Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and calcium/sulfur SAD phasing of the human EF-hand protein S100A2

Michael Koch,[a] Joachim Diez,[b] Armin Wagner[b]‡ and Günter Fritz[a,c]*

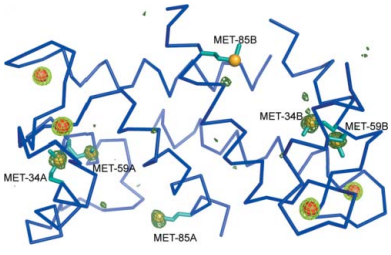
[a]Department of Biology, University of Konstanz, Postfach M665, Universitätsstrasse 10, 78457 Konstanz, Germany, [b]Swiss Light Source at the Paul Scherrer Institute, 5232 Villigen PSI, Switzerland, and [c]Department of Neuropathology, University of Freiburg, Breisacher Strasse 64, 79106 Freiburg, Germany

‡ Current address: Diamond Light Source Ltd, Chilton, Didcot OX11 0DE, England.

Correspondence e-mail: guenter.fritz@uniklinik-freiburg.de

Received 7 May 2010
Accepted 1 August 2010

PDB Reference: Ca^{2+}-bound S100A2, 3n22.



Human S100A2 is an EF-hand protein and acts as a major tumour suppressor, binding and activating p53 in a Ca^{2+}-dependent manner. Ca^{2+}-bound S100A2 was crystallized and its structure was determined based on the anomalous scattering provided by six S atoms from methionine residues and four calcium ions present in the asymmetric unit. Although the diffraction data were recorded at a wavelength of 0.90 Å, which is usually not assumed to be suitable for calcium/sulfur SAD, the anomalous signal was satisfactory. A nine-atom substructure was determined at 1.8 Å resolution using *SHELXD*, and *SHELXE* was used for density modification and phase extension to 1.3 Å resolution. The electron-density map obtained was well interpretable and could be used for automated model building by *ARP*/*wARP*.

1. Introduction

With 24 members, the S100 protein family constitutes the largest subgroup of EF-hand calcium-binding proteins (Fritz & Heizmann, 2004; Marenholz *et al.*, 2004, 2006). These proteins are small acidic proteins with a molecular mass of 10–12 kDa that carry a modified S100-specific EF-hand motif at the N-terminus (14 residues) followed by a classical C-terminal EF-hand motif (12 residues) linked by a hinge region. Upon Ca^{2+} binding, the majority of the S100 proteins undergo a major conformational change involving a 90° rotation of helix III of the C-terminal EF-hand. The movement of the helix opens the structure and exposes several hydrophobic residues to the solvent. These Ca^{2+}-induced structural changes are required for target-protein recognition and binding. Most S100 proteins function as homodimers, but several S100 proteins also form heterodimers and larger oligomers; *e.g.* S100A8/A9 is tetrameric in the Ca^{2+}-bound state (Korndörfer *et al.*, 2007), S100A12 can form hexamers and we have recently reported an octameric form of S100B (Ostendorp *et al.*, 2007). The different S100 proteins are involved in the regulation of diverse cellular processes, including cell-cycle regulation, cell growth, cell differentiation and motility (Marenholz *et al.*, 2004). Accordingly, the dysregulation of S100 proteins is associated with severe diseases such as chronic inflammation, cancer and neurodegenerative and cardiovascular disorders. S100A2 is unique among the S100 proteins with regard to its predominant nuclear localization (Mandinova *et al.*, 1998). S100A2 was identified as a novel tumour suppressor in human mammary epithelial cells (Lee *et al.*, 1991). Further studies suggested that S100A2 binds and activates p53 in a calcium-dependent manner, linking the tumour-suppressing activities of S100A2 with p53, and implied positive regulation of p53 through S100A2 (Mueller *et al.*, 2005; van Dieck *et al.*, 2009; Fernandez-Fernandez *et al.*, 2005, 2008). Besides Ca^{2+}, S100A2 binds Zn^{2+} with high affinity. In a recent study, we have shown that S100A2 contains two different Zn^{2+}-binding sites involving Cys residues whereby the binding of Zn^{2+} dramatically decreases the Ca^{2+} affinity (Koch *et al.*, 2007). The eight Cys residues per homodimer (194 amino acids) are solvent-exposed and oxidize rapidly, resulting in scrambled disulfide-linked protein (Koch *et al.*,

2006). Therefore, for structural studies we replaced all Cys residues by Ser and succeeded in crystallization and structure determination of the Ca^{2+}-free protein (Koch *et al.*, 2006, 2008).

The X-ray structure and molecular-dynamics studies showed that the replacement of the Cys residues by Ser had no impact on the structure of Ca^{2+}-free S100A2. In line with these data, Ca^{2+}-affinity measurements of the Cys-to-Ser mutant revealed no major effect on the Ca^{2+} affinity, showing that the function of the protein was not affected (Koch *et al.*, 2007). The X-ray structure of the Ca^{2+}-free protein provided detailed information on the inactive form of S100A2 (Koch *et al.*, 2008). In order to obtain insight into the Ca^{2+}-induced structural changes required for recognition and binding of target proteins such as p53, we performed crystallization trials on S100A2 in the presence of Ca^{2+}. Here, we report the crystallization of Ca^{2+}-bound S100A2 and the structure determination of S100A2 using sulfur and calcium SAD. The data presented here show that highly redundant data sets recorded at wavelengths that are not assumed to be suitable for calcium/sulfur SAD can contain sufficient anomalous data for structure determination.

Table 1
Data-collection statistics.

Values in parentheses are for the outermost resolution shell.

Space group	$P3_121$
Unit-cell parameters (Å, °)	$a = b = 84.62$, $c = 58.28$, $\alpha = \beta = 90$, $\gamma = 120$
Resolution (Å)	45–1.3 (1.5–1.3)
Wavelength (Å)	0.90
Completeness (%)	100 (99.9)
Total observations	988032 (218520)
Unique reflections	59372 (20501)
Redundancy	16.6 (10.6)
R_{merge}† (%)	4.1 (62.4)
R_{meas}‡ (%)	4.2 (65.5)
$R_{mrgd\text{-}F}$§ (%)	5.4 (38.4)
$R_{p.i.m.}$¶ (%)	0.7 (20.4)
$I/\sigma(I)$	27.4 (4.4)
Wilson B factor (Å^2)	19.8
No. of molecules per ASU	1 homodimer
Matthew coefficient (Å^3 Da^{-1})	2.76
Solvent content (%)	54.4

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{meas} = \sum_{hkl}[N/(N-1)]^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. § $R_{mrgd\text{-}F} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i |F_i(hkl) - \langle F(hkl)\rangle| / \sum_{hkl}\sum_i F_i(hkl)$. ¶ $R_{p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

2. Material and methods

2.1. Protein expression and purification

In previous studies, we have shown that wild-type S100A2 is prone to oxidation (Koch *et al.*, 2007) and used a mutant in which all four Cys residues were replaced by Ser for crystallization of the apo form of the protein (Koch *et al.*, 2006). Structural and molecular-dynamics studies on this mutant showed that the Cys-to-Ser mutation is most likely to have no effect on the structure of S100A2 (Koch *et al.*, 2008). This Cys-to-Ser mutant protein was purified as described previously (Koch *et al.*, 2006). To summarize, the S100A2 Cys-to-Ser mutant was expressed in *Eschericha coli* BL21 (DE3) and purified by Ca^{2+}-dependent interaction on a phenyl Sepharose Fast Flow column (100 ml, GE Healthcare) and size-exclusion chromatography on a Superdex 75 (26/60) column (GE Healthcare). The pure protein was concentrated to 20 mg ml^{-1} by ultrafiltration and aliquots were cryocooled in liquid nitrogen.

2.2. Protein crystallization

Prior to crystallization, the buffer was exchanged to 10 m*M* Tris–HCl, 20 m*M* $CaCl_2$ pH 7.6 on an NAP5 column and the protein concentration was adjusted to 15 mg ml^{-1}. Crystallization trials of Ca^{2+}-bound S100A2 were performed by the sitting-drop or hanging-drop vapour-diffusion method. Initial crystallization trials were prepared using a Cartesian NanoDrop dispenser. For each crystallization solution, a series of three drops with different ratios were mixed and subsequently equilibrated by vapour diffusion against 100 µl of the crystallization solution in the reservoir. Each drop contained 200 nl protein solution and was mixed with 100, 200 or 400 nl crystallization buffer. Although the dispenser was able to produce 50 nl drops, we found that evaporation of water from the protein drop was too rapid within the timescale of plate preparation and drop mixing. We therefore chose a protein drop volume of 200 nl, which we found to be the best compromise between protein savings and reproducibility of the crystallization trials. Crystallization plates were incubated at 293 K. After 10 d small crystals appeared and the condition was refined using hanging-drop vapour diffusion, mixing 1 µl protein solution with 1 µl 0.1 *M* sodium acetate pH 4.0, 40% PEG 400 and equilibrating against 500 µl of the same buffer. Crystals of dimensions of up to 160 × 160 × 200 µm were observed after two weeks (Fig. 1*a*). The crystallization conditions were similar to those

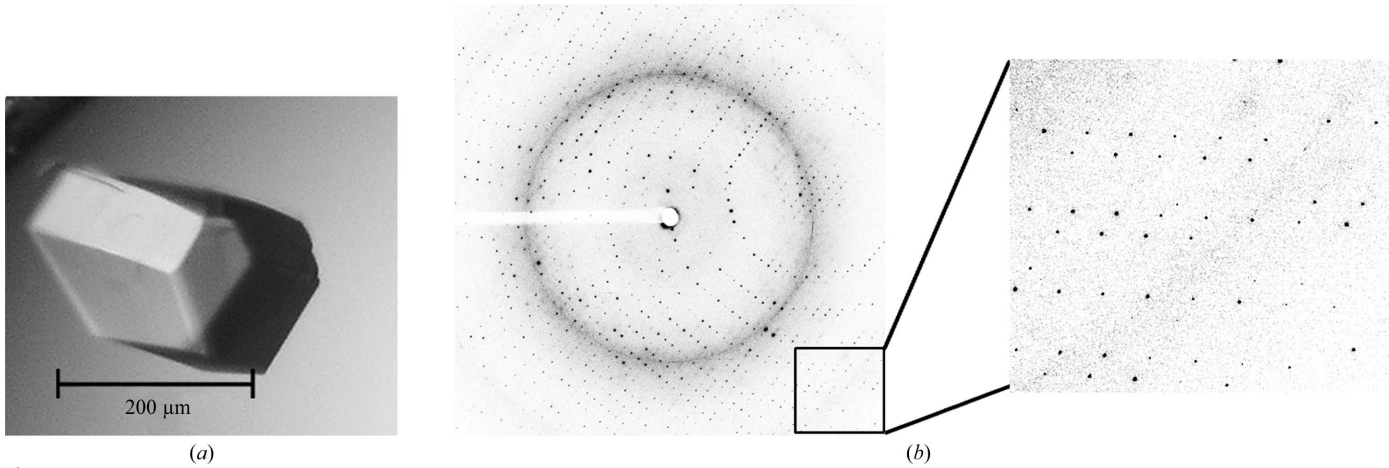


Figure 1
(*a*) Crystal and (*b*) diffraction pattern of Ca^{2+}-bound S100A2. The resolution at the image border is 2.5 Å. A detail of the diffraction pattern with spots at a resolution of up to 1.9 Å is shown on the right.

for apo S100A2, which also contained acetate as buffer ion and PEG 4000 as precipitant. The pH of 4.0 for crystallization is close to the calculated pI of 4.7 for S100A2, *i.e.* the net charge of the protein will be rather low, which may contribute to successful crystallization (Kantardjieff & Rupp, 2004). The preference for PEG for successful crystallization was a negative rather than a positive selection. Since Ca^{2+} is present in the protein solution, any crystallization buffer containing phosphate, sulfate or tartrate will result in the formation of salt crystals and depletion of Ca^{2+} from the protein. Therefore, crystallization buffers containing these ions were omitted from the initial screen.

The high PEG 400 concentration was sufficient as a cryoprotectant for the Ca^{2+}-bound S100A2 crystals. The crystals were cryocooled in a nitrogen cryostream at 100 K since direct freezing in liquid nitrogen led to breakage of the crystals.

2.3. Data collection and processing

Data collection was carried out on beamline X06SA of the Swiss Light Source (Villigen, Switzerland) using a MAR 225 Mosaic CCD detector (MAR Research). The diffraction data were processed with the *XDS* program package (Kabsch, 2010). The crystals of S100A2 diffracted to high resolution using synchrotron radiation (Fig. 1*b*) and a complete high-redundancy data set was recorded to 1.3 Å resolution. To obtain complete data in the low-resolution range, an initial data set was recorded with a strongly attenuated beam and consisted of 500 images over a total φ of 500° and included the low-resolution data up to 2.2 Å resolution. A second data set was recorded from the same crystal over a φ of 180° and contained a high-resolution data set up to 1.3 Å resolution. Data statistics are given in Table 1. The space group was determined as $P3_121$, with unit-cell parameters $a = b = 84.62$, $c = 58.28$ Å. Analysis of the solvent content revealed that two subunits were most likely to be present in the asymmetric unit. The solvent content of the asymmetric unit was calculated as 54.9%, with a Matthews coefficient of 2.73 $Å^3$ Da^{-1}. Analysis of the data revealed very good statistics. Despite the fact that the data were recorded at the short wavelength of 0.90 Å significant anomalous signal could be observed and the data were therefore regarded as being suitable for Ca^{2+}/S SAD phasing. At this wavelength the anomalous contributions f'' to the scattering factor are $f''(Ca^{2+}) = 0.49\ e^-$ and $f''(S) = 0.20\ e^-$, resulting in an estimated anomalous signal (Hendrickson & Teeter, 1981) of $\Delta F/F = 0.57\%$.

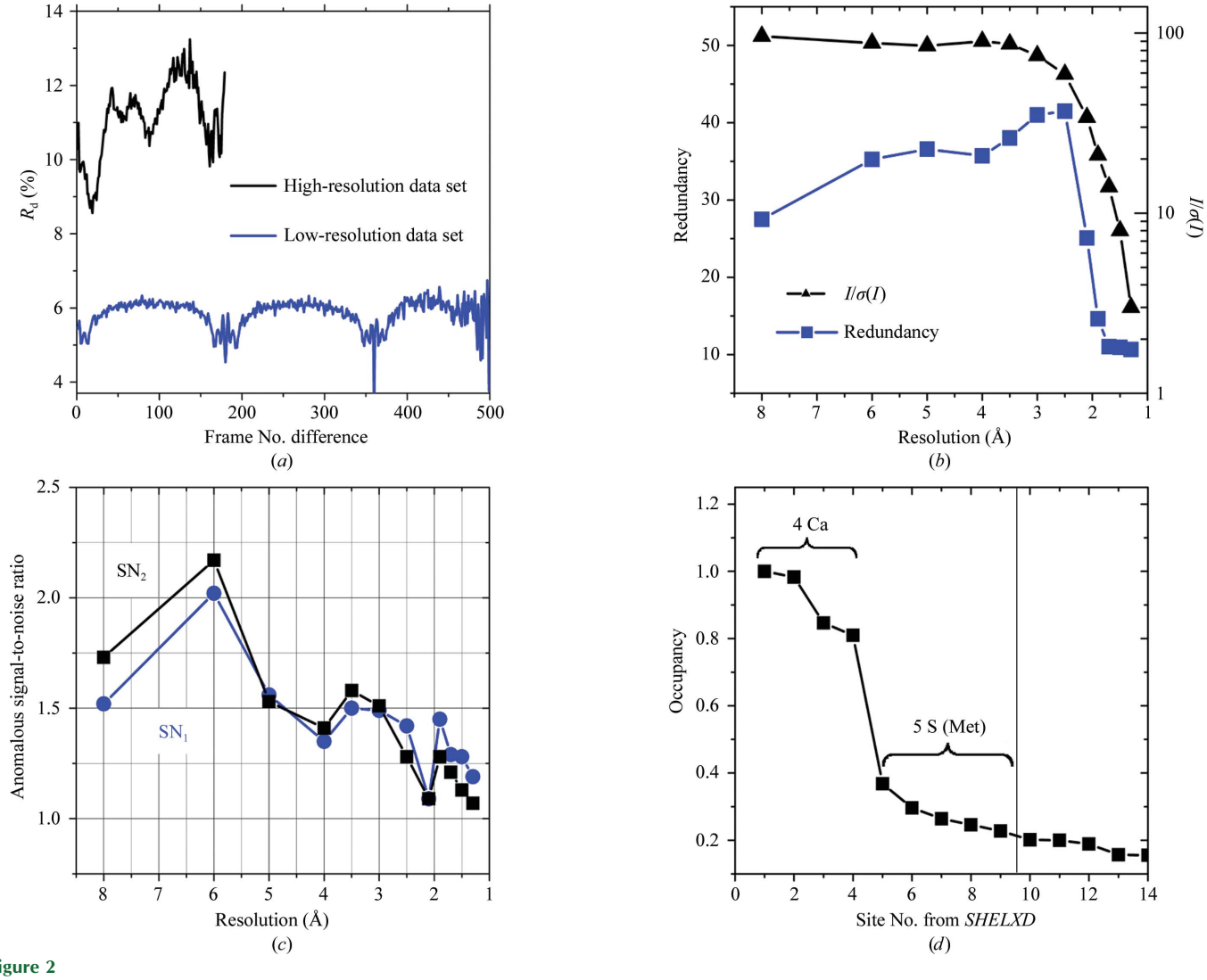


Figure 2
Data statistics. (*a*) Radiation damage monitored by *XDSSTAT* (Diederichs, 2006). (*b*) Redundancy and signal-to-noise ratio. Blue squares, redundancy of data; black triangles, signal-to-noise ratio. (*c*) Anomalous signal-to-noise ratios. Blue circles: SN_1, a signal-to-noise ratio based on input σ-values from *XPREP* (Bruker–Nonius). Black squares: SN_2, a signal-to-noise ratio based on variances of F^+ and F^- from *XPREP*. (*d*) Occupancy of heavy-atom substructure determined by *SHELXD* (Sheldrick, 2008).

Table 2
Anomalous scattering by Ca^{2+} and S in S100A2.

Site No.	Relative occupancy	Peak height in anomalous map	Atom ID
1	1.0	39.7	Ca^{2+}, EF-hand 1, chain *B*
2	0.99	34.6	Ca^{2+}, EF-hand 2, chain *B*
3	0.86	33.9	Ca^{2+}, EF hand 1, chain *A*
4	0.85	30.1	Ca^{2+}, EF-hand 2, chain *A*
5	0.38	10.6	S, Met59, chain *A*
6	0.29	10.2	S, Met85, chain *A*
7	0.28	10.2	S, Met59, chain *A*
8	0.22	10.2	S, Met34, chain *B*
9	0.21	7.9	S, Met34, chain *A*

For comparison, Se SAD phasing with data collected at the peak of an SeMet derivative with scattering factor f''(Se) = 3.8 e^- would result in an estimated anomalous signal of $\Delta F/F = 4.3\%$.

3. Results

3.1. Determination of the heavy-atom substructure

Analysis of the data sets using *XDSSTAT* (Diederichs, 2006) revealed that the crystal did not suffer from radiation damage during recording of the low-resolution data set, as indicated by a low R_d over the course of data collection. The R_d for the high-resolution data set increased significantly, indicating the accumulation of radiation damage, but the quality of the data was still regarded as sufficient for phasing. The mean redundancy was ∼35 for the resolution range between 8 and 2.5 Å but dropped sharply to 10 for data with higher resolution. The signal-to-noise ratio for the 8–2.5 Å resolution range was also high with a value of about 90 and, like the redundancy, dropped to lower values in the higher resolution range. The anomalous signal-to-noise ratios SN_1 and SN_2 calculated by *XPREP* (Bruker–Nonius) are regarded as good indicators of the quality of data used for heavy-atom substructure determination and phasing. As a rule of thumb, a value of 1.5 is assumed to be a reasonable indicator for a cutoff to truncate the data. The SN_1 and SN_2 values dropped sharply to below 1.5 at ∼2.2 Å but increased again at 1.9 Å (Fig. 2). The minimum at 2.2 Å indicates the presence of ice perturbing the data quality, although no ice rings were observed in the diffraction images. We therefore decided to analyze the two data sets individually or combined and truncated at 1.8 Å resolution with *SHELXD* (Sheldrick, 2008) to determine a heavy-atom substructure.

Table 3
Refinement statistics.

Values in parentheses are for the outermost resolution shell.

Resolution (Å)	45–1.3 (1.33–1.30)
Completeness (%)	100 (99.9)
R_{cryst}† (%)	14.8 (25.5)
R_{free}‡ (%)	18.0 (29.4)
Protein atoms	1410
Solvent atoms	191
Ligand/ion atoms	17
B factors (Å^2)	
Overall	31.2
Protein	28.7
Ligand	36.8
Ions	19.5
Solvent	51.0
R.m.s. deviations	
Bond lengths (Å)	0.027
Bond angles (°)	2.23
Ramachandran plot, residues in	
Core region (%)	95.7
Allowed region (%)	4.3
Generously allowed region (%)	0.0
Disallowed region (%)	0.0

† $R_{cryst} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively. ‡ A randomly selected 5% of reflections were excluded from refinement and used for the calculation of R_{free}.

The best results with satisfactory statistics were obtained when the combined data set was used and correct solutions were obtained with a high success rate of 225 correct solutions from 1000 trials. The solution revealed four sites with a high occupancy of between 1.0 and 0.8 and, after a sharp drop in occupancy, six sites with occupancies of between 0.36 and 0.22 and six further sites with occupancies between 0.18 and 0.14. Such a clear drop in occupancy is usually considered as a cutoff criterion to differentiate between heavy-atom sites and noise. The size of the asymmetric unit suggested that there is one S100A2 dimer in the asymmetric unit. The S100A2 dimer contains four EF-hands, each of which binds one calcium ion. Considering the higher scattering factors of calcium, we expected four clear sites as well as six further sites with a relative lower occupancy for the S atoms of methionine. According to the scattering factors, the occupancies of the S atoms relative to the calcium ions were estimated as between 0.4 and 0.3. However, most of the sites with lower occupancies were below these values and a clear assignment to S atoms was not possible (Table 2). Therefore, we used the 12 sites with the highest occupancies for phase calculations to make sure that the weak phase information of the S atoms was also exploited. The sites were used

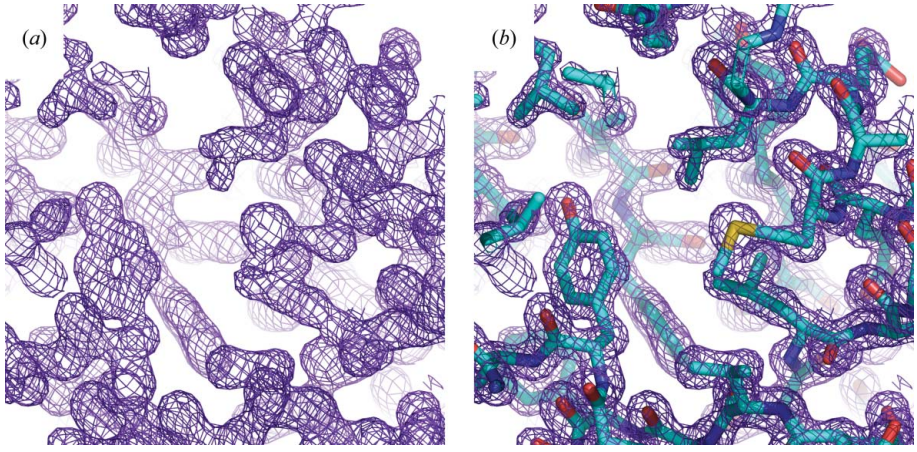


Figure 3
(*a*) Initial experimental F_{obs} map at 2.0σ as calculated by *SHELXE* using the sites determined by *SHELXD* and phase extension to 1.3 Å. (*b*) The refined model of Ca^{2+}-S100A2 is shown as a stick model overlaid with the initial experimental map.

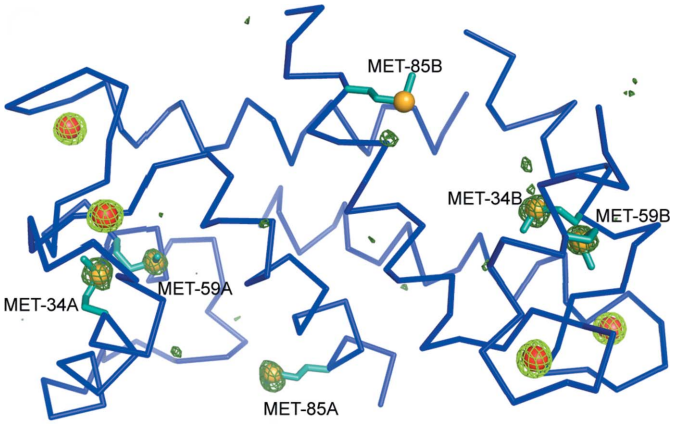


Figure 4
C^{α} trace of the final S100A2 model. Methionine residues are shown as sticks, with SG atoms as yellow spheres. The calcium ions are shown as red spheres. At 0.90 Å the anomalous signal of calcium is ∼2.5-fold stronger than that of sulfur. The heavy-atom substructure was determined with *SHELXD* and the anomalous difference Fourier map as calculated by *SHELXE* is shown at 8σ in bright green around the Ca^{2+} ions and at 3.5σ in dark green around the S atoms. The phases for the anomalous difference Fourier map were calculated from the final model. This figure was generated with *PyMOL* (DeLano, 2002).

without further refinement for phase calculation and extension to 1.3 Å with *SHELXE* (Sheldrick, 2008). The correct solution was identified by the high contrast of 1.07 *versus* 0.44.

3.2. Model building and refinement

The initial experimental map as calculated by *SHELXE* was of extraordinary quality, as shown in Fig. 3(*a*), and was well suited for automatic model building and refinement by *ARP*/*wARP* (Langer *et al.*, 2008; Perrakis *et al.*, 1999), which traced 172 of 194 residues. Subsequent alternating cycles of manual rebuilding were performed with the programs *O* (Jones *et al.*, 1991) and *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010) and refinement was carried out using *REFMAC*5 (Murshudov *et al.*, 1997, 1999). In the last steps of refinement H atoms were added in riding positions. The refinement statistics are summarized in Table 3. Fig. 3(*b*) shows the refined model with the initial experimental map (model coordinates and structure factors have been deposited in the Protein Data Bank with accession code 3n22). In order to clarify which sites identified by *SHELXD* (Sheldrick, 2008) arose from the anomalous signal of calcium and sulfur, the anomalous difference Fourier map was overlaid with the refined structure (Fig. 4). The four highest peaks were clearly assigned to the calcium ions bound to the EF-hands. Five of the eight S atoms from methionine residues were also identified by *SHELXD*. Overall, the data clearly demonstrate that even at a wavelength of 0.90 Å and an anomalous signal of only 0.57% accurate phase information can be gained from high-quality diffraction data from crystals diffracting to atomic resolution.

The financial support of the Deutsche Forschungsgemeinschaft to GF (DFG FR 1488/5-1; FR 1488/3-1) is gratefully acknowledged.

References

DeLano, W. L. (2002). *The PyMOL Molecular Viewer*. http://www.pymol.org.
Dieck, J. van, Fernandez-Fernandez, M. R., Veprintsev, D. B. & Fersht, A. R. (2009). *J. Biol. Chem.* **284**, 13804–13811.
Diederichs, K. (2006). *Acta Cryst.* D**62**, 96–101.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Fernandez-Fernandez, M. R., Rutherford, T. J. & Fersht, A. R. (2008). *Protein Sci.* **17**, 1663–1670.
Fernandez-Fernandez, M. R., Veprintsev, D. B. & Fersht, A. R. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 4735–4740.
Fritz, G. & Heizmann, C. W. (2004). In *Handbook of Metalloproteins*, edited by A. Messerschmidt, R. Huber, T. Poulos & K. Wieghardt, Vol. 2. Chichester: John Wiley & Sons.
Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kantardjieff, K. A. & Rupp, B. (2004). *Bioinformatics*, **20**, 2162–2168.
Koch, M., Bhattacharya, S., Kehl, T., Gimona, M., Vasak, M., Chazin, W., Heizmann, C. W., Kroneck, P. M. & Fritz, G. (2007). *Biochim. Biophys. Acta*, **1773**, 457–470.
Koch, M., Diez, J. & Fritz, G. (2006). *Acta Cryst.* F**62**, 1120–1123.
Koch, M., Diez, J. & Fritz, G. (2008). *J. Mol. Biol.* **378**, 933–942.
Korndörfer, I. P., Brueckner, F. & Skerra, A. (2007). *J. Mol. Biol.* **370**, 887–898.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Lee, S. W., Tomasetto, C. & Sager, R. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 2825–2829.
Mandinova, A., Atar, D., Schäfer, B. W., Spiess, M., Aebi, U. & Heizmann, C. W. (1998). *J. Cell Sci.* **111**, 2043–2054.
Marenholz, I., Heizmann, C. W. & Fritz, G. (2004). *Biochem. Biophys. Res. Commun.* **322**, 1111–1122.
Marenholz, I., Lovering, R. C. & Heizmann, C. W. (2006). *Biochim. Biophys. Acta*, **1763**, 1282–1283.
Mueller, A., Schäfer, B. W., Ferrari, S., Weibel, M., Makek, M., Höchli, M. & Heizmann, C. W. (2005). *J. Biol. Chem.* **280**, 29186–29193.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Ostendorp, T., Leclerc, E., Galichet, A., Koch, M., Demling, N., Weigle, B., Heizmann, C. W., Kroneck, P. M. & Fritz, G. (2007). *EMBO J.* **26**, 3868–3878.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.