doi:10.1016/j.jmb.2006.01.083 J. Mol. Biol. (2006) 357, 1536–1547

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Structure of the Neuronal Protein Calexcitin Suggests a Mode of Interaction in Signalling Pathways of Learning and Memory

P. T. Erskine[1], G. D. E. Beaven[1], R. Hagan[1], I. S. Findlow[1], J. M. Werner[1] S. P. Wood[1] J. Vernon[2], K. P. Giese[2], G. Fox[3] and J. B. Cooper[1]*

[1]*School of Biological Sciences University of Southampton Bassett Crescent East, Southampton SO16 7PX UK*

[2]*Wolfson Institute for Biomedical Research, University College London, Cruciform Building, Gower Street, London WC1E 6BT, UK*

[3]*BM16, European Synchrotron Radiation Facility, BP 220 38043 Grenoble, France*

**Corresponding author*

The three-dimensional structure of the neuronal calcium-sensor protein calexcitin from *Loligo pealei* has been determined by X-ray analysis at a resolution of 1.8 Å. Calexcitin is up-regulated following Pavlovian conditioning and has been shown to regulate potassium channels and the ryanodine receptor. Thus, calexcitin is implicated in neuronal excitation and plasticity. The overall structure is predominantly helical and compact with a pronounced hydrophobic core between the N and C-terminal domains of the molecule. The structure consists of four EF-hand motifs although only the first three EF hands are involved in binding calcium ions; the C-terminal EF-hand lacks the amino acids required for calcium binding. The overall structure is quite similar to that of the sarcoplasmic calcium-binding protein from Amphioxus although the sequence identity is very low at 31%. The structure shows that the two amino acids of calexcitin phosphorylated by protein kinase C are close to the domain interface in three dimensions and thus phosphorylation is likely to regulate the opening of the domains that is probably required for binding to target proteins. There is evidence that calexcitin is a GTPase and the residues, which have been implicated by mutagenesis in its GTPase activity, are in a short but highly conserved region of 3_{10} helix close to the C terminus. This helix resides in a large loop that is partly sandwiched between the N and C-terminal domains suggesting that GTP binding may also require or may cause domain opening. The structure possesses a pronounced electro-positive crevice in the vicinity of the 3_{10} helix, that might provide an initial docking site for the triphosphate group of GTP. These findings elucidate a number of the reported functions of calexcitin with implications for neuronal signalling.

Keywords: X-ray structure; calcium sensing; GTPase; signal transduction; learning and memory

Introduction

The protein calexcitin was originally identified in the photoreceptor neurons of the marine snail *Hermissenda crassicornis* as a protein with modified expression and activity during associative learning.[1] It was initially found as a ~20 kDa cytosolic protein that became up-regulated and phosphorylated following a Pavlovian training protocol. *Hermissenda* are normally attracted towards light but are repelled by movement, which causes them to retract. The snails become conditioned to retract upon exposure to light (normally an attractive stimulus) if they are simultaneously exposed to both light and movement (e.g. on an illuminated turntable). The up-regulation of calexcitin correlated with exposure to both light and movement and not with each separate stimulus alone. Electrophysiological studies showed that micro-injection of the purified calexcitin protein into *Hermissenda* photoreceptors reduced early and late voltage-dependent K^+

Abbreviations used: PKC, protein kinase C; ER, endoplasmic reticulum; ASCP, Amphioxus sarcoplasmic calcium-binding protein.

E-mail address of the corresponding author: j.b.cooper@soton.ac.uk

currents to an extent similar to that caused by the Pavlovian conditioning experiment itself.[2] Thus, calexcitin was shown to regulate the voltage-dependent potassium channel and the calcium-dependent potassium channel and was capable of reproducing the electrophysiological effects of learning when injected into *Hermissenda* neurons. The protein was found to bind calcium and there is evidence that it has GTPase activity.[2]

Calexcitin has been isolated from the optic lobe of the long-finned squid *Loligo pealei*.[3] Monoclonal antibodies to squid calexcitin cross-react with a protein of the same molecular mass in mammalian neuronal cells.[4] Electrophysiological studies showed that low concentrations of squid calexcitin had a strong inhibitory effect on human K^+ channels, suggesting that a functionally related mammalian protein exists and could play a similar role in regulation of these channels.[3] It has been shown that the protein kinase C (PKC)-dependent phosphorylation of calexcitin causes it to translocate to the cell membrane where its effects on membrane excitability are exerted.[5] Accordingly, *L. pealei* calexcitin contains two PKC phosphorylation consensus sequences and is a good substrate for PKC *in vitro*.[5,6] In neuronal cells, the activation of PKC or an increase in intracellular Ca^{2+} can cause long-term inhibition of K^+ currents and it is possible that calexcitin is involved in these processes.

It has been shown that calexcitin causes the release of calcium ions from the endoplasmic reticulum (ER) by binding to and activating ryanodine receptors,[7] the latter form one of the main channels for Ca^{2+} release from the ER. Calexcitin has been shown to co-purify with squid ryanodine receptors and the interaction between these two proteins is calcium-dependent.[8] Phosphorylation of calexcitin by PKC was found to inhibit the interaction with the ryanodine receptor.

Cloning and sequencing of squid calexcitin cDNA showed that the protein consists of 191 amino acid residues and has sequence similarity to several sarcoplasmic calcium-binding proteins (SCPs) and some elements of sequence similarity with GTP-binding proteins of the ADP-ribosylation factor family.[3] There is low but significant sequence identity (31%) between calexcitin and an invertebrate SCP of known X-ray structure, referred to as ASCP, an SCP from the proto-vertebrate marine organism Amphioxus[9] and there is even lower identity (21%) with the SCP from the sandworm *Nereis diversicolour*.[10] The similarity to these proteins of known structure suggests that calexcitin contains up to four EF-hand motifs and calcium-binding assays have shown that two or three calcium ions are bound per monomer.[3,6] The highest affinity Ca^{2+} site of calexcitin, which has a $K_d \sim 40$ nM, is likely to be a calcium buffering site, since it will be occupied even when the neuron is at resting potential ($[Ca^{2+}]$ in the range of 1 nM to 0.1 μM).[11] There is another calcium site with lower affinity, referred to as the regulatory or sensor site, with a K_d value that has been estimated at just below 1 μM and this is in the concentration range for bulk calcium during neuronal activation.

The evidence that calexcitin is a GTPase[3] would make this a remarkable, and possibly unique, protein with dual functions both as a calcium sensor and G-protein. Whilst the early reports of the GTPase activity were contested by another group,[6] recent reports suggest that calexcitin's GTPase activity is stimulated by calcium in several species, most notably *Drosophila*.[12] Calcium ions bind to the protein with sub-micromolar affinity and induce significant conformational changes within the physiological Ca^{2+} concentration range. This has been demonstrated by circular dichroism (CD), fluorescence, NMR and infra-red spectroscopy.[6,11] Calexcitin shares significant sequence identity (almost 40%) with juvenile hormone diol kinase, an insect enzyme that uses ATP or GTP to phosphorylate its polyisoprene substrate.[13–15] Here, we report the crystal structure of *L. pealei* calexcitin, which has been solved at a resolution of 1.8 Å revealing the detailed tertiary fold of the molecule along with its calcium-binding sites and putative GTP-binding regions.

Results and Discussion

Quality of the structure

The final crystal structure consists of two molecules of calexcitin and 685 water molecules in the asymmetric unit. The model, which has been refined using data between 50.0 Å and 1.8 Å, has an *R*-factor of 19.1% and an *R*-free of 24.9% (see Tables 1 and 2). The final structure has 92.2% of its residues within the "most favoured" regions of the Ramachandran plot by the PROCHECK criteria[16] and 6.9% of residues within the so-called "additional allowed" boundary. The remaining three (non-glycine) residues are in regions with poor electron density (at the extreme N and C termini) and hence their conformations are poorly defined. The isotropic temperature factors for the protein are reasonable except for a few residues at the N-terminal ends of both molecules in the asymmetric unit and at the C-terminal end of the second monomer. In these regions the electron density for a small number of residues remains poor despite extensive efforts to rebuild them throughout the refinement process. Both molecules of calexcitin within the asymmetric unit superimpose closely with an rms deviation of 0.7 Å and the average isotropic temperature factor for all protein atoms is 29.8 $Å^2$.

Overall structure of calexcitin

Calexcitin adopts a compact conformation of approximate dimensions 28 Å × 42 Å × 42 Å, which is dominated by α-helical secondary structure elements (Figure 1(a)) packing to form a very

Table 1. Data collection and processing statistics for selenomethionyl and native calexcitin from *L. pealei*

Dataset	Peak	Remote	Inflection	Native
ESRF beam line	BM16	BM16	BM16	ID14-1
λ (Å)	0.9797	0.908	0.9799	0.934
Resolution (Å)	70.0–3.0 (3.2–3.0)	62.1–3.0 (3.2–3.0)	70.0–3.0 (3.2–3.0)	48.3–1.8 (1.9–1.8)
R_{merge}[a] (%)	15.0 (54.2)	8.3 (38.4)	8.4 (47.5)	6.1 (36.9)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	99.9 (100.0)	99.1 (97.5)
Average $I/\sigma(I)$	15.3 (3.5)	14.5 (3.5)	13.8 (3.0)	22.5 (2.3)
Multiplicity	7.7 (7.9)	3.8 (3.9)	3.8 (4.0)	9.0 (3.8)

Values for the outer resolution shell of each dataset are shown in parentheses. The appreciably higher R_{merge} for the peak dataset may arise from the fact that a final "inverse beam" wedge of data was collected at this wavelength after the other MAD datasets had been collected. This has also increased the multiplicity of the peak dataset relative to the other wavelengths.

[a] $R_{merge} = \sum_h \sum_i |(I_{hi} - I_h)| / \sum_h \sum_i (I_{hi})$, where I_h is the mean intensity of the scaled observations I_{hi}.

pronounced hydrophobic core involving numerous aromatic residues. The α-helical segments (numbered 1–8) are formed by residues 8–21, 32–46, 53–73, 83–98, 105–118, 128–136, 142–152 and 162–174, respectively. There are a number of helix-loop-helix motifs in the protein and those formed by helices 1 and 2, 3 and 4, 5 and 6 and finally 7 and 8 are of greatest interest, since they form EF-hand motifs. The EF-hands associate closely together in pairs by antiparallel packing of the α-helices, the first two EF-hands forming the N-terminal domain of the molecule (residues 2–102) and the last two forming the C-terminal domain (residues 103–191) (Figure 1(b) and (c)). The C-terminal region of the molecule beyond the last EF-hand forms a large loop (residues 175–191), which contains a short region of 3_{10} helix (residues 181–185; helix 9) that is partly buried between the N and C-terminal halves of the molecule (Figure 1(d)). This loop also appears to tuck in between the two α-helices (7 and 8) of the last EF-hand motif and it contains residues, which have been implicated by mutagenesis in the GTPase activity of calexcitin.[12]

The first three helix-loop-helix motifs of calexcitin (formed by helices 1–6) are involved in binding calcium ions and the proximity of the first two calcium-binding motifs means that their two calcium ions are separated by a distance of only 12.0 Å. There are parallels with the structure of calmodulin (a four EF-hand protein) in which the first two and last two EF-hands form distinct domains that separate when calcium is bound allowing interactions with target proteins. Most commonly, a segment of the target protein binds in the hydrophobic cleft between the N and C-terminal domains of calmodulin as an α-helix.[17]

In calexcitin, the helices of the helix-loop-helix motifs in the C-terminal domain appear to be more acutely inclined to each other than those in the N-terminal domains and this is reflected in the calculated angles between the exiting and entering helices for each EF hand (shown in Table 3), which were obtained using the method of Yap *et al.*[18,19] Overall the C-terminal domain appears to be more compact than the N-terminal domain due to the helices of the first two EF-hands in calexcitin being splayed apart to a greater extent than those in the C-terminal domain. In gross terms the structure resembles an acorn with the N-terminal domain acting as the acorn cup and the more compact C-terminal domain represented by the acorn itself (Figure 2).

Calcium-binding sites

Of the four helix-loop-helix or EF-hand motifs in calexcitin, the first three have a calcium ion bound to them, which is well defined by the electron density. The last EF-hand has a number of key sequence differences, most notably a lack of carboxyl groups that prevent calcium from binding. Consequently, the loop adopts a significantly different conformation from those of the functional EF-hands in the protein. The functional EF hands in calexcitin adhere to the general sequence consensus for this motif with metal-binding residues at positions 1, 3, 5, 7 and 12. Commonly, the ninth residue in the motif forms a hydrogen bond with a water molecule that binds directly to the metal. The groups directly coordinating the metal ions in the functional EF-hands are shown in Table 3, where it can be seen that each calcium ion is bound by six protein ligands and one water molecule. Thus, in each of calexcitin's EF-hands, the calcium ion is coordinated by seven ligands with pentagonal bipyramidal coordination and Ca^{2+}–O bond lengths of approximately 2.3 Å. The three-dimensional arrangement of residues ligating each metal ion is exemplified by the first calcium-

Table 2. Refinement statistics for the native calexcitin structure

R-factor (%)	19.1
R-free (%)	24.9
Number of reflections	40,860
RMSD bond lengths (Å)	0.007
RMSD 1–3 distances (Å)	0.022
RMSD bumps (Å)	0.008
RMSD chiral tetrahedra (Å)	0.038
RMSD planar groups (Å)	0.021
Mean *B*-factor for main chain atoms (Å^2)	28.3
Mean *B*-factor for side-chain atoms (Å^2)	31.3
Mean *B*-factor for the whole chain (Å^2)	29.8

Data between 50.0 Å and 1.8 Å resolution were used with no $\sigma(F)$ cutoff.

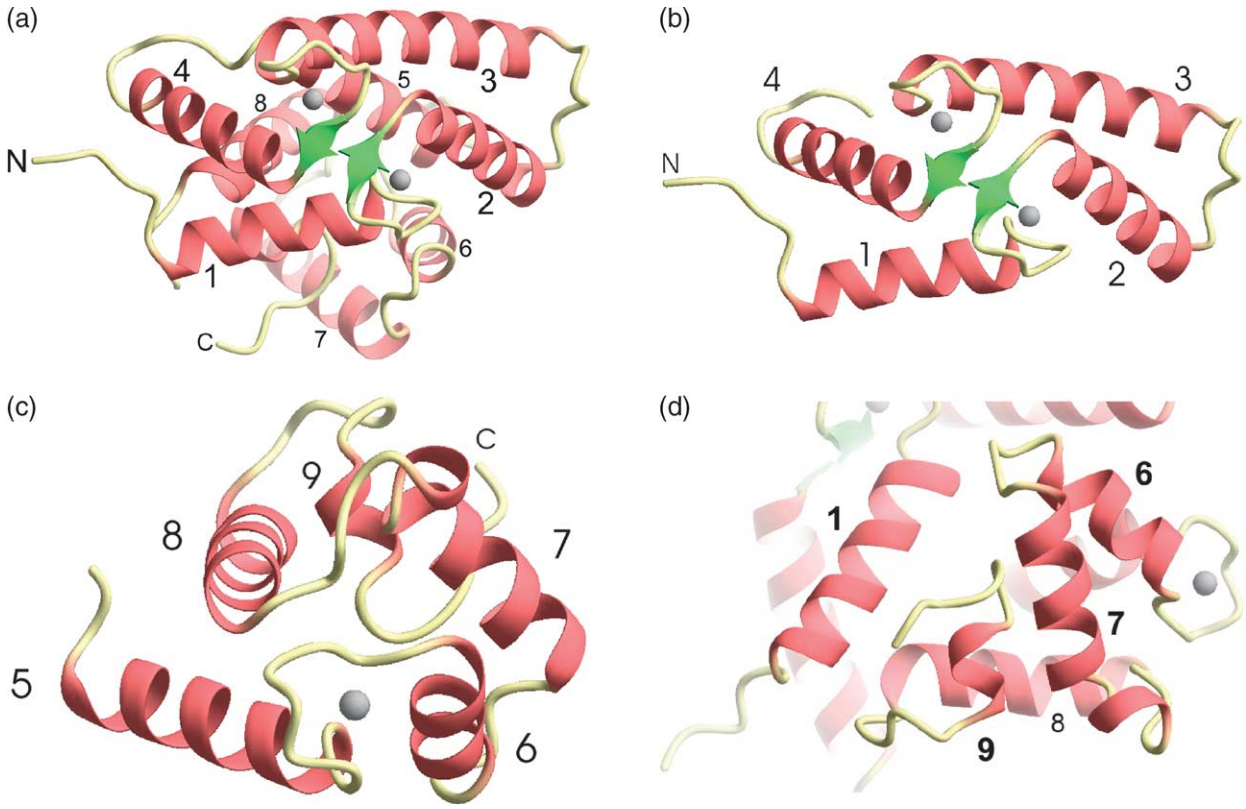

Figure 1. (a) The tertiary structure of calexcitin from *L. pealei* viewed along the pseudo 2-fold axis relating the pairs of EF-hands in the N-terminal and C-terminal domains. The N-terminal domain (residues 2–102) is shown in the foreground with the C-terminal domain behind. The calcium ions are shown in grey and the small regions of β-sheet linking the two EF-hands in the N-terminal domain are visible as green arrows. The N and C-terminal ends of the protein are indicated. (b) The N-terminal domain viewed along the pseudo 2-fold axis and (c) the C-terminal domain viewed along the same axis but from the opposite direction. The presence of only one functional EF-hand in the C-terminal domain, shown in the lower half of (c), renders this domain markedly less symmetric than the N-terminal domain. The non-functional EF-hand in the C-terminal domain, shown in the top half of (c), is followed immediately by a loop region that contains a short 3_{10} helix (helix 9). This loop region has been implicated in the GTPase activity of calexcitins by mutagenesis.[12] (d) The C-terminal 3_{10} helix (helix 9) can be seen at the interface between the N and C-terminal domains packed between helices 1, 7 and 8. The third calcium-binding site can be seen on the right.

binding site, which is shown in Figure 3 along with its electron density at 1.8 Å resolution.

The calcium-binding segments in the N-terminal domain consist of residues 23–34 and residues 74–85. There is a short region of β-sheet interaction between these two calcium-binding loops (Figure 1(b)), which results from residues 29–31 and 80–82 being in extended conformations and the occurrence of main-chain hydrogen bonds between residues Ile30 and Val81. The residues forming these two metal-binding sites are related by an approximate 2-fold axis, which runs between the two central residues forming the β-sheet interaction. The metal ions are coordinated by three acidic residues (the most C-terminal of which ligates the metal ion with both carboxyl oxygen

Table 3. Residues forming the three EF-hand motifs of calexcitin

EF-hand	1	2	3	4
Residues forming the helices	8–21, 32–46	53–73, 83–98	105–118, 128–136	142–152, 162–174
Inter-helical angle (deg.)	111, 105	95, 94	64, 70	71, 78
Metal ligands	Asp23 OD1	Asp74 OD1	Asp119 OD1	n/a
	Asn25 OD1	Asn76 OD1	Ser121 OG	n/a
	Asp27 OD1	Asp78 OD1	Asp123 OD1	n/a
	Val29 O	Gln80 O	Ile125 O	n/a
	Asp34 OD1, OD2	Glu85 OE1, OE2	Glu130 OE1, OE2	n/a
	W1	W2	W3	n/a

The most C-terminal of the motifs (4) is not involved in binding calcium but is included in this Table for completeness. For the other three EF-hands, there are seven ligands to each metal ion, including a solvent molecule indicated by W. The inter-helical angle for each EF-hand is shown for both molecules in the crystallographic asymmetric unit. n/a, not available.

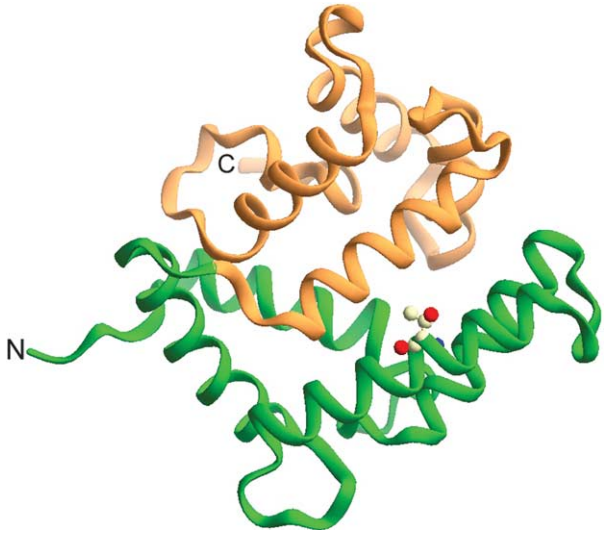

Figure 2. The domain structure of calexcitin with the N-terminal domain shown in green and the C-terminal domain coloured brown. Thr61, one site of phosphorylation by protein kinase C, is shown in ball-and-stick representation in the foreground. The other phosphorylation site (Thr188) is close to the C-terminal end of the molecule (also at the domain interface) which is largely out of view on the far side of the molecule in this view.

atoms), one asparagine and a main-chain oxygen. In addition there is a solvent ligand for the calcium ions at each of these sites that is extremely well defined by the electron density. Whilst this solvent molecule is hydrogen-bonded to the ninth residue in the first EF-hand (Glu31), it makes no such interaction in the second EF-hand where the equivalent residue in smaller and neutral (Thr82).

In the C-terminal domain the only functional EF-hand motif is formed by residues 119–130, which coordinate the metal using residue types similar to

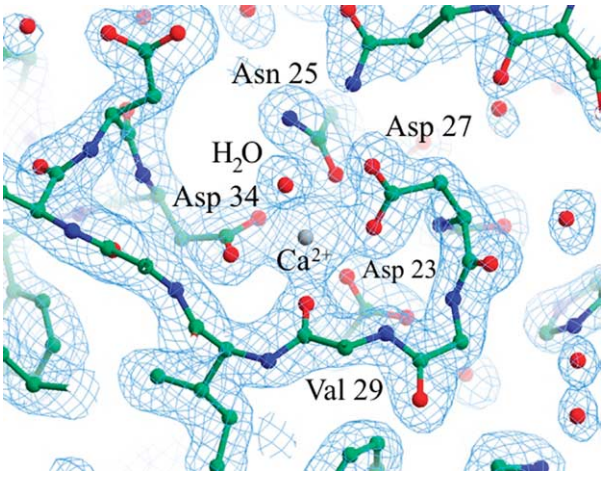

Figure 3. The first calcium-binding site in *L. pealei* calexcitin formed by the loop residues 23–34. The calcium ion is coloured grey. The $2F_o - F_c$ electron density at 1.8 Å resolution is contoured at 1.2σ and is coloured pale blue.

those in the N-terminal calcium-binding sites. The most notable exception is the involvement of a serine residue (Ser121) instead of asparagine. The ninth residue that hydrogen-bonds with the metal-bound water is Asp127. The final and non-functional EF-hand bears no sequence similarity to the other calcium-binding motifs. However, this is a region of strong sequence conservation between different calexcitins and the related protein ASCP, which also possesses a non-functional EF-hand in the equivalent position. The C-terminal end of the helix leading into the non-functional EF-hand appears to protrude significantly from the surface of the protein, forcing the ensuing loop to move away from the equivalent loop of the neighbouring functional EF-hand. This appears to disrupt the β-sheet interaction that occurs between the EF-hands in the N-terminal domain. Nevertheless, one main-chain hydrogen bond does occur between these two loops and this involves Ile126 and Val160. The helices of the two EF-hands in the C-terminal domain are related by the same topological 2-fold axis as the N-terminal domain, which passes in between the functional and non-functional Ca^{2+}-binding loops (Figure 1(c)).

There are reports that calexcitin binds magnesium ions[6,12] but the current structure does not indicate where magnesium would bind, presumably due to the lack of this metal ion in the crystallization conditions. The presence of a magnesium ion at any of the calcium sites would be indicated by negative difference density and/or a high thermal parameter for the calcium ion, but the refined structure lacks these features.

Comparison of calexcitin with ASCP

The low but significant sequence identity of 31% between calexcitin and ASCP suggested that both proteins may have similar tertiary structures and this is confirmed by the current analysis (Figure 4). The structure of ASCP[9] solved at a resolution of 2.4 Å superimposes on calexcitin with an rms deviation of 1.5 Å. Out of a total of 190 residues in the calexcitin structure and 185 in ASCP, 170 C^{α} atoms in both structures lie within 3.5 Å of each other following this superposition. The loop regions forming the calcium-binding sites in the N-terminal domains of both proteins superimpose very closely. However, the structures appear to diverge in the vicinity of the loop connecting the first and second EF-hand motifs where there are also significant local differences in conformation. The structures diverge to an even greater extent in the C-terminal domain where helices 6 and 7 are displaced by approximately 4 Å from their equivalents in ASCP. In contrast helices 5 and 8 of calexcitin superimpose very closely with the ASCP structure. The C-terminal loop of calexcitin (residues 175–191), that tucks back into the buried core of the protein between helices 7 and 8, and the N-terminal domain of the protein, is also present in ASCP and adopts a broadly similar conformation including the short

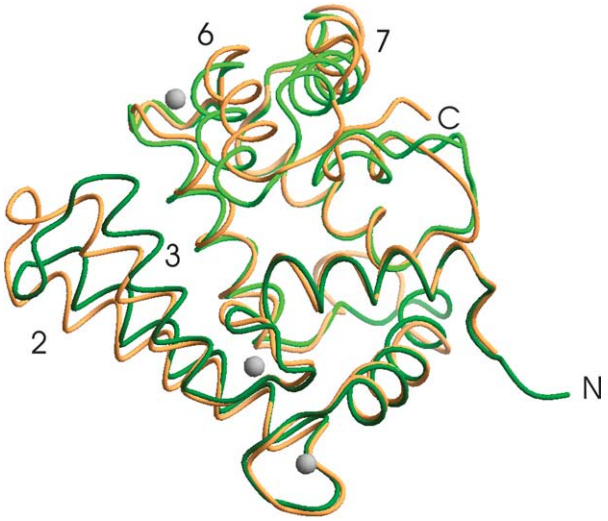

Figure 4. A superposition of calexcitin (shown in green) with the structure of Amphioxus sarcoplasmic calcium binding protein (ASCP; shown in pale brown). The structures are more similar in the N-terminal domains (shown in dark green for calexcitin) than in the C-terminal domains (shown in light green for calexcitin) where there are marked differences in the positions of helices 6 and 7 (shown at the top).

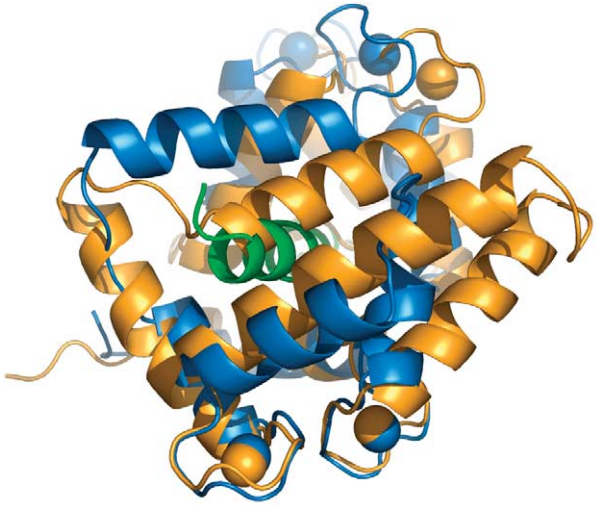

Figure 5. A superposition of calexcitin (shown in yellow) with the structure of a peptide-bound form of calmodulin[20] (shown in dark blue). The peptide bound to calmodulin can be seen as the green α-helical segment between the two domains. The domains of calexcitin would need to separate and rotate in order for it to bind target proteins, since the putative peptide-binding channel is blocked by helix 5 at one end and the C-terminal loop of the protein at the other end.

region of 3_{10} helix (helix 9). The extreme C-terminal ends of the two proteins diverge significantly as do the loops linking the N and C-terminal domains of each protein.

Comparison of calexcitin with calmodulin

Superposition of the calcium-binding loops in the N-terminal domain of calexcitin with those of a peptide-bound calmodulin in the "closed" form[20] (pdb code 1vkr) reveals strong similarities in the tertiary structures of the N-terminal domains of both proteins (Figure 5). However, there are some appreciable differences in the vicinity of the loop linking helices 2 and 3 causing it to adopt positions that are approximately 10 Å apart in the two proteins. This arises partly from the fact that helices 2, 3 and 4 are significantly longer in calexcitin than they are in calmodulin. The C-terminal domains of both proteins exhibit greater differences, most noticeably in the positions of helices 5 and 8. However, helices 6 and 7 superimpose much more closely and it appears that a 30° rigid-body rotation of the C-terminal domain would allow helices 5 and 8 to adopt positions similar to those of the peptide-bound form of calmodulin. This would require some flexibility in the domain linker region of calexcitin (residues 99–104), which is one part of the structure with higher than average *B*-factors. Calmodulin lacks the C-terminal loop region of calexcitin containing the 3_{10} helix (helix 9). Interestingly, this region of calexcitin appears to block one end of the peptide-binding site of calmodulin, suggesting that this loop may undergo major conformational changes to allow calexcitin to bind to target proteins. The other end of the peptide-binding site is blocked by helix 5 emphasising that binding of calexcitin to target proteins must also involve rotation of the domains if the binding mode is similar to that of calmodulin.

Conservation and invariance in the calexcitin family

Sequence comparison of calexcitins from different species suggests that they share the three functional and one non-functional EF-hand sequence motifs that are found in the *L. pealei* structure (Figure 6). Whilst most of the highly conserved or invariant residues occur in the α-helices and calcium-binding loops, the sequences are poorly conserved in the non-functional EF-hand region in the loop linking helices 7 and 8. One of the striking features of the alignment is that the residues in and adjacent to the C-terminal 3_{10} helix in *L. pealei* calexcitin form a group of strongly conserved amino acids (residues 182–187). This emphasises the structural or functional importance of this region of the molecule, which is close to the third EF-hand and has been implicated in calexcitin's GTPase activity.[12] A number of other amino acids are invariant across the family and inspection of the calexcitin structure shows that these are associated with the hydrophobic core of the protein, often in the vicinity of the calcium-binding sites or at the interface of the N and C-terminal domains.

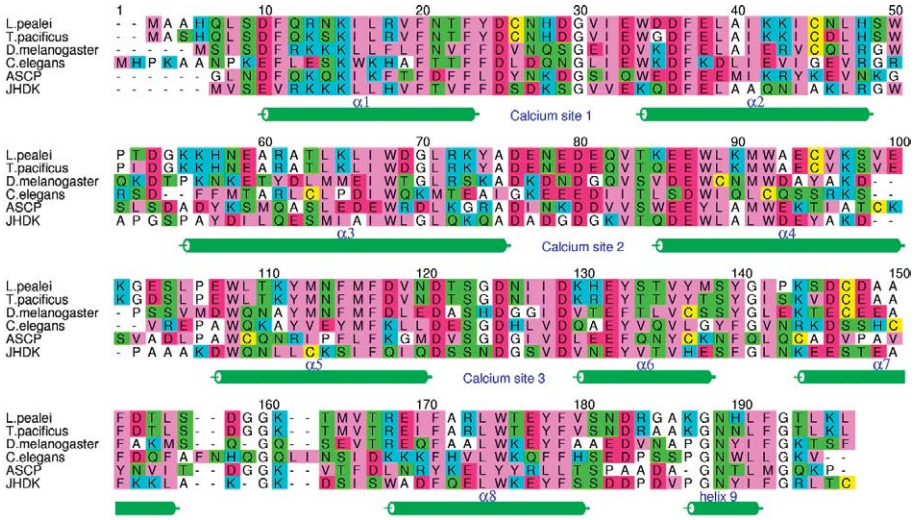

Figure 6. A sequence alignment of *L. pealei* calexcitin with other members of the family. The *L. pealei* protein shares 92%, 45% and 29% sequence identity, respectively, with the other calexcitin sequences shown from the organisms *Todarodes pacificus*, *Drosophila melanogaster* and *Caenorhabditis elegans*. The sequences of Amphioxus sarcoplasmic calcium-binding protein (ASCP) and *Bombyx mori* juvenile hormone diol kinase (JHDK) are also shown. The amino acids are coloured according to the scheme: cyan = basic, red = acidic, green = neutral-polar, pink = bulky hydrophobic, white = Gly, Ala and Pro, and yellow = Cys. The secondary structure elements and calcium-binding sites of the *L. pealei* protein are indicated below the alignment.

Interestingly, a number of amino acid residues that interact with the C-terminal 3_{10} helix (e.g. Phe9 and Lys13) are invariant or strongly conserved, again suggesting that this is a crucial part of the molecule.

Implications for the physiological role of calexcitin

Calcium-binding studies have shown that calexcitin binds two to three Ca ions with apparent dissociation constants in the nanomolar and micromolar ranges.[6,11] Whilst the highest affinity Ca^{2+} sites are likely to be fully occupied even at the resting potential of the neuron, the lower affinity site is likely to be a regulatory or sensor-site that would be occupied only during neuronal activation. This probably triggers a conformational change and/or the binding of calexcitin to a target molecule. CD spectroscopy has shown that calcium binding causes the α-helix content of the protein to increase from 46% to 57%[6] and the latter value agrees very closely with the α-helix content of 59% observed in the crystal structure.

Since the X-ray structure reveals three well-ordered calcium sites that appear to be fully occupied by Ca^{2+}, this raises the question as to which of the sites observed in the structure fulfil a sensory or regulatory role. The second calcium-binding site (residues 74 to 85) has a number of unusual features including the lack of a hydrogen bond to the calcium-bound water molecule (see above) and an abundance of negatively charged residues in the vicinity of the calcium-binding site (Figure 6). The structure shows that most of these acidic side-chains are not matched by positively charged groups, suggesting that even in the presence of a bound calcium ion, the excess of negatively charged groups might cause some instability in this part of the structure. These features may confer this site with a weaker affinity for calcium ions as well as high on and off-rates that would be appropriate for a sensory role. An alternative, and perhaps more elegant proposal, is that site 3 is the low-affinity sensor site due to its lack of a partner calcium-binding loop. In the N-terminal domain, the proximity of the two calcium-binding sites probably means that metal ion binding is a cooperative process and these might therefore be the high-affinity sites of the protein. The proximity of site 3 to the putative GTP-binding sub-domain (see later) might therefore allow the GTPase activity of calexcitin to be activated by calcium ion concentrations in the micromolar range.

Calexcitin is a substrate for PKC and phosphorylation is associated with the translocation of calexcitin to the cell membrane where its effects on potassium channels are exerted.[3] However, phosphorylation of calexcitin has been reported to inhibit its interaction with the ryanodine receptor.[8] Interestingly, phosphorylation of peptides recognised by calmodulin is known to prevent their binding to the protein.[21,22] PKC acts at consensus sequences of the form -[Ser/Thr]-X-[Arg/Lys]- and there is evidence that calexcitin is phosphorylated both at Thr61, which lies in the sequence -Thr-Leu-Lys- and at Thr188, which lies four residues from the C-terminal end of the protein in the same sequence (-Thr-Leu-Lys-).[6] Accordingly, the X-ray structure shows that the side-chains of Thr61 and Thr188 are exposed on the surface of the molecule. The Thr61 residue is approximately in the centre of

helix 3 and is roughly equidistant from the second and third calcium-binding sites (approximately 20 Å). There are no conserved basic residues in the vicinity of Thr61 that might interact with a phosphate group and the distance of this residue from the calcium-binding sites suggests that phosphorylation must exert its effects by some other means. Assuming that calexcitin interacts with target proteins in a manner similar to that of calmodulin, the proximity of Thr61 to the hydrophobic interface between the N and C-terminal domains (Figure 2) suggests that phosphorylation may alter the equilibrium between the open and closed forms of the molecule. The proximity of Thr188 to the C-terminal end of the protein (Figure 1(d)) means that it is also close to the interface between the N and C-terminal domains, and phosphorylation may also have an effect on domain opening. This residue is close to a salt-bridge formed between Asp144 and Lys190 as well as being close to His184 in the C-terminal 3_{10} helix (helix 9) and a number of other basic residues in the vicinity. Thus, phosphorylation of Thr188 could have a significant effect on the local conformation of the protein in a region that has been implicated in its GTPase activity.

There are many reports that calexcitin is a calcium-activated GTPase and recent mutagenesis studies have identified a number of residues that appear to be important in the GTPase activity.[12] These include Gly182, Phe186 and Gly187 and a number of neighbouring amino acids. This region of the calexcitin sequence occurs in the C-terminal loop region that is packed between helices 7 and 8 in the C-terminal domain and helix 1 in the N-terminal domain (Figure 1(c) and (d)). The fact that helices 7 and 8 form the non-functional EF-hand of the protein suggests that during evolution, the calcium-binding properties of this helix-loop-helix motif may have been sacrificed to provide a GTP-binding sub-domain.

The potential GTP-binding residues, as identified by mutagenesis,[12] occur in or immediately adjacent to the region of the 3_{10} helix (helix 9) in the C-terminal loop. The non-polar nature of these residues suggests that they do not have direct catalytic roles but instead probably have structural roles upon GTP binding. Inspection of the native structure of calexcitin does not reveal a suitable pocket in which GTP could bind in the vicinity of this loop. However, the 3_{10} helix occurs at the extremely hydrophobic interface between the two domains of the protein and numerous studies of the related protein calmodulin suggest that this region of the molecule undergoes marked conformational changes upon binding of target molecules. Thus, it is likely that calexcitin would have to undergo a conformational change in order to bind GTP.

There is clear evidence in the structure that part of the buried inter-domain core of calexcitin, adjacent to the 3_{10} helix has a marked propensity for conformational heterogeneity. This is reflected in differences between the conformations of certain side-chains in the two calexcitin molecules making up the crystallographic asymmetric unit. The most pronounced heterogeneity is in the side-chain of the invariant residue Trp169, the indole ring of which adopts diametrically opposite orientations in the two molecules (Figure 7). This appears to force the surrounding side-chains of residues Leu107, Tyr110 (invariant), Met114 (invariant) to adopt significantly different conformations in the two molecules of the asymmetric unit. Visual inspection and use of the program GRASP[23] demonstrates that the buried cores of both calexcitin molecules within the crystallographic asymmetric unit have a number of cavities formed largely by hydrophobic amino acids and in both monomers Trp169 is involved in forming one of these buried cavities. This residue, with its marked conformational heterogeneity, seems to lie in a fluid region of the structure that may be pivotal to conformational changes that take place upon ligand binding to calexcitin. NMR studies have indicated that substantial changes affecting aromatic side-chains, especially those of tryptophan residues, occur upon calcium binding to the protein.[6]

The surface representation of the molecule shown in Figure 8 demonstrates that calexcitin has pronounced crevices between the N and C-terminal domains, which lead into the hydrophobic core of the molecule. Whilst the electrostatic potential at the surface of the molecule is dominated by the acidic residues forming the calcium-binding sites there is a region of significant positive potential in the immediate vicinity of the C-terminal 3_{10} helix (helix 9). This arises from a number of conserved basic residues, in particular Lys13 and Arg16, which reside in helix 1 of calexcitin and interact with residues in the 3_{10} helix. These residues form a positively charged groove between the N and C-terminal domains that might provide an initial docking site for the tri-phosphate group of a GTP molecule (Figure 8).

The amino acid sequence of *L. pealei* calexcitin is 39% identical with juvenile hormone diol kinase from the tobacco hawkmoth, *Manduca sexta*.[13,14] Calexcitin has 36% identity with the recently cloned and sequenced JHDK from *Bombyx mori*[15] and it has similar overall sequence identities with JHDK from other species of moth. The sequence similarity between calexcitin and JHDK extends along the whole length of the two proteins, which are both of similar length, and includes the key residues of the three metal-binding motifs and the C-terminal loop region (Figure 6). The fact that juvenile hormone diol kinase is known to use MgATP or MgGTP to phosphorylate an extremely hydrophobic substrate (juvenile hormone) as the end-point of its metabolism hints at an as yet undiscovered function of calexcitin.

In vivo, certain proteins can activate GTPases by many orders of magnitude by providing substrate-binding/catalytic residues in the functional complex.[24] Thus, the low GTPase activity of some calexcitins *in vitro* may simply reflect on the lack of

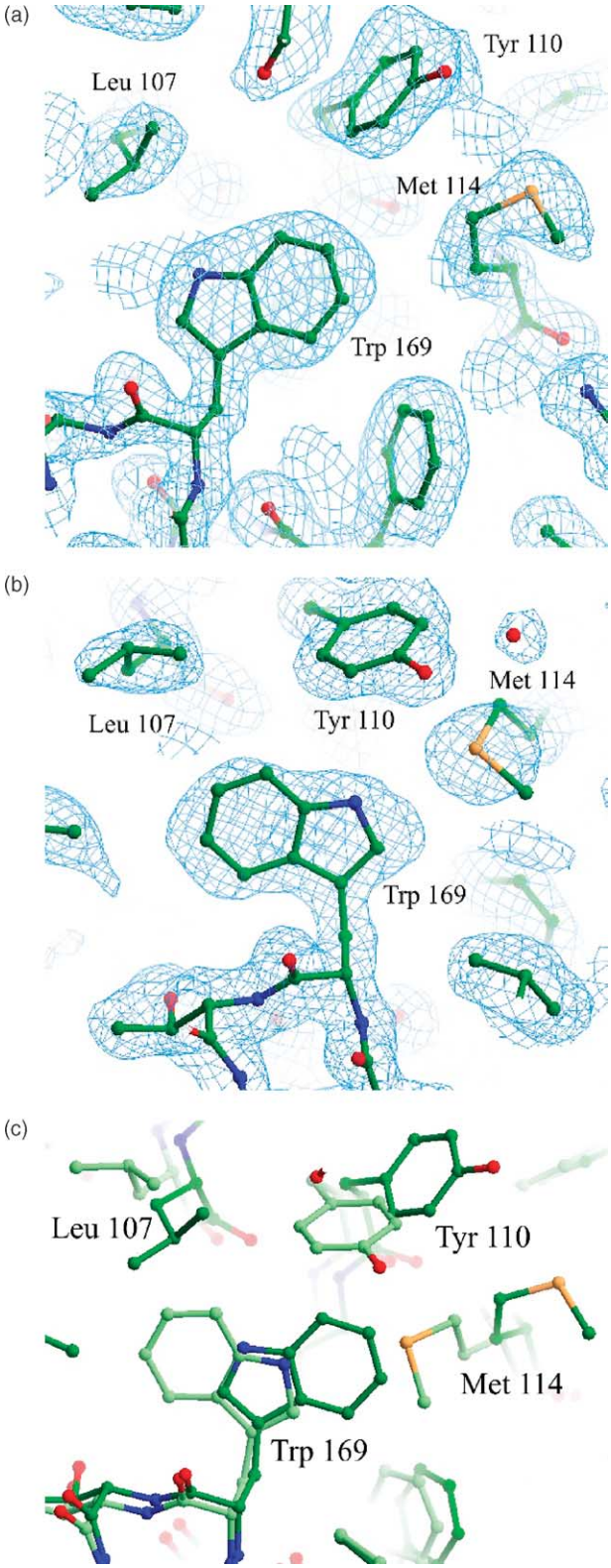

Figure 7. (a) The orientation of Trp169 in the first molecule in the asymmetric unit and (b) its conformation in the second monomer. The $2F_o - F_c$ electron density at 1.8 Å resolution is contoured at 1.1σ and is coloured pale blue. (c) A superposition of Trp169 in both subunits of the asymmetric unit emphasising the conformational heterogeneity that affects the surrounding residues (Leu107, Tyr110 and Met114). The first and second monomer are coloured in dark and light green, respectively.

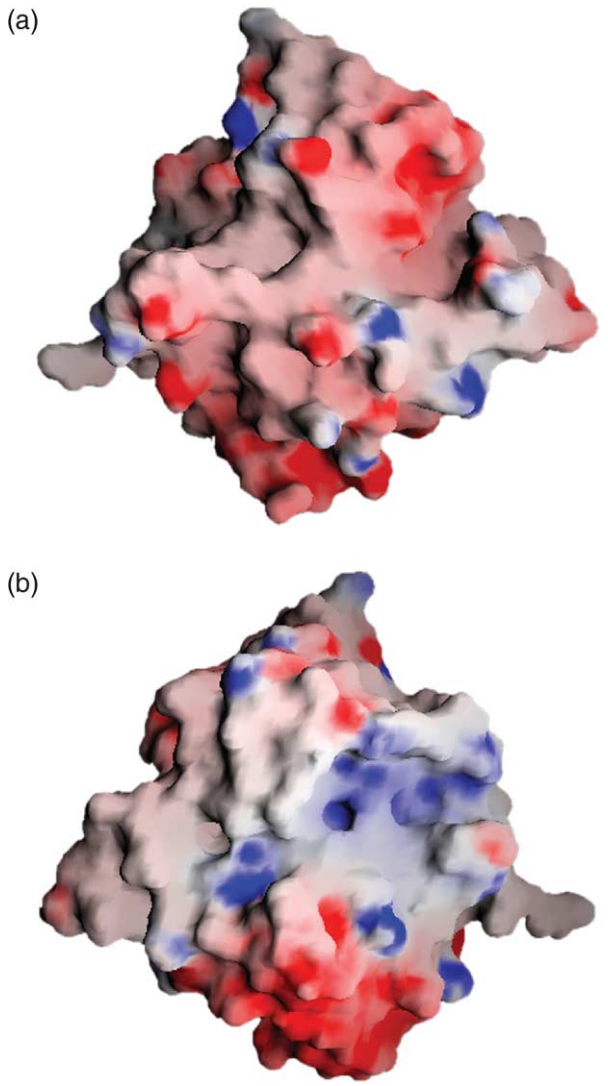

Figure 8. The electrostatic surface potential of *L. pealei* calexcitin. (a) The protein oriented as in Figure 2 emphasises the acidic regions of the protein (red) formed by the calcium-binding loops roughly at the poles of the molecule in this view. The pronounced cavities leading to the hydrophobic interior of the molecule are apparent in the equatorial regions. (b) The molecule viewed from the opposite direction and revealing a positively charged equatorial groove that might provide an initial docking site for the tri-phosphate group of a GTP molecule. The upper right-hand reaches of this channel are formed by residues of the 3_{10} helix (helix 9) implicated in calexcitin's GTPase activity.

an activator protein. Instead, it may be that calexcitin is actually an activator of a partner GTPase and acts in much the same way as calmodulin, which is known to regulate GTPases by binding to the GTPase activator protein GAP.[25] However, the ability of calmodulin to bind certain hetero-aromatic drug molecules in its inter-domain cleft[26] and its sequence similarity with juvenile hormone diol kinase indicates that calexcitin may indeed be able to bind GTP on its own. Unlike these other proteins, calexcitin is a neuron-specific

molecule and the studies reported here form the ground work to further define the structural basis of calexcitin's various roles as well as the signalling proteins that interact with it in learning and memory pathways with potential therapeutic implications.

Materials and Methods

Expression and crystallisation of selenomethionine-substituted calexcitin

Expression and crystallisation of native *L. pealei* calexcitin has been described.[27] For expression of selenomethioninyl calexcitin, a pET11a construct was used with *Escherichia coli* BL21(DE3) cells as the expression host. Cells were grown at 37 °C to mid-log phase in LB medium with 50 μg/ml of ampicillin and were pelleted by centrifugation prior to re-suspension in M9 minimal medium supplemented with 0.4% (w/v) glucose and ampicillin at the above concentration. The cells were then grown at 37 °C for 45 min to use up remaining amino acids prior to the addition of each of the 20 standard amino acids to final concentrations of 40 mg/l with the exception of methionine, which was replaced by L-selenomethionine at the same final concentration. The cultures were supplemented with the vitamins riboflavin, niacinamide, pyridoxine and thiamine, each at a final concentration of 1 mg/l and were shaken at 37 °C for another 20 min. The cultures were then induced by addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 1 mM and grown overnight at 37 °C. Purification of the selenomethionine calexcitin was based on a method reported previously for the native enzyme.[27] The cells were lysed by sonication and subsequent purification involved ammonium sulphate-precipitation. Ammonium sulphate was added to 60% saturation to remove impurities by ultra-centrifugation and calexcitin itself was then precipitated from the supernatant by addition of ammonium sulphate to 90% saturation. Subsequent purification involved gel-filtration using a Superdex 75 column (Amersham Biosciences) and the calexcitin fractions were pooled and concentrated to 3 mg/ml in 20 mM Tris (pH 8.0), 1 mM dithiothreitol and 1 mM $CaCl_2$ for storage at −20 °C. Analysis of the protein by electrospray mass spectrometry showed that the molecular mass of the largest peak corresponded to within 2 Da of the predicted molecular mass of fully selenomethionine-labelled calexcitin lacking the N-terminal methionine, which was previously shown to be removed from the native protein.[27]

Crystals of selenomethionine calexcitin were grown by the vapour diffusion method, which involved mixing 5 μl of protein at a concentration of 2.5 mg/ml with 5 μl of well solution consisting of 35% (w/v) PEG 4000, 100 mM sodium citrate (pH 6.0–7.0) and 200 mM ammonium acetate. Small rectangular needle-like crystals of up to 100 μm in their longest dimension and approximately 30 μm and 10 μm in width and thickness appeared within three weeks. These were cryo-protected by addition of glycerol to a final concentration of 30% (v/v) in the mother liquor prior to mounting in 0.2 mm Litholoops (Molecular Dimensions) on 18 mm CrystalCap pins (Hampton Research). The crystals were flash-cooled in liquid ethane and stored under liquid nitrogen.

X-ray data collection and structure analysis

Multi-wavelength anomalous dispersion X-ray data were collected from one crystal at the European Synchrotron Radiation Facility (ESRF, Grenoble) using beam line BM16 equipped with a MARCCD detector and an Oxford Cryosystems cryo-cooler, which maintained the crystal at a temperature of 100 K. The diffraction quality was improved by re-annealing the crystal in the cryo-stream and reducing the size of the beam to 80 μm. Data to a resolution of 3.0 Å were collected at the selenium fluorescence peak wavelength (0.9797 Å), followed by a remote dataset (0.908 Å) and lastly an inflection point dataset (0.9799 Å). The crystal was found to belong to the $P2_12_12_1$ space group and have unit cell dimensions of $a=47.3$ Å, $b=70.0$ Å and $c=135.2$ Å with two calexcitin molecules per asymmetric unit. The data were processed with MOSFLM[28] and were scaled using SCALA and SCALEIT of the CCP4 suite;[29] the relevant statistics are shown in Table 1. Nine out of the ten expected selenium sites were located by use of the program SOLVE[30] and the tenth site was located from an anomalous difference Fourier. Phase calculation using MLPHARE[29] and density modification using DM[31] yielded a map with promising α-helical regions. Segments of polypeptide were fitted to the map using the automatic map-fitting program MAID[32] and this allowed helical segments of the distantly related protein ASCP to be superimposed by visual inspection using TURBO-FRODO (Biographics, Marseille). Thus, one monomer with the correct sequence of *L. pealei* calexcitin was built into the map and a copy of this model was then fitted by visual inspection to the electron density for the second subunit in the asymmetric unit. The model was then refined using the native calexcitin dataset[27] which extends to 1.8 Å resolution (Table 1) by use of the programs SHELX-97[33] and CNS,[34] interspersed with rounds of manual rebuilding and solvent fitting using TURBO-FRODO. The *R*-free reflection set was chosen using SHELXPRO[33] in thin resolution shells to avoid bias between the working and test sets due to the presence of non-crystallographic symmetry. Except for the initial round of refinement in which rigid body constraints were applied to both monomers in the asymmetric unit, non-crystallographic symmetry (NCS) restraints were not used. The refinement statistics of the final model are shown in Table 2. Figures showing molecular structures were prepared using CueMol† and PyMol‡. The coloured sequence alignment was prepared using Malign[35] and Alscript.[36] The electrostatic surface potential of the protein was calculated and displayed using GRASP.[23]

Protein Data Bank accession codes

The coordinates and structure factors have been deposited with the RCSB Protein Data Bank and assigned accession numbers 2ccm and 2ccmsf, respectively.

Acknowledgements

We thank the BBSRC for a studentship to G.D.E.B. and the Wellcome Trust for infrastructure support.

† www.cuemol.org/en/
‡ www.pymol.org

We are indebted to the ESRF (Grenoble) for beam time and travel support.

References

1. Nelson, T. J., Collin, C. & Alkon, D. L. (1990). Isolation of a G protein that is modified by learning and reduces potassium currents in Hermissenda. *Science*, **247**, 1479–1483.
2. Alkon, D. L. & Nelson, T. J. (1990). Specificity of molecular changes in neurons involved in memory storage. *FASEB J.* **4**, 1567–1576.
3. Nelson, T. J., Cavallaro, S., Yi, C.-L., McPhie, D., Schreurs, B. G., Gusev, P. A. *et al*. (1996). Calexcitin: a signaling protein that binds calcium and GTP, inhibits potassium channels, and enhances membrane excitability. *Proc. Natl Acad. Sci. USA*, **93**, 13808–13813.
4. Nelson, T. J., Yoshioka, T., Toyoshima, S., Fan, Y.-F. & Alkon, D. L. (1994). Characterisation of a GTP-binding protein implicated in both memory storage and inter-organelle vesicle transport. *Proc. Natl Acad. Sci. USA*, **91**, 9287–9291.
5. Nelson, T. J. & Alkon, D. (1995). Phosphorylation of the conditioning-associated GTP-binding protein cp20 by protein kinase C. *J. Neurochem.* **65**, 2350–2357.
6. Gombos, Z., Jeromin, A., Mal, T. K., Chakrabartty, A. & Ikura, M. (2001). Calexcitin B is a new member of the sarcoplasmic calcium-binding protein family. *J. Biol. Chem.* **276**, 22529–22536.
7. Cavallaro, S., Meiri, N., Yi, C.-L., Musco, S., Ma, W., Goldberg, J. & Alkon, D. L. (1997). Late memory-related genes in the hippocampus revealed by RNA fingerprinting. *Proc. Natl Acad. Sci. USA*, **94**, 9669–9673.
8. Nelson, T. J., Zhao, W.-Q., Yuan, S., Favit, A., Pozzo-Miller, L. & Alkon, D. H. (1999). Calexcitin interaction with neuronal ryanodine receptors. *Biochem. J.* **341**, 423–433.
9. 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 Å resolution. *J. Mol. Biol.* **229**, 461–471.
10. Vijay-Kumar, S. & Cook, W. J. (1992). Structure of a sarcoplasmic calcium-binding protein from *Nereis diversicolor* refined at 2.0 Å resolution. *J. Mol. Biol.* **224**, 413–426.
11. Ascoli, G. A., Luu, K. X., Olds, J. L., Nelson, T. J., Gusev, P. A., Bertucci, C. *et al*. (1997). Secondary structure and Ca^{2+}-induced conformational change of calexcitin, a learning-associated protein. *J. Biol. Chem.* **272**, 24771–24779.
12. Nelson, T. J., Quattrone, A., Kim, J., Pacini, A., Cesati, V. & Alkon, D. L. (2003). Calcium-regulated GTPase activity in the calcium-binding protein calexcitin. *Comp. Biochem. Physiol. ser. B*, **135**, 627–638.
13. Maxwell, R. A., Welch, W. H. & Schooley, D. A. (2002). Juvenile hormone diol kinase I. Purification, characterization and substrate specificity of juvenile hormone-selective diol kinase from *Manduca sexta*. *J. Biol. Chem.* **277**, 21874–21881.
14. Maxwell, R. A., Welch, W. H., Horodyski, F. M., Schegg, K. M. & Schooley, D. A. (2002). Juvenile hormone diol kinase II. Sequencing, cloning and molecular modeling of juvenile hormone-selective diol kinase from *Manduca sexta*. *J. Biol. Chem.* **277**, 21882–21890.
15. Li, S., Zhang, Q.-R., Xu, W.-H. & Schooley, D. A. (2005). Juvenile hormone diol kinase, a calcium-binding protein with kinase activity, from the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **35**, 1235–1248.
16. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.
17. Rhoads, A. R. & Friedberg, F. (1997). Sequence motifs for calmodulin recognition. *FASEB J.* **11**, 331–340.
18. Yap, K. L., Ames, J. B., Swindells, M. B. & Ikura, M. (1999). Diversity of conformational states and changes within the EF-hand superfamily. *Proteins: Struct. Funct. Genet.* **37**, 499–507.
19. Yap, K. L., Ames, J. B., Swindells, M. B. & Ikura, M. (2001). Vector geometry mapping: a method to characterise the conformation of helix-loop-helix calcium binding proteins. *Methods Mol. Biol.* **173**, 317–324.
20. Mirzoeva, S., Weigand, S., Lukas, T. J., Shuvalova, L., Anderson, W. F. & Watterson, D. M. (1999). Analysis of the functional coupling between calmodulin's calcium binding and peptide recognition properties. *Biochemistry*, **38**, 3936–3947.
21. Minakami, R., Jinnai, N. & Sugiyama, H. (1997). Phosphorylation and calmodulin binding of the metabotropic glutamate receptor subtype 5 (mGluR5) are antagonistic *in vitro*. *J. Biol. Chem.* **272**, 20291–20298.
22. Nakajima, Y., Yamamoto, T., Nakayama, T. & Nakanishi, S. (1999). A relationship between protein kinase C phosphorylation and calmodulin binding to the metabotropic glutamate receptor subtype 7. *J. Biol. Chem.* **274**, 27573–27577.
23. Nicholls, A., Sharp, K. & Honig, B. (1991). Protein folding and association–insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281–296.
24. Helmreich, E. J. M. (2004). Structural versatility of small GTPases Can it explain their functional versatility?. *Biol. Chem.* **385**, 1121–1136.
25. Bahler, M. & Rhoads, A. (2002). Calmodulin signaling via the IQ motif. *FEBS Letters*, **513**, 107–113.
26. Vandonselaar, M., Hickie, R. A., Quail, J. W. & Delbaere, L. T. J. (1994). Trifluoperazine-induced conformational change in Ca^{2+}-calmodulin. *Nature Struct. Biol.* **1**, 795–801.
27. Beaven, G. D. E., Erskine, P. T., Wright, J. N., Mohammed, F., Gill, R., Wood, S. P. *et al*. (2005). Crystallisation and preliminary X-ray diffraction analysis of calexcitin from *Loligo pealei*–a neuronal protein implicated in learning and memory. *Acta Crystallog. sect. F*, **61**, 879–881.
28. Leslie, A. G. W. (2006). The integration of macromolecular diffraction data. *Acta Crystallog. sect. D*, **62**, 48–57.
29. Collaborative Computing Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
30. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallog. sect. D*, **55**, 849–861.
31. Cowtan, K. (1994). DM: an automated procedure for phase improvement by density modification. In *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, vol. 31, pp. 34–38, Daresbury Laboratory, Warrington, UK.

32. Levitt, D. G. (2001). A new software routine that automates the fitting of protein X-ray crystallographic electron density maps. *Acta Crystallog. sect. D*, **57**, 1013–1019.
33. Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high-resolution refinement. *Methods Enzymol.* **277**, 319–343.
34. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
35. Johnson, M. S., Overington, J. P. & Blundell, T. L. (1993). Alignment and searching for common protein folds using a data bank of structural templates. *J. Mol. Biol.* **231**, 735–752.
36. Barton, G. J. (1993). ALSCRIPT–a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40.

Edited by I. Wilson

(Received 15 November 2005; received in revised form 22 January 2006; accepted 23 January 2006)
Available online 8 February 2006