Review

Target selectivity in EF-hand calcium binding proteins

Shibani Bhattacharya, Christopher G. Bunick, Walter J. Chazin[*]

Departments of Biochemistry and Physics, Center for Structural Biology, 5140 BIOSCI/MRBIII, Vanderbilt University, Nashville, TN 37232-8725, United States

Abstract

EF-hand calcium binding proteins have remarkable sequence homology and structural similarity, yet their response to binding of calcium is diverse and they function in a wide range of biological processes. Knowledge of the fine-tuning of EF-hand protein sequences to optimize specific biochemical properties has been significantly advanced over the past 10 years by determination of atomic resolution structures. These data lay the foundation for addressing how functional selectivity is generated from a generic ionic signal. This review presents current ideas about the structural mechanisms that provide the selectivity of different EF-hand proteins for specific cellular targets, using S100 and calmodulin family proteins to demonstrate the critical concepts. Three factors contribute significantly to target selectivity: molecular architecture, response to binding of Ca^{2+} ions, and the characteristics of target binding surfaces. Comparisons of calmodulin and S100 proteins provide insights into the role these factors play in facilitating the variety of binding configurations necessary for recognizing a diverse set of targets.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Selectivity; Calcium; Binding protein

1. Introduction

Intracellular calcium signaling relies on the superfamily of EF-hand calcium binding proteins, which regulate diverse cellular activities ranging from exocytosis, muscle contraction, metabolism, transcription, fertilization and cell proliferation, to Ca^{2+} buffering and homeostasis [7,8,67]. EF-hand calcium sensors such as S100 and calmodulin family proteins translate the physiological changes in calcium levels into specific cellular responses by undergoing a large conformational change that exposes a binding site recognized by downstream effectors. The sensitivity of these proteins to temporal and spatial changes in calcium concentrations in the range of 0.1–10 μM is achieved by fine-tuning their calcium affinities into precisely the appropriate range.

The basic structural/functional unit of EF-hand proteins is a pair of EF-hand motifs [34] that together form a stable four-helix bundle domain (Fig. 1) [48,61]. The pairing of EF-hands enables cooperativity in the binding of Ca^{2+} ions, which is essential for generating a clean response to the relatively modest change in Ca^{2+} concentration during active signaling [37]. Studies have shown that the entire EF-hand domain must be treated as a single, globally cooperative unit [48,61]. In order to understand transduction of Ca^{2+} signals by EF-hand proteins, it is also important to recognize that the affinity for Ca^{2+}, conformational response to the binding of ions, and interaction with targets are all energetically coupled [6].

Many EF-hand proteins contain more than a single EF-hand domain. Structural variations in how the domains are organized contribute significantly to selectivity for targets
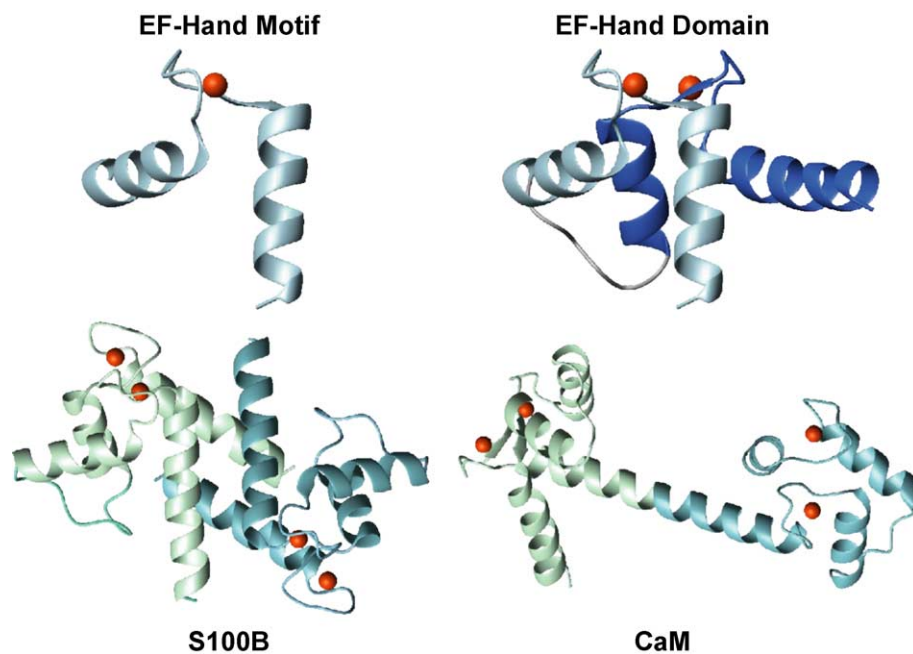

Fig. 1. Basic structural features of EF-hand Ca^{2+} binding proteins. The panels display structures of the isolated EF-hand motif (1CLL) [9], EF-hand domain from CaM (1CLL) [9], intact Calmodulin (1CLL) [9] and S100B (1MHO) [41]. Different colors are used to distinguish between the pairs of EF-hands, the two subunits of S100B and the two EF-hand domains of CaM. All figures in the paper were generated in MOLMOL [33].

and the diversity in EF-hand protein function. The well-studied S100 and calmodulin family proteins provide a clear example of differences in the organization of two domains (Fig. 1).

The interaction between any two proteins ultimately comes down to the characteristics of their respective binding surfaces. Consequently, understanding the Ca^{2+}-dependent regulation of signaling pathways by EF-hand proteins requires atomic level knowledge of the structures of the complex with the target protein. Over the past ~10 years, the number of three-dimensional structures of complexes has increased at a very rapid pace. Considerable progress has been made in elucidating guiding principles of target interactions, particularly for calmodulin [13,71]. This review focuses on the structural basis for the selectivity of EF-hand proteins for different targets. Calmodulin and S100 family proteins are used to demonstrate basic principles. We highlight the differences in the overall architecture of these proteins and the malleability of the target binding surface as part of a combined mechanism to generate target selectivity amongst members of these two families.

2. Calmodulin family proteins

The calmodulin family is very large and has been extensively characterized amongst the EF-hand Ca^{2+} sensor proteins [7,8,67]. Much of what is known today in regards to the mode of action of EF-hand Ca^{2+} sensors is based on the structural analysis of the parent protein, calmodulin (CaM) and its interactions with various targets. Hence, CaM

is used as an example for discussing the principal determinants of target selectivity.

2.1. Structural overview of CaM

Calmodulin, the archetypal EF-hand calcium sensor, consists of two canonical EF-hand domains tethered by a flexible helical linker (Fig. 1). The two domains share high overall sequence homology (75%), as well as structural similarity in the presence [9,11,21] and absence of Ca^{2+} ions [22,35,74]. However, the differences in the two domains are significant enough to result in distinct biochemical properties. For example, the calcium affinity of CaM-C ($K_d \sim 10^{-6}$ M) is 10-fold stronger than CaM-N ($K_d \sim 10^{-5}$ M) [23,37].

The consequences of structural changes within each domain induced by the chelation of Ca^{2+} ions by side-chain (and one backbone) oxygen atoms in each of the two EF-hands creates a large solvent exposed hydrophobic surface that interacts with targets. The presence of four methionines in each binding pocket is critical, as the unique physiochemical properties of this side chain helps stabilize the open conformation and provides a high degree of structural malleablility for adapting to the many different targets regulated by calmodulin [13,48,49,73]. Remarkably, the electrostatic character of the binding surface of the two domains is significantly different. CaM-N possesses acidic patches that are partly neutralized by inclusion of basic residues in and around the binding pocket. In contrast, the binding pocket of CaM-C is predominantly acidic. In addition to the fundamental differences between the two domains, the flexible linker between them enables a great

deal of variability in the relative orientations of the two domains [2], which is a critical factor in the ability of CaM to interact with such a large number of targets.

Structural characterization of CaM by NMR has also revealed significant differences in the conformational dynamics of the N- and C-domains, which appears to influence their independent functions. The NMR structure of apo-CaM suggests a closed conformation for the N-domain, with the hydrophobic core largely sequestered from solvent [11,35,74]. In contrast, there is clear evidence for dynamic exchange between a closed and one or more partially open conformations in the C-domain [19,20,35,39]. A dynamic equilibrium involving conformations with a partially exposed hydrophobic core has been proposed to play a critical role in CaM function, by enabling pre-association of CaM with targets before a calcium signal is generated [10,62]. In this hypothesis, canonical signal transduction by CaM corresponds to transition from a pre-associated state mediated by the C-domain to full engagement by both domains. This mechanism solves the timing problem caused by the need for activated CaM to diffuse through the crowded intracellular environment to find the target once a Ca^{2+} signal is generated. An excellent illustration of this idea is provided by the recently published structure of apo-CaM complexed with the gating domain of SK Potassium Channel [57]. In this structure the target is pre-associated exclusively with a closed conformation of the C-domain of apo-CaM using a limited set of intermolecular contacts. There are no detectable interactions with the N-domain of CaM in the absence of calcium. This is unlike the situation at elevated levels of calcium when both domains of CaM are engaged in interactions with the intracellular gating domain [58]. A "semi-open" conformation of CaM-C has

also been observed for Ca^{2+}-free EF-hand domains of myosin light chains (essential and regulatory) in complex with an IQ-motif from the heavy chain [28].

2.2. Target interactions

Calmodulin mediates a myriad of cellular processes such as cell division and differentiation, gene transcription, ion transport by channels, membrane fusion and muscle contraction [7,8,67]. The list of targets regulated by CaM is extensive and includes adenylyl cyclases, CaM kinases, ion channels and various cell surface receptors. This diversity is matched by variation in the way CaM interacts with its many targets [76].

2.2.1. Canonical wrap-around mode

The best known mode of interaction of CaM with target involves the binding of both CaM domains to a single binding region, as exemplified by the structure of CaM in complex with the binding peptide of smMLCK (Fig. 2) [44]. This binding mode is observed for all peptides derived from the autoinhibitory domains of myosin light chain kinases and other CaM-dependent kinases [43,50]. The peptide adopts a helical conformation and is enveloped by the two domains, hence the term "wrap-around". The intrinsic flexibility in the linker enables a large reduction in the separation of the two domains from the 50 Å in the crystal structure of free CaM to less than 10 Å when wrapped around the target peptide. The complex is anchored by bulky hydrophobic side chains and further stabilized by basic side chains from the peptide that interact favorably with acidic residues in and around the binding site. The helical conformation of the peptide facilitates the correct
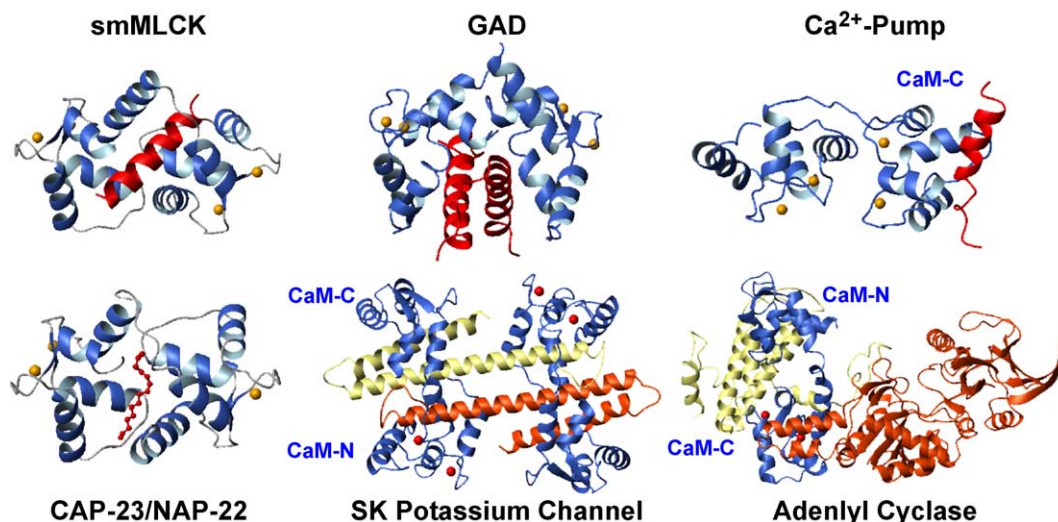
Fig. 2. Structures of calmodulin in complex with targets showing different binding modes. The panels show the wrap-around CaM binding modes for complexes containing peptides derived from smooth muscle Myosin Light Chain Kinase (smMLCK, 1CDL) [44] and the N-terminal myristoylation group from CAP-23/NAP-22 (1L7Z) [40]; CaM-induced dimerization of the N-terminal domain of Petunia Glutamate Decarboxylase (1NWD) [72] and cytoplasmic CaMBD from SK Potassium Channel (1G4Y) [58]; extended binding mode for the complex with the Ca^{2+}-Pump (1CFF) [18] and anthrax edema factor (1K93) [17]. In each of the figures the ribbon representation of CaM is colored blue and the peptide is a shade of red or yellow. The calcium atoms are represented as orange colored spheres.

spacing of key anchor points at the surface of the target, which is important for interactions with both domains.

Based on the spacing of bulky hydrophobic groups that serve as anchor points, the CaM binding motifs have been classified into four categories described as 1–5–10, 1–8–14, 1–16 and IQ motifs, respectively [13,54,71] (Fig. 3). Typically the first position is bound to the C-terminal domain of CaM while the residues at position 10 or 14 interact with CaM-N. In the 1–16 motif, the alignment of the peptide is reversed and the residue at the first position interacts with CaM-N instead. The orientation of the peptide is determined by steric factors and the presence of a basic patch near the N-terminus (classes 1–5–10, 1–8–14) or C-terminus (class 1–16) of the target sequence, respectively [50].

An interesting variation on the wrap-around binding mode involves CAP-23/NAP-22, a brain specific PKC substrate involved in axon regeneration that lacks a consensus CaM-binding motif [40]. An intriguing hint to the unique nature of this interaction was that the binding of CAP-23 is dependent on myristoylation at its N-terminus. Insight into the structural mechanism was obtained from a recent X-ray structure, which showed the myristoyl group in the same site as target peptides that bind in the wrap-around mode (Fig. 2). In this case, the fatty acid was enveloped within the groove between the CaM-N and CaM-C domains.

The wrap-around mode was initially assumed to be the predominant structural mechanism for CaM binding to its targets. However, the publication of several structures in the past few years in which the target has posttranslational modifications or contains non-consensus CaM binding motifs has greatly expanded the CaM binding repertoire. These structures confirm the important roles of conformational flexibility within each CaM domain and variability in the relative positioning of the two domains for adapting to different targets. For example, the intrinsic malleability of CaM enables it to compensate for the absence of what were assumed to be critically positioned anchor residues. Some of these structures have also presented the first evidence supporting a role for CaM dimerization in biological function.

2.2.2. Extended binding modes

Although the wrap-around binding mode of target binding to CaM is the best characterized, many members of the CaM family, including CaM itself, are distinguished by binding in an extended mode such that their domains interact with different regions of the target. This mode of binding is best characterized for troponin C (TnC), which has two EF-hand domains similar to CaM [26,60,59,27]. The recently published high-resolution X-ray structure of the core domain of the cardiac troponin complex has provided details of interactions involving TnC, the Ca^{2+}-binding subunit in the ternary complex containing TnI, the inhibitory subunit and TnT, the tropomyosin-binding sub-unit in muscle contraction [64]. The mode of action of the complex involves the sequential calcium triggered conformational response of the two domains from TnC. The C-domain of TnC has such high Ca^{2+} affinity, that it is Ca^{2+}-loaded even in the resting state and hence constitutively bound to a portion of TnI. At elevated levels of calcium, the regulatory N-domain of TnC binds Ca^{2+} ions and undergoes a conformational change, which in turn engages a different part of TnI. TnC is a classic example where the calcium response is generated by coupling the functionality of the CaM-like domains in the extended mode.

Similarly, the structure of CaM in complex with a fragment of the plasma membrane Ca^{2+} pump revealed that at basal levels of Ca^{2+}, the channel is partially activated by interactions of CaM-C with an auto-inhibitory domain that releases the active site of the ATP-binding domain from the Ca^{2+} Pump [18] (Fig. 2). However, maximum pumping efficiency is achieved only after Ca^{2+} levels are elevated and CaM-N is activated and engaged with a different region of the ATP-binding domain.

The most direct example of CaM binding in the extended mode comes from the structure of the complex with the anthrax exotoxin, Edema Factor [17] (Fig. 2). Although the Edema factor carries a 1–8–14 CaM consensus sequence, the structure of the intact catalytic domain (EFCD) complexed with CaM revealed an extended mode of binding. The EFCD surrounds and makes extensive contacts with both domains of CaM, which stabilizes the open conformation of the active state of EFCD. In the structure, CaM-C is Ca^{2+}-loaded and occupies an open conformation that engages the EFCD in a conventional manner through an exposed hydrophobic patch. In contrast, CaM-N is Ca^{2+}-free and occupies a closed conformation, which causes it to contact the EFCD through it exterior surface. This binding at the surface of the domain is reminiscent of a report on CaM interaction with a MARCKS peptide, in which interactions were proposed to be mediated by CaM-N in a closed or partially open conformation [70]. The differences in how the two domains interact with the EFCD is consistent with NMR studies that reveal the Ca^{2+} affinity of CaM-N to be much weaker than that of CaM-C in the presence of the EFCD [66].

The extended binding mode is also used for targets that bind to apo-CaM. Many of these targets, such as

Canonical CaM binding Motifs

1-5-10	CaM-C	CaM-N
CaMKI	I K K N F A **K** S **K** W K Q A **F** N A T A **V** V R H - - M R K - - - - - - - -
CaMKII	L K K F N A **R R** W K K L K G A **I** L T T M **L** A T R N - F S - - - - - - - - -

1-8-14		
skMLCK	- - - - - - **K R R** W K K N F I A **V** S A A N R - - **F** K K I S S S G A L
smMLCK	- - - - - A **R R K** W Q K T G H A **V** R A I G R - - **L** S S - - - - - - - -

1-16	CaM-N	CaM-C
CaMKK	- - - V K L I P S **W** T T V I L V K S M L **R K R** S **F** G N P F - - - - -

Extended CaM binding Motifs

	CaM-C	CaM-N
CaPump	L **R R** - - G Q I L **W** F R G **L** N R I Q T Q I K - - - - - - - - - - - - -
Kar1p	**K K R** E L I E S K **W** H R L **L** F H D K K M V E K K L E S L - - - - - -

Fig. 3. Comparison of the amino acid sequences of CaM binding motifs. For canonical CaM binding motifs the Trp side chain is colored green and remaining hydrophobic anchor points are colored red. Basic residues are indicated in blue.

neuromodulin and neurogranin, interact through the IQ motif [32], which contains the consensus sequence IQxxxRGxxxR (reviewed in Ref. [1]). Some IQ motifs bind to CaM in both the absence and presence of Ca^{2+} (e.g., insulin receptor substrate-1, myosin) and in some cases the IQ motif is combined with other CaM binding sequences (e.g., the cardiac L-type voltage gated calcium and the cardiac voltage gated sodium channels [65,69]). As noted above, no structural details for CaM bound to IQ motif targets are available. However, a model for CaM/IQ motif interactions has been proposed in which the IQxxxR segment interacts strongly with a semi-open CaM-C, with support from limited contacts between the GxxxR segment and the surface of CaM-N in a standard closed conformation [14,28].

2.2.3. Dimerization binding modes

Several examples of CaM-induced dimerization of the target have been reported. The structure of CaM in complex with the gating domain of the Ca^{2+}-activated K^+ membrane channel is an interesting example where dimerization of the channel has been achieved by CaM adopting the extended configuration of the EF-hand domains (Fig. 2). At low levels of Ca^{2+}, the pore of the channel is closed with apo-CaM bound to the intracellular gating domain of the channel via the C-domain [57]. At elevated levels of Ca^{2+}, CaM-C still retains its apo-state and is associated with the target.

However, under these conditions CaM-N is loaded with the Ca^{2+} ions and engages a regulatory domain from a second molecule leading to the opening of the channel [58].

CaM-induced dimerization plays a vital role in triggering the activity of glutamate decarboxylase (GAD), an enzyme required for normal plant growth through regulation of γ-aminobutyrate and glutamate metabolism. The structure of the complex revealed a novel arrangement of CaM wrapped around two GAD peptides, which interact with each other and also make contact with one of the domains from CaM [72] (Fig. 2).

A third type of dimerization binding mode was found for CaM interacting with the N-terminal basic region from the homodimeric basic-helix-loop-helix (bHLH) motif-containing transcription factor SEF2-1/E2-2. The family of bHLH transcription factors is activated by a common CaM-dependent mechanism [12], where two molecules of CaM interact in an extended mode with the SEF2-1 dimer [36].

2.3. Target selectivity within the CaM family

The observation of fundamentally different modes of interaction and binding to different conformations of CaM domains implies that elucidating the structural basis for target selectivity will be difficult. Moreover, the disclosure of many of the extended and dimerization mode structures is fairly recent. The one area where excellent progress has
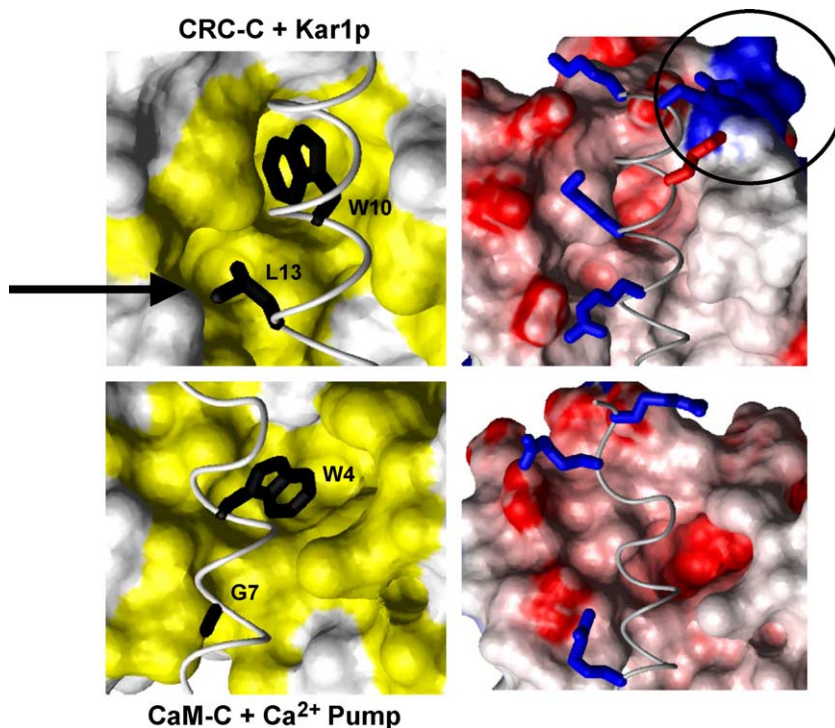
Fig. 4. Fine-tuning of the binding surface in CaM family of proteins. Molecular surface representations of the complex of CRC-C bound to a peptide from Kar1p (1OQP) [29] and CaM-C bound to the peptide from Ca^{2+}-Pump (1CFF) [18]. For each complex two separate representations of the hydrophobic surface (yellow) and electrostatics (red/blue) were generated in MOLMOL. The arrow (size of hydrophobic pocket) and circle (basic patch) indicate points of difference between the binding surfaces that impart target specificity. Hydrophobic (black) and charged side chains (basic/blue and acidic/red) from the target peptide are shown in neon representation.

been made is the detailed analysis of structures of target complexes with CaM in the wrap-around binding mode. As noted above, the classification based on the spacing of hydrophobic anchors is well established [71].

At the atomic level, the differences in sequence at key hydrophobic sites lead to a specific pattern of knobs and holes for each protein. This pattern at the surface needs to be matched by the target in order to attain high affinity binding. The flexible nature of the Met side chains at the binding surface has been discussed as a critical factor in facilitating this surface complementarity between CaM and the target [13,49,73]. Here, we provide one example of fine-tuning of hydrophobic anchors, by comparing the binding surfaces of CaM and a very close homolog, Centrin. This comparison is made using the structure of the complex of CaM-C with a peptide from the plasma membrane Ca^{2+}-pump (PMCP) [18], which is strikingly similar to the structure of the complex of the C-terminal domain of centrin (CRC-C) with a peptide fragment from the centrosomal protein Kar1p [29].

Centrin is an essential component of the centrosome and other microtubule organizing centers with substantial sequence homology to calmodulin. There is striking similarity between CRC-C and CaM-C (82% homology, 48% identity), which extends even to the positioning of side chains in the hydrophobic core of the protein [29]. Despite this rather extensive similarity, a limited number of differences at critical places in the sequence have significant effects on their respective binding surfaces, sufficient to substantially alter the nature of target interactions. The surface representations in Fig. 4 reveal the importance of these differences in creating complementary shape and charge features to match the target sequence. The Trp side chain is the primary hydrophobic anchor in Kar1p, and is buried in a deep pocket, not unlike the peptide from the calcium pump. However, the replacement of two Met residues with smaller Ala and Ile residues in CRC-C enlarges the holes in the CRC-C binding surface, thereby enabling interaction with a target that has larger hydrophobic side chains (Leu in Kar1p instead of a Gly in PMCP) than a target optimized to fit in the CaM-C binding site [29]. The presence of a basic patch in CRC-C near the N-terminus of the target peptide is yet another factor that imparts specificity to CRC interactions with Kar1p relative to CaM binding of PMCP. Similar conclusions on the origin of target selectivity can be made based on atomic level comparisons of the CaM-C binding site containing other canonical binding motifs [29].

3. S100 family proteins

The ~23 members of the S100 proteins identified in humans constitute a major family of EF-hand calcium sensor proteins. S100s are associated with multiple targets that promote cell growth and differentiation, cell cycle regulation, transcription and cell surface receptor activities

(reviewed in Refs. [16,25]). The calcium signaling activities of these proteins are tailored in part by a distinct pattern of subcellular localization and tissue specific expression [75]. The detection of abnormal S100 gene expression in different disease states including chronic inflammation, tumor progression, cardiomyopathy, Alzheimer's and psoriasis and has fueled further clinical interest in this multigenic family of proteins [25].

3.1. Structural overview

The basic structural and functional unit of the S100 proteins is a symmetric dimer comprised of two EF-hand domains, which are organized into an eight-helix bundle [51]. While all S100 proteins form homodimers, some are known to preferentially form heterodimers with other S100 isoforms, e.g., S100A8/S100A9 and S100A1/S100B (reviewed in Ref. [15]). At the primary sequence level, S100 proteins are distinguished from all other EF-hand proteins by a unique 14-residue S100-specific Ca^{2+}-binding loop in the N-terminal EF-hand of each domain [63].

The conformational change triggered by calcium binding to S100 proteins is vastly different in the two EF-hands. The N-terminal EF-hand undergoes a relatively small change in conformation, whereas the change in the C-terminal EF-hand is much larger, although not as dramatic as in CaM [38,56]. As a consequence, the depth of the hydrophobic pocket created upon ion binding is very shallow in S100 proteins compared to CaM. Another characteristic feature of S100 proteins is that they have two symmetrically disposed target binding sites per dimeric globular domain. This implies that the structural mode of target interaction is fundamentally different between the S100 and CaM families.

One additional distinctive characteristic of S100 proteins is their binding of additional divalent metal ions such as Zn^{2+} and Cu^{2+} at sites that are remote from the Ca^{2+} binding sites. Binding of these "other" divalent ions with high affinity plays a role in chemotactic activity and the homeostasis of toxic metal ions (reviewed in Ref. [24]). The structures of zinc-bound S100A7 [5], S100B [68] and a copper-bound S100A12 [46] revealed symmetrically disposed metal binding sites near the dimer interface. However, the effect of these ions on the structure is very limited, with the most significant changes due to rearrangement of side chains that chelate the ions. Most importantly, it has not been firmly established how the binding of divalent ions supports or antagonizes interaction with targets.

A final point of interest regarding S100 protein structure is their potential to form higher order organizational states ranging from disulfide cross-linked forms to higher order oligomers [31,45]. Although there is no structural information available, oxidation of S100B into a disulfide cross-linked form has been shown to promote neurotropic effects in the extracellular space [3], and higher order oligomerization states of S100A12 homodimers are proposed to mediate function of the multimeric cell surface receptor for advanced

glycation end products [45,47]. Interestingly, the structure of a S100A12 hexamer was determined at millimolar concentrations of Ca^{2+} found only in the extracellular milieu. S100A9 was crystallized in an octameric state and the structure revealed an entirely different mode of oligomerization using CHAPS molecule to bridge the dimers via interactions with the S100 target-binding site [31]. The functional relevance of S100 protein oligomerization remains an area of intense interest.

3.2. Target interactions

S100 proteins activate multiple targets that are involved in regulating a diverse range of cellular activities, including cell growth and differentiation, cell cycle regulation, transcription and cell surface signaling (reviewed in Refs. [16,25]). The calcium signaling activities of S100 proteins are tailored in part by a distinct pattern of subcellular localization and tissue-specific expression [75]. These characteristics suggest that S100 proteins are used in parallel to the traditional EF-hand Ca^{2+} sensors, to provide very tightly regulated and specific calcium signaling pathways.

Several high-resolution structures are available for S100-target peptide complexes (Fig. 5). Because the hydrophobic groove in S100 proteins is rather shallow compared to CaM family proteins, the S100 target peptides do not insert themselves deeply into a binding pocket as do CaM binding proteins, but rather splay across the wider S100 protein binding surface [4]. As noted above the dimeric architecture also clearly distinguishes S100 complexes from CaM complexes, in particular, because there are two identical binding sites per S100 dimer. Theoretically, S100 proteins could serve to link different target molecules, although there is no known example reported to date.

The typical S100-binding region from the target is a short helical segment rich in basic and hydrophobic residues. These characteristics complement the hydrophobic and acidic S100 target-binding site formed from side chains in the two helices of the C-terminal EF-hand and the linker connecting the two EF-hands (Figs. 5 and 6). Unlike CaM targets, the S100 target sequences display greater variation in the content and position of anchor residues. Some

attempts have been made to extract a consensus sequence '+OXO*XOO' (+=basic, O=hydrophobic, *=hydrophilic, X=variable) for S100B targets but accurate predictions based on this motif have not been made for all the known biological targets of S100B [42].

Compared to CaM, the binding surface of S100 proteins is flatter, more open and characterized by a variability of sequence that suggests a different structural mechanism for modulating the shape of the binding surface [4]. Of note, there is a significant difference in the ratio of polar and hydrophobic residues at the binding surface of the CaM and S100 protein families (Fig. 7). Thus, the specificity appears not to be regulated as strictly by the spacing of bulky hydrophobic groups as in the case of the CaM family. Other factors like the ability to complement the polar and charged groups at the surface of the S100s appear to be of greater importance in target specificity [4]. The validity of these hypotheses remains to be tested as more high-resolution structural information on S100 proteins in complex with binding partners become available.

3.3. Target selectivity within the S100 protein family

There is considerably less information available on S100-target interactions as compared to CaM family proteins. However, since the 3D structures of S100 proteins are so similar yet target binding is known to be specific, it is already possible to analyze some of the factors contributing to target selectivity. In fact, considerable progress has been made in understanding how the sequences of S100 proteins are fine-tuned to generate unique target binding surfaces.

The flatter binding surface of S100 proteins relative to CaM seems to result in a remarkable *lack* of uniformity in the orientation of the target with respect to the S100 protein (Fig. 5). Even more surprising is the observation that individual S100 proteins can bind to different targets in different ways [4]. For example, the structures of peptides from p53 [55], CapZ (not shown in figure) [30] and NDR kinase [4] in complex with S100B all have different orientations in the binding site (Fig. 5).

These structures show that different modes of binding are generated by variability in S100 sequence (15–65%

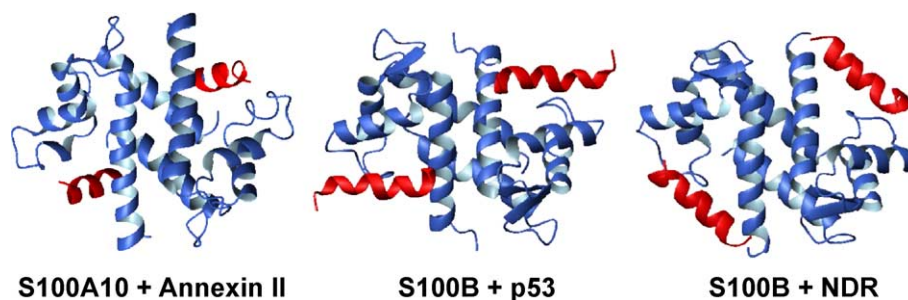**S100A10 + Annexin II   S100B + p53   S100B + NDR**

Fig. 5. Variation in the target binding modes of S100 family of proteins. The panels display the complexes of S100A10 bound to N-terminal Annexin II peptide (1BT6) [53] and S100B bound to peptides from p53 (1DT7) [55] and NDR kinase (1PSB) [4]. The ribbon representation of the protein is colored blue and the peptide is red.
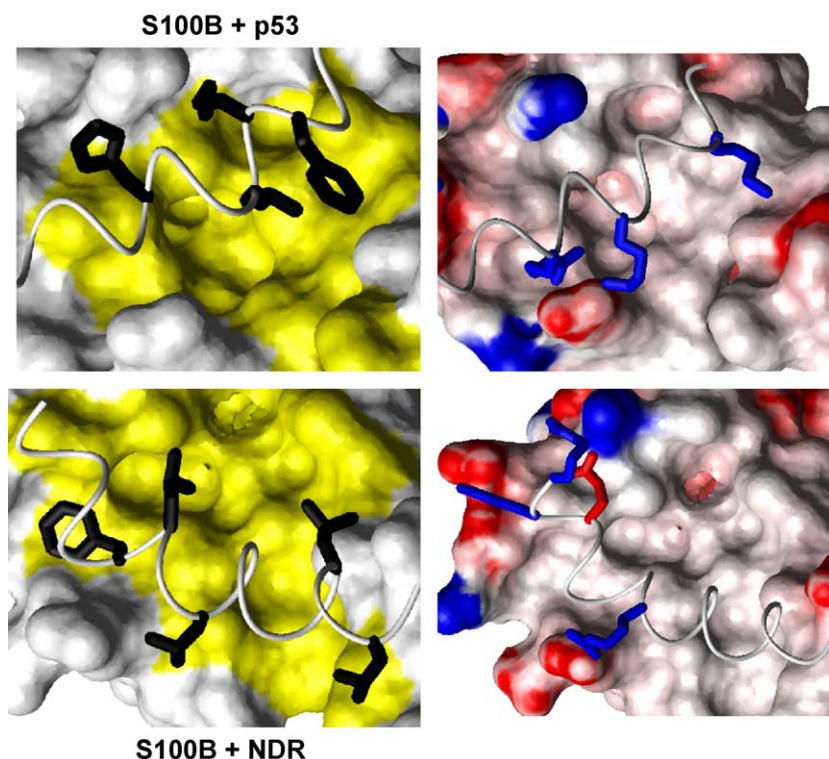
Fig. 6. Variability in the target-binding surface of S100 family of proteins. Molecular surface representations of the complex of S100B bound to peptides from p53 (1DT7) [55] and NDR kinase (1PSB) [4]. For each complex two separate representations of the hydrophobic surface (yellow) and electrostatics (red/blue) were generated in MOLMOL. Hydrophobic (black) and charged side chains (basic/blue and acidic/red) from the target peptide are shown in neon representation.

	A6 (1K9K)	A7 (3PSR)	A8 (1MR8)	A12 (1E8A)	B (1MHO)	P (1J55)	CaM-N (1CLL)	CaM-C (1CLL)
PHE	0.5	0.3	–	–	14.7	8.7	6.6	3.0
VAL	< 0.1	0.2	4.8	2.1	5.7	4.4	2.3	7.9
LEU	14.5	1.7	3.6	0.6	2.6	12.3	16.3	11.1
ILE	8.3	2.1	4.2	5.5	1.8	2.4	0.8	2.6
MET	--	0.8	8.0	–	3.7	7.5	18.0	18.1
CYS	1.3	0.3	0.1	–	0.7	1.8	--	–
ALA	15.0	7.6	2.6	7.0	5.2	8.7	4.6	9.7
% Hydrophobic	39.6	13.0	23.3	15.2	34.4	45.8	48.6	52.5
TRP	--	--	–	–	–	–	--	–
PRO	1.9	3.4	0.4	–	–	2.2	1.3	–
TYR	3.1	2.4	2.4	7.3	–	5.1	--	–
THR	5.2	4.2	1.0	5.6	2.9	2.4	--	6.8
SER	2.9	16.0	9.9	2.2	3.6	0.6	--	–
ASN	0.1	7.1	2.2	10.3	–	–	--	–
GLN	8.2	5.8	4.0	3.3	3.5	–	6.2	–
HIS	--	4.5	3.8	8.5	5.8	0.6	--	–
% Polar	21.4	43.7	23.7	37.2	15.8	10.9	7.5	6.8
ARG	6.6	4.3	3.2	–	–	–	--	–
LYS	9.9	20.3	26.9	27.4	14.2	19.4	6.5	–
% Basic	16.5	24.6	30.1	27.4	14.2	19.4	6.5	0.0
ASP	7.0	9.7	7.0	3.7	4.3	12.9	6.5	–
GLU	10.5	2.9	14.8	13.8	31.1	5.7	30.9	40.7
% Acidic	17.5	12.6	21.8	17.5	35.4	18.6	37.4	40.7
GLY	4.9	6.2	0.8	2.5	–	5.4	--	–
Total Area	2696	2975	2987	2841	3088	2972	1468	1665

Fig. 7. Comparison of the % accessible surface area (ASA) of residues at the target-binding site of each subunit of S100 proteins and the isolated EF-hand domains of calmodulin. The four-letter PDB codes are indicated for each structure used in the analysis. The ASA of individual residues (>10%) and surface properties such as hydrophobic, polar or charged (>30%) that contribute more than the cutoff are highlighted in color. The surface areas were calculated in Insight II (Accelrys) using the Lee and Richards algorithm.

homology), which alters the hydrophobic and polar character of the target binding sites adapted for a wide range of targets. Each target appears to use its particular hydrophobic and charged side chains to match a complementary surface of the S100 protein, maximizing energetically favorable contacts (Fig. 6). This conclusion also supports the observed variation in the binding modes for S100B, where a common interaction surface is employed to generate a slightly different set of contacts to match the properties of each target peptide [4].

The obvious differences in orientation of the target peptide with respect to the protein mask the fact that a number of key hydrophobic contacts to residues in the C-terminal EF-hand are conserved in each of these structures [4]. The biggest differences and hence the controlling factor appears to arise from contacts with the linker region, which contains a number of charged and polar residues. The functional importance of multiple binding modes available to S100 proteins has still to be established in relation to the intact target protein.

4. Conclusions

The basic mechanism for calcium driven signal transduction pathways controlled by CaM and S100 proteins relies on conformational change in the EF-hands, which exposes a hydrophobic surface for interacting with the target. However, understanding the structural origins of functional specificity requires deeper insight into the structural basis for interaction. The specificity of an EF-hand protein for a given target is derived from the physicochemical properties of the binding surface, as well as structural features like the flexible nature of the dumbbell structure of CaM and the oligomeric states of S100 proteins. Although the binding of targets to CaM and S100 proteins is driven largely by hydrophobic interactions, a comparison of the binding surfaces has revealed that the mechanism dictating target selectivity involves a complex mixture of factors.

The most critical obstacle to developing a complete understanding of the structural basis for calcium dependent signal transduction by EF-hand proteins is the dearth of structures with intact proteins or protein domains. The large majority of structures to date are of an EF-hand protein with a peptide fragment corresponding to the binding region from the target. While these structures reveal the requisite atomic level detail to understand the mode of action of the EF-hand protein, they do not provide insight into the means by which intact target proteins are activated. For instance, the X-ray structures of N-terminal peptides of annexin II bound to S100A10 [53] and of annexin I bound to S100A11 [52] provide only a partial glimpse of how these interactions promote membrane fusion. To achieve further understanding of this mechanism, it is necessary to determine structures of complexes in multiple states and utilize

spectroscopic approaches to characterize the transitions and dynamics between states. Excellent progress has been made in a few selected systems (e.g., the troponin complex [64]), and this set the stage for several exciting new breakthroughs that are anticipated in the next few years.

References

[1] M. Bahler, A. Rhoads, Calmodulin signaling via the IQ motif, FEBS Lett. 513 (2002) 107–113.

[2] G. Barbato, M. Ikura, L.E. Kay, R.W. Pastor, A. Bax, Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible, Biochemistry 31 (1992) 5269–5278.

[3] S.W. Barger, S.R. Wolchok, L.J. Van Eldik, Disulfide-linked S100 beta dimers and signal transduction, Biochim. Biophys. Acta 1160 (1992) 105–112.

[4] S. Bhattacharya, E. Large, C.W. Heizmann, B. Hemmings, W.J. Chazin, Structure of the Ca^{2+}/S100B/NDR kinase peptide complex: insights into S100 target specificity and activation of the kinase, Biochemistry 42 (2003) 14416–14426.

[5] D.E. Brodersen, J. Nyborg, M. Kjeldgaard, 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, Biochemistry 38 (1999) 1695–1704.

[6] R.D. Brokx, M.M. Lopez, H.J. Vogel, G.I. Makhatadze, Energetics of target peptide binding by calmodulin reveals different modes of binding, J. Biol. Chem. 276 (2001) 14083–14091.

[7] E. Carafoli, C. Klee, Calcium as a Cellular Regulator, Oxford University Press, New York, 1999.

[8] M.R. Celio, Guidebook to the Calcium-Binding Proteins, Oxford University Press, New York, 1996.

[9] R. Chattopadhyaya, W.E. Meador, A.R. Means, F.A. Quiocho, Calmodulin structure refined at 1.7 A resolution, J. Mol. Biol. 228 (1992) 1177–1192.

[10] D. Chin, A.R. Means, Calmodulin: a prototypical calcium sensor, Trends Cell Biol. 10 (2000) 322–328.

[11] J.J. Chou, S. Li, C.B. Klee, A. Bax, Solution structure of Ca(2+)-calmodulin reveals flexible hand-like properties of its domains, Nat. Struct. Biol. 8 (2001) 990–997.

[12] B. Corneliussen, M. Holm, Y. Waltersson, J. Onions, B. Hallberg, A. Thornell, T. Grundstrom, Calcium/calmodulin inhibition of basic-helix-loop-helix transcription factor domains, Nature 368 (1994) 760–764.

[13] A. Crivici, M. Ikura, Molecular and structural basis of target recognition by calmodulin, Annu. Rev. Biophys. Biomol. Struct. 24 (1995) 85–116.

[14] Y. Cui, J. Wen, K. Hung Sze, D. Man, D. Lin, M. Liu, G. Zhu, Interaction between calcium-free calmodulin and IQ motif of neurogranin studied by nuclear magnetic resonance spectroscopy, Anal. Biochem. 315 (2003) 175–182.

[15] J.C. Deloulme, B.J. Gentil, J. Baudier, Monitoring of S100 homodimerization and heterodimeric interactions by the yeast two-hybrid system, Microsc. Res. Tech. 60 (2003) 560–568.

[16] R. Donato, S100: a multigenic family of calcium-modulated proteins of the EF-hand type with intracellular and extracellular functional roles, Int. J. Biochem. Cell Biol. 33 (2001) 637–668.

[17] C.L. Drum, S.Z. Yan, J. Bard, Y.Q. Shen, D. Lu, S. Soelaiman, Z. Grabarek, A. Bohm, W.J. Tang, Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin, Nature 415 (2002) 396–402.

[18] B. Elshorst, M. Hennig, H. Forsterling, A. Diener, M. Maurer, P. Schulte, H. Schwalbe, C. Griesinger, J. Krebs, H. Schmid, T. Vorherr, E. Carafoli, NMR solution structure of a complex of

calmodulin with a binding peptide of the Ca^{2+} pump, Biochemistry 38 (1999) 12320–12332.

[19] J. Evenas, A. Malmendal, M. Akke, Dynamics of the transition between open and closed conformations in a calmodulin C-terminal domain mutant, Structure 9 (2001) 185–195.

[20] J. Evenas, A. Malmendal, E. Thulin, G. Carlstrom, S. Forsen, Ca^{2+} binding and conformational changes in a calmodulin domain, Biochemistry 37 (1998) 13744–13754.

[21] J.L. Fallon, F.A. Quiocho, A closed compact structure of native Ca(2+)-calmodulin, Structure 11 (2003) 1303–1307.

[22] B.E. Finn, J. Evenas, T. Drakenberg, J.P. Waltho, E. Thulin, S. Forsen, Calcium-induced structural changes and domain autonomy in calmodulin, Nat. Struct. Biol. 2 (1995) 777–783.

[23] S. Forsen, S. Linse, T. Drakenberg, J. Kordel, M. Akke, P. Sellers, C. Johansson, E. Thulin, I. Andersson, P. Brodin, et al., Ca^{2+} binding in proteins of the calmodulin superfamily: cooperativity, electrostatic contributions and molecular mechanisms, Ciba Found. Symp. 161 (1991) 222–236.

[24] C.W. Heizmann, J.A. Cox, New perspectives on S100 proteins: a multi-functional Ca(2+)-, Zn(2+)- and Cu(2+)-binding protein family, BioMetals 11 (1998) 383–397.

[25] C.W. Heizmann, G. Fritz, B.W. Schafer, S100 proteins: structure, function and pathology, Front. Biosci. 7 (2002) 1356–1368.

[26] O. Herzberg, M.N. James, Structure of the calcium regulatory muscle protein troponin-C at 2.8 A resolution, Nature 313 (1985) 653–659.

[27] A. Houdusse, M.L. Love, R. Dominguez, Z. Grabarek, C. Cohen, Structures of four Ca^{2+}-bound troponin C at 2.0 A resolution: further insights into the Ca^{2+}-switch in the calmodulin superfamily, Structure 5 (1997) 1695–1711.

[28] A. Houdusse, M. Silver, C. Cohen, A model of Ca(2+)-free calmodulin binding to unconventional myosins reveals how calmodulin acts as a regulatory switch, Structure 4 (1996) 1475–1490.

[29] H. Hu, W.J. Chazin, Unique features in the C-terminal domain provide caltractin with target specificity, J. Mol. Biol. 330 (2003) 473–484.

[30] K.G. Inman, R. Yang, R.R. Rustandi, K.E. Miller, D.M. Baldisseri, D.J. Weber, Solution NMR structure of S100B bound to the high-affinity target peptide TRTK-12, J. Mol. Biol. 24 (2002) 1003–1014.

[31] H. Itou, M. Yao, I. Fujita, N. Watanabe, M. Suzuki, J. Nishihira, I. Tanaka, The crystal structure of human MRP14 (S100A9), a Ca(2+)-dependent regulator protein in inflammatory process, J. Mol. Biol. 316 (2002) 265–276.

[32] L.A. Jurado, P.S. Chockalingam, H.W. Jarrett, Apocalmodulin, Physiol. Rev. 79 (1999) 661–682.

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

[34] R.H. Kretsinger, C.E. Nockolds, Carp muscle calcium-binding protein: II. Structure determination and general description, J. Biol. Chem. 248 (1973) 3313–3326.

[35] H. Kuboniwa, N. Tjandra, S. Grzesiek, H. Ren, C.B. Klee, A. Bax, Solution structure of calcium-free calmodulin, Nat. Struct. Biol. 2 (1995) 768–776.

[36] G. Larsson, J. Schleucher, J. Onions, S. Hermann, T. Grundstrom, S.S. Wijmenga, A novel target recognition revealed by calmodulin in complex with the basic helix-loop-helix transcription factor SEF2-1/E2-2, Protein Sci. 10 (2001) 169–186.

[37] S. Linse, S. Forsen, Determinants that govern high-affinity calcium binding, Adv. Second Messenger Phosphoprot. Res. 30 (1995) 89–151.

[38] L. Maler, M. Sastry, W.J. Chazin, A structural basis for S100 protein specificity derived from comparative analysis of apo and Ca(2+)-calcyclin, J. Mol. Biol. 317 (2002) 279–290.

[39] A. Malmendal, J. Evenas, S. Forsen, M. Akke, Structural dynamics in the C-terminal domain of calmodulin at low calcium levels, J. Mol. Biol. 293 (1999) 883–899.

[40] M. Matsubara, T. Nakatsu, H. Kato, H. Taniguchi, Crystal structure of a myristoylated CAP-23/NAP-22 N-terminal domain complexed with Ca(2+)/calmodulin, EMBO J. 23 (2004) 712–718.

[41] H. Matsumura, T. Shiba, T. Inoue, S. Harada, Y. Kai, A novel mode of target recognition suggested by the 2.0 A structure of holo S100B from bovine brain, Structure 6 (1998) 233–241.

[42] K.A. McClintock, G.S. Shaw, A logical sequence search for S100B target proteins, Protein Sci. 9 (2000) 2043–2046.

[43] W.E. Meador, A.R. Means, F.A. Quiocho, Modulation of calmodulin plasticity in molecular recognition on the basis of X-ray structures, Science 262 (1993) 1718–1721.

[44] W.E. Meador, A.R. Means, F.A. Quiocho, Target enzyme recognition by calmodulin: 2.4 A structure of a calmodulin–peptide complex, Science 257 (1992) 1251–1255.

[45] O.V. Moroz, A.A. Antson, E.J. Dodson, H.J. Burrell, S.J. Grist, R.M. Lloyd, N.J. Maitland, G.G. Dodson, K.S. Wilson, E. Lukanidin, I.B. Bronstein, The structure of S100A12 in a hexameric form and its proposed role in receptor signalling, Acta Crystallogr., D Biol. Crystallogr. 58 (2002) 407–413.

[46] O.V. Moroz, A.A. Antson, S.J. Grist, N.J. Maitland, G.G. Dodson, K.S. Wilson, E. Lukanidin, I.B. Bronstein, Structure of the human S100A12–copper complex: implications for host–parasite defence, Acta Crystallogr. D59 (2003) 859–867.

[47] O.V. Moroz, G.G. Dodson, K.S. Wilson, E. Lukanidin, I.B. Bronstein, Multiple structural states of S100A12: A key to its functional diversity, Microsc. Res. Tech. 60 (2003) 581–592.

[48] M.R. Nelson, W.J. Chazin, Structures of EF-hand Ca(2+)-binding proteins: diversity in the organization, packing and response to Ca^{2+} binding, BioMetals 11 (1998) 297–318.

[49] K.T. O'Neil, W.F. DeGrado, How calmodulin binds its targets: sequence independent recognition of amphiphilic alpha-helices, Trends Biochem. Sci. 15 (1990) 59–64.

[50] M. Osawa, H. Tokumitsu, M.B. Swindells, H. Kurihara, M. Orita, T. Shibanuma, T. Furuya, M. Ikura, A novel target recognition revealed by calmodulin in complex with Ca^{2+}-calmodulin-dependent kinase kinase, Nat. Struct. Biol. 6 (1999) 819–824.

[51] B.C. Potts, J. Smith, M. Akke, T.J. Macke, K. Okazaki, H. Hidaka, D.A. Case, W.J. Chazin, The structure of calcyclin reveals a novel homodimeric fold for S100 Ca(2+)-binding proteins, Nat. Struct. Biol. 2 (1995) 790–796.

[52] S. Rety, D. Osterloh, J.P. Arie, S. Tabaries, J. Seeman, F. Russo-Marie, V. Gerke, A. Lewit-Bentley, Structural basis of the Ca(2+)-dependent association between S100C (S100A11) and its target, the N-terminal part of annexin I, Struct. Fold Des. 8 (2000) 175–184.

[53] S. Rety, J. Sopkova, M. Renouard, D. Osterloh, V. Gerke, S. Tabaries, F. Russo-Marie, A. Lewit-Bentley, The crystal structure of a complex of p11 with the annexin II N-terminal peptide, Nat. Struct. Biol. 6 (1999) 89–95.

[54] A.R. Rhoads, F. Friedberg, Sequence motifs for calmodulin recognition, FASEB J. 11 (1997) 331–340.

[55] R.R. Rustandi, D.M. Baldisseri, D.J. Weber, Structure of the negative regulatory domain of p53 bound to S100B(betabeta), Nat. Struct. Biol. 7 (2000) 570–574.

[56] M. Sastry, R.R. Ketchem, O. Crescenzi, C. Weber, M.J. Lubienski, H. Hidaka, W.J. Chazin, The three-dimensional structure of Ca(2+)-bound calcyclin: implications for Ca(2+)-signal transduction by S100 proteins, Structure 6 (1998) 223–231.

[57] M.A. Schumacher, M. Crum, M.C. Miller, Crystal structures of apocalmodulin and an apocalmodulin/SK potassium channel gating domain complex, Structure 12 (2004) 849–860.

[58] M.A. Schumacher, A.F. Rivard, H.P. Bachinger, J.P. Adelman, Structure of the gating domain of a Ca^{2+}-activated K+ channel complexed with Ca^{2+}/calmodulin, Nature 410 (2001) 1120–1124.

[59] S.K. Sia, M.X. Li, L. Spyracopoulos, S.M. Gagne, W. Liu, J.A. Putkey, B.D. Sykes, Structure of cardiac muscle troponin C unexpectedly reveals a closed regulatory domain, J. Biol. Chem. 272 (1997) 18216–18221.

[60] C.M. Slupsky, B.D. Sykes, NMR solution structure of calcium-saturated skeletal muscle troponin C, Biochemistry 34 (1995) 15953–15964.

[61] N.C. Strynadka, M.N. James, Crystal structures of the helix-loop-helix calcium-binding proteins, Annu. Rev. Biochem. 58 (1989) 951–998.

[62] M.B. Swindells, M. Ikura, Pre-formation of the semi-open conformation by the apo-calmodulin C-terminal domain and implications binding IQ-motifs, Nat. Struct. Biol. 3 (1996) 501–504.

[63] D.M. Szebenyi, K. Moffat, The refined structure of vitamin D-dependent calcium-binding protein from bovine intestine. Molecular details, ion binding, and implications for the structure of other calcium-binding proteins, J. Biol. Chem. 261 (1986) 8761–8777.

[64] S. Takeda, A. Yamashita, K. Maeda, Y. Maeda, Structure of the core domain of human cardiac troponin in the Ca(2+)-saturated form, Nature 424 (2003) 35–41.

[65] W. Tang, D.B. Halling, D.J. Black, P. Pate, J.Z. Zhang, S. Pedersen, R.A. Altschuld, S.L. Hamilton, Apocalmodulin and Ca^{2+} calmodulin-binding sites on the CaV1.2 channel, Biophys. J. 85 (2003) 1538–1547.

[66] T.S. Ulmer, S. Soelaiman, S. Li, C.B. Klee, W.J. Tang, A. Bax, Calcium dependence of the interaction between calmodulin and anthrax edema factor, J. Biol. Chem. 278 (2003) 29261–29266.

[67] L. Van Eldik, D.M. Watterson, Calmodulin and Signal Transduction, Academic Press, 1998.

[68] P.T. Wilder, D.M. Baldisseri, R. Udan, K.M. Vallely, D.J. Weber, Location of the Zn(2+)-binding site on S100B as determined by NMR spectroscopy and site-directed mutagenesis, Biochemistry 42 (2003) 13410–13421.

[69] T.L. Wingo, V.N. Shah, M.E. Anderson, T.P. Lybrand, W.J. Chazin, J.R. Balser, An EF-hand in the sodium channel couples intracellular calcium to cardiac excitability, Nat. Struct. Mol. Biol. 11 (2004) 219–225.

[70] E. Yamauchi, T. Nakatsu, M. Matsubara, H. Kato, H. Taniguchi, Crystal structure of a MARCKS peptide containing the calmodulin-binding domain in complex with Ca^{2+}-calmodulin, Nat. Struct. Biol. 10 (2003) 226–231.

[71] K.L. Yap, J. Kim, K. Truong, M. Sherman, T. Yuan, M. Ikura, Calmodulin target database, J. Struct. Funct. Genomics 1 (2000) 8–14.

[72] K.L. Yap, T. Yuan, T.K. Mal, H.J. Vogel, M. Ikura, Structural basis for simultaneous binding of two carboxy-terminal peptides of plant glutamate decarboxylase to calmodulin, J. Mol. Biol. 328 (2003) 193–204.

[73] T. Yuan, H.J. Vogel, Substitution of the methionine residues of calmodulin with the unnatural amino acid analogs ethionine and norleucine: biochemical and spectroscopic studies, Protein Sci. 8 (1999) 113–121.

[74] M. Zhang, T. Tanaka, M. Ikura, Calcium-induced conformational transition revealed by the solution structure of apo calmodulin, Nat. Struct. Biol. 2 (1995) 758–767.

[75] D.B. Zimmer, E.H. Cornwall, A. Landar, W. Song, The S100 protein family: history, function, and expression, Brain Res. Bull. 37 (1995) 417–429.

[76] K.P. Hoeflich, M. Ikura, Calmodulin in action: diversity in target recognition and activation mechanisms, Cell 108 (2002) 739–742.