J Biomol NMR (2015) 62:233–238
DOI 10.1007/s10858-015-9927-6

NMR STRUCTURE NOTE

The solution structure of human calcium-bound S100A4 mutated at four cysteine loci

Ching Chang Cho[1] · Kuo-Wei Hung[1] · Dhilli Rao Gorja[1] · Chin Yu[1]

Received: 2 March 2015 / Accepted: 1 April 2015 / Published online: 9 April 2015

Biological context

The S100 proteins constitute the largest family within the EF-hand protein superfamily, shown to bind and control various proteins involved in several cellular functions such as proliferation, differentiation, apoptosis, Ca^{2+} homeostasis, and energy metabolism (Donato et al. 2013; Hermann et al. 2012; Leclerc and Heizmann 2011). They regulate a wide range of important cellular processes via protein–protein interactions (Schafer and Heizmann 1996). Calcium interactions with the EF-hand motifs result in a conformation change of target protein binding by exposing hydrophobic regions in S100 proteins (Smith and Shaw 1998). The calcium binding EF-hand motif starts the action with structural changes in the S100 proteins, allowing them to interact through target selectivity (Yap et al. 1999; Zimmer and Weber 2010). The protein S100A4 was first derived from both tumor and stroma. It is a homodimeric protein in solution and has been shown to function as a metastasis-promoting protein (Ambartsumian et al. 2005; Ismail et al. 2010). Its presence has now been well documented in many cancers including breast, colorectal, gastric, pancreatic, and bladder cancers. It plays a role in tumor formation and angiogenesis (Ambartsumian et al. 2001; Barraclough et al. 2009; Ford and Zain 1995). The hinge region and C-terminal EF-hand of the S100A4 protein are unique compared to other S100 proteins; however most S100 proteins are involved in target protein binding. The binding of calcium ions produces conformational changes in proteins resulting in exposure of the hydrophobic pocket of residues in helices 3 and 5, the hinge region, and the C-terminal EF-hand (Malashkevich et al. 2008; Mishra et al. 2012; Semov et al. 2005).

Previous results have provided insights into the dynamic mechanism of the C-terminal in S100A4 as a mediator of S100A4-driven metastasis, and they highlight its role in tuning the Ca^{2+}-binding affinity of S100A4. These results also suggest that locking the C-terminus to the core domain may be an alternative strategy for inhibiting its metastasis-promoting activities (Duelli et al. 2014). In this study, we mutated four cysteine residues (Cys3, Cys76, Cys81, and Cys86) on S100A4 to serine. The Cys3 residue is located very near the N-terminus, and the remaining three are in helix 4 (H4) (Pathuri et al. 2008). We labeled this protein 'mutant S100A4' and determined the three-dimensional structure of the calcium-bound of this protein using NMR. With this structure solved, we could explain why full length S100A4 has weaker calcium-binding affinity than its truncated form, where the last 13 amino acids in the C-terminal are deleted.

Methods and results

Protein expression and purification

Wild type human S100A4 contains four cysteine residues. One is at the N-terminal position of the amino acid sequence, and the remaining three are in the H4 region in the presence of DTT as a reducing agent, necessary in NMR buffer conditions. To study the three-dimensional structure

Electronic supplementary material The online version of this article (doi:10.1007/s10858-015-9927-6) contains supplementary material, which is available to authorized users.

✉ Chin Yu
cyu.nthu@gmail.com

[1] Department of Chemistry, National Tsing Hua University, Hsinchu 30013, Taiwan

of calcium-bound mutant S100A4 in NMR buffer conditions, Cys3, Cys76, Cys81, and Cys86 were mutated to serine because the structure of mutant S100A4 is stabilized in the presence of reducing agent. Recombinant mutant protein was over-expressed in *E. coli* of the BL21 (DE3) strain using pET-20b (Novagen) containing the cDNA of human mutant S100A4 (C3S, C76S, C81S, and C86S mutations). Double-labeled $^{15}N/^{13}C$ samples were grown in M9 minimal media with 0.5 g/L $^{15}NH_4Cl$ as the only nitrogen source and 2.4 g/L ^{13}C-labeled glucose as the only carbon source. Bacterial cultures were grown at 37 °C until an optical density of 0.7–0.9 was reached. They were then induced with 1.0 mM IPTG and grown for 16–20 h at 25 °C. The cells were harvested and lysed using a French press in resuspension buffer containing 50 mM Tris, 300 mM KCl, 1 mM DTT, 1 mM EDTA, and 10 % glycerol at pH 7.5. The lysed fraction was further sonicated for 20 min and centrifuged. A salting buffer containing 50 mM Tris, 300 mM KCl, 1 mM DTT, 1 mM EDTA, 10% glycerol, and 4.32 M $(NH_4)_2SO_4$ at pH 7.5 was added until the buffer content was 2.16 M $(NH_4)_2SO_4$. The mixture was then stored at 4 °C for 12 h. After ultra-centrifugation at 16,000 rpm and 4 °C for 30 min, the mutant S100A4 protein was found in the soluble fraction of the cell lysate and was then purified using a HiPrep 16/60 Phenyl FF column (GE Healthcare). The column was washed with 2 column volumes of A1 buffer (pH 7.5, containing 20 mM Tris, 300 mM KCl, 2 mM $CaCl_2$, 1 mM DTT, 1 mM EDTA, and 1 M $(NH_4)_2SO_4$) then washed with 2 column volumes of A2 buffer (pH 7.5, containing 20 mM Tris, 300 mM KCl, 2 mM $CaCl_2$, 1 mM DTT, and 1 mM EDTA) and eluted using a linear gradient with buffer B (pH 7.5, containing 20 mM Tris, 1 mM DTT, and 10 mM EDTA). The elution fractions were further purified using a HiLoad Superdex 200 column (GE Healthcare) equilibrated with the NMR buffer (pH 6.0, containing 16 mM Tris, 8 mM NaCl, 6 mM $CaCl_2$, 0.1 mM EDTA, 6 mM $CaCl_2$, and 0.34 mM NaN_3). The identity of the purified mutant S100A4 was checked using SDS-PAGE and confirmed using ESI-MASS.

NMR spectroscopy

The NMR spectra were collected on a Varian NMR 700 MHz spectrometer with a cryogenically cooled probe head. The double-labeled($^{15}N/^{13}C$)mutant S100A4 (0.5 mM) was prepared in 16 mM Tris-d_6 (pH 7.5, 90 % H_2O and 10 % D_2O) containing 8 mM NaCl, 6 mM $CaCl_2$, 0.1 mM EDTA, and 0.34 mM NaN_3. Samples for the measurement of intermolecular NOEs were prepared by mixing refolding of denatured $^{15}N/^{13}C$-labeled and unlabeled mutant S100A4 at 1:1 molar ratio. The combined information obtained from ^{1}H-^{15}N HSQC, HNCA, HNCO, HN(CO)CA, CBCA(CO)NH, HNCACB CC(CO)NH, HCC(CO)NH, ^{15}N-edited TOCSY-HSQC, HCCH-TOCSY, and HBHA(CO)NH spectra were utilized for the resonance assignment of the protein backbone and side chains. The NMR data were processed using VnmrJ/Topspin software and analyzed using SPARKY software (T. D. Goddard and D. G. Kneller, SPARKY 3.114, University of California, San Francisco).

Structure calculation

All NOE-based distance constraints used in the structure calculation were extracted from ^{13}C,^{15}N-edited NOESY-HSQC and ^{13}C-edited NOESY data sets. The angle restraints of the protein backbone dihedral angles were predicted using the TALOS + software (Shen et al. 2009). Hydrogen–bond distance restraints were derived using CSI (Wishart and Sykes 1994). Non-bonded interactions were calculated using the PARALLHDG force field (Linge and Nilges 1999). The initial folding of the dimeric form of the mutant S100A4 was generated based on manually-assigned intramolecular and intermolecular NOEs. The unambiguous NOE crosspeak is correctly described using outputs of previous ARIA runs and were analyzed and utilized as inputs for the structure calculation. Lower ambiguous cutoff parameters and slow-cooling standard simulated annealing protocols were used in the refinement steps of these calculations. The structures with the lowest overall energy were further calculated using water refinement in ARIA (Rieping et al. 2007). The quality of calculated conformers was checked by analyzing violations using the MOLMOL and PROCHECK software programs (Koradi et al. 1996; Laskowski et al. 1996).

Structure determination of calcium-bound mutant S100A4

The recombinant human mutant S100A4 exists as a dimer in solution. The ^{15}N-^{1}H HSQC spectra of human mutant S100A4 at pH 6.0 contains a single set of protein resonance with well-dispersed peaks (Fig. 1a), indicating that the recombinant protein is well-structured. The ^{1}H, ^{15}N, and ^{13}C resonances of calcium-bound mutant S100A4 at pH 6.0 and 25 °C have been completely assigned and deposited in the BMRB under accession number 25069.

The solution structure of human calcium-bound mutant S100A4 was further computed by ARIA based on distance geometry calculations and energy minimization with 2467 experimental and empirical NMR restraints, including 2014 NOE distance restraints, 46 hydrogen-bond distance restraints, and 160 dihedral angel restraints. Figure 1b shows the superimposition of backbone traces of an ensemble of 20 structures (selected from a set of 100

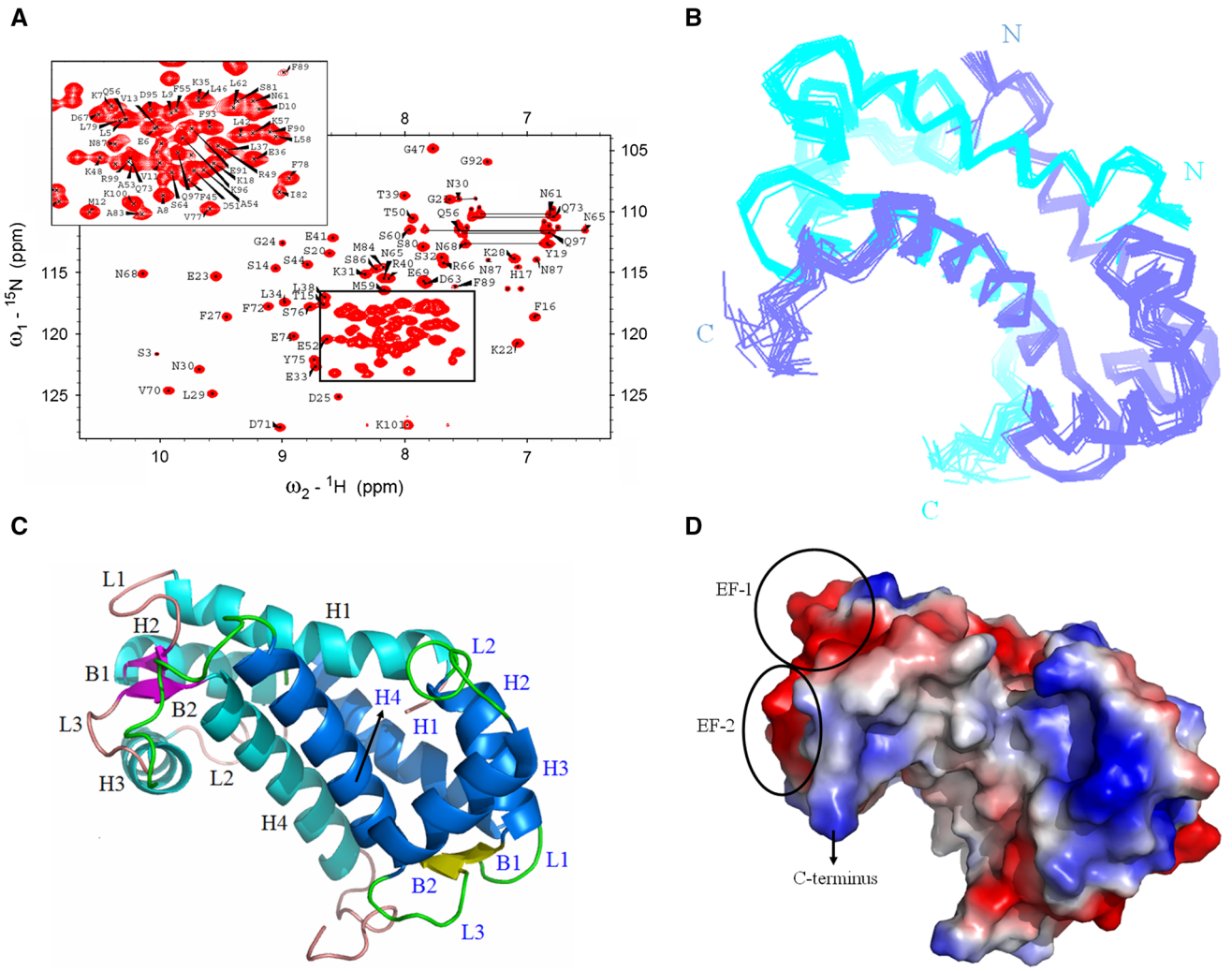

Fig. 1 **a** ^{1}H-^{15}N HSQC spectra of 0.5 mM of recombinant mutant S100A4 acquired at 298 K, and pH 6.0. The reduced spectrum was obtained in the presence of 10 mM DTT and enlargement of the spectrum in the box. **b** Overlay of the backbone traces of 20 lowest-energy NMR structures of the human mutant S100A4 dimer (*cyan* and *blue*). The N- and C-terminal ends are indicated as N and C, respectively. Ribbon representation of the tertiary structure of the human mutant S100A4 dimer. **c** The secondary structure elements of monomers, helix (*cyan* and *blue*), beta-sheet (*magenta* and *yellow*), loop and terminus (*pink* and *green*); they are labeled in black and blue respectively. **d** Surface charge distribution of the human mutant S100A4 dimer shown in the same orientation as (**c**). The surfaces that are negatively charged, positively charged, and hydrophobic are colored in *red*, *blue*, and *white*, respectively

structures) with the lowest energies, revealing a good agreement with NMR restraints. A summary of structure statistics for the best 20 conformers is given Table 1. The root mean square deviations (RMSDs) of these structures were 0.35 ± 0.12 and 0.82 ± 0.20 Å for the backbone atoms in structured regions and whole protein, respectively. The Ramachandran plot analysis indicated that 89 % of residues are in the most favored region, 6.6 % in the additionally allowed region, 4.4 % in the generously allowed region and 0.0 % of residues are in the disallowed regions. The NMR restraints and coordinates of 20 best structures of the human calcium-bound mutant S100A4 dimer have been deposited in Protein Data Bank under code number 2MRD. The last 11 residues (C-termianl) in mutant S100A4 structure as show in Fig. 1b contain 540 NOEs (among them, there are 134 long range NOEs) as show in Table 2. With so many NOEs detected in C-terminal by NMR, we concluded that the C-terminal region is not flexible with little structure, as detected in X-ray. It should be rigid as we reported.

Overview of the structure of human calcium-bound mutant S100A4 protein

The NMR solution structure of the human calcium-bound mutant S100A4 dimer reveals that each monomer comprises two β-strands (B1, a.a. 28–30 and B2, a.a. 69–71) and four α-helixes (H1, a.a. 3–21; H2, a.a. 31–42; H3, a.a.

Table 1 Structural statistics of human calcium-bound mutant S100A4

Number of NMR restraints per protomer	
Intraresidual NOEs (|i−j = 0|)	609
Sequential NOEs (|i−j = 1|)	406
Short-range NOEs (|3 ≥ i−j ≥ 2|)	308
Medium-range NOEs (|5 ≥ i−j ≥ 4|)	177
Long-range NOEs (|i−j > 5|)	221
Intermolecular NOEs	87
Hydrogen bonds	46
Dihedral angles	160
CNS energies (kcal mol^{-1})[a]	
E_{total}	−7081.74 ± 225.60
E_{bond}	82.10 ± 6.57
E_{angle}	440.35 ± 32.82
E_{impr}	994.41 ± 156.36
E_{dihed}	1095.85 ± 19.52
E_{vdw}	−822.60 ± 35.49
E_{elec}	−8871.85 ± 215.62
R.M.S.D. from experimental constraints	
Distances (Å)	0.045 ± 0.010
Dihedral angles (deg)	1.11 ± 0.31
R.M.S.D. from mean structure	
Backbone in structured region (Å)[b]	0.35 ± 0.12
Heavy atoms in structured region (Å)[b]	0.89 ± 0.11
Backbone in whole protein (Å)	0.82 ± 0.20
Heavy atoms in whole protein (Å)	1.33 ± 0.18
Ramachandran plot	
Most favored region (%)	89.0
Additionally allowed (%)	6.6
Generously allowed (%)	4.4
Disallowed (%)	0.0

[a] Residues in the structured regions: 3–21, 28–30, 31–42, 51–63, 69–71, and 72–91

Four α-helixes (H1, a.a. 3–21; H2, a.a. 31–42; H3, a.a. 51–63; and H4, a.a. 72–91) and two β-strands (B1, a.a. 28–30 and B2, a.a.69–71) in an H1–B1–H2–H3–B2–H4 topology. Two EF-hand motifs (helix–loop–helix), including H1–B1–H2 and H3–B2–H4, are connected by loop L2

51–63; and H4, a.a. 72–89) in an H1–B1–H2–H3–B2–H4 topology. Two EF-hand motifs (helix-loop-helix), including H1–L1–B1–H2 (pseudo EF-hand, EF-1) and H3–L3–B2–H4 (canonical EF-hand, EF-2), are connected by loop L2. B1 and B2 are short β-strands forming an anti-parallel β-sheet between two EF-hand motifs (Fig. 1c). Residues near EF-1 and EF-2 possess highly negatively charged surfaces for specific calcium interactions, and the last few residues of the C-terminal carry a positive charge (Fig. 1d).

Table 2 List of C-terminal NOE restraints

Residue	Number of NOEs	Long range NOEs
Phe90	18	2
Glu91	52	16
Gly92	30	2
Phe93	10	0
Pro94	62	18
Asp95	32	14
Lys96	74	42
Gln97	94	22
Pro98	50	14
Arg99	54	4
Lys100	40	0
Lys101	24	0
Total NOEs in C-terminus	540	134

The conformation and properties of the C-terminus in human mutant S100A4

The sequence comparison of mutant S100A4 proteins were analyzed by Multalin software (Corpet 1988). The alignment results indicate that the side chains of residues involved in the calcium binding is well conserved among three mammalian species with sequence identity from 64 to 73 %. It shows that negative residues (close to L1 and L3), including Glu23, Asp25, and Glu33 in EF-1 (N-terminal Ca^{2+} binding site) and Asp63, Asp67, Glu69, and Glu74 in EF-2 (C-terminal Ca^{2+} binding site), are well-conserved as calcium binding sites.

Calcium-binding affinity is important for the biological activity of S100A4. The long C-terminal region has a hydrophobic stem and some basic regions (Lys96, Arg99, Lys100, and Lys101; Fig. 2a). Binding partners of S100A4 (for example p53, annexin A2, CD16, and non-muscle myosin II) are mediated by Ca^{2+}-induced conformational changes. It has been widely recognized that the conformational change upon Ca^{2+} binding opens up a hydrophobic cleft on the protein surface (Grigorian et al. 2001; Pathuri et al. 2008; Semov et al. 2005).

It was reported that the truncated S100A4 (1-87 a.a.) had stronger calcium binding affinity than the full length form (1-101 a.a.) (Duelli et al. 2014). Figure 2b shows the overlapped structure of truncated S100A4 from an X-ray crystal (shown in yellow color) and full length structure from our NMR data (shown in green color). The two structures are similar except that the C-terminal of the full length form has 13 additional amino acids (shown in blue color). The orientation of these amino acids compels the

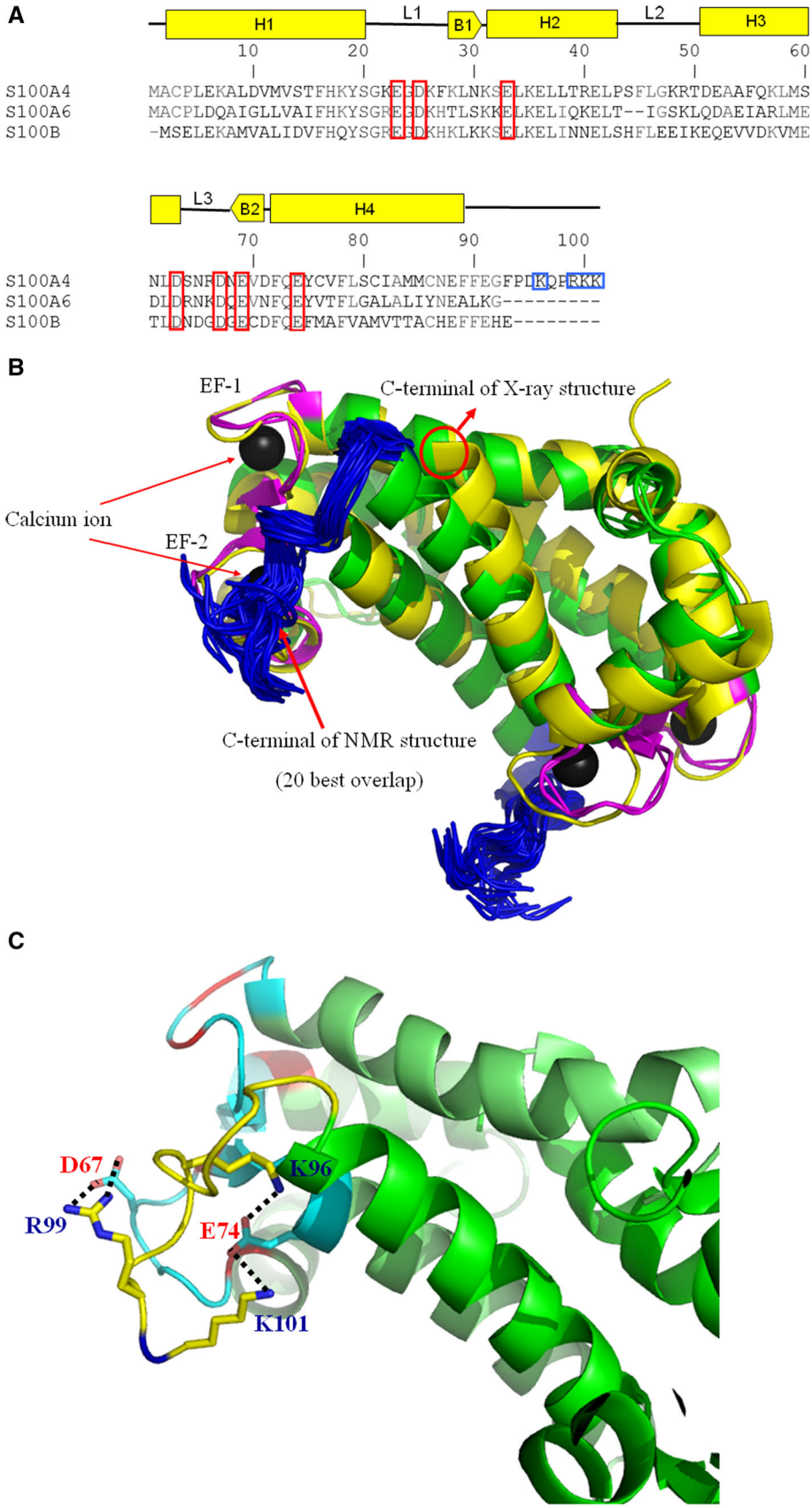

Fig. 2 **a** Sequence alignments of the human mutant S100A4 with others in the S100 proteins family; the *red boxes* are important negative charge residues that interact with Ca^{2+}; positive charge residues in the C-terminal are boxed in *blue*. **b** Overlap crystal structure of mutant human S100A4 dimer (*yellow*; PDB ID:4CFQ) with the solution structure of the calcium–bound mutant S100A4 dimer (*green*; PDB ID: 2MRD); calcium ions in crystal structure are depicted with black balls; EF-motifs, and the C-terminal of the solution structure are colored in magenta and blue. **c** The positive residues (*blue*) of the C-terminal (*yellow*) interact with the negative residues (*red*) in EF-hand (*cyan*). The interactions were indicated by *dot lines*

formation of EF-2, rendering the full length form weaker in calcium binding affinity. The charge–charge interaction between C-terminal and EF-2 provides the second reason why the full length form has weaker calcium binding affinity. On the C-terminal, positively charged residues Lys96, Arg99 and Lys101 are associated with the negatively charged residues Asp67 and Glu74 in EF-2. As show in Fig. 2c, the positive residues (blue) of the C-terminal (yellow) interact with the negative residues (red) in EF-2 (cyan). The interactions were indicated by dot lines. This interaction will neutralize the negative charge of EF-2 motif. After the negative charge of EF-2 motif had been neutralized, the binding affinity of calcium ion will be decreased.

In summary, the C-terminal (90–101 a.a.) of the full length form will destabilize the EF-2 through both steric and charge–charge interactions. As a result, the calcium binding affinity of the full length form of S100A4 is less than that of the truncated form.

Acknowledgments We acknowledge financial support from the National Science Council (NSC) Taiwan (Grant number NSC 103-2113-M-007-017-MY3). We appreciate having use of the 700 MHz Nuclear Magnetic Resonance Facility in the Chemistry Department, National Tsing Hua University.

Conflict of interest The authors declare no conflict of interest.

Ethical standards Research does not involve human participants and/or animals.

References

Ambartsumian N, Klingelhofer J, Grigorian M, Christensen C, Kriajevska M, Tulchinsky E, Georgiev G, Berezin V, Bock E, Rygaard J, Cao R, Cao Y, Lukanidin E (2001) The metastasis-associated Mts1(S100A4) protein could act as an angiogenic factor. Oncogene 20:4685–4695

Ambartsumian N, Grigorian M, Lukanidin E (2005) Genetically modified mouse models to study the role of metastasis-promoting S100A4(mts1) protein in metastatic mammary cancer. J Mol Graph 72:27–33

Barraclough DL, Platt-Higgins A, de Silva Rudland S, Barraclough R, Winstanley J, West CR, Rudland PS (2009) The metastasis-associated anterior gradient 2 protein is correlated with poor survival of breast cancer patients. Am J Pathol 175:1848–1857

Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16:10881–10890

Donato R, Cannon BR, Sorci G, Riuzzi F, Hsu K, Weber DJ, Geczy CL (2013) Functions of S100 proteins. Curr Mol Med 13:24–57

Duelli A, Kiss B, Lundholm I, Bodor A, Petoukhov MV, Svergun DI, Nyitray L, Katona G (2014) The C-terminal random coil region tunes the Ca(2)(+)-binding affinity of S100A4 through conformational activation. PLoS ONE 9:e97654

Ford HL, Zain SB (1995) Interaction of metastasis associated Mts1 protein with nonmuscle myosin. Oncogene 10:1597–1605

Grigorian M, Andresen S, Tulchinsky E, Kriajevska M, Carlberg C, Kruse C, Cohn M, Ambartsumian N, Christensen A, Selivanova G, Lukanidin E (2001) Tumor suppressor p53 protein is a new target for the metastasis-associated Mts1/S100A4 protein: functional consequences of their interaction. J Biol Chem 276: 22699–22708

Hermann A, Donato R, Weiger TM, Chazin WJ (2012) S100 calcium binding proteins and ion channels. Front Pharmacol 3:67

Ismail TM, Zhang S, Fernig DG, Gross S, Martin-Fernandez ML, See V, Tozawa K, Tynan CJ, Wang G, Wilkinson MC, Rudland PS, Barraclough R (2010) Self-association of calcium-binding protein S100A4 and metastasis. J Biol Chem 285:914–922

Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 14:51–55, 29–32

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8:477–486

Leclerc E, Heizmann CW (2011) The importance of Ca^{2+}/Zn^{2+} signaling S100 proteins and RAGE in translational medicine. Front Biosci 3:1232–1262

Linge JP, Nilges M (1999) Influence of non-bonded parameters on the quality of NMR structures: a new force field for NMR structure calculation. J Biomol NMR 13:51–59

Malashkevich VN, Varney KM, Garrett SC, Wilder PT, Knight D, Charpentier TH, Ramagopal UA, Almo SC, Weber DJ, Bresnick AR (2008) Structure of Ca^{2+}-bound S100A4 and its interaction with peptides derived from nonmuscle myosin-IIA. Biochemistry 47:5111–5126

Mishra SK, Siddique HR, Saleem M (2012) S100A4 calcium-binding protein is key player in tumor progression and metastasis: preclinical and clinical evidence. Cancer Metastasis Rev 31:163–172

Pathuri P, Vogeley L, Luecke H (2008) Crystal structure of metastasis-associated protein S100A4 in the active calcium-bound form. J Mol Biol 383:62–77

Rieping W, Habeck M, Bardiaux B, Bernard A, Malliavin TE, Nilges M (2007) ARIA2: automated NOE assignment and data integration in NMR structure calculation. Bioinformatics 23:381–382

Schafer BW, Heizmann CW (1996) The S100 family of EF-hand calcium-binding proteins: functions and pathology. Trends Biochem Sci 21:134–140

Semov A, Moreno MJ, Onichtchenko A, Abulrob A, Ball M, Ekiel I, Pietrzynski G, Stanimirovic D, Alakhov V (2005) Metastasis-associated protein S100A4 induces angiogenesis through interaction with Annexin II and accelerated plasmin formation. Trends Biochem Sci 280:20833–20841

Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS + : a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR 44:213–223

Smith SP, Shaw GS (1998) A change-in-hand mechanism for S100 signalling. Biochem Cell Biol 76:324–333

Wishart DS, Sykes BD (1994) The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. J Biomol NMR 4:171–180

Yap KL, Ames JB, Swindells MB, Ikura M (1999) Diversity of conformational states and changes within the EF-hand protein superfamily. Proteins 37:499–507

Zimmer DB, Weber DJ (2010) The calcium-dependent interaction of S100B with its protein targets. Cardiovasc Psychiatry Neurol 2010. doi:10.1155/2010/728052