Available online at www.sciencedirect.com
ScienceDirect
Comparative Biochemistry and Physiology, Part B 146 (2007) 551–559
CBP
www.elsevier.com/locate/cbpb

Characterization of a novel EF-hand homologue, CnidEF, in the sea anemone *Anthopleura elegantissima*

Laura L. Hauck *, Wendy S. Phillips, Virginia M. Weis

Department of Zoology, 3029 Cordley Hall, Oregon State University, Corvallis, Oregon 97331, USA

Received 25 May 2006; received in revised form 1 December 2006; accepted 3 December 2006
Available online 12 December 2006

Abstract

The superfamily of EF-hand proteins is comprised of a large and diverse group of proteins that contain one or more characteristic EF-hand calcium-binding domains. This study describes and characterizes a novel EF-hand cDNA, CnidEF, from the sea anemone *Anthopleura elegantissima* (Phylum Cnidaria, Class Anthozoa). CnidEF was found to contain two EF-hand motifs near the C-terminus of the deduced amino acid sequence and two regions near the N-terminus that could represent degenerate EF-hand motifs. CnidEF homologues were also identified from two other sea anemone species. A combination of bioinformatic and molecular phylogenetic analyses was used to compare CnidEF to EF-hand proteins in other organisms. The closest homologues identified from these analyses were a luciferin binding protein (LBP) involved in the bioluminescence of the anthozoan *Renilla reniformis*, and a sarcoplasmic calcium-binding protein (SARC) involved in fluorescence of the annelid worm *Nereis diversicolor*. Predicted structure and folding analysis revealed a close association with bioluminescent aequorin (AEQ) proteins from the hydrozoan cnidarian *Aequorea aequorea*. Neighbor-joining analyses grouped CnidEF within the SARC lineage along with AEQ and other cnidarian bioluminescent proteins rather than in the lineage containing calmodulin (CAM) and troponin-C (TNC).

Keywords: Cnidarian; EF-hand; *Anthopleura elegantissima*; Calcium-binding protein; Molecular evolution

1. Introduction

Calcium is used to mediate a large variety of biological functions in animals, including bioluminescence, muscle contraction, neurotransmitter release, cell growth and development, and signal transduction. The EF-hand motif is often present in calcium-binding proteins. Familiar examples of these proteins include calmodulin (CAM), troponin-C (TNC), sarcoplasmic calcium-binding protein (SARC), and aequorin (AEQ), which functions in bioluminescence in hydrozoan cnidarians. The EF-hand domain consists of a helix-loop-helix structure formed by a highly conserved 12 residue calcium-binding loop, flanked on both sides by α-helices (Nelson and Chazin, 1998). This loop binds calcium with the coordination of seven ligands (Fig. 1). Five are from side-chain carboxylate oxygens donated by residues 1, 3, 5, and 12. One is a backbone carbonyl oxygen from residue 7. And finally an indirect association with residue 9, which is mediated by the seventh ligand of a water-molecule hydrogen bound to its side chain and also often associated with residue 3 (Malmendal et al., 1998; Nelson and Chazin, 1998). A variety of EF-hand subfamilies show conservation of the first position D (D_1) which provides a ligand binding group, a G_6 allowing for a 90° turn in the loop, and an E_{12} which provides two ligand binding groups (Jamieson et al., 1980; Malmendal et al., 1998; Yuasa et al., 2001).

There are at least 45 distinct subfamilies of EF-hand proteins which contain up to eight EF-hand domains that usually occur in pairs (Nakayama and Kretsinger, 1994; Kawasaki et al., 1998). The function of only 13 of these 31 subgroups is known. For the well-studied EF-hand proteins, in most cases, adjacent EF-hand domains cross-associate to form a calcium-binding pocket. The most significant differences between EF-hand proteins occur in the regions outside of the EF-hand domains. These can be catalytic, such as an oxygenase region encoded in AEQ, or structural, such as CAM which

* Corresponding author. Tel.: +1 541 737 4358; fax: +1 541 737 0501.
E-mail address: hauckl@science.oregonstate.edu (L.L. Hauck).

1096-4959/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.cbpb.2006.12.004

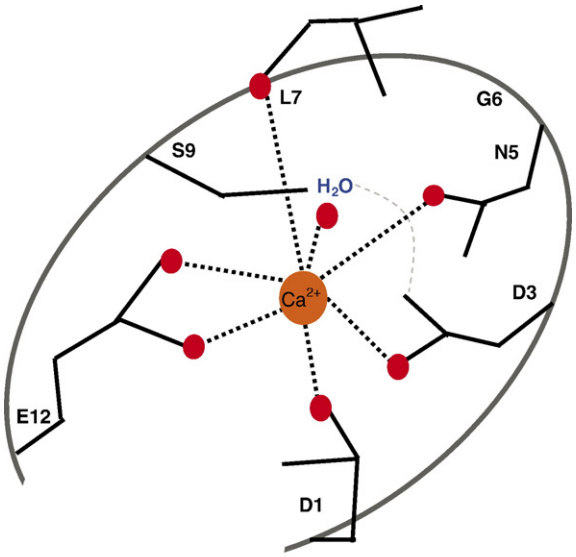

Fig. 1. Schematic representation of an EF-hand calcium-binding loop domain depicting the positions of the seven ligands that contribute to coordinated calcium binding.

transduces intracellular calcium signals through a conformational change (Nakayama and Kretsinger, 1994). The conformational change in CAM causes activation of the enzymatic or structural portion of the protein. Given the variety and importance of the many calcium-related biological functions, it is not surprising that this family of proteins has been well-studied, including its molecular evolution and phylogenetics (Nakayama and Kretsinger, 1994).

Several EF-hand proteins have been described in cnidarians, and are reviewed by Tsuji et al. (1995). These include AEQ, obelin (OBL), mitrocomin, and cyclin, all isolated from different hydrozoans, and all proteins that function in bioluminescence. Luciferin-binding protein (LBP), another bioluminescence protein, as well as a CAM have also been described in the anthozoan *Renilla reniformis*.

In this study, we describe a novel EF-hand protein, CnidEF, from another anthozoan, the North American Pacific coast temperate anemone *Anthopleura elegantissima*. CnidEF was first identified from *A. elegantissima* during screening of a subtracted library that was generated to characterize anemone genes expressed specifically as a function of its symbiosis with the dinoflagellate alga *Symbiodinium muscatinei* (Muscatine, 1971; LaJeunesse and Trench, 2000). Symbiosis between *A. elegantissima* and *S. muscatinei* is facultative; in high light environments anemones harbor symbionts in their tissues but in low light environments, such as caves and rock crevices, anemones are symbiont-free (Weis and Levine, 1996). This study describes comprehensive expression studies that revealed no correlation of CnidEF levels to symbiotic state. In addition, we have identified CnidEF homologues from two other anemone species and performed structural and phylogenetic analyses of CnidEF, both of which place CnidEF within the SARC lineage as a novel addition to the EF-hand family of calcium-binding proteins.

2. Methods

2.1. Isolation of CnidEF from A. elegantissima

CnidEF was originally isolated from a subtracted cDNA library made from *A. elegantissima* RNA and designed to identify sequences enhanced in the symbiotic state. See supplementary online information for details of the subtracted library.

2.2. RNA extraction and cDNA synthesis

Total RNA was extracted from animals as described in Weis and Reynolds (1999). RNA was reverse transcribed using 0.5 μg oligo $(dT)_{12–18}$ and the Superscript Preamplification System (Life Technologies) to synthesize cDNA.

2.3. Northern hybridizations

To determine the relative level of CnidEF expression in field collected samples, Northern hybridizations were conducted. Total RNA (6.75 μg) from 56 animals was resolved on a 1% formaldehyde denaturing gel in 1× MOPS buffer (0.1 M MOPS, 40 mM sodium acetate, 5 mM EDTA). RNA was then transferred to a nylon membrane and hybridized overnight at 68 °C in DIG Easy Hyb (Roche) containing 100 μg/mL salmon sperm DNA and a 307 nucleotide digoxigenin-labeled RNA probe transcribed from a CnidEF clone with specific primers (see Fig. 2 for primer locations). The membrane was washed twice at room temperature with 2× SSC, 0.1% SDS. The hybridized RNA was then immunodetected using anti-digoxigenin antibody conjugated to alkaline phosphatase (Roche). To quantify the relative expression levels, a pixilation analysis of the blot film was conducted using the ImageQuaNT phosphoimager (Molecular Dynamics, Inc.) according to the manufacturer's instructions. Optical density values were subjected to a 2-way ANOVA to test for differences in CnidEF expression between months and among groups (symbiotic versus aposymbiotic). They were also subjected to a post hoc Tukey HSD multiple comparison test to identify the difference in expression by month. All statistical analyses were performed using SPSS V. 9.0.0 (1989).

2.4. Motif recognition

Motif searches of homologues from the closest EF-hand families, and all CnidEF sequences, were conducted at the ExPASy Website hosted by The Swiss Institute of Bioinformatics (Bucher and Bairoch, 1994; Hofmann et al., 1999). Many EF-hand proteins, including those used in the phylogenetic analysis described below, contain more than two EF-hands. Motif searches were conducted on all sequence groups of the final alignment to determine the number and location of EF-hands. This information was used to verify the alignment parameters, and to search for evidence of evolutionarily lost EF-hands within the CnidEF sequence. General EF-hand conservation patterns, such as D_1, G_6, and E_{12} (Jamieson et al., 1980; Malmendal et al., 1998; Yuasa et al., 2001) were used to identify the possible locations of degenerate EF-hands.

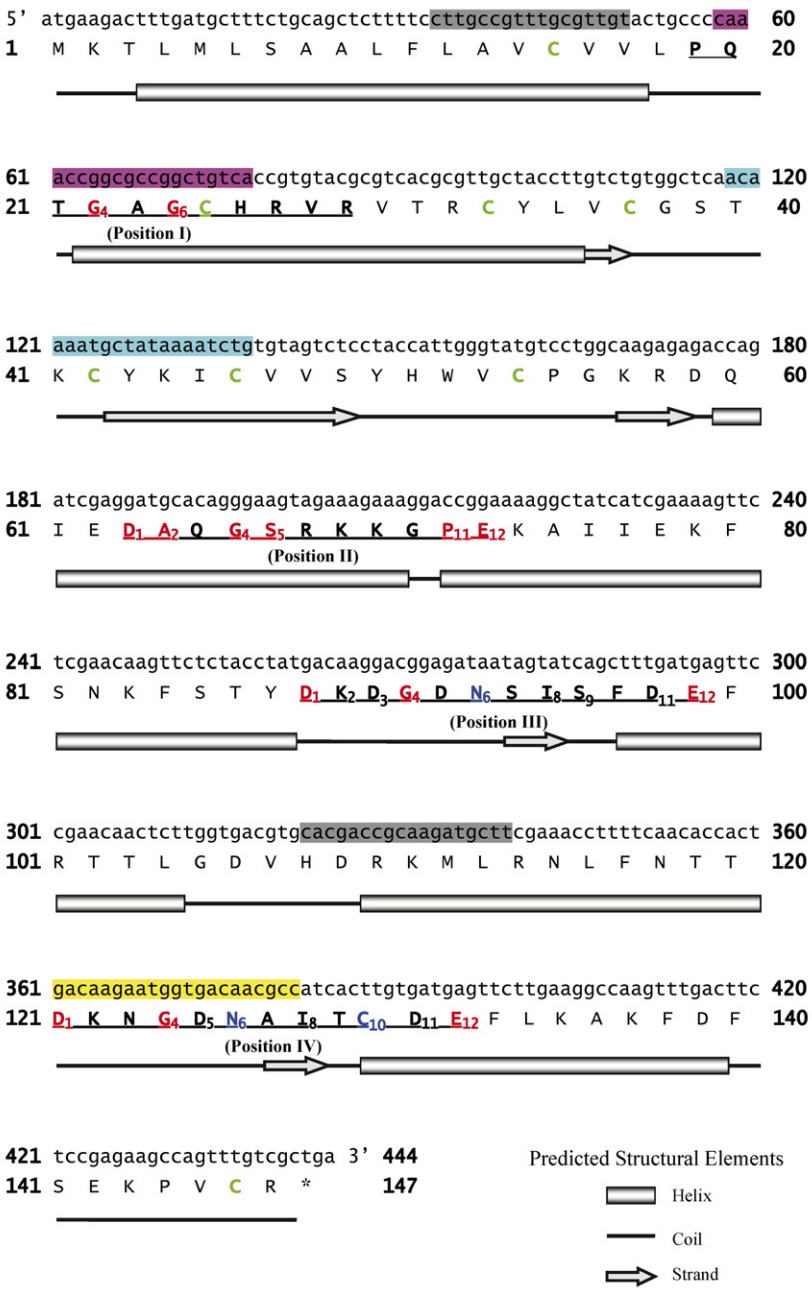

Fig. 2. The nucleotide, predicted amino acid sequence and predicted secondary structure of *Anthopleura elegantissima* CnidEF (Genbank Accession no. ABD16201). Divergent EF-hand regions (position I and position II) and true EF-hand regions (position III and position IV) are in bold and underlined in the amino acid sequence. Conserved EF-hand residues are in red with numerical annotation, and unique CnidEF conserved residues are in blue with numerical annotation. The numerous cysteine residues are green. Location of predicted secondary structural elements of helix, coil and strand is depicted below the amino acid sequence. Primer locations are as follows: specific forward and reverse primers for Northern hybridization probe construction are shaded gray, degenerate forward primer CnidEFDF1, shaded magenta, used with an oligo d(T) primer 1-rev (see text) for amplification of a CnidEF fragment from *A. artemisia*, degenerate forward primer CnidEFDF2, shaded blue, used together with degenerate reverse primer CnidEFDR1, shaded yellow, for amplification of a CnidEF fragment from *Aiptasia pallida*.

2.5. Identification of CnidEF homologues from other anemone species

CnidEF homologues were identified from other anemone species using degenerate primers and reverse transcriptase PCR. EF-hand domains were the targets for degenerate primer design (see Fig. 2 for primer locations), and a modified oligo-dT primer 1-rev (5′-CTCTAGAACTAGTCT$_{18}$-3′, Weis and Reynolds, 1999) was used for reverse priming off of the poly-A tail. A fragment of a CnidEF homologue from *Anthopleura artemisia* was amplified with degenerate forward primer CnidEFDF1 and 1-rev. CnidEF amplification from *Aiptasia pallida* was performed in stages. The first fragment was amplified from a degenerate forward primer CnidEFDF2, and a degenerate reverse primer CnidEFDR1. A specific forward primer (5′-CTGACTACGACAAGAACAAG-3′) was made from this sequence and

used with 1-rev to amplify the 3′ end of the cDNA. The PCR Reagent System (Invitrogen) was used for all PCR reactions.

Single band PCR products were TA cloned into pGEM-T vectors (Promega). Following amplification in DH5α *Escherichia coli* cells (Invitrogen), the plasmid DNA was sequenced with a Terminator v3.1 cycle sequencing kit (BigDye®) using a 3100 Genetic Analyzer (ABI Prism®), 3100 data collection software v.1.1 and DNA sequencing analysis software v.3.7 (ABI Prism®). Sequences were aligned in ClustalX with the original *A. elegantissima* CnidEF nucleotide sequence to determine if the isolated sequence was a CnidEF homologue (Thompson et al., 1997).

2.6. Structural analyses of sequences

To predict the secondary structure of CnidEF, the deduced amino acid sequence was analyzed by the Protein Structure Prediction Server (PSIPRED, McGuffin et al., 2000). The predicted structure was then compared with known crystal structures through a fold recognition analysis using the (m) GenTHREADER server (McGuffin and Jones, 2003). *R. reniformis* LBP was also threaded to this database because there is no crystallized structure of LBP available for direct comparison with CnidEF. The next closest homologue, SARC from *Nereis diversicolor*, is in the crystallized structure database thus allowing for a structural comparison of all three homologues.

2.7. Phylogenetic analyses of sequences

Deduced amino acid sequences were compared with known protein sequences using BlastP and PsiBlast network databases at the National Center for Biotechnology Information (NCBI, Altschul et al., 1990, 1997). Additional EF-hand sequences from other subfamilies that did not appear in the BlastP and PsiBlast searches were also identified for use in the phylogenetic analysis to explore the molecular evolution of CnidEF. These included parvalbumin (PARV), oncomodulin, S100, calbindin, secretegogin, calsequestrin, calregulin, calcyclin, metastacin, and calreticulin.

The EF-hand subfamily sequences were included in alignments conducted in ClustalX. Sequence lengths were not adjusted, nor were they masked during the alignment. Neighbor-Joining (NJ) analyses were conducted from these alignments using PAUP software (Swofford, 2000), to generate a tree with bootstrap values of 1000. According to Nakayama and Kretsinger (1994), EF-hand proteins are thought to have evolved from a "UR" precursor, after which the SARC and CAM/TNC lineages diverged from one another. Due to grouping of CnidEF within the SARC lineage, CAM and TNC sequences were retained in the final analysis to root the tree.

3. Results

3.1. Northern hybridization

A Northern blot analysis was conducted to examine expression patterns of the CnidEF gene in anemones collected from the field at various times throughout the year. The results show a highly variable level of expression, with aposymbiotic and symbiotic samples collected in close proximity to one another having almost identical expression levels (Fig. 3). The ANOVA showed no difference in expression between symbiotic and aposymbiotic animals. According to the Tukey HSD, the month of May had statistically significant elevations in gene expression for both groups ($p=0.05$).

3.2. A. elegantissima CnidEF sequence description

The CnidEF cDNA (Genbank Accession no. ABD16201) is 635 bp in length, with a 441 bp open reading frame (ORF) corresponding to 147 amino acids and an approximate molecular mass of 16.6 kDa (Fig. 2). Motif searches of the deduced amino acid sequence of CnidEF revealed the presence of EF-hand calcium-binding loop motifs. Prosite identified two such domains near the C-terminus. In an alignment of multiple EF-hand proteins (Fig. 4) there were four possible EF-hand locations, with CnidEF having only positions III and IV.

It is possible that a CnidEF progenitor had additional EF-hands that have since diverged beyond recognition. Analysis of the CnidEF sequence near the N-terminus (Fig. 2), where the position I EF-hand is present in other sequences of the alignment, shows G_4 and G_6 of the binding loop may have been maintained. These two residues are also found in position I EF-hand of CAM. These are the only possible retained residues and they do not provide convincing evidence for the existence of an ancestral EF-hand at this location.

The EF-hand loop region in position II of CnidEF shifted by several residues among the different alignments, most likely due to high sequence divergence in this region, and a 5 residue insertion–deletion (indel) mutation present in the AEQ and OBL sequences. According to one alignment, CnidEF contains

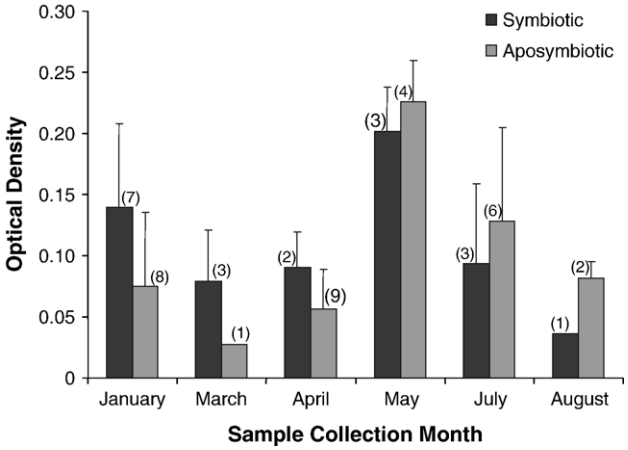

Fig. 3. Northern blot analysis comparing CnidEF expression between symbiotic and aposymbiotic *Anthopleura elegantissima* sampled at various times throughout the year. Number of individuals sampled is in parentheses above the column, error bars represent standard deviation. There was a statistically significant ($p>0.05$) elevation in expression according to the Tukey HSD analysis for the month of May.

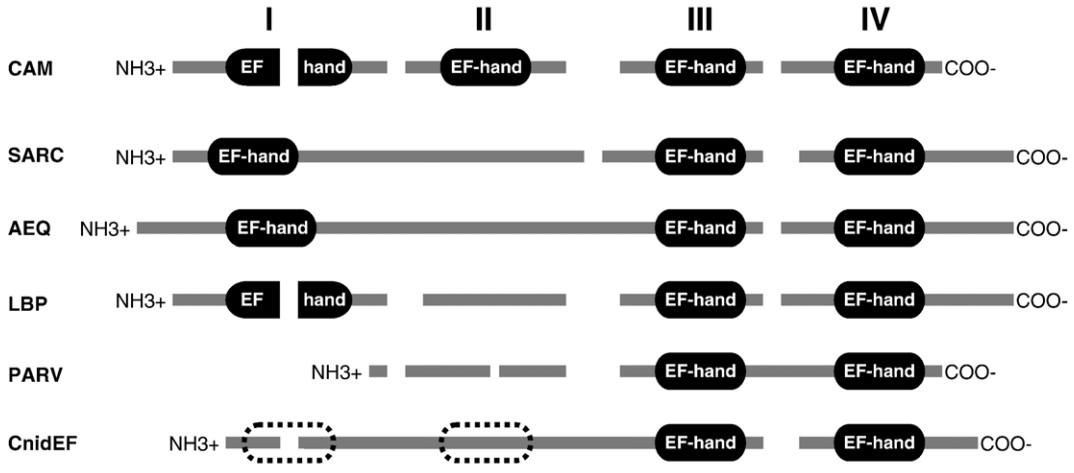

Fig. 4. Schematic diagram depicting location and number of EF-hand domains in *Anthopleura elegantissima* CnidEF and other EF-hand protein subfamilies. The grey bar indicates the amino acid strand of the entire protein, EF-hands are represented by black ovals, gaps as determined in a ClustalX alignment, are shown as breaks in the grey bars. The possible degenerate EF-hand locations in CnidEF are indicated by dotted ovals. The presence of all EF-hands was verified in Scan Prosite.

a position II DA_G near the beginning of this EF-hand loop region that may relate to the D_1, A_2, G_4 pattern present in CAM, and also contains $P_{11}E_{12}$, which is shared with AEQ, OBL, and CAM. In addition, it shares a position II S_5 with most of the PARV sequences, and with TNC of *Danio rerio*. These six conserved residues provide evidence that there was once a Ca^{2+} binding EF-hand loop in position II.

Comparison of the true EF-hands in positions III and IV of CnidEF to other EF-hand sequences in the alignment revealed substantial conservation in the calcium-binding loop. Position III shares D_1, K_2, D_3, G_4, I_8, S/T_9, D_{11} and E_{12} with most sequences, and position IV shares D_1, G_4, D_5, I/L_8, D_{11} and E_{12} with most subfamilies and K_2 and N_3 only with TNC. CnidEF also has an N_6, in place of the more widely conserved G_6, which appears to be an uncommon EF-hand feature that in this alignment is shared only with N_6 of the EF-hand loop in position III of SARC from the annelid *N. diversicolor*. CnidEF has an unusually high number of cysteine residues with a total of nine cysteines, seven of which are concentrated in a region forty amino acids in length (Fig. 2).

3.3. Predicted secondary structure and mGen-THREADER analysis

The secondary structure of *A. elegantissima* CnidEF was predicted by the PSIPRED server (McGuffin et al., 2000) and is shown in Fig. 2. Despite the divergence of all but two residues in the position I EF-hand loop region of CnidEF, the predicted secondary structure retained the EF-hand pattern (α-helix, coil loop, α-helix), although the second helix begins midway through the area where the calcium-binding loop should have been. The position II divergent EF-hand region, which showed the most potential for having an ancestral EF-hand origin, contained no structural similarity to true EF-hands. The true EF-hands (III and IV) showed the expected EF-hand structure: α-helix, coil loop, α-helix with a small β-strand leading into the

Table 1
The most significant hits from the (m)GenTHREADER analysis for CnidEF from *Anthopleura elegantissma* and LBP from *Renilla reniformis*

	Protein	Organism	Expect value	Confidence	Protein function
Hits for *Anthopleura elegantissima* CnidEF	*AEQ	*Aequorea aequorea*	0.010	Medium	Bioluminescence
	*Myosin regulatory domain	*Aequipecten irradians*	0.011	Medium	Skeletal muscle contraction
	*SARC	*Nereis diversicolor*	0.025	Medium	Fluorescence
	*Calcineurin	*Homo sapiens*	0.035	Medium	CAM-stimulated phosphatase
Hits for *Renilla reniformis* LBP	*AEQ	*Aequorea aequorea*	0.001	High	Bioluminescence
	*SARC	*Nereis diversicolor*	0.002	High	Fluorescence
	SARC	*Branchiostoma lanceolatum*	0.002	High	Possibly muscle contraction
	Recoverin	*Bos taurus*	0.005	High	Control of rhodopsin kinase activity
	*Calcineurin	*Homo sapiens*	0.008	High	CAM-stimulated phosphatase
	Calcium-binding protein	*Arabidopsis thaliana*	0.009	High	Unknown
	*Myosin regulatory domain	*Aequipecten irradians*	0.010	Medium	Skeletal muscle contraction
	Apoptosis-linked protein (alg-2)	*Mus musculus*	0.010	Medium	Apoptosis signaling
	EhCaBP	*Entamoeba histolytica*	0.013	Medium	Signal transduction in pathogenesis

Asterisks denote hits common to both lists.

second α-helix. For comparative purposes *R. reniformis* LBP and AEQ, both cnidarian EF-hand sequences, were also submitted for analysis by PSIPRED. The predicted structure among the three representatives was similar, to include a common small β-strand leading into the second helix of the position IV EF-hands (data not shown).

A. elegantissima CnidEF and *R. reniformis* LBP were threaded to a database of crystallized EF-hand homologues at the PSIPRED site to allow for structural comparison between the two anthozoan EF-hand sequences (Table 1). The results show a close structural relationship between these two homologues. The four most significant hits for CnidEF were

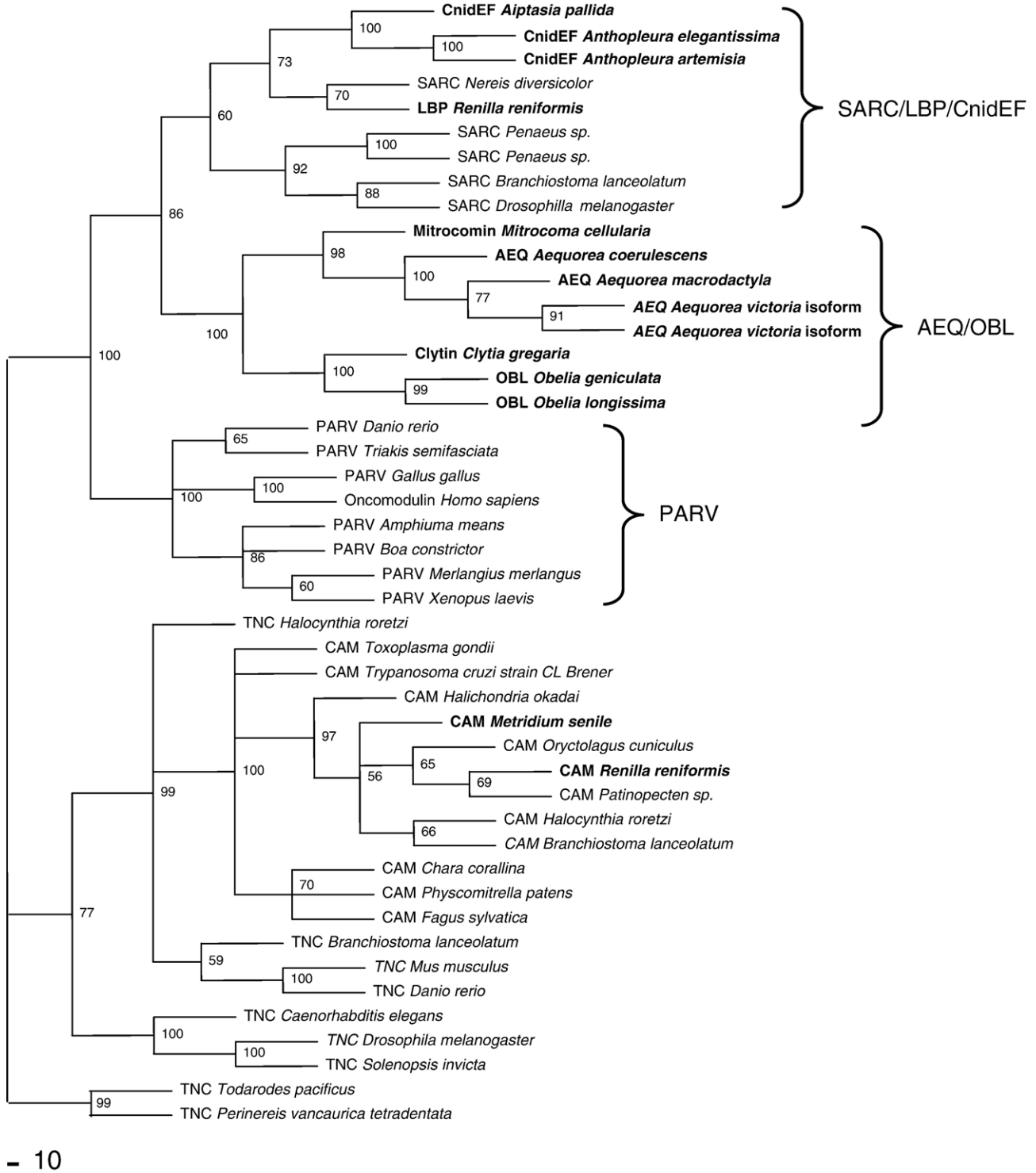

Fig. 5. Phylogenetic relationships of EF-hand proteins including CnidEF from *Anthopleura elegantissima* inferred by a Neighbor-Joining analysis. Percent of 1000 bootstrap replicates supporting the topology is given at each node. The tree is rooted with the CAM/TNC group. Cnidarian sequences are in bold. GenBank entries for the included sequences are listed in Table 2.

all also high hits for LBP and the best hit for both sequences was AEQ, another cnidarian homologue. In addition, both sequences threaded to a SARC homologue from *N. diversicolor*, a sequence of significance in phylogenetic analyses (see ahead).

3.4. Identification of CnidEF homologues from other anemone species

RT-PCR was used to identify CnidEF homologues from other anemones. A 476 bp fragment of a CnidEF was isolated from the closely related sympatric nonsymbiotic anemone *A. artemisia* (Genbank Accession no. ABD16203). This sequence displayed a 91% sequence identity to the *A. elegantissima* homologue. In addition, a 478 bp fragment of a CnidEF was identified from the more distantly related symbiotic tropical anemone *A. pallida* (Genbank Accession no. ABD16202). This sequence was just 44% identical to the *A. elegantissima* CnidEF, considerably more divergent than the *A. artemisia* sequence. For the EF-hands in positions III and IV, in the two *Anthopleura* species, the sixth residue is an N instead of the highly conserved G present in most other organisms. In *A. pallida* it remains a G in the EF-hand of position III, but is an R in position IV. All three CnidEF sequences contain a unique C_{10} in the fourth EF-hand loop.

3.5. Phylogenetic analyses of sequences

The majority of BlastP returned hits to *A. elegantissima* CnidEF were CAMs from a variety of different organisms, whereas PsiBlast and literature searches identified a variety of homologues from other EF-hand subfamilies including LBP, AEQ, OBL, SARC and TNC. In initial analyses, using all identified EF-hand sequences (listed in Methods), the only subfamilies with which CnidEF associated were LBP, AEQ, OBL and SARC (data not shown). Therefore, the final alignment used for phylogenetic tree generation (Fig. 5) used only these sequences and CAM and TNC. The CAM and TNC sequences grouped outside of the SARC lineage with strong bootstrap support and were used to root the tree. The distance analysis divided the tree into two major clades with a very strong support, one consisting of PARV homologues and the other containing SARCs, AEQs and OBLs. This clade was further subdivided with moderate support into AEQ/OBL and SARC/LBP/CnidEF lineages. Within this SARC/LBP/CnidEF lineage, the analysis provided strong support for a terminal CnidEF and two SARC clades. A lineage containing *R. reniformis* LBP and *N. diversicolor* SARC grouped with the CnidEF sequences with very weak support. Further, there was very weak support for the deepest branching node, suggesting that this analysis cannot confidently resolve the relationships within the SARC/LBP/CnidEF lineage.

Table 2
GenBank accession ID

Organism	Gene name	Accession ID #
Anthopleura elegantissima	CnidEF	ABD16201
Anthopleura artemisia	CnidEF	ABD16203
Aiptasia pallida	CnidEF	ABD16202
Renilla reniformis	LucBP	P05938
Aequorea victoria	AEQ	P02592
Aequorea victoria	AEQ	P07164
Aequorea macrodactyla	AEQ	AAK02061
Aequorea coerulescens	AEQ	AAO91813
Mitrocoma cellularia	Mitrocomin	P39047
Clytia gregaria	Clytin	CAA49754
Obelia geniculata	Obelin	AAL86372
Obelia longissima	Obelin	Q27709
Neanthes diversicolor	SARC	P04571
Drosophila melanogaster	SARC	AF014952.1
Penaeus sp.	SARC	P02636
Branchiostoma lanceolatum	SARC	P04569
Penaeus sp.	SARC	P02635
Danio rerio	Parv	AAH81523
Triakis semifasciata	Parv	P30563
Gallus gallus	Parv	P43305
Merlangius merlangus	Parv	1A75B
Amphiuma means	Parv	P02616
Boa constrictor	Parv	P02615
Xenopus laevis	Parv	AAA49925
Homo sapiens	Oncomodulin	NP_006179
Chara corallina	CAM	BAA96536
Physcomitrella patens	CAM	CAA62150
Renilla reniformis	CAM	P62184
Metridium senile	CAM	BAB61796
Halichondria okadai	CAM	BAB61797
Patinopecten sp.	CAM	P02595
Halocynthia roretzi	CAM	BAA19788
Branchiostoma lanceolatum	CAM	BAA19786
Fagus sylvatica	CAM	Q39752
Oryctolagus cuniculus	CAM	1003191A
Toxoplasma gondii	CAM	CAA69660
Trypanosoma cruzi strain CL Brener	CAM	XP_808089
Drosophila melanogaster	TC	P47947
Solenopsis invicta	TC	AAL57489
Caenorhabditis elegans	TC	BAB84566
Todarodes pacificus	TC	BAB40597
Perinereis vancaurica tetradentata	TC	BAB18898
Mus musculus	TC	AAH24390
Danio rerio	TC	AAH64284
Branchiostoma lanceolatum	TC	BAA13733
Halocynthia roretzi	TC	BAA13630

4. Discussion

CnidEF is a novel addition to the large EF-hand family of proteins, possessing two EF-hands near the C-terminus, and two upstream regions that might be degenerate EF-hands. The retention of key residues in the position II region of CnidEF indicates that this protein may have once contained an EF-hand in this position that has since been evolutionarily lost. Because EF-hands usually occur in pairs, and the predicted secondary structure in the position I region reflects the expected α-helix, coil loop, α-helix pattern, it is also possible that there was once a position I EF-hand that has since diverged beyond recognition.

As discussed earlier, all EF-hand proteins evolved from a common UR ancestor which then, through a series of gene duplications and fusions (or lack there of), formed several distinct evolutionary clades (Nakayama and Kretsinger, 1994). According to the phylogenetic analysis presented here (Fig. 5), the evolutionary origin of CnidEF, as well as AEQ and LBP,

appears to be within the SARC lineage. CAM and TNC, although close BLAST homologues to CnidEF, distinctly group outside the SARC lineage, as does PARV. This is an interesting phenomenon as AEQ and SARC have previously been considered separate subfamilies, and have not been joined by other molecular analyses. Structural analysis through the PSIPRED server (Table 1) supports the placement of CnidEF with AEQ and SARC. Within the SARC lineage, anthozoan LBP and CnidEF do not group with the bioluminescent hydrozoan EF-hand proteins, such as AEQ and OBL, but rather with the SARCs from other organisms. There was a lack of association with CAM proteins isolated from the anthozoans *Metridium senile* and *R. reniformis* (Fig. 5) supporting the conclusion that CnidEF is a novel EF-hand protein that is not in the CAM lineage. The placement of CnidEF by sequence and structural analysis with LBP from the anthozoan *R. reniformis*, SARC from the annelid *N. diversicolor*, and with AEQ and OBL from hydrozoans is intriguing because all of these proteins function in the production or modulation of light. Upon binding three calcium ions, the photoprotein AEQ undergoes a conformational change that converts it into an oxygenase, which in turn catalyzes the oxidative decarboxylation of tightly bound luciferin. It then uses green fluorescent protein as a secondary emitter to produce green light (Lewis and Daunert, 2000). The bioluminescent function of AEQ is made possible by several specific residues: a C-terminal proline, three cysteines in positions 145, 152 and 180, a histidine at position 169, and several tryptophan residues in positions 12, 79, 86, 108, 129 and 173 (Lewis and Daunert, 2000). LBP function is much like that of AEQ, however unlike AEQ, LBP requires, in addition to Ca^{2+}, the presence of dissolved oxygen and a luciferase (Charbonneau and Cormier, 1979). The binding of Ca^{2+} to the LBP photoprotein initiates a conformational change that exposes the bound luciferin and oxygen to the luciferase, thus allowing substrate oxidation to occur and light to be emitted. The function of *N. diversicolor* SARC is fluorescence, and the fluorescent properties of this SARC are linked to the presence of three tryptophan residues in positions 4, 57 and 170 (Sillen et al., 2003). SARCs, in general, have four typical helix-loop-helix EF-hands, however, like AEQ, only sites I, III and IV have retained Ca^{2+} or Mg^{2+} binding capacity (Cook et al., 1993; Sillen et al., 2003). Comparison of the *N. diversicolor* SARC to other invertebrate SARCs has shown a high variability among the proteins, with only 9 residues in common to all species (Collins et al., 1988).

There is no evidence that *A. elegantissima*, *A. artemisia* and *A. pallida*, all of which express a form of CnidEF, are bioluminescent organisms. Comparison of CnidEF sequences to the residues important for AEQ light emission also disputes any link of CnidEF function to bioluminescence. CnidEF lacks tryptophan residues, which are responsible for coelentrazine substrate binding and the formation of a hydrophobic pocket in AEQ. CnidEF does contain many cysteine residues, although not in the same locations as those in AEQ, and lacks the H_{169} responsible for oxygen binding. CnidEF does, however, have a proline very close to the C-terminus, which aligns with a proline from AEQ. All three anemones expressing CnidEF do fluoresce (Buchsbaum, 1968; Rice, 1999; Mazel, 2006), however all three forms of CnidEF lack the tryptophan residues that impart the fluorescent properties of *N. diversicolor* SARC (Sillen et al., 2003). In addition, the work presented in this study shows that CnidEF expression is not linked to the symbiotic state. Samples collected in close proximity to one another had nearly identical expression (data not shown). This suggests that environmental factors played a role in CnidEF expression. The emergence of CnidEF as a symbiosis-linked gene in the subtracted library was likely an artifact of the collection process. Therefore, the function of CnidEF remains unclear. It has been hypothesized that the actual function of bioluminescent proteins is not bioluminescence at all; rather that bioluminescence is a byproduct of a different biological function (Shimomura, 1985). Perhaps, due to the lack of key residues necessary for light production or modulation, CnidEF never developed the side effect of bioluminescence and only performs the original (and still unknown) biological function of this curious group of proteins.

Acknowledgements

We thank Steve Giovannoni for assistance with bioinformatic analysis, and Andrew Karplus and Elisha Wood-Charlson for assistance with protein structure analysis. We also thank the members of the Weis Lab for comments on the manuscript. This work was supported by grants from the National Science Foundation (MCB0237230 and IOB0542452) to V.M.W. and by a National Science Foundation Fellowship to L.L.H.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cbpb.2006.12.004.

References

Altschul, S.F., Gish, S.F., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tools. J. Mol. Biol. 215, 403–410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Bucher, P., Bairoch, A., 1994. A generalized profile syntax for biomolecular sequence motifs and its function in automatic sequence interpretation. In: S. D. (Ed.), Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. AAAIPress, Menlo Park, pp. 53–61.

Buchsbaum, V.M., 1968. Behavioral and physiological responses to light by the sea anemone *Anthopleura elegantissima* as related to its algal endosymbionts. Department of Biological Sciences. Stanford University, Stanford, CA, p. 123.

Charbonneau, H., Cormier, M.J., 1979. Ca2+-induced bioluminescence in *Renilla reniformis*. J. Biol. Chem. 254, 769–780.

Collins, J.H., Cox, J.A., Theibert, J.L., 1988. Amino acid sequence of a sarcoplasmic calcium-binding protein from the sandworm *Nereis diversicolor*. J. Biol. Chem. 263, 15378–15385.

Cook, W.J., Jeffrey, L.C., Cox, J.A., Vijay-Kumar, S., 1993. Structure of a sarcoplasmic calcium-binding protein from amphioxus refined at 2.4 A resolution. J. Mol. Biol. 229, 461–471.

Hofmann, K., Bucher, P., Flaquet, L., Bairoch, A., 1999. The PROSITE database, its status in 1999. Nucleic Acids Res. 27, 215–219.

Jamieson Jr., G.A., Bronson, D.D., Schachat, F.H., Vanaman, T.C., 1980.

Structure and function relationships among calmodulins and troponin C-like proteins from divergent eukaryotic organisms. Ann. N.Y. Acad. Sci. 356, 1–13.

Kawasaki, H., Nakayama, S., Kretsinger, R.H., 1998. Classification and evolution of EF-hand proteins. BioMetals 11, 227–295.

LaJeunesse, T.C., Trench, R.K., 2000. Biogeography of two species of *Symbiodinium* (Freudenthal) inhabiting the intertidal sea anemone *Anthopleura elegantissima* (Brandt). Biol. Bull. 199, 126–130.

Lewis, J.C., Daunert, S., 2000. Photoproteins as luminescent labels in binding assays. Fresenius' J. Anal. Chem. 366, 760–768.

Malmendal, A., Carlstrom, G., Hambraeus, C., Drakenberg, T., Forsen, S., Akke, M., 1998. Sequence and context dependence of EF-hand loop dynamics. An 15N relaxation study of a calcium-binding site mutant of calbindin D9k. Biochemistry 37, 2586–2595.

Mazel, C., 2006. NightSea Website: Posting of Photos by Charles Mazel and Jack Sullins Depicting *Anthopleura artemisia* Fluorescence. http://nightsea.com/home.htm.

McGuffin, L.J., Jones, D.T., 2003. Improvement of the GenTHREADER method for genomic fold recognition for genomic sequences. Bioinformatics 19, 874–881.

McGuffin, L., Bryson, K., Jones, D., 2000. The PSIPRED protein structure prediction server. Bioinformatics 16, 404–405.

Muscatine, L., 1971. Experiments on green algae coexistent with zooxanthellae in sea anemones. Pac. Sci. 25, 13–21.

Nakayama, S., Kretsinger, R.H., 1994. Evolution of the EF-hand family of proteins. Annu. Rev. Biophys. Biomol. Struct. 23, 473–507.

Nelson, M.R., Chazin, W.J., 1998. Chapter 2: Calmodulin as a Calcium Sensor. Academic Press, New York. 17–64 pp.

Rice, J., 1999. Writhe and shine. Nat. Hist. 108, 76–77.

Shimomura, O., 1985. Bioluminescence in the sea: photoprotein systems. Symp. Soc. Exp. Biol. 39, 351–372.

Sillen, A., Verheyden, S., Delfosse, L., Braem, T., Robben, J., Volckaert, G., Engelborghs, Y., 2003. Mechanism of fluorescence and conformational changes of the sarcoplasmic calcium binding protein of the sand worm *Nereis diversicolor* upon Ca2+ or Mg2+ binding. Biophys. J. 85, 1882–1893.

Swofford, D.L., 2000. PAUP∗: Phylogenetic Analysis Using Parsimony and Other Methods (Software). Sinauer Associates, Sunderland, MA.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX-Windows interface: flexible strategies for multiple sequence alignment aided by quantity analysis tools. Nucleic Acids Research 25, 4876–4882.

Tsuji, F.I., Ohmiya, Y., Fagan, T.F., Toh, H., Inouye, S., 1995. Molecular evolution of the Ca2+-binding photoproteins of the hydrozoa. Photochem. Photobiol. 62, 657–661.

Weis, V.M., Levine, P.R., 1996. Differential protein profiles reflect the different lifestyles of symbiotic and aposymbiotic *Anthopleura elegantissima*, a sea anemone from temperate waters. J. Exp. Biol. 199, 883–892.

Weis, V.M., Reynolds, W.S., 1999. Carbonic anhydrase expression and synthesis in the sea anemone *Anthopleura elegantissima* are enhanced by the presence of dinoflagellate symbionts. Physiol. Biochem. Zool. 72, 307–316.

Yuasa, H.J., Suzuki, T., Yazawa, M., 2001. Structural organization of lower marine nonvertebrate calmodulin genes. Gene 279, 205–212.