J Biol Inorg Chem
DOI 10.1007/s00775-008-0402-7

ORIGINAL PAPER

Statistical analysis of structural characteristics of protein Ca^{2+}-binding sites

Michael Kirberger · Xue Wang · Hai Deng · Wei Yang · Guantao Chen · Jenny J. Yang

Received: 30 January 2008 / Accepted: 17 June 2008

Abstract To better understand the biological significance of Ca^{2+}, we report a comprehensive statistical analysis of calcium-binding proteins from the Protein Data Bank to identify structural parameters associated with EF-hand and non-EF-hand Ca^{2+}-binding sites. Comparatively, non-EF-hand sites utilize lower coordination numbers (6 ± 2 vs. 7 ± 1), fewer protein ligands (4 ± 2 vs. 6 ± 1), and more water ligands (2 ± 2 vs. 1 ± 0) than EF-hand sites. The orders of ligand preference for non-EF-hand and EF-hand sites, respectively, were H_2O (33.1%) > side-chain Asp (24.5%) > main-chain carbonyl (23.9%) > side-chain Glu (10.4%), and side-chain Asp (29.7%) > side-chain Glu (26.6%) > main-chain carbonyl (21.4%) > H_2O (13.3%). Less formal negative charge was observed in the non-EF-hand than in the EF-hand binding sites (1 ± 1 vs. 3 ± 1). Additionally, over 20% of non-EF-hand sites had formal charge values of zero due to increased utilization of water and carbonyl oxygen ligands. Moreover, the EF-hand sites presented a narrower range of ligand distances and bond angles than non-EF-hand sites, possibly owing to the highly conserved helix–loop–helix motif. Significant differences between ligand types (carbonyl, side chain, bidentate) demonstrated that angles associated with each type must be classified separately, and the EF-hand side-chain Ca–O–C angles exhibited an unusual bimodal quality consistent with an Asp distribution that differed from the Gaussian model observed for non-EF-hand proteins. The results of this survey more accurately describe differences between EF-hand and non-EF-hand proteins and provide new parameters for the prediction and design of different classes of Ca^{2+}-binding proteins.

Keywords Calcium binding · EF-hand · Metalloprotein · Structure · Prediction

Michael Kirberger, Xue Wang, and Hai Deng contributed equally to this article.

Electronic supplementary material The online version of this article (doi:10.1007/s00775-008-0402-7) contains supplementary material, which is available to authorized users.

M. Kirberger · W. Yang · J. J. Yang (✉)
Department of Chemistry,
Center for Drug Design and Biotechnology,
Georgia State University,
Atlanta, GA 30303, USA
e-mail: chejjy@langate.gsu.edu

X. Wang · H. Deng · G. Chen
Department of Computer Science,
Georgia State University,
Atlanta, GA 30303, USA
e-mail: matgtc@langate.gsu.edu

G. Chen
Department of Mathematics and Statistics,
Georgia State University,
Atlanta, GA 30303, USA

Introduction

Metalloproteins play significant roles in numerous biological processes, and approximately 40% of all natural proteins are known to bind metals [1–7]. Metal-binding sites in general are characterized by a central shell of hydrophilic ligands to chelate the ion, with a surrounding shell of hydrophobic residues [8, 9]. The most common biologically important metals (e.g., Mg^{2+}, Zn^{2+}, Ca^{2+}, Mn^{2+}) frequently bind proteins selectively in different geometric configurations, utilizing different ligand atoms, mainly oxygen, nitrogen, and sulfur from side-chain groups, and oxygen from main-chain carbonyls [10–19].

Ca^{2+} is demonstrably one of the more relevant metal ions associated with biological functions. Ca^{2+} binds to various Ca^{2+}-binding proteins (CaBPs) inducing functions such as muscle contraction, neurotransmitter release, enzyme activation, and blood clotting [20–22]. Ca^{2+} may also interact with proteins to participate in various functions related to cell division, apoptosis, and intracellular signaling [23–26], as well as to enhance protein stability [27, 28].

Various studies have evaluated structural properties of Ca^{2+} binding in small molecules and proteins, indicating that coordination of Ca^{2+} is dominated by oxygen ligands from side-chain carboxyl groups (Asp, Glu) carboxamide groups (Asn, Gln), hydroxyl groups (Ser, Thr) main-chain carbonyl groups, and water, while the overall geometry is frequently configured as either pentagonal bipyramidal (Fig. 1a) or octahedral [5, 17–19, 26, 29–34]. The coordination number for Ca^{2+} binding in proteins has been reported at five to eight ligands, with an average between six and seven [35, 36]. The mean Ca–O ligand binding distance has been reported, from different studies, as either 2.4 ± 0.1 or 2.42 Å, with a range of 2.01–3.15 Å [5, 10, 21, 37–40].

According to the nature of liganding residues, Ca^{2+}-binding sites are often classified either as continuous, where ligands originate in a continuous short sequence (Fig. 1b), or discontinuous, where multiple ligands are distant from each other within the sequence (Fig. 1d) [35, 36, 41, 42].

The EF-hand motif, first described by Kretsinger and Nockolds [43] in 1973, is a continuous binding site that has been studied extensively. It is the most common motif associated with Ca^{2+} binding, characterized by a helix–loop–helix structure composed of approximately 30 amino acids, and can be subdivided into canonical EF-hand (e.g., calmodulin) and pseudo EF-hands found in the S100 protein N-termini [44–47]. For canonical EF-hands, calcium ions bind in a 12-residue central loop, utilizing side-chain oxygen ligands from loop positions 1, 3, 5, 9, and 12, as well as a main-chain carbonyl oxygen from position 7. Ligands associated with EF-hands are typically Asp at position 1, Asp or Asn at position 3, Asp, Ser, or Asn at position 5, a water molecule at position 9, and a bidentate Glu at position 12 [30]. Pseudo EF-hands coordinate the Ca^{2+} ion predominantly with main-chain carbonyl oxygen atoms in a 14-residue loop.

In one of the most recent statistical analyses, Pidcock and Moore [35] attempted to present a more comprehensive summary of non-EF-hand binding sites without a bias towards EF-hand proteins. Their dataset comprised 60 calcium-binding sites from 44 crystal structures of CaBPs deposited in the Protein Data Bank (PDB) from 1994 to 1999, with resolutions within the range 1.0–2.5 Å, and data from the Cambridge Structural Database (http://www.ccdc.cam.ac.uk/). However, of the more than 3,500 X-ray diffraction structures with identified Ca^{2+} ions currently in the PDB, over 2,200 of them were deposited after the Pidcock and Moore study in 2001. This increase is partially due to advances in structural genomics and rapid development of methods for structural determination [44, 48–51].

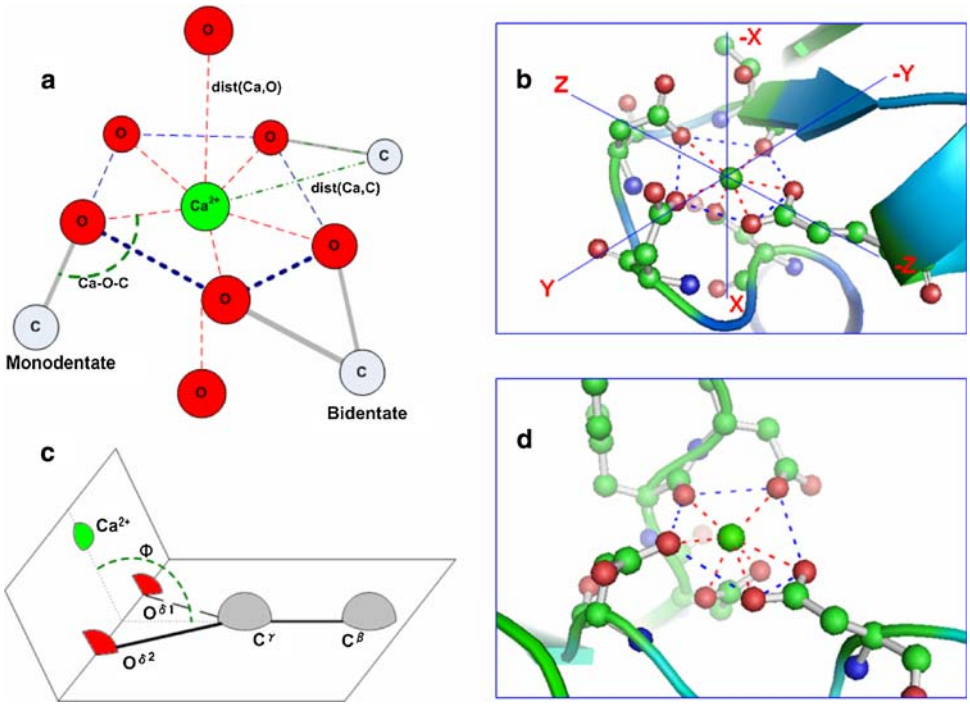

Fig. 1 **a** Illustration of key structural characteristics. The physical relationships between the Ca^{2+} ion, the ligand oxygen, and the ligand oxygen atoms covalently bound to carbon are defined by the angle Ca–O–C and distances dist(Ca,C) and dist(Ca,O). **b** Pentagonal-bipyramidal geometry of EF-hand binding site surrounding Ca^{2+} ion (Protein Data Bank, PDB, sequence ID 1168) from calmodulin (3cln.pdb). The water molecule at $-X$ is not shown. **c** Dihedral angle of bidentate ligands. **d** Pseudo-EF-hand binding site surrounding Ca^{2+} ion (PDB sequence ID 1420) from synaptotagmin I C2B domain (1uow.pdb). Pentagonal-bipyramidal geometry is formed by ligands distant from one another in the sequence

To enhance the capabilities of software to predict binding sites for Ca^{2+} and other metals, and to improve our ability to design CaBPs with high specificity and tailored functions, we report in this paper efforts to more precisely define the structural parameters associated with Ca^{2+} binding. A revised statistical analysis of the structural characteristics of both EF-hand and non-EF-hand proteins was conducted to determine the physical parameters and key characteristics associated with Ca^{2+} binding in different classes of CaBPs. The dataset used to build a comprehensive and representative database for our statistical analysis comprised a total of 1,605 binding sites from 558 PDB files with resolution $R \leq 2.0$ Å, based on a cutoff distance of 3.5 Å as the maximum ligand distance for oxygen and nitrogen, and following removal of all structures with 90% or greater sequence homology. We have shown that clear differences were apparent in the distribution of ligand types and charged ligand residues between the two classes. Quantifiable differences were also apparent in bond angles between EF-hand and non-EF-hand binding sites. New statistical parameters were introduced to correlate the relationship between ligand distances and bond angles.

Methods

Selection of Ca^{2+}-binding protein dataset

All data files were downloaded from the PDB, and relevant data were extracted using MATLAB (The MathWorks, Natick, MA, USA). The PDB file ID and sequence ID for the Ca^{2+} ion associated with each binding sites are summarized in Table S1. All statistical results were divided into two datasets (Table 1): one each for non-EF-hand and EF-hand proteins, where the non-EF-hand protein dataset contains 1,468 binding sites and the EF-hand protein dataset contains 137 calcium-binding sites, respectively. The preponderance of non-EF-hand proteins was attributable to the fact that more non-EF-hand protein structures were available at or below the resolution cutoff of 2.0 Å. A number of EF-hand proteins were not identified as such by their Structural Classification of Proteins (SCOP) classification, but were later identified using pattern motifs and related software developed in our laboratory [44], and less than 10% of sites classified as EF-hand sites for the statistical analysis were pseudo-EF-hand sites belonging to the S100 family. Summarized values presented in Table 1 are discussed later in the appropriate sections.

The cutoff distance of 3.5 Å was selected for several reasons. First, various studies have evaluated first-shell Ca–O binding up to 4.0 Å, and reported a limited number of Ca–O bonds within the range 3.4–3.8 Å, although the majority of bond lengths fall within the range 2.2–2.9 Å [10, 29, 34]. Dudev et al. [52] evaluated first shell–second shell interactions for metal binding with a second-shell cutoff distance of 3.5 Å. Additionally, Nayal and Di Cera [53] reported statistical results using a cutoff distance of 3.5 Å, and previous work in our laboratory has demonstrated that this cutoff distance is valid for rapid and accurate prediction of Ca^{2+}-binding sites [54].

Table 1 Summary of binding distance and angle values

	Non-EF-hand	EF-hand
Total binding sites	1,468	137
FC	1 ± 1	3 ± 1
Mean CN PLW	6 ± 2	7 ± 1
Mean CN PL	4 ± 2	6 ± 1
In Hull (%)	72	100
Bidentate dihedral angle	168.1 ± 9.7	170.6 ± 7.1
Mean Ca–O distance (Å)		
MC carbonyl	2.4 ± 0.2	2.3 ± 0.1
SC	2.4 ± 0.2	2.4 ± 0.2
Bidentate	2.6 ± 0.3	2.5 ± 0.2
Ca–O distance range (Å)		
MC carbonyl	2.0–3.5	2.0–2.6
SC	1.6–3.5	1.8–3.5
Bidentate	1.8–3.5	2.2–3.5
Mean Ca–C distance (Å)		
MC carbonyl	3.5 ± 0.2	3.5 ± 0.1
SC	3.5 ± 0.2	3.4 ± 0.1
Bidentate	2.9 ± 0.2	2.9 ± 0.1
Ca–C distance range (Å)		
MC carbonyl	3.0–4.6	3.1–3.9
SC	2.8–4.6	2.9–3.9
Bidentate	2.4–3.7	2.6–3.4
Mean Ca–O–C angle (°)		
MC carbonyl	151.5 ± 15.8	159.8 ± 12.5
SC	140.4 ± 15.2	136.7 ± 16.0
Bidentate	93.6 ± 11.3	92.9 ± 6.8
Ca–O–C angle range (°)		
Carbonyl	81–180	126–180
SC	56–180	116–170
Bidentate	61–140	66–120

FC negative formal charge, *CN* coordination number, *PLW* protein and water ligands, *PL* protein ligands, *MC* main chain, *SC* side chain

Parameter analysis

The mean Ca–O $(\overline{Ca-O})$ and Ca–C $(\overline{Ca-C})$ distance values were calculated using Eqs. 1 and 2, respectively.

$$\overline{\text{Ca}-\text{O}} = \frac{1}{k}\sum_{1}^{k} \text{dist(Ca,O)} \quad (1)$$

$$\overline{\text{Ca}-\text{C}} = \frac{1}{m}\sum_{1}^{m} \text{dist(Ca,C)} \quad (2)$$

In Eq. 1, k is the number of ligands in one site. In Eq. 2, m is the number of bonded carbon atoms, and $k \geq m$. When k equals m in a single binding site, this indicates that only monodentate ligands appear in this site, otherwise k must be greater than m for polydentate ligands.

Bidentate ligands commonly originate from residues Glu and Asp, and to a lesser extent from Gln and Asn. The extent to which a side-chain residue is bidentate is dependent on the relative position of the metal ion to the ligand atoms. To our knowledge, previous statistical analyses identifed bidentate ligands only on the basis of a cutoff distance, so to more accurately report coordination numbers for this analysis, bidentate ligands were identified on the basis of a bidentate ligand propensity property L_β, as defined by Eq. 3, which predicts bidentate property as a function of deviation from an idealized symmetry.

$$L_\beta = (d_1/d_2) \quad (3)$$

To evaluate this, we first considered a theoretically ideal, symmetrical bidentate ligand (Fig. 2a) where the Ca–O distances (d_1, d_2) and Ca–O–C angles (θ_1, θ_2) for each potential ligand are equivalent: therefore $d_1/d_2 = 1$, and $\theta_1/\theta_2 = 1$. A geometric relationship (described in the supplementary electronic material) exists between these ratios which allows us to use d_1/d_2 as a measure of the deviation from ideal symmetry. As the position of the metal ion shifts relative to that of the ligands (Fig. 2b), the binding character becomes increasingly monodentate, and the ratio d_1/d_2 increases or decreases proportionally, describing a range of valid values to distinguish bidentate from monodentate ligands. To establish a valid range for L_β to identify bidentate ligands, 61 potential bidentate pairs obtained from the Pidcock and Moore [35] dataset were visually inspected using PyMOL (http://pymol.sourceforge.net/). From this dataset, a valid range of L_β for identified Ca^{2+}-binding bidentate ligand pairs was calculated as 1.07 ± 0.34. This range was then used as a filter to identify bidentate ligands for the EF-hand and non-EF-hand sites.

Dihedral angles were also calculated for bidentate ligands, defined as the angle between the plane formed by the side-chain carboxyl group (–COO), and the plane formed by the two carboxyl oxygen atoms and the Ca^{2+} ion (Fig. 1c). Finally, an additional property, Hull (Fig. 2c–f), was examined to describe the spatial relationship of the Ca^{2+} ion to the interior volume of the inner shell binding ligands. This property functioned as a Boolean operator, indicating only whether or not the ion was enclosed in the defined volume.

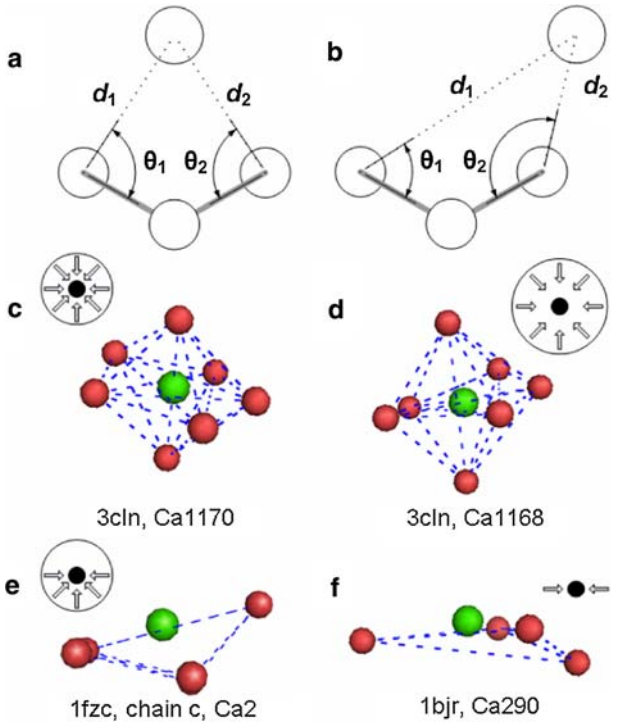

Fig. 2 **a** Symmetrical bidentate structure and **b** monodentate structure where the ion is bound to only one ligand atom. The relationship between each potential ligand oxygen and the Ca^{2+} ion is defined by distances d_1 and d_2, and angles θ_1 and θ_2. **c** Tight holospheric binding and **d** loose holospheric binding where the Ca^{2+} ion is enclosed in a volume defined by binding ligands. **e** Hemispheric binding where the Ca^{2+} ion is exposed on one hemispheric surface and **f** planar binding where the Ca^{2+} ion is bound in a ring structure with exposure above and below the plane

Analysis of bimodal peak distribution for EF-hand ligands

An observed bimodal distribution of Ca–O–C angles for side-chain and main-chain EF-hand ligands was further analyzed on the basis of ligand distribution and protein family. Ligands comprising the Ca–O–C angle distribution were plotted for comparative analysis. The distribution of Ca–O–C angles was subdivided into two regions, R1 and R2, corresponding to angle ranges 116.00°–138.49° and 138.50°–170.00° for the side chain, and 116.00°–163.49° and 163.50°–180.00° for the main chain, respectively. The ligand contribution to these regions from each protein structure was determined. A multiple-sequence alignment was conducted with ClustalW, using a gap open penalty of 10 and a gap extension penalty of 0.5 [55] for all chains of all EF-hand PDB structures. The resulting output file was edited to remove redundant chains, so the final data file

contained only multiple chains from a structure if the chains were unique. This file was then read using TreeView software [56] to generate unrooted neighbor-joining phylogenic trees (Fig. S1). Chains were further labeled on the phylogenic trees with their appropriate SCOP [57] family classifications using data from the PDB, except for the calmodulin-like SCOP families, where the label was excluded to improve readability. Further detailed analysis was then performed to determine potential correlation between the ligands within the Ca–O–C angle data and protein family evolution based on the ligand distribution in regions R1 and R2.

Results and discussion

Coordination number and geometric configuration

Figure 3a summarizes mean coordination number values reported both with and without contribution from water oxygen ligands, where protein ligands may be oxygen or nitrogen. In all, 9,507 ligands were identified, and of these, 2,915 were water oxygen ligands. Interspersed within this group of identified ligands were 137 nitrogen atoms, approximately 42% of which fell within 2.9 Å of the Ca^{2+} ion.

As seen in Table 1, the mean coordination number for EF-hand sites, including water, was 7 ± 1, compared with 6 ± 2 for non-EF-hand sites. All coordination number values were reported as whole integers only. When water was excluded, coordination number values for EF-hand and non-EF-hand sites were reduced to 6 ± 1 and 4 ± 2, respectively.

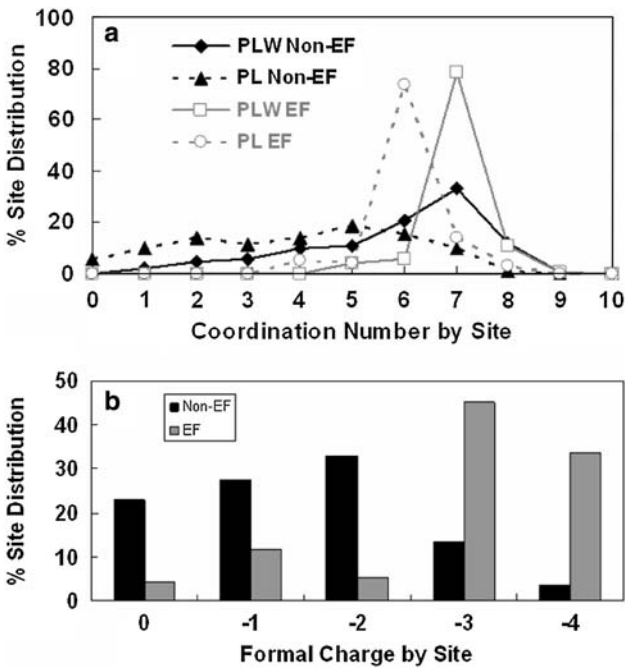

Fig. 3 **a** Coordination number and **b** distribution of formal charge by site for non-EF-hand and EF-hand protein classes

Figure 3a summarizes the percent site distribution of coordination number values for EF-hand and non-EF-hand sites, both with and without water. It is evident that fewer protein ligands are involved in the coordination of the non-EF-hand binding sites, which also coordinate with more water ligands. The significant impact of the water molecules in the non-EF-hand proteins is clearly evident in Fig. 3a, where the distribution of ligands reaches a maxima at a coordination number of 7 only with the inclusion of water.

On the other hand, the mean coordination number values reported in this study for EF-hand proteins were consistent with previously published results indicating a coordination number of six to eight ligands for Ca^{2+} binding, with a mean of 7.3, or between six and seven [5, 36]. Additionally, we identified 13 EF-hand sites having fewer than six protein ligands (Fig. 3a). Closer inspection of the PDB structures associated with these sites revealed that they are either S100-type sites classified as EF-hands, or included an additional water oxygen ligand, although the binding sites were still located within a flexible loop region flanked by two helices.

The space defined by edges connecting the ligand vertices of each binding site, termed Hull, was examined to determine the relative position of the Ca^{2+} ion to the binding-site center of volume. EF-hand proteins tend to have well-defined pentagonal-bipyramidal geometric structure, as seen in Fig. 1b; however, the results of this analysis indicated a wider variety of low-coordination structures associated with Ca^{2+} binding in non-EF-hand proteins. In Fig. 2c–f, four generalized structures of Ca^{2+}-binding sites are described. Figure 2c shows a holospheric site, with the ion enclosed in the Hull volume, and closer ligand interactions between the ion and the first shell. Figure 2d also shows a holospheric site, but with increasing distance between the first-shell ligands and Ca^{2+}. Figure 2e describes a hemispheric site, with the Ca^{2+} ion partially or fully removed from the defined volume, whereas Fig. 2f describes a roughly planar binding site. The results of the statistical analysis (Table 1) indicated that 72 and 100% of evaluated ions were positioned within the Hull, for non-EF-hand and EF-hand, respectively. These results suggest that most Ca^{2+} binding involves a holospheric coordination geometry, which corresponds to an increasing coordination number value. This geometric configuration may contribute to strong metal selectivity for calcium over other metal ions.

Ligand type

Ligand distribution is summarized in Fig. 4, which illustrates major differences in ligand type between EF-hand sites and non-EF-hand sites. Side-chain groups account for the highest proportion of ligands in both classes, but represent a much higher percentage for EF-hand (65.3%) than for non-EF-hand (42.9%). For EF-hand, the ligand distribution follows the order side-chain Asp > side-chain Glu > main-chain carbonyl, with the percentages of Asp and Glu being nearly equivalent (29.7 and 26.6%; Fig. 4a). For non-EF-hand, water is the dominant ligand, comprising 33.1%, followed by side-chain Asp (24.5%), main-chain carbonyl (23.9%), and side-chain Glu (10.4%). It can be seen that the ratio of Asp to Glu for non-EF-hand (23.9:10.4) is nearly identical to that reported by Pidcock and Moore [35] (2.4:1) but is not so for EF-hand (29.7:26.6). Moreover, as seen in Fig. 4, the presence of side-chain hydroxyl and amide groups (Asn, Gln, Ser, Thr, Tyr) is minimal in both classes.

Consistent with previous analyses on the canonical EF-hand motifs based on either sequence or structure [30, 35], Glu is much more common to EF-hand binding sites than to the non-EF-hand sites, but this finding differs significantly from ligand preferences of 65 and 21% for Asp and Glu, respectively, reported by Falke et al. [58] based on 567 canonical EF-hand sequences analyzed. This common presence of Glu within EF-hand sites is likely due to its strategic importance as a bidentate, anchoring ligand for Ca^{2+} [59], and its strong propensity towards helix formation.

In our analysis, we have shown that EF-hand proteins, including both canonical and pseudo EF-hand motifs, have similar distributions of Asp and Glu side chains. The presence of carbonyl oxygen ligands is similar for EF-hand (21.4%) and non-EF-hand (23.9%), so the major contributor in non-EF-hand to replace the missing carboxyl ligands comes from the increasing presence of available water oxygen atoms, as previously noted.

The data for EF-hand sites follow the trends reported by Dudev and Lim [10] (carboxylates > carbonyls > water ≫ hydroxyl atoms), but vary for the non-EF-hand proteins, owing to the increased presence of water. In fact, the 33.1% distribution of water oxygen ligands for non-EF-

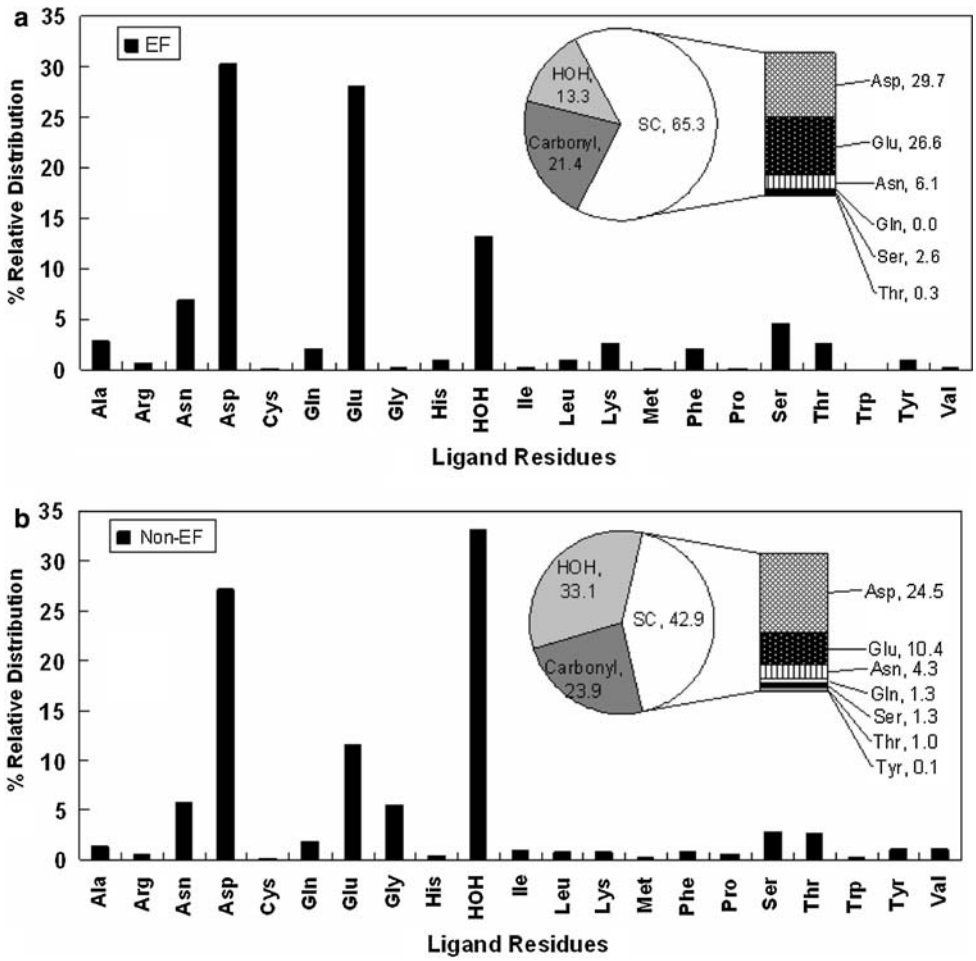

Fig. 4 Frequency distribution of ligand residues for **a** EF-hand and **b** non-EF-hand proteins. The pie graphs show the distribution between water, carbonyl, and side-chain-residue oxygen ligands. *SC* side chain

hand sites is much higher than the values reported by McPhalen et al. [31] (20%) and Dudev et al. [52] (22.4%). In the case of Dudev et al., this difference may in part be attributed to the shorter distance cutoff constraint in their study (2.9 vs. 3.5 Å). When water oxygen ligands with Ca–O distances in the range 2.9–3.5 Å are eliminated from our statistics, the distribution of water oxygen ligands drops to 25.0%, which is only a modest increase from the value reported by Dudev et al., but is still significantly different from the 13.3% reported for EF-hand sites (Fig. 4a), which all fall within 2.9 Å.

Charge analysis by site

Formal charge by site was simplified to account only for negatively charged side-chain carboxyl groups (−1) from Glu and Asp [30]. Mean negative formal charge values of 1 ± 1 and 3 ± 1 were found for non-EF-hand and EF-hand sites, respectively. The difference in distribution between formal charge in non-EF-hand and EF-hand sites is shown in Fig. 3b, with increased negative charge (3–4) being more apparent in the case of EF-hand sites. The charge data for EF-hand sites are consistent with those of previous studies of Ca^{2+} indicating that a microenvironment containing three to four negative charges likely represents an optimal charge configuration [30, 60]; however, it is surprising that only a small percentage of non-EF-hand sites exhibit negative charge greater than 2 (Fig. 3b).

Calcium-binding sites with high negative formal charges are likely located in flexible loop regions of the protein. Figure 5 shows examples of canonical EF-hand proteins and non-EF-hand proteins with −4 formal charge values. Figure 5a shows the EF-hand binding site with a helix–loop–helix from the crystal structure of coelenterazine-binding protein from *Renilla muelleri* (2hq8.pdb) formed by the side chains of Asp112, Asp114, Asp116, and Glu123 and the main chain of Tyr118. Figure 5b shows Ca^{2+} binding in a flexible loop region between two β-strands from family 9 carbohydrate-binding module of *Thermotoga maritima* xylanase10A (1i8a.pdb). Figure 5c shows Ca^{2+} binding between two loop regions discontinuous in the primary sequence from calcium-binding domain 2B (1uow.pdb) [61].

Binding sites without side chains were identified within both classes. For the EF-hand proteins, six binding sites were identified which had zero formal charge in the binding site, all from the protein calprotectin (1xk4.pdb) [62]. These are S100/pseudo EF-hand sites predominantly binding with main-chain carbonyl oxygen atoms. Surprisingly, over 20% of the non-EF-hand binding sites (338 sites from 153 proteins) had reported formal charge values of zero.

Detailed structural analyses of the protein environments of these charge-deficient sites reveals that charge–charge stabilization beyond the chelated metal ion leads to the exclusion of available negatively charged side-chain residues and facilitates the binding of Ca^{2+} with carbonyl oxygen atoms. Figure 5d shows the Ca^{2+}-binding site (Ca^{2+} sequence ID 3012) of Drp35 (2dg1.pdb), a protein induced by cell-wall-affecting antibiotics or detergents, which possesses calcium-dependent lactonase activity [63]. Ca^{2+} is bound with six oxygen and two nitrogen ligands from a loop region in chain F. Binding ligands for this site are summarized in Table S2. It is interesting to note that, except for OG1 of Thr133, all other protein oxygen ligands are from the main chain, despite the apparent availability of side-chain oxygen atoms from Tyr135, Asp130, and Ser110. A closer examination (Fig. 5d) of the structure reveals several probable charge interactions (Table S3) between Asp130 OD2 and Lys86 NZ and between Ser110 N and Asp107 OD2. These interactions with ligands beyond the second shell likely stabilize the holoprotein and account for the use of main-chain carbonyl oxygen ligands in the binding site. These types of interactions are even more apparent with the structure from 2exh.pdb, a family 43 β-xylosidase from *Geobacillus stearothermophilus* [64]. Ca^{2+} site 2003 from 2exh is coordinated by three main-chain oxygen ligands, two from Asp residues 333 and 528, where 528 originates in a β-strand, and three water oxygen ligands (Fig. 5e). As was observed with 2dg1 of Drp35, charge interactions are apparent between Asp333 N and Lys331 O, Asp333 OD2 and Lys331 NZ, Asp528 OD2 and Lys395 NZ, and Asp528 OD1 and His363 NE2 (Table S3). For these residues, both side-chain and main-chain charge interactions with ligands in the second shell would stabilize the structure, enabling the observed Ca^{2+} coordination by available main-chain oxygen atoms.

Of these sites, 244 lacking charged ligand residues had fewer than three protein ligands. The absence of charge and low coordination number is unusual for Ca^{2+}-binding sites, and may represent incidental effects of crystallization, rather than naturally occurring binding sites. As such, data collected from these sites, while still meaningful with respect to individual ligand interactions, are less desirable for characterization of the total binding site. An example of a very unusual site (Ca^{2+} 503 from 1ava.pdb) is shown in Fig. 5f. This site shows a hydrated Ca^{2+} ion bound between barley α-amylase and its endogenous protein inhibitor barley α-amylase/subtilisin inhibitor [65]. As seen in Fig. 5f, no protein ligands are associated with the Ca^{2+} ion: however, the hydrated ion is surrounded by an outer shell of protein ligand oxygen atoms orientated towards the hydrated ion, probably as a result of hydrogen-bonding with the inner shell of water molecules.

As shown in Fig. 6a, the non-EF-hand site (site 145) from the isolated N-terminal domain of protein S from *Myxococcus xanthus* (1nps.pdb) is coordinated by three

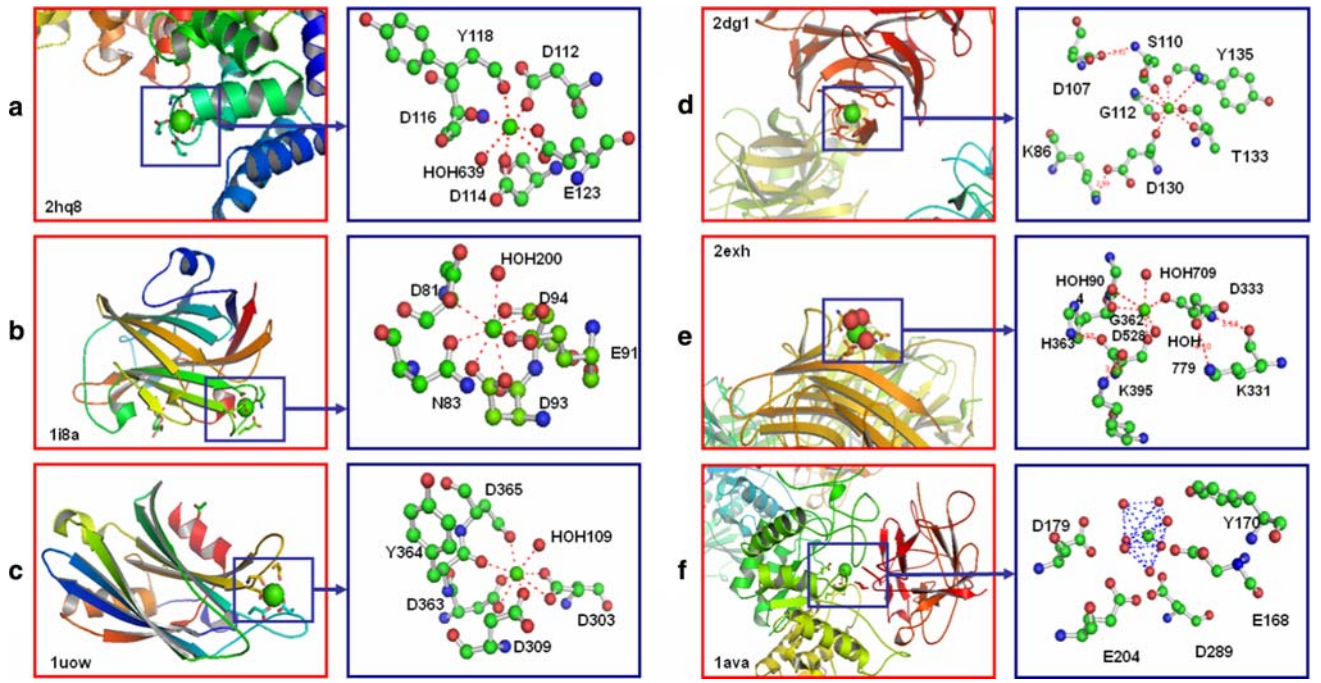

Fig. 5 **a** Canonical EF-hand binding site from coelenterazine-binding protein from *Renilla muelleri*. **b** C-terminal module of the thermostable *Thermotoga maritima* xylanase 10A (1i8a.pdb). Four negative charges are found in the site. **c** Synaptotagmin I C2B domain (1uow.pdb). Four negative charges are found in the site. **d** Binding site (PDB sequence ID 3012) of Drp35 (2dg1.pdb). Ca^{2+} is bound with five main-chain and one hydroxyl oxygen atoms, and two nitrogen atoms from the loop region of chain F. Formal charge in site is zero. **e** Binding site (PDB sequence ID 2003) from β-xylosidase (2exh.pdb). Ca^{2+} is coordinated by three protein ligands and three water oxygen ligands. Formal charge in site is zero. **f** Binding site (PDB sequence ID 503) from 1ava.pdb. The hydrated Ca^{2+} is bound between barley α-amylase and its endogenous protein inhibitor barley α-amylase/subtilisin inhibitor. Formal charge is zero

protein ligands, Asn36, Asn77, and Gln53 (Fig. 6c), with three additional water oxygen ligands. The absence of charge in this region is illustrated in the electrostatic potential map (Fig. 6b) generated using DelPhi and GRASP [66–70], where the surface region corresponding to the binding site appears neutral (i.e., surface area represented in gray and white contours, with no red contouring indicating the presence of negative charge). Unlike the majority of CaBPs, Ca^{2+} binding was not shown to increase protein stability in the case of the isolated N-terminal domain of protein S, although the intact protein was reported by Wenk et al. [71] to be very stable across a wide pH range (2–10), and to be resistant to both urea and thermal unfolding. Wenk et al. also reported an unfolding intermediate where the N-terminal domain remains folded during unfolding of the C-terminal domain. Cooperativity between these domains contributes to the overall stability, and it is evident from their study that the absence of charge in the binding site (site 145) has no apparent correlation with stability. Studies of Ca^{2+} have indicated higher binding affinities are observed in the presence of a net negative charge within 5–15 Å of the ion [72, 73]; however, the relationship between charge and thermal stability is less easily evaluated. Vogt et al. [74] presented a summary of studies related to thermal stability where charge interaction and metal binding were cited as factors contributing to increased stability. Recent work in our laboratory has suggested that increasing the number of localized charge residues in a relatively constrained calcium-binding site while increasing both metal binding affinity and stability of the calcium-loaded form also leads to decreased stability of the apo form [75, 76]. It is clear that such effects are context-dependent; i.e., strongly depend on the protein environment. Consequently, no attempt was made to evaluate thermal stability with respect to charge availability on the basis of the data presented in this study.

Nitrogen, which is associated with Ca^{2+} binding in studies of small molecules, was not observed to a significant extent in the protein structures analyzed. Of the 137 nitrogen ligands identified in the study, only two were found within EF-hand binding sites. Interestingly, 85 of the 135 nitrogen ligands identified in non-EF-hand binding sites are reported in the zero-charge sites. This suggests that nitrogen may potentially play a secondary role in Ca^{2+} binding, and that this role may increase in the absence of negative charge, possibly owing to a reduction in charge repulsion forces.

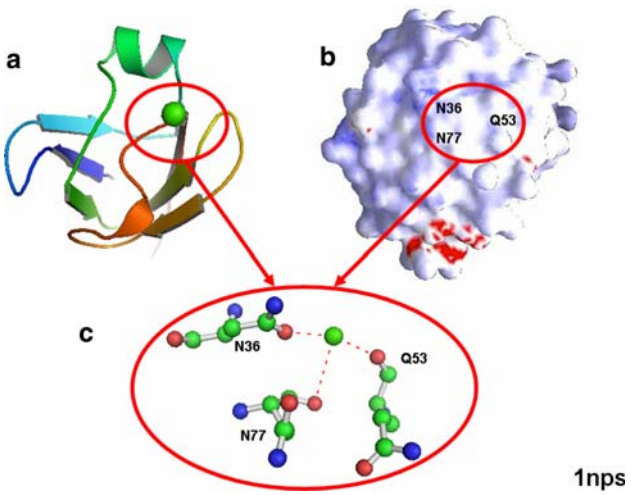

Fig. 6 **a** Ca^{2+}-binding site for protein S from *Myxococcus xanthus* (1nps.pdb), where the binding site around calcium 145 has zero assigned formal charge. **b** Electrostatic potential map for 1nps showing neutral charge (*gray* and *white contours*) in the site. **c** Expanded view of binding-site protein ligands

Distance parameters

Mean distance parameters are summarized in Table 1 and in Fig. 7. The ranges of reported Ca–O and Ca–C distances are also summarized in Table 1, based on a 0.1-Å-interval bin.

A clear delineation between these distances is apparent for both carbonyl and side-chain ligands in the non-EF-hand and EF-hand datasets (Fig. 7a, c) at or near 3.0 Å. The mean Ca–O distance values reported in this study for EF-hand and non-EF-hand indicate very little difference between carbonyl and side-chain oxygens, and between the different classes (Table 1) for each ligand type. These results are identical to those of previously cited studies; however, our data separate bidentate ligand distances, which are slightly longer for both EF-hand (2.5 ± 0.2 Å) and non-EF-hand (2.6 ± 0.3 Å) than for the carbonyl and side-chain ligands, in their respective classes. Moreover, a more pronounced change was observed for the bidentate mean Ca–C distances, which were 0.5–0.6 Å shorter than the distances found for carbonyl and side-chain ligand oxygen atoms, resulting in overlap between the Ca–O and Ca–C shells (Fig. 7b, d). Also, the majority of reported Ca–O distance values for carbonyl and side-chain oxygen ligands, as seen in Fig. 7a (non-EF-hand) and Fig. 7c (EF-hand) indicate a narrower range of ligand distances (2.0–2.8 and 2.1–2.6 Å, respectively) than previously reported values approximately between 2.0−3.2 Å [31].

These effective binding ranges suggest that our cutoff of 3.5 Å may be arbitrarily long for structural parameterization, given the narrow distributions seen in Fig. 7a and c. Additionally, nearly 60% of the identified nitrogen ligands fell within the range 2.9–3.5 Å, and because nitrogen is accepted as only a marginal binding ligand for Ca^{2+}, it is probable that some of these identified ligands do not interact with the Ca^{2+} ion. However, only 7% of the non-EF-hand ligands, and 1% of the EF-hand ligands, fell within the range 2.9–3.5 Å, and these data did not significantly alter the statistical results, except in the case of water molecules for non-EF-hand proteins, where approximately 25% of the water oxygen ligands were identified in

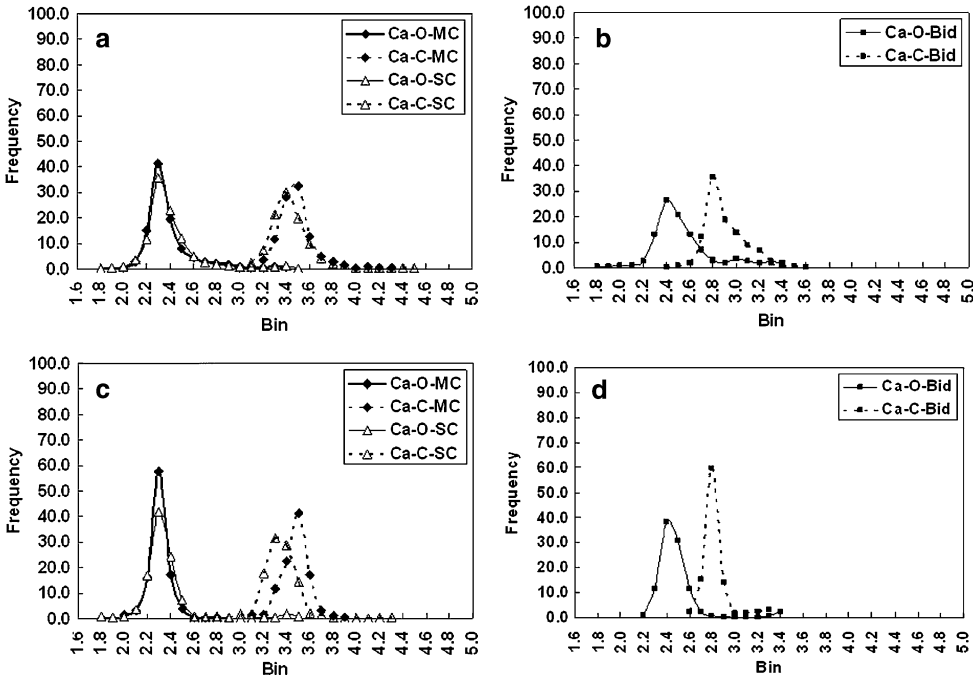

Fig. 7 Ca–O distance comparisons. **a** Non-EF-hand carbonyl and side-chain oxygen, **b** non-EF-hand bidentate oxygen, **c** EF-hand carbonyl and side-chain oxygen, and **d** EF-hand bidentate oxygen. *Bid* bidendate, *MC* main chain

this range. It is likely that not all of these identified water oxygen atoms are ligands; however, it is clear from Fig. 3a that binding in the non-EF-hand sites with 7 ± 1 ligands is possible only with inclusion of the water molecules, including those more than 2.9 Å from the Ca^{2+} ion. We conclude from this that the longer cutoff of 3.5 Å, while valid for this large statistical analysis, can likely be reduced to 2.9 Å for future studies without loss of relevant data, with a possible exception for the special case of water.

Analysis of bond angles

The mean Ca–O–C angles (Fig. 1a) are summarized in Table 1, where differences are observed between carbonyl, side-chain, and bidentate oxygen ligands. The Ca–O–C angles were largest for carbonyl (151.5 ± 15.8° and 159.8 ± 12.5°), followed by side chain (140.4 ± 15.2° and 136.7 ± 16.0°), and bidentate (93.6 ± 11.3° and 92.9 ± 6.8°) for non-EF-hand and EF-hand, respectively.

Figure 8a and b shows the distribution of Ca–O–C angles for non-EF-hand and EF-hand Ca^{2+}-binding ligands, respectively. It can be seen that a Gaussian distribution of angle values is associated with non-EF-hand ligands, and the range values for both classes are nearly identical for carbonyl, side chain, and bidentate (Table 1). For the bidentate ligands, the mean and standard deviation values for dihedral angles are also summarized in Table 1. For non-EF-hand and EF-hand these values were found to be 168.1 ± 9.7° and 170.6 ± 7.1°, respectively.

To understand the observed bimodal distribution associated with Ca–O–C angles for side-chain, and to a lesser extent, main-chain EF-hand ligands in Fig. 8b, two additional analyses were performed. The results of these analyses are summarized in Fig. 8c and Table S4. In Fig. 8c, the ligands represented by the angle data in Fig. 8b were further divided by residue type. The majority of these ligands (262 of 366) originated from Asp, and it is clear that the bimodal distribution seen in Fig. 8b mirrors the distribution of Asp in Fig. 8c. A similar analysis (not shown) of the data for main-chain carbonyl residues associated with the Ca–O–C angles in Fig. 8a did not reveal a residue-specific origin for the apparent bimodal main-chain distribution.

Table S4 summarizes data for the EF-hand ligands comprising the Ca–O–C angle distribution in Fig. 8b, and the associated protein families. From the data presented in Table S4, protein structures were identified where a majority of the ligands contributed preferentially to one peak region or the other. These regions, mapped to the unrooted neighbor-joining phylogenic trees in Fig. S1, indicated two major phylogenic branches where the peak-region side-chain Ca–O–C angle distribution correlated with related protein families. In Fig. S1a, peak-region R1 angles corresponded with the highly conserved parvalbumin and penta-EF SCOP families (shaded oval), while angles from peak-region R2 corresponded with the S100 family of proteins (shaded rectangle). However, no consistent trends were observed along the calmodulin-like

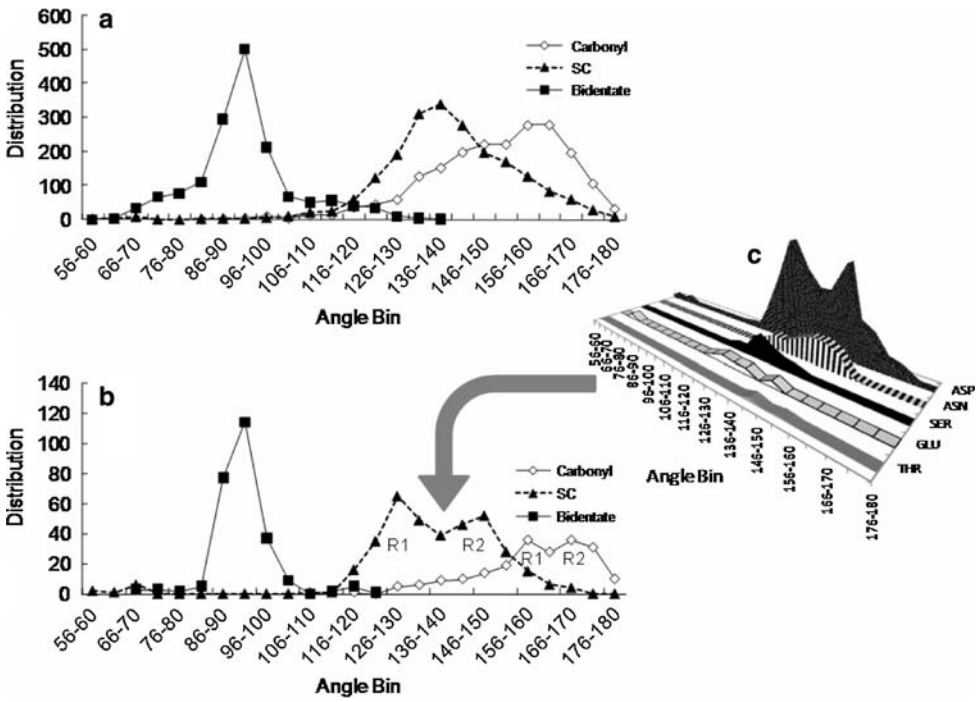

Fig. 8 Ca–O–C angle distribution for **a** non-EF-hand and **b** EF-hand binding sites. **c** Ligands comprising bimodal side-chain distribution for EF-hand

SCOP branches. Interestingly, the data in Table S4 suggest an inverse correlation between the side-chain and main-chain angles and the related protein families, where the smaller side-chain angles in peak-region R1 (shaded oval in Fig. S1a) are coupled, by the protein chain, with larger main-chain angles in peak-region R2 (shaded rectangle in Fig. S1b) and vice versa. However, this reversal was not observed in all cases (Table S4), so the correlation is considered weak.

These results suggest two possible explanations for the observed side-chain bimodal distribution in Fig. 8b. First, the emergence of two quasi-discrete peaks for Asp may result from charge interactions between the Ca^{2+} ion and the nonligand carboxyl oxygen from Asp, such that proximity of the second oxygen, while outside the ligand cutoff distance of 3.5 Å, is sufficient to effectively reduce the Ca–O–C angle of the primary oxygen ligand. Second, based on the data presented in Table S4 and Fig. S1, it is possible that discrete angles are conserved along evolutionary lines. A third possibility is that the observed bimodal distribution relates to different secondary properties associated with the ligand residues, and further work is currently in progress to evaluate this.

Conclusions

The data presented in this study are based on the most comprehensive statistical analysis of higher-resolution Ca^{2+}-binding structures available to date. While certain data presented here with respect to EF-hand proteins are generally consistent with those of previously reported studies, a clear distinction can be made between EF-hand and non-EF-hand proteins, on the basis of the physical properties assessed. It is apparent from the data that non-EF-hand CaBPs coordinate with fewer ligands, on average, than the EF-hand proteins, and with a higher proportion of bound water molecules. Less formal charge is evident in the non-EF-hand binding sites, which is expected given the lower proportion of charged side-chain ligands. The EF-hand sites additionally exhibit a bimodal distribution of side-chain Ca–O–C angles, which may be due to the abundant presence of Asp as a chelating ligand residue, which in turn may be conserved along evolutionary lines. In both classes, the majority of Ca^{2+} ions are surrounded by a holospheric binding geometry. In the case of EF-hand proteins, this frequently involves a pentagonal-bipyramidal geometry, whereas the non-EF-hand binding sites exhibit less regular structure. The Ca–O bond lengths for both classes were generally equivalent, but discrete differences were apparent in the bond angles, and in both cases the range of bond angles was narrower than previously assumed (Table 1). Additionally, the dihedral angles for non-EF-hand and EF-hand binding sites were generally equivalent, with low standard deviations, indicating that these values (168.1 ± 9.7° and 170.6 ± 7.1°) may be utilized as input parameters for computational design.

The significant differences between ligand types (carbonyl, side chain, bidentate) demonstrate the necessity of classifying these angles separately. Moreover, the small standard deviation in each case provides a narrower range of ideal angles for each ligand type, thus improving our input parameters used to design proteins with specific Ca^{2+}-binding characteristics.

The physical parameters and key characteristics associated with Ca^{2+} binding in different classes of CaBPs identified from our analysis have twofold significance. First, structural parameters derived from a more current, comprehensive dataset provide a more accurate representation of Ca^{2+} binding, particularly between different classes of CaBPs. Second, these data will provide input parameters to both improve the accuracy of prediction algorithms and facilitate the design of engineered CaBPs with high selectivity and affinity for Ca^{2+}. The data compiled in this analysis have been directly applied to define weighted coefficients used in a graph-theory-based prediction algorithm developed in our laboratory which can predict Ca^{2+}-binding sites, within 0.20 Å of the documented site, with 94% sensitivity and 93% selectivity (data to be published). The algorithm also correctly identifies only those ligands comprising the binding site in 45 out of 48 test sites. These results are in part attributable to refinement of the algorithm based on the availability of more precise structural parameters obtained from the statistical analysis reported in this article.

Owing to the ubiquitous presence of CaBPs in biological processes, and the roles of Ca^{2+} imbalance in different diseases, the ability to predict and identify Ca^{2+}-binding sites using computational methods can accelerate our understanding of these processes and problems, and subsequently improve our ability to alter Ca^{2+}-dependent functions for therapeutic purposes, and design CaBPs with tailored functions for medical diagnostics.

Acknowledgments We would like to thank Yubin Zhou, Kun Zhao, and other members of J.J.Y.'s laboratory for their critical review of this manuscript and helpful discussions, and Wei Yang from NIH for additional suggestions. This work is supported in part by the following grants: NIH GM081749, NIH GM 62999, and NSF MCB-0092486 to J.J.Y., and DMS-0500951 to G.C.

References

1. Ibers JA, Holm RH (1980) Modeling coordination sites in metallobiomolecules. Science 209(4453):223–235
2. Bertini I, Sigel A, Sigel H (2001) Handbook on metalloproteins, vol xxvii. Marcel Dekker, New York, p 1182

3. Holm RH, Kennepohl P, Solomon EI (1996) Structural and functional aspects of metal sites in biology. Chem Rev 96(7):2239–2314
4. Tainer JA, Roberts VA, Getzoff ED (1992) Protein metal-binding sites. Curr Opin Biotechnol 3(4):378–387
5. Glusker JP (1991) Structural aspects of metal liganding to functional groups in proteins. Adv Protein Chem 42:1–76
6. Silva JRRFd, Williams RJP (1991) The biological chemistry of the elements: the inorganic chemistry of life, vol xxi. Clarendon Press/Oxford University Press, Oxford/New York, p 561
7. Lippard SJ, Berg JM (1994) Principles of bioinorganic chemistry, vol xvii. University Science Books, Mill Valley, p 411
8. Yamashita MM et al (1990) Where metal ions bind in proteins. Proc Natl Acad Sci USA 87(15):5648–5652
9. Bagley SC, Altman RB (1995) Characterizing the microenvironment surrounding protein sites. Protein Sci 4(4):622–635
10. Dudev T, Lim C (2003) Principles governing Mg, Ca, and Zn binding and selectivity in proteins. Chem Rev 103(3):773–787
11. Babu CS et al (2003) A combined experimental and theoretical study of divalent metal ion selectivity and function in proteins: application to *E. coli* ribonuclease H1. J Am Chem Soc 125(31):9318–9328
12. Dudev T, Lim C (2007) Effect of carboxylate-binding mode on metal binding/selectivity and function in proteins. Acc Chem Res 40(1):85–93
13. Chakrabarti P (1990) Geometry of interaction of metal ions with histidine residues in protein structures. Protein Eng 4(1):57–63
14. Chakrabarti P (1989) Geometry of interaction of metal ions with sulfur-containing ligands in protein structures. Biochemistry 28(14):6081–6085
15. Chakrabarti P (1990) Interaction of metal ions with carboxylic and carboxamide groups in protein structures. Protein Eng 4(1):49–56
16. Glusker JP, Lewis M, Rossi M (1994) Crystal structure analysis for chemists and biologists, vol xvii. VCH, New York, p 854
17. Harding MM (2001) Geometry of metal–ligand interactions in proteins. Acta Crystallogr D Biol Crystallogr 57(Pt 3):401–411
18. Harding MM (1999) The geometry of metal–ligand interactions relevant to proteins. Acta Crystallogr D Biol Crystallogr 55(Pt 8):1432–1443
19. Harding MM (2000) The geometry of metal–ligand interactions relevant to proteins. II. Angles at the metal atom, additional weak metal–donor interactions. Acta Crystallogr D Biol Crystallogr 56(Pt 7):857–867
20. Mann KG et al (1990) Surface-dependent reactions of the vitamin K-dependent enzyme complexes. Blood 76(1):1–16
21. Herzberg O, Moult J, James MN (1986) A model for the Ca^{2+}-induced conformational transition of troponin C. A trigger for muscle contraction. J Biol Chem 261(6):2638–2644
22. Holmes KC et al (1990) Atomic model of the actin filament. Nature 347(6288):44–49
23. Carafoli E (1994) The signaling function of calcium and its regulation. J Hypertens Suppl 12(10):S47–S56
24. Santella L, Carafoli E (1997) Calcium signaling in the cell nucleus. FASEB J 11(13):1091–1109
25. Martin R (1984) Bioinorganic chemistry of calcium. In: Sigel H (ed) Metal ions in biological systems, vol 17. Marcel Dekker, New York, pp 1–49
26. Nelson MR, Chazin WJ (1998) An interaction-based analysis of calcium-induced conformational changes in Ca^{2+} sensor proteins. Protein Sci 7(2):270–282
27. Kuroki R et al (1989) Design and creation of a Ca^{2+} binding site in human lysozyme to enhance structural stability. Proc Natl Acad Sci USA 86(18):6903–6907
28. Tajima M et al (1976) Role of calcium ions in the thermostability of thermolysin and *Bacillus subtilis* var. *amylosaccharitic*us neutral protease. Eur J Biochem 64(1):243–247
29. Harding MM (2004) The architecture of metal coordination groups in proteins. Acta Crystallogr D Biol Crystallogr 60(Pt 5):849–859
30. Marsden BJ, Shaw GS, Sykes BD (1990) Calcium binding proteins. Elucidating the contributions to calcium affinity from an analysis of species variants and peptide fragments. Biochem Cell Biol 68(3):587–601
31. McPhalen CA, Strynadka NC, James MN (1991) Calcium-binding sites in proteins: a structural perspective. Adv Protein Chem 42:77–144
32. Strynadka NC, James MN (1989) Crystal structures of the helix-loop-helix calcium-binding proteins. Annu Rev Biochem 58:951–998
33. Einspahr H, Cook WJ, Bugg CE (1981) Conformational flexibility in single-stranded oligonucleotides: crystal structure of a hydrated calcium salt of adenylyl-(3′–5′)-adenosine. Biochemistry 20(20):5788–5794
34. Harding MM (2002) Metal–ligand geometry relevant to proteins and in proteins: sodium and potassium. Acta Crystallogr D Biol Crystallogr 58(Pt 5):872–874
35. Pidcock E, Moore G (2001) Structural characteristics of protein binding sites for calcium and lanthanide ions. J Biol Inorg Chem 6:479–489
36. Yang JJ, Yang W (2005) In: King RB (ed) Encyclopedia of inorganic chemistry, 2nd edn. Wiley, New York
37. Harding MM (2006) Small revisions to predicted distances around metal sites in proteins. Acta Crystallogr D Biol Crystallogr 62(Pt 6):678–682
38. Einspahr H, Bugg CE (1984) Crystal structure studies of calcium complexes and implications for biological systems. In: Sigel H (ed) Calcium and its role in biology, vol xxiii. M. Dekker, New York, p 532
39. Swain AL, Kretsinger RH, Amma EL (1989) Restrained least squares refinement of native (calcium) and cadmium-substituted carp parvalbumin using X-ray crystallographic data at 1.6-A resolution. J Biol Chem 264(28):16620–16628
40. Vyas MN, Jacobson BL, Quiocho FA (1989) The calcium-binding site in the galactose chemoreceptor protein. Crystallographic and metal-binding studies. J Biol Chem 264(34):20817–20821
41. Hill E et al (2001) Cadherin superfamily proteins in *Caenorhabditis elegans* and *Drosophila melanogaster*. J Mol Biol 305(5):1011–1024
42. Truong K, Ikura M (2002) The cadherin superfamily database. J Struct Funct Genomics 2(3):135–143
43. Kretsinger RH, Nockolds CE (1973) Carp muscle calcium-binding protein. II. Structure determination and general description. J Biol Chem 248(9):3313–3326
44. Zhou Y et al (2006) Prediction of EF-hand calcium-binding proteins and analysis of bacterial EF-hand proteins. Proteins 65(3):643–655
45. Nelson MR, Chazin WJ (1997) The EF-Hand Calcium-Binding Proteins Data Library
46. Yap KL et al (2000) Calmodulin target database. J Struct Funct Genomics 1(1):8–14
47. Kawasaki H, Nakayama S, Kretsinger RH (1998) Classification and evolution of EF-hand proteins. Biometals 11(4):277–295
48. Andreini C, Bertini I, Rosato A (2004) A hint to search for metalloproteins in gene banks. Bioinformatics 20(9):1373–1380
49. Reeves GA, Thornton JM (2006) Integrating biological data through the genome. Hum Mol Genet 15(Spec No 1):R81–R87
50. Dupont CL et al (2006) Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. Proc Natl Acad Sci USA 103(47):17822–17827

51. Passerini A et al (2007) Predicting zinc binding at the proteome level. BMC Bioinformatics 8:39
52. Dudev T et al (2003) First–second shell interactions in metal binding sites in proteins: a PDB survey and DFT/CDM calculations. J Am Chem Soc 125(10):3168–3180
53. Nayal M, Di Cera E (1994) Predicting Ca^{2+}-binding sites in proteins. Proc Natl Acad Sci USA 91(2):817–821
54. Deng H et al (2006) Predicting calcium-binding sites in proteins—a graph theory and geometry approach. Proteins 64(1):34–42
55. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22(22):4673–4680
56. Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12(4):357–358
57. Murzin AG et al (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247(4):536–540
58. Falke JJ et al (1994) Molecular tuning of ion binding to calcium signaling proteins. Q Rev Biophys 27(3):219–290
59. Yang W et al (2002) Structural analysis, identification, and design of calcium-binding sites in proteins. Proteins 47(3):344–356
60. Yang W, Jones LM, Isley L, Ye Y, Lee HW, Wilkins AL, Liu Z, Hellinga HW, Malchow R, Ghazi M, Yang JJ (2003) Rational design of a calcium-binding protein. J Am Chem Soc 125:6165–6171
61. Cheng Y et al (2004) Crystallographic identification of Ca^{2+} and Sr^{2+} coordination sites in synaptotagmin I C_2B domain. Protein Sci 13(10):2665–2672
62. Korndorfer IP, Brueckner F, Skerra A (2007) The crystal structure of the human (S100A8/S100A9)2 heterotetramer, calprotectin, illustrates how conformational changes of interacting alpha-helices can determine specific association of two EF-hand proteins. J Mol Biol 370(5):887–898
63. Tanaka Y et al (2007) Structural and mutational analyses of Drp35 from *Staphylococcus aureus*: a possible mechanism for its lactonase activity. J Biol Chem 282(8):5770–5780
64. Brux C et al (2006) The structure of an inverting GH43 beta-xylosidase from *Geobacillus stearothermophilus* with its substrate reveals the role of the three catalytic residues. J Mol Biol 359(1):97–109
65. Vallee F et al (1998) Barley alpha-amylase bound to its endogenous protein inhibitor BASI: crystal structure of the complex at 1.9 A resolution. Structure 6(5):649–659
66. Honig B, Nicholls A (1995) Classical electrostatics in biology and chemistry. Science 268(5214):1144–1149
67. Rocchia W, Alexov E, Honig B (2001) Extending the applicability of the nonlinear Poisson–Boltzmann equation: multiple dielectric constants and multivalent ions. J Phys Chem B 105(28):6507–6514
68. Yang AS et al (1993) On the calculation of p*k*(a)s in proteins. Proteins Structure Function Genetics 15(3):252–265
69. Gilson MK, Honig B (1988) Calculation of the total electrostatic energy of a macromolecular system: solvation energies, binding energies, and conformational analysis. Proteins 4(1):7–18
70. Nicholls A, Honig B (1991) A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation. J Comp Chem 12:435–445
71. Wenk M et al (1999) The domains of protein S from *Myxococcus xanthus*: structure, stability and interactions. J Mol Biol 286(5):1533–1545
72. Linse S et al (1988) The role of protein surface charges in ion binding. Nature 335(6191):651–652
73. Linse S, Forsen S (1995) Determinants that govern high-affinity calcium binding. Adv Second Messenger Phosphoprotein Res 30:89–151
74. Vogt G, Woell S, Argos P (1997) Protein thermal stability, hydrogen bonds, and ion pairs. J Mol Biol 269(4):631–643
75. Jones LM et al (2008) Rational design of a novel calcium-binding site adjacent to the ligand-binding site on CD2 increases its CD48 affinity. Protein Sci 17(3):439–449
76. Maniccia AW et al (2006) Using protein design to dissect the effect of charged residues on metal binding and protein stability. Biochemistry 45(18):5848–5856