However, calcium binding does not reduce the fast motion detected

for residues in the hinge loop, the observed $^1H–^{15}N$-NOE being even smaller than in the apo form (the average $^1H–^{15}N$-NOE in the hinge loop is 0.60 and 0.40 for the apo and calcium forms, respectively).

Solution structures of apo and Ca^{2+}-bound S100A5

The solution structures of the human S100A5 in the apo and calcium-loaded forms were obtained. A total of 2,752 and 2,530 meaningful upper distance limits per dimer, including 184 and 190 intersubunit upper distance limits for the apo and the calcium forms, respectively, were used (Tables 1, 2). Few NOE patterns were detected for residues in the hinge loop and at the C-terminus, consistent with the observed mobility of such regions. In the calcium form, the Ca^{2+} ions were restrained to be within 3 Å from the oxygen ligand atoms (O of Ser-20, Glu-23, Ser-25, Thr-28 and OE1, OE2 of Glu-33 for the first binding site, and OD1 of

Table 1 Structural restraints and statistical analysis of apo-S100A5

Structural restraints

NOE upper distance limits		
Intrasubunit	1,284	
Intraresidue	358	
Interresidue		
Sequential ($|i − j| = 1$)	395	
Medium range ($|i − j| < 4$)	320	
Long range ($|i − j| > 5$)	211	
Intersubunit	184	
Dihedral angle restraints		
φ	96	
ψ	96	

Statistical analysis

	Family	Mean
RMS violations per meaningful distance restraints (Å)		
Intraresidue	0.0183 ± 0.0028	0.0178
Sequential	0.0179 ± 0.0025	0.0176
Medium range	0.0149 ± 0.0023	0.0093
Long range	0.0090 ± 0.0020	0.0087
RMS violations per meaningful dihedral angle restraints (°)		
φ	4.87 ± 1.23	4.00
ψ	4.04 ± 1.42	2.82
Average number of restraints per residue	14.96	14.96
Average number of violations per conformer		
φ	6.93 ± 2.00	8.00
ψ	6.63 ± 2.20	4.00
NOE violations between 0.1 and 0.3 Å	9.40 ± 2.59	8.0
NOE violations larger than 0.3 Å	0	0
Average RMSD from the mean (Å)		
Backbone	1.00 ± 0.09[a]	
	0.78 ± 0.09[b]	
Heavy	1.47 ± 0.09[a]	
	1.25 ± 0.09[b]	
Residual CYANA target function (Å^2)	1.18 ± 0.31	
Structural analysis		
Residues in most favorable regions (%)	82.7[a]/88.1[b]	92.6
Residues in allowed regions (%)	13.2[a]/10.0[b]	6.6
Residues in generously allowed regions (%)	2.2[a]/0.9[b]	0.0
Residues in disallowed regions (%)	1.9[a]/1.0[b]	0.8

NOE nuclear Overhauser effect, *RMS* root mean square, *RMSD* root mean square deviation

[a] RMS deviation values were calculated in the sequence range 3–82

[b] RMS deviation values were calculated excluding flexible loop 41–52 of both subunits

Asp-60, Asn-62, OD1, OD2 of Asp-64, O of Glu-66, and OE1, OE2 of Glu-71 for the second binding site). No symmetry constraint was used.

Table 2 Structural restraints and statistical analysis of Ca$_2$-S100A5

Structural restraints

NOE upper distance limits		
Intrasubunit	1,170	
Intraresidue	380	
Interresidue		
Sequential ($|i − j| = 1$)	325	
Medium range ($|i − j| < 4$)	280	
Long range ($|i − j| > 5$)	185	
Intersubunit	190	
Dihedral angle restraints		
φ	120	
ψ	120	

Statistical analysis

	Family	Mean
RMS violations per meaningful distance restraints (Å)		
Intraresidue	0.0106 ± 0.0027	0.0123
Sequential	0.0092 ± 0.0016	0.0077
Medium range	0.0085 ± 0.0018	0.0083
Long range	0.0056 ± 0.0018	0.0053
RMS violations per meaningful dihedral angle restraints (°)		
Phi	1.97 ± 0.23	1.6246
Psi	0.63 ± 0.42	0.5218
Average number of restraints per residue	13.75	13.75
Average number of violations per conformer		
φ	9.74 ± 1.76	9.0
ψ	1.83 ± 1.23	2.0
NOE violations between 0.1 and 0.3 Å	4.03 ± 1.93	6.0
NOE violations larger than 0.3 Å	0	0
Average RMSD from the mean (Å)		
Backbone	0.93 ± 0.11[a]	
	0.83 ± 0.10[b]	
Heavy	1.40 ± 0.10[a]	
	1.29 ± 0.09[b]	
Residual CYANA target function (Å^2)	0.31 ± 0.03	
Structural analysis		
Residues in most favorable regions (%)	86.0[a]/90.4[b]	90.7
Residues in allowed regions (%)	11.2[a]/7.5[b]	9.3
Residues in generously allowed regions (%)	1.9[a]/1.5[b]	0.0
Residues in disallowed regions (%)	0.9[a]/0.6[b]	0.0

[a] RMSD values were calculated in the sequence range 3–90

[b] RMSD values were calculated excluding flexible loop 41–49 of both subunits

The calculated families of structures are shown in Fig. 3. In both forms, the eight individual helices of the two EF-hand motifs of each subunit present in the dimeric structure are very well defined; the four calcium binding loops are less well defined, whereas the linker regions

between the two EF-hand calcium binding domains are poorly defined. These results are in line with what was previously found for other EF-hand proteins [36, 40]. Each calcium binding loop contains a short antiparallel β strand.

The root mean square deviation (RMSD) from the mean subunit structure for the structured regions of the protein is 0.73 ± 0.10 Å (backbone) and 1.22 ± 0.09 Å (heavy atoms) for apo-S100A5 (residues 3–40, 53–82) and 0.71 ± 0.09 Å (backbone) and 1.20 ± 0.07 Å (heavy atoms) for Ca_2-S100A5 (residues 3–40, 50–90). The RMSD from the mean dimeric structure for the structured regions of the protein is 0.78 ± 0.09 Å (backbone) and 1.25 ± 0.09 Å (heavy atoms) for apo-S100A5 (residues 3–40, 53–82 of both subunits) and 0.83 ± 0.10 Å (backbone) and 1.29 ± 0.09 Å (heavy atoms) for Ca_2-S100A5 (residues 3–40, 50–90 of both subunits). More than 95% of the residues (including those in the poorly defined regions) in all structures were located in the allowed regions of the Ramachandran plot. The conformational and energetic analyses of both structures are reported in Tables 1 and 2.
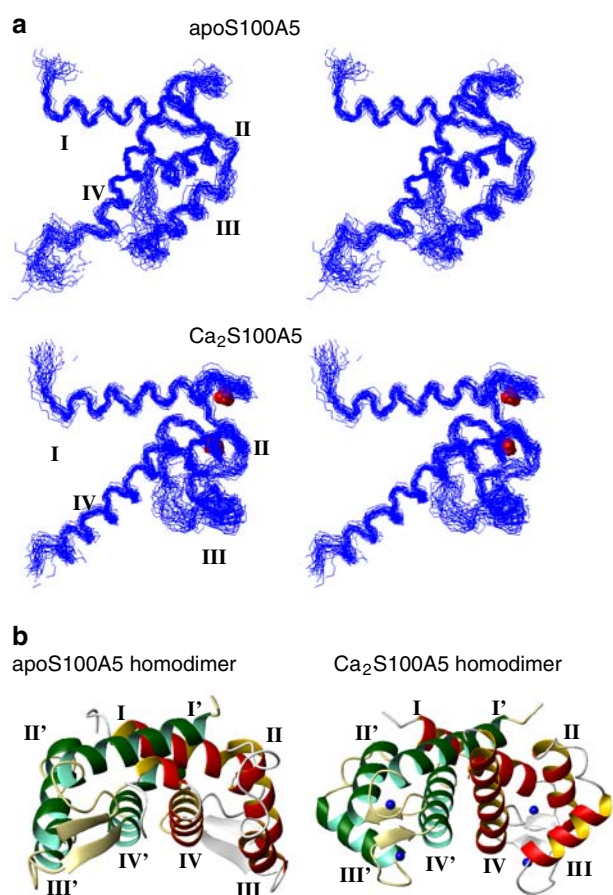

Fig. 3 Stereoview of the families of the solution structures of the S100A5 subunit in the apo and calcium(II)-loaded forms (**a**) and ribbon representation of the homodimer mean structures (**b**) obtained after restrained energy minimization

The relaxation rates were then calculated using HYDRONMR [34] and the minimized mean structures obtained, and are reported in Fig. 2 as bars. An overall agreement is observed between calculated and experimental values for the residues located on the protein helices; on the other hand, the differences between calculated and observed values make it easier to appreciate the presence of mobility in some residues of the loops.

In both the apo and the calcium-loaded forms, S100A5 forms homodimers owing to the interactions between helices I and I′ and between helices IV and IV′ of the two subunits. There is a symmetry relationship between the subunits consisting in a twofold rotational axis passing through the dimer interface approximately perpendicular to helix I and helix I′ and parallel to helix IV and helix IV′. At the dimer interface, residues in the hinge loop between helix II and helix III make contacts with residues near the N-terminus of helix I of the other subunit. Residues Phe-69, Lys-70, Ser-73, and Cys-80 in helix IV also make several contacts with helix I′ and helix IV′ of the other subunit. All these interactions align helix I and helix IV in opposite directions to helix I′ and helix IV′, respectively, in the dimer.

Discussion

The overall structures of both the apo and the calcium(II)-loaded forms of S100A5 are in good agreement with those obtained for other S100 proteins, such as S100A1, S100A4, S100A6, S100A8, S100A12, S100A13, or S100B [6, 39, 41]. The comparison of the apo and calcium-loaded S100A5 structures shows that the N-terminal EF-hands (residues 5–41) are similar to one another (the backbone RMSD is 2.0 Å), thus indicating that there is no large conformational rearrangement upon calcium binding. In contrast, the C-terminal EF-hand (residues 49–82) undergoes a major conformational change upon calcium binding, the backbone RMSD between the two forms increasing to 4.2 Å. This conformational rearrangement includes a quite different orientation of helix III and nonnegligible changes in helix IV and in the hinge loop (Fig. 4). These rearrangements upon calcium binding are similar to those observed for other S100 proteins [5, 6, 39, 42, 43], with the exception of S100A10, which is known to have a "calcium-ready state" in both the N-terminal and the C-terminal EF-hands although it does not bind calcium(II) [4]. In apo-S100A6 (1K9P) and apo-S100A13 (1YUR), for instance, helix III is almost antiparallel to helix IV, but opens by 30–40° upon calcium binding (1K9K and 1YUT). The same degree of opening is observed in other EF-hand proteins, such as calmodulin [21], not belonging to the S100 family. In S100A5 the angle between helices III and
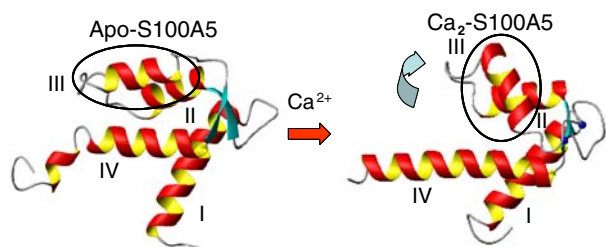
Fig. 4 Major structural differences upon calcium binding: the different angle between helices III (enclosed in *circles*) and IV, and the longer α-helical structure of helix IV

IV changes from 168° to 118° on passing from the apo to the calcium-bound form, so the two helices are almost perpendicular in the latter form.

Analogously to most S100 proteins, helices IV and IV′ in the apo form tend to be antiparallel (forming an angle of 152°), whereas they form an angle of about 130° in the calcium-bound form, while helices I and I′ form a similar angle (147–142°) in both forms.

Structural changes within the EF-hand family can be monitored through a principal component analysis of the six interhelix angles representing the reciprocal orientation of the four helices [21]. It was shown that the EF-hand proteins can be clustered according to subgroups and metal content using the first two principal components, which concentrate the information distributed throughout the six interhelix angles. The values of the first two principal components also permit us to identify whether S100 proteins have a structure typical of the apo or the calcium-loaded form. The principal component values were thus calculated for the two forms of S100A5, and were plotted together with the values previously calculated for the S100 proteins (Fig. 5), and with the values relative to other S100 proteins deposited in the Protein Data Bank in the meantime, by using the same coefficients for the interhelix angles (calculated with the program MOLMOL) reported by Babini et al. [21]. The figure shows that S100A5 is regularly positioned with respect to the other S100 proteins in both the apo and the calcium-loaded forms, thus pointing to the occurrence of similar structures, and thus of similar overall rearrangement upon calcium(II) binding. It is to be noted that the only two S100 proteins not regularly placed are calbindin D$_{9k}$ and S100A10 in the apo form.

The concomitant 50° reorientation of helix III with respect to helix IV and the reorientation and translation of helices IV and IV′ in S100A5 upon calcium(II) binding result in an increased solvent-exposed surface of the hinge loop and of some positively charged residues of helix II and helix III in the calcium-loaded form. In fact, several hydrophobic residues on helix III (Ile-52, Leu 55, Met-56, and Leu-59), helix II (Ile-38), and helix IV (Phe-75,
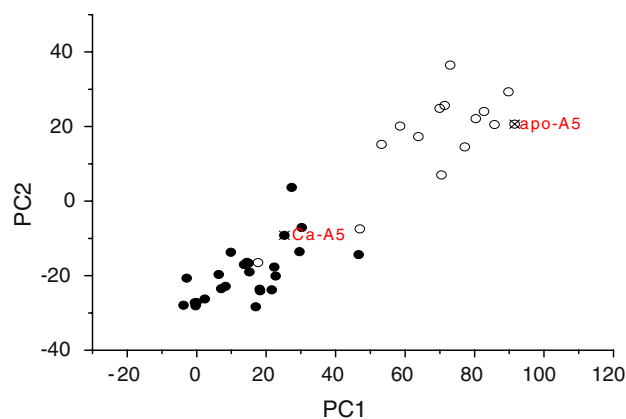

Fig. 5 Principal component plot for the S100 proteins derived from principal components analysis of the six interhelix angles [21]. Apoproteins (S100A1, S100A2, S100A3, S100A4, S100A5, S100A6, S100A10, S100A11, S100A13, S100B, calbindin D$_{9k}$) are indicated with *open circles* and calcium-loaded proteins (S100A1, S100A4, S100A5, S100A6, S100A7, S100A8, S100A9, S100A12, S100A13, S100B, calbindin D$_{9k}$, S100P) are indicated with *solid circles*. The two *open symbols* not regularly placed with respect to the others correspond to calbindin D$_{9k}$ and S100A10 in the apo form. The data are based on the structural information reported in the supporting information in [21] as well as on more recent structural information reported in the electronic supplementary material

Met-78, Tyr-83) are constrained in a hydrophobic cluster in apo-S100A5, which is loosened upon calcium(II) binding. On the other hand, calcium binding results in a decrease in the exposure of the metal ligand residues Asp-60, Asn-62, Asp-64, and Glu-71 in the C-terminal calcium binding loop. Similarly to what was found for other S100 proteins, the structural differences induced by calcium(II) binding in the homodimer thus lead to an exposure of two symmetrically positioned clefts, defined by helix III, helix IV, the hinge loop, and the last C-terminal residues, where target proteins can be accommodated [14].

In the apo form, residue Ile-52 was identified by relaxation measurements to experience mobility on the millisecond to microsecond time scale. The side chain of this residue experiences NOE contacts with both Met-56 and Tyr-83, which are positioned in opposite directions. As anticipated in "Relaxation measurements," this may be due to the occurrence of conformational exchange, and these data thus indicate that the side chain of Ile-52 can rotate along an axis perpendicular to helix III, so a conformational exchange is also affecting the backbone amide group. As a consequence, the residues forming a hydrophobic patch with Ile-52, and particularly the ones localized on the protein surface such as residue Tyr-83, may also experience sizable mobility on the side-chain and/or backbone atoms.

The slow time scale motion detected for residues Ile-52 and Tyr-83 in the apo form is absent in the calcium form, as a consequence of the conformational rearrangement of both helix III and helix IV. In the apo form, in fact, the
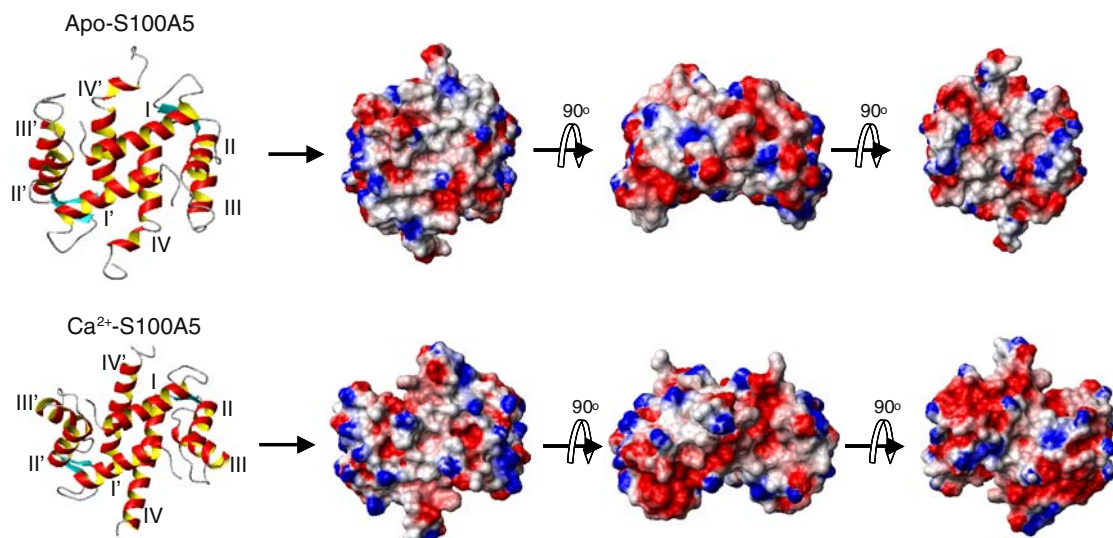
Fig. 6 Electrostatic surface representation of the S100A5 dimer

aromatic ring of Tyr-83 of each subunit participates in a hydrophobic cluster including the side chains of residues Leu-44, Met-47, Ile-52, and Leu-79 of the same subunit. As already noted [6], calcium binding overcomes the hydrophobic interactions that keep this cluster together, so the side chain of Tyr-83 changes orientation and forms new hydrophobic contacts with the side chains of Leu-9, Val-13, and Thr-14 of helix I of the other subunit of the dimeric structure. This rotation is experimentally confirmed by the ^1H–^1H-NOEs observed between Tyr-83 and residues Leu-44, Lys-48, Ile-52, and Leu-79 of the same subunit in the apo form and with residues Leu-9, Thr-10, and Val-13 of the other subunit in the calcium(II) form. The loosening of the hydrophobic cluster including Leu-44 and Met-47 may be responsible for the larger fast motion deduced for the hinge loop residues from the lower ^1H–^{15}N-NOE values measured for the calcium form with respect to the apo form.

Another difference between apo-S100A5 and Ca$_2$-S100A5 is that the C-terminal helix IV is shorter in the apo form. This very same difference has already been observed for S100A6 and S100B [6, 40, 44]. This is due to the unwinding of helix IV in apo-S100A5 at Tyr-83. The different orientation of the side chain of Tyr-83 is in fact responsible for a break in the α-helical structure, being consistent with a regularly formed α helix only in the calcium-loaded form [6].

The hydrophobic residues at the extreme C-terminus (Phe-87 and Leu-88) are important for stabilizing both the apo-S100A5 and the Ca$_2$-S100A5 homodimer. However, these residues form a hydrophobic cluster with different partners. In the apo form, they are in contact with Leu-27 in the first calcium binding loop of the other subunit, whereas in the calcium form they have hydrophobic interactions with Val-13 and Thr-14 of helix I of the other subunit.

Analogously to relaxation studies of apo-S100B and apo-S100A4 [35, 37], relaxation studies indicate that helices I and IV are quite rigid, whereas helix III is somewhat more flexible. On the other hand, in S100A5 the loop experiencing a very large mobility is only the hinge loop, whereas in S100A4 both the hinge loop and the calcium binding loops are quite mobile, and in S100B the mobility of the calcium binding loops is even larger than that of the hinge loop [35, 37]. A larger mobility for the hinge loop with respect to the calcium binding loops was also observed for S100A1 [36].

The combination of the structural differences results in a change of the global shape and distribution of surface charges of the S100A5 homodimer upon calcium binding (Fig. 6), whereas no major differences in motion are seen in the two forms. The change in the shape of the protein on passing from the apo to the calcium-loaded form is common to most S100 proteins [14, 39]. On the other hand, the change in the charge distribution seems to depend largely on the particular S100 protein [45]. S100A5 shows a number of charged residues, both positive and negative, on the protein surface in both the apo and the calcium-loaded forms. Interestingly, upon calcium binding, some more exposed positive residues (Lys-48, Lys-57) are moved away from the inner part of the opened cleft, which becomes slightly more hydrophobic, and the negative electrostatic surface is smaller and more clustered around the calcium binding sites. Other S100 proteins show a different change in the surface charge and hydrophobic distribution upon calcium binding: for instance, in S100A6 a larger increase of the hydrophobic surface was observed; in S100B a larger negative charged surface is exposed; in S100A4 the large hydrophobic surface present in the apo form remains exposed also upon calcium binding; in
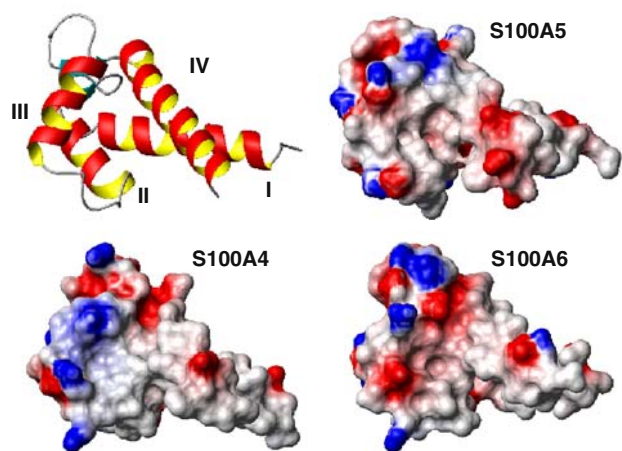
Fig. 7 Electrostatic surface representation of the reported ligand binding area typical of S100 proteins, for S100A4 (2Q91), S100A5 (present work), and S100A6 (1K9K)

S100A13 the charge distribution remains as large as in the apo form, but appears quite different. Overall, the present data provide a further example of the diversity of the exposed protein surface upon calcium(II) binding, which is likely to be reflected in a diversity in target protein(s).

S100A4, S100A5, and S100A6 are the most closely related S100 proteins according to phylogenetic trees constructed on the basis of multiple sequence alignments of S100 proteins [5]. Upon calcium binding, in all these proteins the opened cleft exposes hydrophobic residues, and positive residues become more exposed in the hinge loop and at the end of the third helix, i.e., in the typical protein target binding region. S100A6 also shows some more exposed negative charges at the end of the third helix. The charged groups surrounding the hydrophobic patch, as well as the shape of the surface, are however distinctly different for these proteins in the calcium-bound form (Fig. 7). Such differences observed for S100A4 and S100A6 reflect their different target specificity [46]. This suggests that the function of S100A5 may also be different from that of the other two proteins, and especially from that of S100A6, owing to the different charge pattern.

References

1. Donato R (1999) Biochim Biophys Acta 1450:191–231
2. Bhattacharya S, Bunick CG, Chazin WJ (2004) Biochim Biophys Acta 1742:69–79
3. Bhattacharya S, Large E, Heizmann CW, Hemmings B, Chazin WJ (2003) Biochemistry 42:14416–14426
4. Santamaria-Kisiel L, Rintala-Dempsey AC, Shaw GS (2006) Biochem J 396:201–214
5. Marenholz I, Heizmann CW, Fritz G (2004) Biochem Biophys Res Commun 322:1111–1122
6. Otterbein L, Kordowska J, Witte-Hoffmann C, Wang CL, Dominguez R (2002) Structure 10:557–567
7. Pathuri P, Vogeley L, Luecke H (2008) J Mol Biol 383:62–77
8. Schäfer BW, Fritschy J-M, Murmann P, Troxler H, Durussel I, Heizmann CW, Cox JA (2000) J Biol Chem 275:30623–30630
9. Hancq S, Salmon I, Brotchi J, De Witte O, Gabius H-J, Heizmann CW, Kiss R, Decaestecker C (2004) Neuropathol Appl Neurobiol 30:178–187
10. Leclerc E, Fritz G, Vetter SW, Heizmann CW (2009) Biochim Biophys Acta 1793:993–1007
11. Ridinger K, Ilg EC, Niggli FK, Heizmann CW, Schäfer BW (1998) Biochim Biophys Acta 1448:254–263
12. Kluger Y, Yu H, Qian J, Gerstein M (2003) BMC Genomics 4:49
13. Mijalski T, Harder A, Halder T, Kersten M, Horsch M, Strom TM, Liebscher HV, Lottspeich F, Hrabe de Angelis M, Beckers J (2005) Proc Natl Acad Sci USA 102:8621–8626
14. Donato R (2001) Int J Biochem Cell Biol 33:637–668
15. Orrenius S, Zhivotovsky B, Nicotera P (2003) Nat Rev Mol Cell Biol 4:552–565
16. Nelson MR, Chazin WJ (1998) Biometals 11:297–318
17. Finn BE, Evenäs J, Drakenberg T, Waltho J, Thulin E, Forsén S (1995) Nat Struct Biol 2:777–783
18. Zhang M, Tanaka T, Ikura M (1995) Nat Struct Biol 2:758–767
19. Kuboniwa H, Tjandra N, Grzesiek S, Ren H, Klee CB, Bax A (1995) Nat Struct Biol 2:768–776
20. Ikura M (1996) Trends Biochem Sci 21:14–17
21. Babini E, Bertini I, Capozzi F, Luchinat C, Quattrone A, Turano M (2005) J Proteome Res 4:1961–1971
22. Capozzi F, Luchinat C, Micheletti C, Pontiggia F (2007) J Proteome Res 6:4245–4255
23. Fragai M, Luchinat C, Parigi G (2006) Acc Chem Res 39:909–917
24. Bertini I, Gupta YK, Luchinat C, Parigi G, Peana M, Sgheri L, Yuan J (2007) J Am Chem Soc 129:12786–12794
25. Keller R (2004) The computer aided resonance assignment tutorial. CANTINA, Goldau
26. Cornilescu G, Delaglio F, Bax A (1999) J Biomol NMR 13:289–302
27. Guntert P (2004) Methods Mol Biol 278:353–378
28. Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozell S, Tsui V, Gohlke H, Mongan J, Hornak V, Cui G, Beroza P, Schafmeister CE, Caldwell JW, Ross WS, Kollman PA (2008) AMBER 10. University of California, San Francisco
29. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM (1996) J Biomol NMR 8:477–486
30. Kay LE, Torchia DA, Bax A (1989) Biochemistry 28:8972–8979
31. Barbato G, Ikura M, Kay LE, Pastor RW, Bax A (1992) Biochemistry 31:5269–5278
32. Goddard TD, Kneller DG (2000) SPARKY 3. University of California, San Francisco
33. Lipari G, Szabo A (1982) J Am Chem Soc 104:4546–4559
34. Garcia de la Torre JG, Huertas ML, Carrasco B (2000) J Magn Reson 147:138–146
35. Inman KG, Baldisseri DM, Miller KE, Weber DJ (2001) Biochemistry 40:3439–3448
36. Zhukov I, Ejchart A, Bierzynski A (2008) Biochemistry 47:640–650
37. Dutta K, Cox CJ, Basavappa R, Pascal SM (2008) Biochemistry 47:7637–7647

38. Bertini I, Fragai M, Luchinat C, Parigi G (2000) Magn Reson Chem 38:543–550
39. Arnesano F, Banci L, Bertini I, Fantoni A, Tenori L, Viezzoli MS (2005) Angew Chem Int Ed 44:6341–6344
40. Smith SP, Shaw GS (1998) Structure 6:211–222
41. Drohat AC, Baldisseri DM, Rustandi RR, Weber DJ (1998) Biochemistry 37:2729–2740
42. Maler L, Sastry M, Chazin WJ (2002) J Mol Biol 317:279–290
43. Bhattacharya S, Chazin WJ (2003) Structure 11:738–739
44. Kilby PM, Van Eldik LJ, Roberts GC (1996) Structure 4:1041–1052
45. Koch M, Diez J, Fritz G (2008) J Mol Biol 378:933–942
46. Gingras AR, Basran J, Prescott A, Kriajevska M, Bagshaw CR, Barsukov IL (2008) FEBS Lett 582:1651–1656