Zinc-Binding Site of an S100 Protein Revealed. Two Crystal Structures of Ca^{2+}-Bound Human Psoriasin (S100A7) in the Zn^{2+}-Loaded and Zn^{2+}-Free States[†,‡]

D. E. Brodersen, J. Nyborg, and M. Kjeldgaard*

Macromolecular Crystallography, Institute of Molecular and Structural Biology, Aarhus University, Gustav Wieds Vej 10c, DK-8000 Aarhus C, Denmark

Received October 19, 1998; Revised Manuscript Received November 30, 1998

ABSTRACT: The crystal structure of human psoriasin (S100A7) in the native, calcium-bound form has been determined from two crystal forms of the protein crystallized with and without divalent zinc. The overall structures of the dimeric protein closely resemble the previously determined holmium-substituted structure. The structures also reveal a zinc-binding site of the protein, which is formed by three histidines and an aspartate residue. Together, these residues coordinate the zinc ion in a way similar to the pattern seen in certain metalloproteases and in particular the collagenase family of proteins. Sequence comparison suggests that this zinc site is present in a number of the remaining members of the S100 family. The structure of S100A7 crystallized in the absence of zinc further shows that loss of zinc results in a reorganization of the adjacent empty and distorted EF-hand loop, causing it to resemble a calcium-loaded EF-hand.

Psoriasin (S100A7) is a small, homodimeric protein of 22.7 kDa belonging to the S100 class of calcium binding EF-hand proteins. The protein is abundant in many human cell types, but most remarkably, it was found to be at least 5 times upregulated in the keratinocytes of patients suffering from the chronic and hyperproliferative skin disease psoriasis (*1*−*3*). Even though much scientific research has been carried out in establishing the molecular basis of psoriasis, very little is known presently about this complex disease, and the exact function of S100A7 is currently unknown. It is, however, believed that the protein responds to transient changes in the cellular calcium concentration by binding yet unidentified receptor molecules.

The three-dimensional crystal structure of S100A7 was recently determined at atomic resolution using a holmium-substituted variant of the protein to solve the crystallographic phase problem (*4*). Earlier work suggests a great degree of chemical similarity between calcium and the lanthanide elements with respect to binding by proteins (*5*, *6*), and it is believed that the derivative S100A7 structure reflects the physiological and calcium-loaded form of the protein with high accuracy. The overall structure of S100A7 is largely similar to other S100 proteins for which the structure is known, but several important differences exist with respect to, for example, interhelical angles and the layout of the individual helices. Most importantly, S100A7 seems to lack the ability to bind calcium in the N-terminal EF-hand, consistent with a three-residue shortening of the loop region that should otherwise accommodate the ion.

While calcium is known to stimulate many cell types, zinc often has an antagonistic effect on cell growth. Important proteins such as calmodulin and the S100 proteins are believed to take part in this regulation (*7*). The existence of distinct EF-hand calcium binding motifs is well established for all members of the S100 family, but numerous studies also imply that divalent zinc should be considered a major regulator of their function (*8*−*12*). In fact, it has been postulated that the S100 proteins should be considered as primarily zinc-regulated proteins and not merely as proteins that bind zinc for structural reasons. The latter conclusion is based on the fact that protein stability does not seem to be altered upon binding or loss of the zinc ion (*13*).

A number of the S100 proteins have been shown to bind zinc with significant affinity, but there seems to be no consensus with respect to the number of distinct sites present. For most of the studied S100 proteins, a relatively large number (six to eight) of zinc sites have been reported (*8*, *10*, *12*−*14*), whereas only one zinc site per monomer was reported for S100A12 (*15*). In all cases, however, it is found that the calcium- and zinc-binding sites are distinct and that interaction with the two ions conveys different structural changes in the proteins. For example, it has been suggested that the binding of zinc by S100B results in the exposure of hydrophobic groups (*16*). In some cases, these conformational changes supposedly lower the affinity of the protein for calcium, whereas in other cases, the affinity increases. With respect to the location of the zinc-binding site, it was proposed that a His-X-X-X-His motif which is often found near the C terminus in a number of the proteins might be involved, provided the motif is present in a helix (*12*, *15*, *17*). Moreover, some of the S100 proteins contain high amounts of cysteine residues, and it has been claimed that

[†] This work was funded by a special program for biotechnology under the Danish Natural Science Research Council and a Hallas-Møller fellowship from the Novo Nordisk Foundation to M.K.

[‡] Atomic coordinates for the structures described in this paper have been deposited in the Protein Data Bank under entries 2psr (high-zinc structure) and 3psr (low-zinc structure), along with structure factor amplitudes.

* To whom correspondence should be addressed. E-mail: mok@imsb.au.dk.

10.1021/bi982483d CCC: $18.00
Published on Web 01/20/1999

these could possibly participate in a zinc finger-like structure (*9*, *10*).

In a previous study, the divalent cation binding properties of S100A7 were studied (*18*) and the protein was found to bind several calcium and zinc atoms per monomer, the affinity for calcium being considerably higher than that for zinc. The binding constant for the first zinc ion was measured to 100 μM, which is in the typical range for the S100 proteins. The existence of one or more zinc-binding sites has thus been described previously in a number of biochemical studies for members of the S100 family, but the location of the site and the zinc ligands have not been identified so far.

In this paper, we describe for the first time the exact structure of the regulatory zinc site in an S100 protein, and it is shown how zinc binding is similar to what is seen for another important class of proteins. We report two crystal structures of the native, calcium-loaded form of S100A7 in the presence and absence of divalent zinc. The structures reveal the location and conformation of the zinc-binding site of S100A7 as well as structural changes imposed by the binding of the ion.

EXPERIMENTAL PROCEDURES

Purification and Crystallization. Recombinant S100A7 was expressed in *Escherichia coli* as a fusion protein, purified, and crystallized as previously reported (*19*). Different crystal forms of the protein were obtained depending on whether crystallization was carried out in the absence (type I) or in the presence (type II) of divalent zinc ions. Rod-shaped crystals of type I were obtained for a protein concentration of approximately 5.5 mg/mL in 22% (w/v) polyethylene glycol (PEG)[1] 4000, 5−10% (v/v) glycerol, and 5−10 mM $CaCl_2$ in a pH range of 6.5−8.0 at 20 °C. In contrast, crystals of type II were bipyramidic and appeared at pH 6.7 and 4 or 20 °C in 5−12% (w/v) polyethylene glycol (PEG) 4000, 10% (v/v) glycerol, and 20 mM $ZnSO_4$ using a protein concentration of approximately 14 mg/mL.

Data Collection and Processing. For crystal form II, monochromatic X-ray diffraction data ($\lambda = 0.862$ Å) were collected to high resolution at a cryogenic temperature (100 K) at the BW7B wiggler beamline (EMBL-Hamburg, Germany) on a single crystal mounted in a loop (*19*). These crystals could be frozen directly from the mother liquor and diffracted to at least 2.0 Å. Native data collected from a single crystal were integrated and processed with the DENZO/SCALEPACK suite (*20*). The crystals were tetragonal in space group $P4_x2_12$ with the following dimensions: $a = b = 51.96$ Å, $c = 116.03$ Å, and $\alpha = \beta = \gamma = 90°$; however, the identity of the 4-fold axis (proper or screw) could not be determined because of missing reflections along the *c*-axis (see Table 1 for data collection statistics). Structure factor amplitudes were calculated with the CCP4 program TRUNCATE (*21*, *22*).

Crystals of type I did not diffract as well as the tetragonal crystal form. For these crystals, room-temperature data were collected to 2.5 Å on a single crystal mounted in a glass capillary on an R-AXIS II detector at the H. C. Ørsted Institute (Copenhagen, Denmark), operating a Rigaku rotating anode emitting CuKα radiation at 1.5418 Å (*19*). A highly redundant data set was recorded in frames of $\Delta\varphi = 2.5°$. The crystals are orthorhombic and belong to space group $P2_12_12_1$ with the following cell dimensions: $a = 52.15$ Å, $b = 56.67$ Å, $c = 76.38$ Å, and $\alpha = \beta = \gamma = 90°$. The data were integrated with DENZO and reduced with the CCP4 program SCALA (*22*−*24*). Structure factors were calculated with TRUNCATE.

Table 1: Summary of Crystallographic Data Collection

	crystal form I	crystal form II
resolution range (Å)	38.4−2.5	19.9−2.05
detector	R-AXIS II	MAR30
temperature (K)	298	100
space group	$P2_12_12_1$	$P4_32_12$
cell dimensions		
a (Å)	52.15	51.96
b (Å)	56.67	51.96
c (Å)	76.38	116.03
no. of unique reflections	8199	10791
average redundancy[a]	5.4 (5.3)	7.0 (6.9)
completeness (%)[a]	99.3 (99.3)	98.9 (100.0)
R_{sym} (%)[a]	12.2 (54.4)	11.3 (42.3)
I/σ_I[a]	5.4 (1.4)	15.9 (5.0)

[a] Numbers in parentheses correspond to the outermost resolution shell, 2.64−2.50 and 2.12−2.05 Å for crystal forms I and II, respectively.

Molecular Replacement. Both structures were solved by molecular replacement (*25*) using the program AMORE (*22*, *26*). Because the structures were believed to be similar to the structure of holmium-substituted S100A7 previously determined in our laboratory using MAD phasing (*4*), a search model was constructed that included all main chain and side chain atoms from one monomer (residues A1-A100), but not solvent or the heavy metal ion. The location of the molecule in the data from crystal form II was determined successfully using space group $P4_32_12$, thus fixing the unknown nature of the 4-fold axis. The resulting *R*-factor was 40.0% and the correlation coefficient 55.0%. The crystals contain a single monomer in the asymmetric unit with the dimer dyad placed along the proper 2-fold crystallographic axis.

For the orthorhombic data, four clear solutions to the rotation function appeared for the monomer which two by two corresponded to symmetry-related solutions for the dimer. This crystal form thus contained both monomers in the asymmetric unit. The final *R*-factors were 47.8% and 49.4% for the two monomers with correlation coefficients of 32.1% and 25.2%, respectively.

Phasing. Both molecular replacement solutions were initially subjected to 10 cycles of rigid-body refinement with the program REFMAC (*22*, *27*, *28*) using data to the maximum resolution. To minimize bias from the calculated phases of the search model, solvent flattening and histogram matching were applied for several cycles in the program DM (*22*, *29*), exploiting the high solvent contents of 60% for the tetragonal crystal form and 50% for the orthorhombic crystal form. This significantly improved the initial molecular replacement phases, and investigation of the resulting electron density map revealed several cases of reorganization

[1] Abbreviations: PEG, polyethylene glycol; rms, root-mean-square; rmsd, root-mean-square deviation; MMP, matrix metalloprotease; NMR, nuclear magnetic resonance; asu, asymmetric unit; NCS, noncrystallographic symmetry.

Table 2: Summary of Crystal Structure Refinement

	crystal form I	crystal form II
no. of reflections used (work set)	7757	10236
no. of molecules in the asu	2	1
no. of parameters refined	6442	3509
no. of atoms in the model		
protein	1538	769
solvent	68	105
NCS restraint used	yes	no
crystallographic residual		
R-factor (%)[a]	22.4 (18.9)	21.9 (20.1)
free *R*-factor (%)[a]	29.3 (26.6)	27.0 (25.1)
B-factor model	individual, isotropic	individual, isotropic
average *B*-factor		
protein (Å^2)	51.3	25.1
solvent (Å^2)	56.4	41.0
rms deviations from ideality		
bond lengths (Å)	0.004	0.005
bond angles (deg)	0.015	0.019
Ramachandran plot statistics[b]		
most favored regions (%)	90.8	94.3
additional allowed regions (%)	8.6	5.7
generously allowed regions (%)	0.6	0.0
disallowed regions (%)	0.0	0.0
overall *G*-factor	0.16	0.18

[a] The numbers in parentheses correspond to the *R*-factor calculated for reflections that are more significant than 4σ only. [b] Ramachandran statistics and the overall *G*-factor were calculated with ProCheck (*49*).

with respect to the search model. Also, omit maps showed density for the parts that were left out from the calculation of the phases.

Refinement. Initial models from the two crystal forms were manually fitted to the electron density map using the graphics program O (*30*). These models were both subjected to least-squares refinement using the program SHELXL of the SHELX-97 program suite (*31*) with the standard protein geometry restraints of Engh and Huber (*32*). The "free" reflection set was selected as a random 5% pool of reflections. For the tetragonal crystal form, the final model includes 96 of 100 residues in the asymmetric unit, one Ca^{2+} and one Zn^{2+} with refined occupancies of 0.87 and 1.00, respectively, and 105 fully occupied solvent positions (see Table 2 for a summary of the refinement).

For the orthorhombic crystal form, rebuilding of side chains and a part of the main chain in the N-terminal EF-hand loop was carried out to fit both $2mF_o - DF_c$ and $mF_o - DF_c$ density maps. The final model includes residues 1−96 for both monomers, the last couple of residues being too flexible to trace properly in the density. Additionally, two Ca^{2+} ions, a single Zn^{2+} ion, and 68 full occupancy solvent molecules have been included. The occupancies of the calcium ions are very high and approach unity during the refinement process, whereas the zinc ion has a somewhat smaller occupancy of 0.55. A noncrystallographic symmetry (NCS) restraint was used to relate the two monomers in the asymmetric unit for the orthorhombic data.

RESULTS

Overall Structure. The overall structure of both Ca^{2+}-bound S100A7 dimers is very similar to the previously determined structure of the Ho-substituted protein in that both structures describe a homodimer that consists of mainly α-helical elements. Each monomer contains, in the nomenclature used for holmium-substituted S100A7, five helices numbered I (Gln4−Lys18), II (Lys28−Phe39), II′ (Asn41−Lys48), III (Tyr53−Lys61), and IV (Phe71−Gln88). I and II are part of the N-terminal EF-hand, while III and IV are part of the C-terminal EF-hand (Figure 1). The linker helix, II′, is a regular $2^1/_2$-turn helix, whereas helix III (that takes part in the C-terminal EF-hand) is short and distorted.

The structure of the protein crystallized in the presence of excess amounts of divalent zinc (denoted the high-zinc structure from hereon) reveals clear and high-occupancy sites for both zinc and calcium, even though only zinc and not calcium was present in the crystallization buffer. Calcium, however, is bound very strongly to the C-terminal EF-hand of the protein, and very harsh and denaturing conditions are necessary to remove it. Thus, calcium seems to be present in the protein throughout the purification regardless of its concentration in the surrounding buffer. In the structure from the orthorhombic crystal form (which shall be denoted the low-zinc structure from hereon), no zinc was present during crystallization and only one of the monomers binds the ion in the crystal structure whereas both monomers bind calcium strongly. When the crystal packing of the molecule is inspected, it is evident that the loss of zinc from one monomer but not the other is due to different solvent accessibilities. The vacant zinc site is positioned near a large solvent cavity, whereas the occupied zinc site is tightly packed against a helix from a neighboring molecule. The crystal packing interactions between the two zinc sites and neighboring molecules, however, are similar for both monomers.

At the bottom of the dimer (near the C termini in Figure 1), the putative receptor-binding cleft is evident from all three structures. The width of the cleft, however, varies depending on whether zinc is bound, in such a way that in the high-zinc structure the distance between the two canonical EF-hand calcium ions is 27.8 Å whereas this increases to 28.7 Å for the low-zinc structure.

Calcium-Binding Sites. In accordance with the predictions made from the structure of Ho-bound S100A7, only the C-terminal EF-hand (termed the "canonical" EF-hand) of the calcium-loaded form binds an ion, whereas the N-terminal EF-hand (the "variant" EF-hand) does not. The affinity for calcium in the canonical EF-hand is supposedly very high in that the occupancy refines to approximately unity for both the high-zinc and low-zinc structures. The occupancy is somewhat higher than that for at least one of the ions in the holmium-substituted S100A7 structure, probably reflecting the difference in ionic radii (Ca^{2+} radius, 0.99 Å; Ho^{3+} radius, 0.894 Å) or varying solvent accessibilities due to crystal packing. In both structures, the calcium ion in the canonical EF-hand is coordinated by residues Asp62 (†), Asn64 (†), Asp66 (†), Lys68 (†), Glu73(‡), and a solvent molecule as expected from the canonical Tufty and Kretsinger EF-hand motif (see the legend of Figure 3 for a description of the ligand notation). The positions of both main chain and side chain atoms in the canonical EF-hand are very similar for both the holmium-substituted and calcium-bound structures with an overall root-mean-square difference (rmsd) of about 0.3 Å between all main chain atoms of the helix−loop−helix motif. A similar difference is found when the two calcium-bound structures are compared. Thus, the substitu-

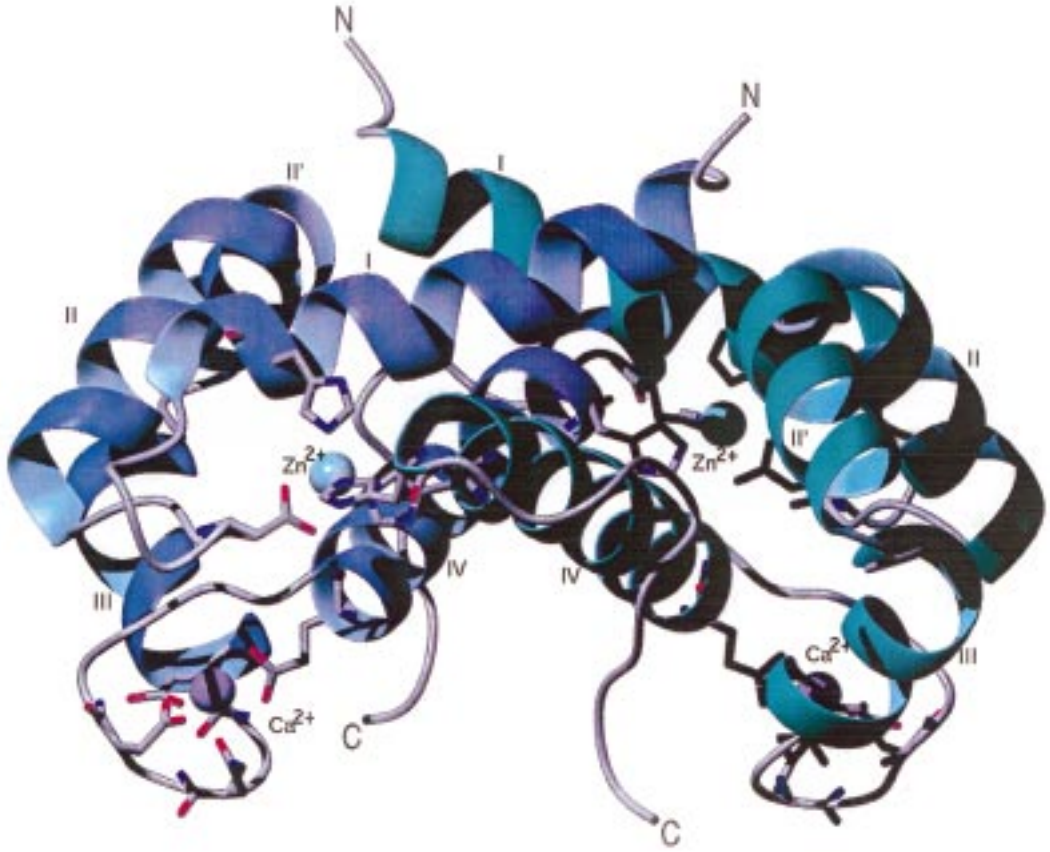

FIGURE 1: Cartoon representation of the overall structure of the zinc-bound S100A7 homodimer (the high-zinc structure). The monomers have been colored differently but represent the same structure due to the crystallographic 2-fold axis. The helices are labeled I, II, II′, III, and IV from the N terminus, for which I and II form the variant EF-hand, II′ is the linker helix, and III and IV form the canonical EF-hand motif. The calcium ions in the canonical EF-hand (colored violet) are shown along with their ligands Asp62, Asn64, Asp66, Lys68, and Glu73. The zinc ion is cyan and is surrounded by the ligands HisB86, HisB90, HisA17, and AspA24. Figures 1, 2, and 6 have been produced with MOLMOL (*50*).

tion of calcium for holmium does not appear to change the ion coordination in the canonical EF-hand.

High-Zinc Structure. The strong binding of a single zinc ion per monomer in the tetragonal crystal form is achieved through three histidines each donating their $N\epsilon_2$ atom and an aspartate donating both its terminal oxygen atoms, thus forming an overall five-coordinate conformation (Figure 2a). The zinc-binding site furthermore ties the two monomers together, because two of the coordinating histidines (HisB86 and HisB90) originate from the C terminus of one monomer whereas the other histidine and the aspartate (HisA17 and AspA24) are located in the N terminus of the other. Thus, S100A7 binds a total of two zinc ions per dimer, effectively involving all histidine residues in the protein. In Figure 3, the four amino acids that participate in binding of the zinc ion have been labeled in a way similar to the classic Tufty and Kretsinger notation for calcium in which † denotes a monodentate ligand and ‡ a bidentate ligand. We suggest that this notation be used to describe the coordination of zinc by the S100 proteins. The coordination sphere is thus formed by the three monodentate histidines (HisA17, HisB86, and HisB90) and a single bidentate aspartate (AspA24). In addition to the metal ion ligands, the zinc site contains a number of highly ordered solvent molecules (Figure 2a). Four water molecules are directly involved in hydrogen bonds to the zinc ligands (two to AspA24, one to HisB86, and one to HisB90), whereas at least eight other water molecules are located in a complex solvent network in the vicinity of the zinc site. This hydrogen bond network extends as far as the side chain of ThrB83 and the main chain oxygen of AspA24 in one end, and the side chain of AspB14 and SerB89 in the other end and possibly further.

Low-Zinc Structure. In the orthorhombic crystal form, one of the two monomers in the asymmetric unit clearly binds a zinc ion whereas the other does not. The conformation of the zinc-loaded site is largely identical to the equivalent position in the tetragonal crystal form, but the empty zinc site shows several interesting differences. With respect to side chain orientation, the three histidines of the coordination sphere move only slightly upon loss of the ion, but the adjacent aspartate has turned and is now pointing away from the site (Figure 2b). Furthermore, loss of the zinc ion also has a profound impact on the main chain of the empty variant EF-hand loop in which this aspartate is located. When we examined the resulting electron density map after molecular replacement as well as the $mF_o - DF_c$ difference map through the course of refinement, it was evident that the most significant structural changes of the main chain with respect to the search model occur in this particular region. The result of this structural rearrangement is that the variant EF-hand loop becomes more wide and open compared to the high-zinc structure (Figure 4), whereas the side chains of the three histidines remain more or less in the same position. Figure 5 shows the distance between the main chain atoms in the variant EF-hand loop for the two structures. It is clear that while helices I and II which are stabilized by secondary

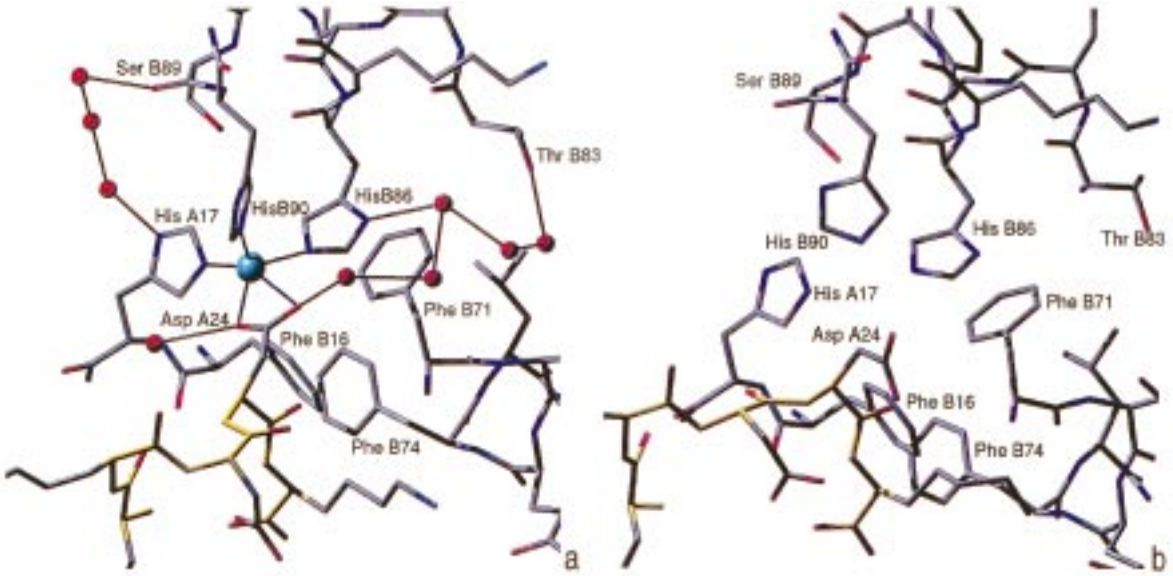

FIGURE 2: S100A7 zinc site. (a) The zinc site in the high-zinc structure with the ion (cyan sphere), surrounding residues, and solvent molecules (red spheres). The zinc ion is tightly coordinated by the three histidines and an aspartate residue from the variant EF-hand loop, for which the main chain has been colored yellow for clarity. Also, the zinc ligands participate in an extensive hydrogen bond network as shown by the thin red lines. The zinc ion is close to a hydrophobic cluster comprised of three phenylalanines. (b) The empty zinc site from the low-zinc structure. Here, the aspartate that otherwise coordinates zinc has moved away, resulting in a remodeling of the variant EF-hand loop of which it is part (yellow main chain). Notice the remarkable widening of this adjacent loop toward the left in the figure. Also, significant reorganization of the nearby hydrophobic cluster is evident compared with the zinc-bound case. Due to loss of the ion or perhaps the somewhat lower resolution of this data set, no clear solvent network is visible.

structure interactions remain more or less unchanged, large structural changes take place for the main chain of residues 21−23 as well as for the zinc-coordinating His17. Moreover, it seems that the histidine main chain in its conformational change also brings along its neighboring residues to some extent (Figures 2 and 5). Interestingly, however, the main chain of Asp24 that is also part of the zinc coordination does not shift as much as neighboring residues 21−23. The overall rmsd between the main chain atoms in the two conformations of the loop is about 1.5 Å. In addition to this dramatic, local conformational change, significant global structural changes from the loss of the zinc ion are also evident. Thus, in the low-zinc structure, one end (the zinc binding) appears to be more compact than the other. It seems as if the protein has the ability to contract asymmetrically when a single ion is lost. The highly ordered solvent structure observed in the zinc-bound state is absent in the low-zinc structure, but this could also be due to the fact that the data from the orthorhombic crystals are somewhat poorer.

DISCUSSION

Overall Structure. The fact that the overall shapes of the two calcium-loaded S100A7 structures reported in this paper resemble the previously determined holmium-bound form to a great extent first of all confirms the suitability of lanthanide elements as substitutes for calcium for at least the EF-hand proteins. Most of the remainder of the S100 proteins are good candidates for application of this powerful method of structure determination, but since similar experiments have been carried out for other calcium-binding proteins (*5*, *6*), it is likely to apply in a wider range of cases. This procedure is thus a valuable tool for structure determination, due to the significant anomalous dispersion signal from the heavy atom.

The presence of a crystallographic 2-fold axis relating the two parts of the homodimer in the tetragonal crystal form shows that the sequence identity of the two monomers is reflected at the structural level under conditions where an equal number of ions is bound. However, as the low-zinc structure confirms, the protein is also perfectly capable of attaining an overall asymmetric conformation. Whether this conformation is physiologically relevant or due to crystal packing is not certain and will have to be investigated further.

When the high-zinc structure of S100A7 is structurally aligned with the low-zinc structure, a small but visible breath-like movement of the dimer is evident. Upon binding of zinc, the molecule contracts around the zinc sites but also around the potential substrate binding cleft, whereas it expands around the top of the molecule (Figure 1). Furthermore, the structure of the hydrophobic core of the protein is significantly altered upon binding of zinc. This result confirms the previous observation that the absorption characteristics of the S100 proteins are modified upon interaction with metal ions (*8*, *11*, *13*). Thus, the binding of zinc not only causes a local structural change in the neighborhood of the zinc site but also has a profound impact on the overall protein conformation. This could indicate that the binding or the release of the intrinsic zinc ion is involved in regulation of substrate binding affinity.

Calcium Binding. Calcium coordination by the variant EF-hand is generally achieved by main chain oxygen atoms in addition to a conserved bidentate glutamate, but this motif shows much variability for members of the S100 family. Because S100A7 lacks the important glutamate and because the EF-hand loop is three residues shorter (Figure 3), it was originally believed that the protein could not bind calcium in the classical way (*4*). This conclusion was based on the holmium-substituted structure only, and did not take into account the fact that the protein's affinity is likely to be

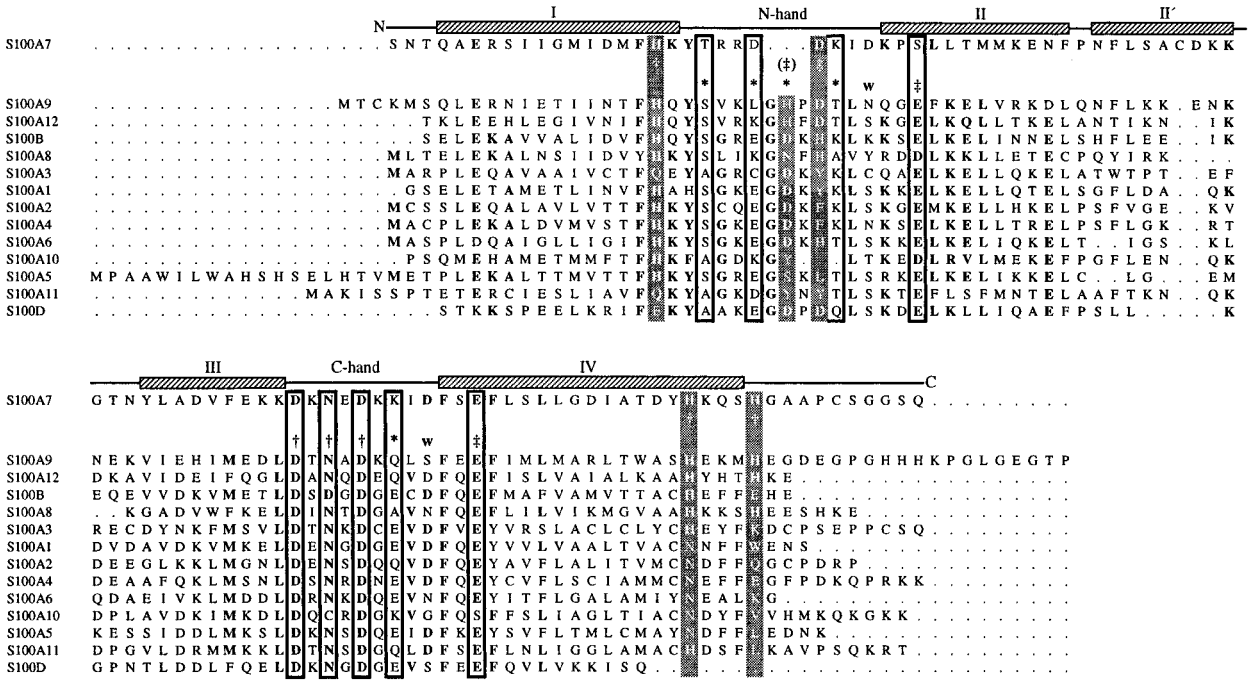

FIGURE 3: Sequence alignment of the known members of the S100 family. With psoriasin (S100A7) at the top, the rest of the S100 proteins have been placed in order of decreasing similarity with respect to the zinc-binding site. The secondary structure (helices as hatched bars and loop regions as black lines) as well as the helix nomenclature is shown above the S100A7 sequence. Residues that are supposed to participate in calcium binding are in outline boxes, whereas proposed zinc ligands are in gray boxes. The first line below the S100A7 sequence represents the proposed zinc coordination by a given residue, and the second line represents the calcium ligands in the Tufty and Kretsinger notation († for a monodentate ligand, ‡ for a bidentate ligand, w for solvent molecules, and * for a main chain oxygen). Only S100A7 and S100A10 lack the three amino acids in the variant EF-hand which should coordinate calcium with main chain oxygen atoms. For the remainder of the family, the residues in this region that could potentially act as zinc ligands are shown in the box denoted (‡). All sequences are from the human genome, except for S100A6 and S100B for which the sequences correspond to the protein used for structure determination (rabbit and bovine, respectively). Protein accession numbers in the SwissProt database are P31151 (S100A7), P06702 (S100A9), P80511 (S100A12), P02638 (S100B), P05109 (S100A8), P33764 (S100A3), P23297 (S100A1), P29034 (S100A2), P26447 (S100A4), P30801 (S100A6), P08206 (S100A10), P33763 (S100A5), P31949 (S100A11), and P29377 (S100D).

different due to the different chemical characteristics (such as ionic radii) of holmium with respect to calcium. For none of the crystal forms presented here is calcium bound in the distorted N-terminal EF-hand of S100A7. These results therefore confirm the finding that S100A7, in contrast to the remaining S100 proteins, only binds one calcium ion per monomer. This result indicates that it is much more likely that the physiological function of S100A7 is controlled through binding and release of the zinc ion, since loss of calcium from the C-terminal EF-hand leads to protein denaturation (M. Etzerodt, personal communication).

Zinc Binding. The results reported in this paper establish the existence of a zinc site for S100A7 that is likely to be physiologically relevant and outline the residues involved in zinc binding. Zinc-binding motifs similar to the one seen for S100A7 have been reported for other proteins, even though these sites often consist of only two histidines in addition to the acidic residue. The two histidines from the C-terminal part of S100A7 that participate in zinc binding are located near the end of helix IV and are placed four residues apart, a very common feature that corresponds to a complete helical turn (the His-X-X-X-His motif). Less consensus is found with respect to the sequence separation of the other two residues participating in the coordination. In the case of S100A7, the coordinating residues in the N terminus are even located on a different molecule. His17 is at the very end of long helix I, whereas Asp24 is located in the middle of the empty variant EF-hand loop right after the three-amino acid gap in the sequence alignment with the rest of the members of this family (Figure 3). Thus, the fact that the zinc site is made up of residues His86 and His90 from one monomer and residues His17 and Asp24 from the other results in binding together the empty variant EF-hand loop of one monomer with the C-terminal helix of the other. Significant stabilization of the dimer structure is likely to result from this interaction, but the low-zinc structure (as well as the holmium-bound structure) also shows that zinc is not necessary for protein stability. The location of the zinc-binding site is clearly at the surface of the dimer and close to the putative receptor molecule binding cleft. This strongly suggests that zinc binding is involved in the physiological function of the protein and does not merely represent a structural feature. The notion that the affinity for zinc is lower than that for calcium is in accordance with several biochemical analyses that all conclude that the S100 proteins should be considered zinc-regulated proteins as well as calcium-regulated proteins (*13*).

Reorganization of an EF-Hand Loop. Upon the discovery that S100A7 was probably not able to bind calcium with its variant EF-hand, it was also concluded that the three missing amino acids in the sequence (with respect to the remainder of the S100 family) forced the main chain to "cut short" and continue in the next helix before the loop was "complete". When viewed along helix I (Figure 4b), this loop appears narrower and much flatter than that of the corresponding calcium-loaded EF-hand (Figure 4a). Evidently, the

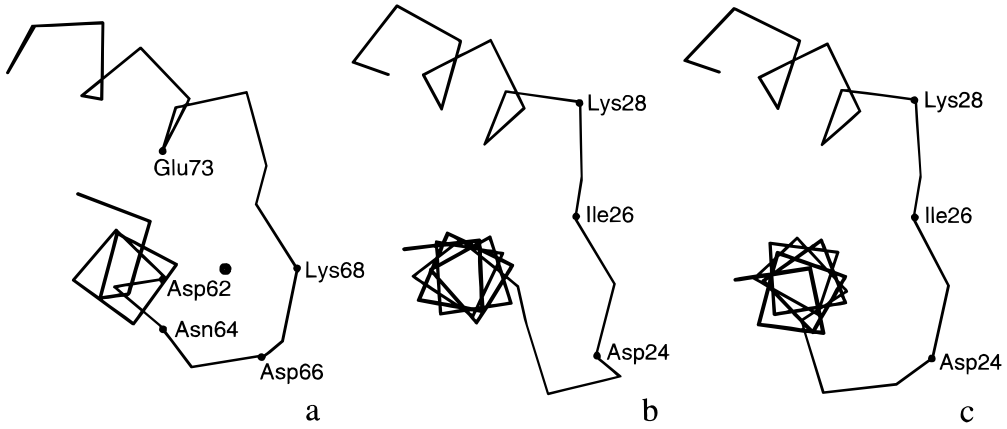

FIGURE 4: Structural alignment of the EF-hand motifs in some of the S100A7 structures. (a) The canonical EF-hand in S100A7 from the high-zinc structure. Notice the significant width of the loop that surrounds the calcium ion. (b) The variant EF-hand as it appeared in the holmium-substituted S100A7 structure and also in the high-zinc structure from crystal form II. This EF-hand loop is much narrower and thus does not support an ion. (c) The remodeled variant EF-hand loop from the low-zinc structure in crystal form I. The remodeling of the EF-hand loop makes it wider, and it closely resembles the conformation found in the canonical hand.

loop is not large enough to accommodate calcium because the interatomic distances between the main chain oxygen atoms that were supposed to coordinate the ion are not large enough. When we compare the variant EF-hand loop from the high-zinc structure (Figure 4b) with the reorganized loop that results from the loss of zinc in the nearby binding site (Figure 4c), it is clear that the latter closely resembles the open loop conformation found in the calcium-loaded EF-hand. This suggests (even though no calcium has yet been identified in the variant loop) that structural changes resulting from binding or loss of zinc could trigger or regulate calcium binding by the variant EF-hand. Although the missing amino acids in the case of S100A7 could impede the binding of calcium in this particular protein, such a zinc-triggered calcium binding could perhaps be responsible for regulation of several of the other members of the S100 family.

A Recurring Theme. In Figure 3, the amino acid sequence of psoriasin (S100A7) has been aligned with those of the other members of the S100 family. As seen from the figure, S100A9 and S100A12 most closely resemble the pattern seen in S100A7 with respect to zinc binding, in that both the position and type of the four amino acids are unchanged. It is therefore very likely that these two proteins both contain high-affinity zinc sites such as the one seen in S100A7. Moreover, none of them exhibit the gap of three amino acids in the variant EF-hand and also contain the important conserved glutamate necessary for binding calcium in this loop. Thus, it is clear that these two members of the family most probably bind calcium in the variant EF-hand and that this could be regulated by binding or release of zinc. The situation is different for proteins S100B and S100A8, which lack the bidentate aspartate involved in zinc binding. This residue has instead been substituted for a histidine. Another aspartate, however, can be found two residues earlier, and this could play a role in the zinc binding. In the sequences below S100A8 in Figure 3, the zinc-binding motif gradually vanishes. The most conserved parts are the N-terminal histidine and the aspartate, although this has been shifted by two residues in the sequence (shown with the box marked (‡) in Figure 3). In contrast, the two C-terminal histidines are missing for the remainder of the S100 family, one of them being substituted for an asparagine in several of the sequences. Thus, to summarize, the members of the family that are most likely to contain a zinc-binding site like the one observed in S100A7 include S100A9, S100A12, and possibly S100B and S100A8. Structural investigation of the exact functional relationship between these two types of ligand sites is of the utmost importance for the further understanding of the function and regulation of the dimeric S100 proteins.

Structural Alignment of S100B and S100A6. Apart from S100A7, the homodimeric members of the S100 family for which three-dimensional structures are known comprise S100B and S100A6, which both have been determined in the apo and calcium-loaded states by NMR (*33*−*36*). Recently, a structure of S100B in the calcium-loaded state also determined by X-ray crystallography has been published (*37*). S100A6 (calcyclin) is less interesting in this context, because it completely lacks the two histidines in the C terminus and thus only contains two histidines in the variant EF-hand. Whether zinc binding can be achieved by other means for this protein cannot be judged from the present data. In the apo structure of S100B, however, most of the important residues are in place. When we compare the available structures with the primary sequence alignment mentioned above, it is clear that residues HisA15, HisA25, HisB90, and HisB85 are all localized close to each other in the structure. Moreover, both glutamates B89 and B91 seem to be properly placed for participation in a possible zinc-binding motif. The side chains are too far apart in the structural models, however, for actual binding of zinc to be possible. But here it must be taken into account that the available models represent the apo structures of the proteins and that significant structural changes have been observed upon calcium binding. Exactly which of the residues that would interact in a possible zinc site is unclear, but it seems most likely that the coordination involves HisA15, HisA25, HisB90, and either GluB89 or GluB91.

Comparison with the Metalloproteases. Using the four ligands of the zinc-binding site of S100A7 as a template, a search for structural homologues was carried out with the program SPASM (*38*). Interestingly, it turned out that the zinc site in S100A7 is clearly reminiscent of zinc sites found in certain metalloproteases, especially the collagenase family that contains both a structural and a catalytic zinc site similar to the one in S100A7 (PDB entries 1kbc, 1mmb, 1cge, and

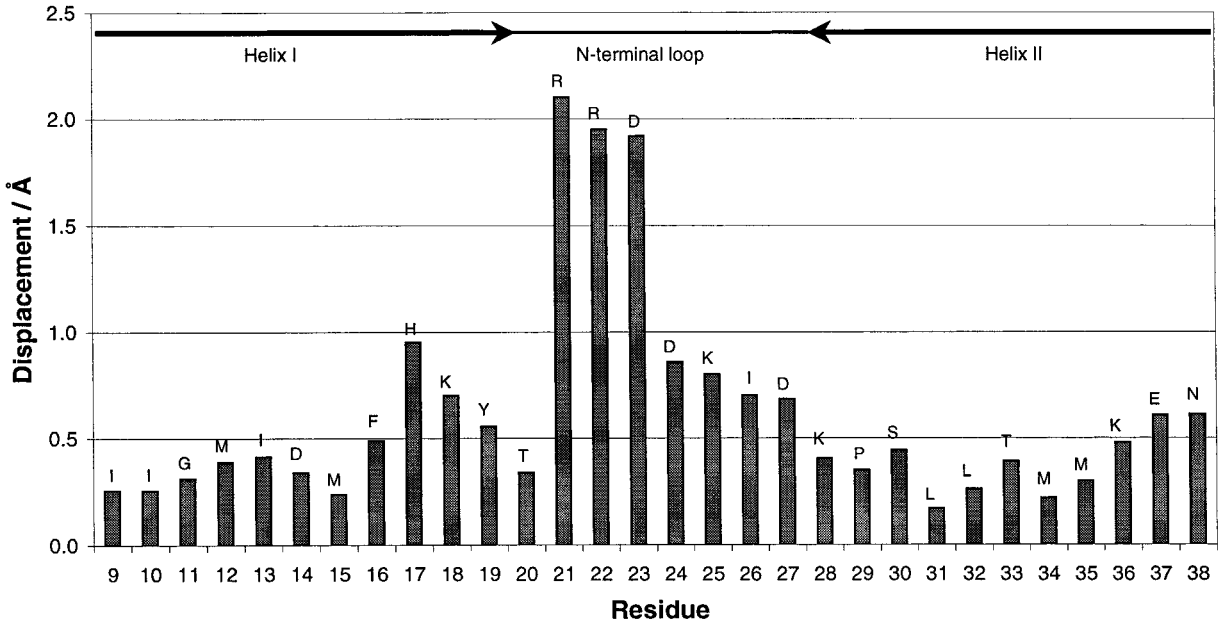

FIGURE 5: Average distance between main chain atoms in the variant EF-hand in the zinc-bound and zinc-free cases. The upper part of the graph shows the secondary structure corresponding to this region of the monomer. Helices I and II are represented by thick arrows, and the variant EF-hand loop is represented by a thin line. Whereas the helix regions generally are quite rigid with a displacement of less than 0.5 Å, both His17 and Asp24 that take part in the zinc coordination are more flexible (with an about 1.0 Å change). Notice how the residues in the neighborhood of these zinc ligands also are displaced more than average. The largest changes, however, are seen for residues Arg21, Arg22, and Asp23 in the beginning of the variant EF-hand loop.

1jap) (*39*−*43*). Similar motifs have been observed for other proteins, including matrilysin (PDB entries 1mmp, 1mmq, and 1mmr) (*44*) and stromelysin (PDB entries 1hfs, 1slm, and 1sln) (*45*−*47*). Common to these proteins is the fact that they all belong to the family of matrix metalloproteases (MMPs). This group of proteins is specifically involved in the degradation of extracellular matrix components such as fibrillar collagens, proteoglycans, fibronectin, laminin, and gelatin in such diverse processes as wound healing and embryonic development. Also, they have been reported to be universally expressed during tumor progression and metastasis (*48*). These proteases are generally larger than the S100 proteins and consist of several domains that are important for their substrate specificity. The overall folding of the MMPs is also very different from that of the S100 proteins, which are pure α-domain structures. Common between the two groups of proteins, however, is their shared ability to bind both calcium and zinc. Calcium is bound specifically in loop regions of the MMPs, but not by an EF-hand motif as is the case for the S100 proteins. In contrast, the similarities are very clear with respect to the zinc site. The MMPs all contain two zinc sites per molecule, one being directly involved in substrate binding (referred to as the "catalytic site") and one that is not involved in proteolysis, but which stabilizes the tertiary structure of the protein (the "structural site"). In both these sites, zinc is coordinated by three histidines but whereas the structural site clearly has a monodentate aspartate, the acidic side chain is located about 4 Å from the catalytic site (see Figure 6b,c). The fact that the aspartate involved in the structural zinc site is monodentate implies that its side chain is tilted with respect to the zinc ion and that it thus only contributes one of its two Oδ atoms to the coordination sphere. The structural zinc is therefore four-coordinate, while the catalytic zinc is five-coordinate when bound to an inhibitor. In Figure 6, the zinc site of S100A7 has been structurally aligned with the inhibitor-bound catalytic zinc site (Figure 6b) and the structural zinc site (Figure 6c) of neutrophil collagenase. The zinc site in S100A7 shares with the structural zinc site of collagenase the nature of the acidic side chain in the coordination, which is an aspartate in both cases. The orientation of this aspartate residue is clearly different, however, in that the aspartate in S100A7 donates both its Oδ atoms to the coordination. Therefore, the 5-fold coordination in S100A7 more closely resembles the catalytic zinc site of collagenase when bound to a substrate.

Implications for Substrate Binding. The location of the zinc site in S100A7 near the putative substrate-binding cleft for this type of proteins and the fact that coordination of the zinc ion clearly resembles that of the MMPs has implications for the present theory of substrate binding to the S100 proteins. The fact that the zinc can be lost from S100A7 without denaturation of the protein makes it unlikely that the ion should be present for solely structural reasons. S100A7 (and the rest of the members of the S100 family) does not, however, contain the large catalytic domain necessary for proteolytic activity. But the close proximity of the zinc site to the substrate-binding cleft and the location of the site on the surface of the molecule strongly suggest that the ion be somehow functionally related to the binding of receptor molecules. The binding of inhibitor molecules similar to real substrates is well documented for the MMPs also from a structural point of view (*39*, *40*, *42*, *45*, *46*). In these cases, the proximal glutamate in the catalytic zinc site turns away, leaving room for the substrate which then will coordinate zinc in a similar bidental manner (see Figure 6b). The interatomic distances between zinc and its ligands in S100A7 (with the His $N\epsilon_2$−Zn distance being 2.0−2.2 Å, the Asp $O\delta_1$−Zn distance being 2.2 Å, and the Asp $O\delta_2$−Zn distance being 2.6 Å) very much resemble the distances

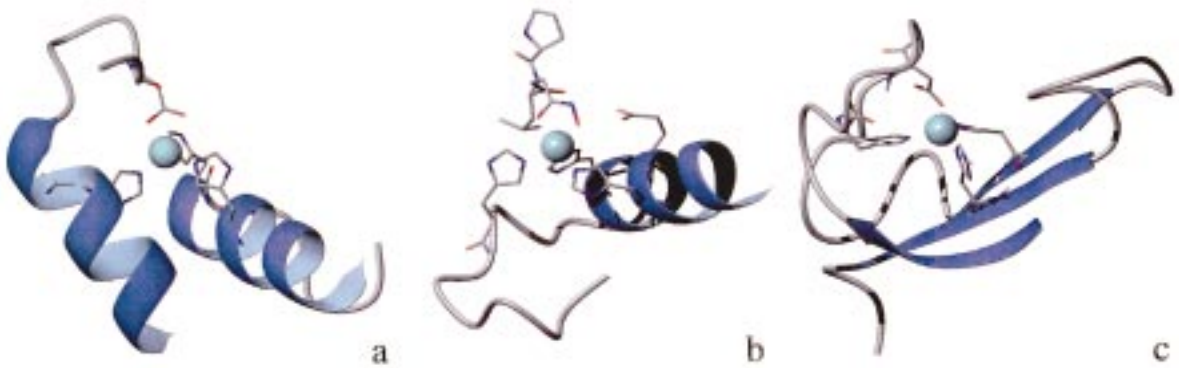

FIGURE 6: Structural alignment with zinc-binding motifs from the metalloproteases. (a) The zinc-binding site in S100A7 consists of three histidines from helical regions and an aspartate from the variant EF-hand loop. (b) The catalytic zinc-binding site of neutrophil collegenase with bound inhibitor (PDB entry 1jap) (*39*). The catalytic zinc site in collagenase also consists of two histidines near each other in a helix, whereas the third is in a loop region. The hydroxamate terminus of the inhibitor peptide simulates the coordination of zinc by a carboxyl group. (c) The structural zinc-binding site of neutrophil collagenase (PDB entry 1cge) (*42*). This resembles the zinc site in S100A7 with respect to the number and type of coordinating residues but differs in that they originate primarily from β-strands.

found in the collagenase catalytic zinc site between the protease and its inhibitor (the His Nϵ_2−Zn distance being 2.1 Å, the inhibitor carboxylate O_1−Zn distence being 2.2 Å, and the O_2−Zn distance being 2.6 Å). Thus, the coordination of the zinc ion in S100A7 more closely resembles that of the inhibitor-bound collagenase than that of the collagenase with an empty catalytic site. Conclusively, the observations leave open the question of whether zinc regulates the function of the S100 proteins by binding and thereby imposing a conformational change or whether the zinc site is directly involved in binding of substrate molecules in a way similar to the situation observed for the MMPs.

Conclusion. The reported structures of the calcium-bound form of S100A7 suggest that the high-affinity EF-hand calcium mainly is a structural site, since loss of the ion leads to denaturation of the protein. Also, we have so far not been able to identify any bound calcium in the variant EF-hand; therefore, regulation of the protein's function is not likely to be controlled by this site. Thus, regulation of receptor or substrate affinity must be achieved partly through binding or loss of the zinc ion. The location of the zinc-binding site close to the putative substrate-binding cleft strongly supports this theory. An important regulatory role for zinc is also suggested from the reorganization of the variant EF-hand that results from loss of the ion. Even though calcium binding is not observed, the possibility that conformational changes resulting from loss of zinc can result in binding of calcium by the variant EF-hand under suitable conditions cannot be ruled out. From an alignment of the structures now available of two different native states of S100A7, part of the internal mobility and possible reorganization of the molecule has become evident. The alignment reveals a highly flexible macromolecule with the ability of adapting different shapes under different cellular conditions. Research is underway to establish the identity of the receptor molecule of S100A7, and we are hopeful that future experiments will reveal in detail how metal ion regulation affects complex formation.

ACKNOWLEDGMENT

We thank Sigrun Nolsøe, Søren Thirup, Lasse Jenner, Ole Kristensen, Poul Nissen, Galina Polekhina, Morten Geday, and Trine Koch for data collection on BW7B at EMBL-Hamburg as well as Sigrun Nolsøe for collecting data on the R-AXIS facility at the H. C. Ørsted Institute. Also, we thank Michael Etzerodt and H. C. Thøgersen at the Laboratory of Gene Expression (IMSB, Aarhus University) for providing the recombinant protein and for helpful discussions.

REFERENCES

1. Madsen, P., Rasmussen, H. H., Leffers, H., Honoré, B., Dejgaard, K., Olsen, E., Kiil, J., Walbum, E., Andersen, A. H., Basse, B., Lauridsen, J. B., Ratz, G. P., Celis, A., Vandekerckhove, J., and Celis, J. E. (1991) *J. Invest. Dermatol. 97*, 701−712.
2. Celis, J. E., Crüger, D., Kiil, J., Lauridsen, J. B., Ratz, G., Basse, B., and Celis, A. (1990) *FEBS Lett. 262*, 159−164.
3. Celis, J. E., Cruger, D., Kiil, J., Dejgaard, K., Lauridsen, J. B., Ratz, G. P., Basse, B., Celis, A., Rasmussen, H. H., and Bauw, G. (1990) *Electrophoresis 11*, 242−254.
4. Brodersen, D. E., Etzerodt, M., Madsen, P., Celis, J. E., Thøgersen, H. C., Nyborg, J., and Kjeldgaard, M. (1998) *Structure 6*, 477−489.
5. Weis, W. I., Kahn, R., Fourme, R., Drickamer, K., and Hendrickson, W. (1991) *Science 254*, 1608−1615.
6. Burling, F. T., Weis, W. I., Flaherty, K. M., and Brünger, A. T. (1996) *Science 271*, 72−77.
7. Baudier, J., Haglid, K., Haiech, J., and Gerard, D. (1983) *Biochem. Biophys. Res. Commun. 114*, 1138−1146.
8. Baudier, J., and Gerard, D. (1983) *Biochemistry 22*, 3360−3369.
9. Filipek, A., Heizmann, C. W., and Kuznicki, J. (1990) *FEBS Lett. 264*, 263−266.
10. Föhr, U. G., Heizmann, C. W., Engelkamp, D., Schäfer, B., and Cox, J. A. (1995) *J. Biol. Chem. 270*, 21056−21061.
11. Leung, I. K. M., Mani, R. S., and Kay, C. M. (1987) *FEBS Lett. 214*, 35−40.
12. Raftery, M. J., Harrison, C. A., Alewood, P., Jones, A., and Geczy, C. L. (1996) *Biochem. J. 316*, 285−293.
13. Baudier, J., Glasser, N., and Gerard, D. (1986) *J. Biol. Chem. 261*, 8192−8203.
14. Franz, C., Durussel, I., Cox, J. A., Schäfer, B. W., and Heizmann, C. W. (1998) *J. Biol. Chem. 273*, 18826−18834.
15. Dell'Angelica, E. C., Schleicher, C. H., and Santomé, J. A. (1994) *J. Biol. Chem. 269*, 28929−28936.
16. Baudier, J., Holzscherer, C., and Gerard, D. (1982) *FEBS Lett. 148*, 231−234.
17. Clohessy, P., and Golden, B. E. (1996) *J. Leukocyte Biol. 60*, 674.
18. Vorum, H., Madsen, P., Rasmussen, H. H., Etzerodt, M., Svendsen, I., Celis, J. E., and Honoré, B. (1996) *Electrophoresis 17*, 1787−1796.
19. Nolsøe, S., Thirup, S., Etzerodt, M., Thøgersen, H. C., and Nyborg, J. (1997) *Acta Crystallogr. D53*, 119−121.

20. Otwinowski, Z. (1993) in *Proceedings of the CCP4 Study Weekend: Data Collection and Processing* (Sawyer, L., Isaacs, N., and Bailey, S., Eds.) pp 56−62, Daresbury Laboratory, Warrington, U.K.
21. French, G. S., and Wilson, K. S. (1978) *Acta Crystallogr. A34*, 517.
22. CCP4 (1994) *Acta Crystallogr. D50*, 760−763.
23. Evans, P. R. (1993) *Proceedings of the CCP4 Study Weekend on Data Collection and Processing*, pp 114−122, Daresbury Laboratory, Warrington, U.K.
24. Evans, P. R. (1997) *Proceedings of the CCP4 Study Weekend on Recent Advances in Phasing*, Daresbury Laboratory, Warrington, U.K.
25. Rossmann, M. G., and Blow, D. M. (1962) *Acta Crystallogr. 15*, 24−31.
26. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
27. Murshudov, G., Vagin, A., and Dodson, E. (1996) *Proceedings of the Daresbury Study Weekend on Refinement of Protein Structures*, Daresbury Laboratory, Warrington, U.K.
28. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D53*, 240−255.
29. Cowtan, K. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, Vol. 31, pp 34−38, Daresbury Laboratory, Warrington, U.K.
30. Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol. 277*, 173−207.
31. Sheldrick, G. M., and Schneider, T. R. (1997) in *Methods in Enzymology* (Carter, C. W., and Sweet, R. M., Eds.) pp 319−343, Academic Press, Orlando, FL.
32. Engh, R. A., and Huber, R. (1991) *Acta Crystallogr. A47*, 392−400.
33. Potts, B. C. M., Smith, J., Akke, M., Macke, T. J., Okazaki, K., Hidaka, H., Case, D. A., and Chazin, W. J. (1995) *Nat. Struct. Biol. 2*, 790−796.
34. Kilby, P. M., Eldik, L. J. V., and Roberts, G. C. K. (1996) *Structure 4*, 1041−1052.
35. Sastry, M., Ketchem, R. R., Crescenzi, O., Weber, C., Lubienski, M. J., Hidaka, H., and Chazin, W. J. (1998) *Structure 6*, 223−231.
36. Smith, S. P., and Shaw, G. S. (1998) *Structure 6*, 211−222.
37. Matsumura, H., Shiba, T., Inoue, T., Harada, S., and Kai, Y. (1998) *Structure 6*, 233−241.
38. Kleywegt, G. J., and Jones, T. A. (1998) *Acta Crystallogr. D54*, 1119−1131.
39. Bode, W., Reinemer, P., Huber, R., Kleine, T., Schnierer, S., and Tschesche, H. (1994) *EMBO J. 13*, 1263−1269.
40. Grams, F., Reinemer, P., Powers, J. C., Kleine, T., Pieper, M., Tschesche, H., Huber, R., and Bode, W. (1995) *Eur. J. Biochem. 228*, 830−841.
41. Li, J., Brick, P., O'Hare, M. C., Skarzynski, T., Lloyd, L. F., Curry, V. A., Clark, I. M., Bigg, H. F., Hazleman, B. L., Cawston, T. E., and Blow, D. M. (1995) *Structure 3*, 541−549.
42. Lovejoy, B., Cleasby, A., Hassell, A. M., Longley, K., Luther, M. A., Weigl, D., McGeehan, G., McElroy, A. B., Drewry, D., Lambert, M. H., and Jordan, S. R. (1994) *Science 263*, 375−377.
43. Betz, M., Huxley, P., Davies, S. J., Mushtaq, Y., Pieper, M., Tschesche, H., Bode, W., and Gomis-Rueth, F. X. (1997) *Eur. J. Biochem. 247*, 356.
44. Browner, M. F., Smith, W. W., and Castelhano, A. L. (1995) *Biochemistry 34*, 6602−6610.
45. Gooley, P. R., O'Connell, J. F., Marcy, A. I., Cuca, G. C., Salowe, S. P., Bush, B. L., Hermes, J. D., Esser, C. K., Hagmann, W. K., Springer, J. P., and Johnson, B. A. (1994) *Nat. Struct. Biol. 1*, 111−118.
46. Becker, J. W., Marcy, A. I., Rokosz, L. L., Axel, M. G., Burbaum, J. J., Fitzgerald, P. M. D., Cameron, P. M., Esser, C. K., Hagmann, W. K., Hermes, J. D., and Springer, J. P. (1995) *Protein Sci. 4*, 1966−1976.
47. Esser, C. K., Bugianesi, R. L., Caldwell, C. G., Chapman, K. T., Durette, P. L., Girotra, N. N., Kopka, I. E., Lanza, T. J., Levorse, D. A., Maccoss, M., Owens, K. A., Ponpipom, M. M., Simeone, J. P., Harrison, R. K., Niedzwiecki, L., Becker, J. W., Marcy, A. I., Axel, M. G., Christen, A. J., McDonnell, J., Moore, V. L., Olszewski, J. M., Saphos, C., Visco, D. M., Shen, F., Colletti, A., Krieter, P. A., and Hagmann, W. K. (1997) *J. Med. Chem. 40*, 1026.
48. Werb, Z. (1997) *Cell 91*, 439−442.
49. Laskowski, R., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283.
50. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.

BI982483D