Calcium binding activity of the epidermal growth factor-like domains of the apicomplexan microneme protein EtMIC4

Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular & Biochemical Parasitology 143 (2005) 192–199
MOLECULAR & BIOCHEMICAL PARASITOLOGY

Calcium binding activity of the epidermal growth factor-like domains of the apicomplexan microneme protein EtMIC4

Javier Periz [a], Andrew C. Gill [a], Vroni Knott [b], Penny A. Handford [b], Fiona M. Tomley [a,*]

[a] *Institute for Animal Health, Division of Molecular Biology, Compton, Newbury, Berkshire RG20 7NN, UK*
[b] *Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK*

Received 2 March 2005; accepted 2 June 2005
Available online 1 July 2005

Abstract

Microneme proteins are secreted from apicomplexan parasites during invasion of host cells and they play crucial roles in parasite–host cell adhesion. EtMIC4 is a 240 kDa transmembrane protein from *Eimeria tenella* that contains 31 tandemly arranged epidermal growth factor (EGF), like repeats within its extracellular domain. The majority of these repeats have calcium binding (cb) consensus sequences. Little is known about cbEGFs in apicomplexan parasites but their presence in microneme proteins suggests that they may contribute to parasite–host interactions. To investigate the potential role of cbEGFs we have expressed and correctly refolded a cbEGF triplet from EtMIC4 ($cbEGF_{7–9}$) and demonstrated that this triplet binds calcium. Circular dichroism spectroscopic analysis of $cbEGF_{7–9}$ demonstrates that the molecule undergoes a gradual change in conformation with increasing levels of calcium. In the presence of calcium, the triplet becomes resistant to proteolytic degradation by a variety of proteases, a characteristic feature of cbEGF repeats from higher eukaryotic proteins, such as fibrillin, suggesting that calcium binding induces the formation of a rigid conformation. Moreover, mass spectrometric mapping of the cleavage sites that are protected by calcium shows that these sites are located both close to and distant from the calcium binding sites, indicating that protection is not due to steric hindrance by calcium ions, but rather due to the overall conformation adopted by the triplet in the presence of calcium. Thus, the tandemly-arranged cbEGF repeats within EtMIC4 provide a mechanism whereby, in the calcium-rich extracellular environment, the molecule could adopt a protease-resistant, rigid structure that could favour its interaction with host cell ligands.

Keywords: Apicomplexa; Host cell invasion; Adhesion; Motility; Microneme

1. Introduction

Eimeria tenella is an intracellular protozoan parasite that belongs to the phylum Apicomplexa, a large group of organisms that cause important diseases of medical and veterinary importance, such as malaria, toxoplasmosis, neosporosis, cryptosporidiosis and, for *Eimeria* species, coccidiosis. Apicomplexan parasites share a mechanism of invasion that is linked to their unique form of locomotion, gliding motility, and is dependent on the sequential secretion of material from specialised secretory organelles called micronemes and rhoptries [1]. Micronemes store collections of soluble and transmembrane proteins that are secreted from the apical end of the parasite onto the parasite surface, where they interact with host cell ligands [2]. The energy required for invasion is provided by an actin–myosin motor that lies just under the parasite plasma membrane and which is linked to the parasite surface through dynamic multi-protein complexes that include the C-terminal domains of the surface-bound transmembrane microneme proteins [3,4].

Several apicomplexan microneme proteins have been described that contain sequences with homology to epidermal growth factor (EGF)-like domains. A small number of these contain specialised EGFs that are predicted to bind calcium ions. These domains share low overall amino acid homol-

Abbreviations: EtMIC4, *Eimeria tenella* microneme protein 4; cbEGF, calcium binding epidermal growth factor like domain

* Corresponding author. Tel.: +44 1635 577 276; fax: +44 1635 577 263.
E-mail address: Fiona.Tomley@bbsrc.ac.uk (F.M. Tomley).

0166-6851/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2005.06.003

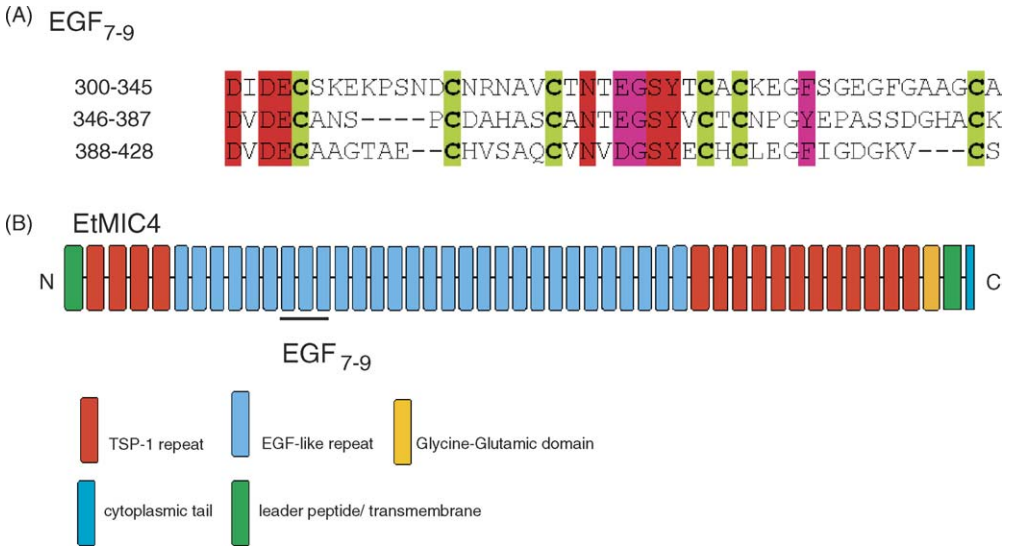

Fig. 1. Structure of cbEGF$_{7–9}$. (A) Amino acid sequence. For each cbEGF, cysteine residues involved in disulphide bond formation are embolded, and the canonical calcium binding consensus sequences are highlighted in red. Residues for hydrophobic packing of cbEGF pairs are highlighted in magenta. (B) Overall domain structure of EtMIC4.

ogy but maintain the classical EGF pattern of six cysteine residues together with a short calcium binding (cb) consensus sequence (D/N)-X-(D/N)-E/Q-X$_m$(D/N*)X$_n$(Y/F), where *m* and *n* are variable and the asterisk indicates possible β-hydroxylation [5]. EGF-like repeats are widely distributed in nature and cbEGFs within extracellular proteins are involved in mediating protein–protein and protein–carbohydrate binding interactions [6]. They are also structurally important in functionally diverse processes, such as maintenance of extracellular matrix architecture, blood coagulation and cholesterol uptake [7].

Although cbEGF domains are present in a number of secreted, microneme proteins, little is known about the function of these domains and their calcium binding activity in the biology and disease processes of apicomplexan parasites. In this study, we have expressed and purified a prototypical cbEGF triplet (Fig. 1) from the *E. tenella* EtMIC4 protein, a 240 kDa multi-modular, transmembrane protein that contains 31 tandemly arrayed EGFs [8]. We have used the recombinant protein to investigate the calcium binding properties of this cbEGF triplet, which gives an insight into the likely functional role of cbEGFs within EtMIC4 and other apicomplexan microneme proteins.

2. Experimental procedures

2.1. Materials

Unless otherwise stated, all chemicals and reagents were purchased from Sigma. The *E. coli* strain NMR554, transformed by plasmid pRep4 has been described previously [9].

2.2. Cloning and expression of cbEGF$_{7–9}$

The coding sequences for cbEGF$_{7–9}$ (amino acids 300–429) were PCR-amplified from plasmid pFT1 [8] using *Pfu* polymerase and oligonucleotide primers EGFfw$_{7–9}$ (5′-TAGTAGGGATCCATAGAAGGACGATCAGCAGTGGATATCGACGAGTGCTCCAAG-3′), which incorporates a *Bam* HI site (underlined), factor Xa cleavage site (double underlined) and two spacer amino acids (A, V) upstream of the start of the cbEGF$_7$ sequence; and EGFrv$_{7–9}$ (5′-TAGTAGAAGCTTCTATTAACTGCACACCTTTCCGTCGCC-3′), which incorporates a *Hind* III site (underlined) and two stop codons downstream of the cbEGF$_9$ sequence. The authenticity of the amplified product was confirmed by DNA sequencing and it was ligated into plasmid expression vector pQ30 (Qiagen) in frame with the N-terminal (His)$_6$ tag. Recombinant plasmid pEGF$_{7–9}$pQ30 was transformed into *E. coli* NMR554 [10], which contains the *lac* repressor on plasmid pRep4. Bacteria were grown at 37 °C with shaking in 2 l of 2xTY medium containing 25 μg ml^{-1} kanamycin and 100 μg ml^{-1} ampicillin until reaching an OD A$_{600}$ of 0.8 as determined by UV spectroscopy. Protein expression was induced by the addition of 2 mM IPTG to the culture for a period of 3 h.

2.3. Purification of the cbEGF$_{7–9}$ triplet

Bacteria were pelleted by centrifugation, then suspended in 50 ml of 6 M guanidine–HCl, 50 mM sodium phosphate, pH 6, 5 mM 2-mercaptoethanol (lysis buffer) and incubated for 2 h at room temperature. Insoluble material was removed by centrifugation at 100,000 × *g* for 1 h and the supernatant was loaded onto a nickel-sepharose column (Pharmacia).

After washing once with lysis buffer, protein was eluted from the column in lysis buffer containing 20 mM EDTA. The eluate was adjusted to 0.1 mM DTT and 0.1 M Tris–HCl, pH 8.3, incubated for 2 h at room temperature then dialysed overnight against 50 mM Tris–HCl, pH 8.3. After dialysis, the buffer was acidified to pH 3 and the protein was further purified by reverse-phase (RP)-HPLC as previously described [11]. The protein was then lyophilised and suspended at 2 mg ml^{-1} in an in vitro refolding buffer consisting of 50 mM Tris, pH 8.3, 3 mM L-cysteine, 0.3 mM L-cystine and incubated for 72 h at 4 °C. The refolded protein was re-purified by RP-HPLC, lyophilised, then suspended in 50 mM Tris–HCl, pH 7.5, 100 mM NaCl, 1 mM $CaCl_2$ and digested with bovine factor X (Amersham Pharmacia) overnight at 37 °C at a 1:1000 enzyme to protein ratio by weight to remove the $(His)_6$ tag. The cleaved $cbEGF_{7–9}$ was purified over a MonoQ column (Pharmacia) using a gradient of 0–0.5 M NaCl in 50 mM Tris–HCl, pH 7.3 and subjected to a final round of RP-HPLC. The concentration of $cbEGF_{7–9}$ was calculated from its theoretical molar extinction coefficient and absorbance at 280 nm wavelength.

2.4. Circular dichroism

Recombinant $cbEGF_{7–9}$ was suspended at 0.6 mg ml^{-1} in 20 mM Tris–HCl, pH 7.5 containing $CaCl_2$ at 1 mM, 250 μM, 62 μM or 7 μM or containing 10 mM EGTA. Circular dichroism spectra were recorded from 260 to 190 nm by use of a Jasco J-710 spectropolarimeter. A cell of pathlength 0.2 mm was used and all spectra were acquired at room temperature. The final CD spectra represent the sum of 20 individual scans. Control samples included $cbEGF_{7–9}$ in 20 mM Tris–HCl, pH 7.5 without added $CaCl_2$. CD spectra of all buffers were recorded and subtracted from the final spectra of the triplet.

2.5. Protease digestions

Recombinant $cbEGF_{7–9}$ was suspended at 0.5 mg ml^{-1} in 100 mM NaCl, 50 mM Tris–HCl, pH 8 supplemented with 1 mM $CaCl_2$, 1 mM $MgCl_2$ or 10 mM EGTA and samples were incubated at room temperature for 30 min to equilibrate. Enzymes were added to samples at a 1:4 (elastase) or 1:12 (trypsin or chymotrypsin) enzyme to protein ratio by weight, and incubated for 1 h at 37 °C. Digestion was terminated by the addition of Laemmli reducing SDS-PAGE loading buffer and heating to 95 °C for 5 min. All samples were analysed by SDS-gel electrophoresis on 4–15% acrylamide gradient gels and proteins were visualised by staining with Coomassie brilliant blue.

2.6. Mass spectrometry

All mass spectrometry was performed by use of a Quattro II tandem quadrupole instrument (Micromass UK Ltd., Warrington, UK) in the Proteomics Facility at IAH Compton. Samples were analysed by online HPLC–MS. Briefly, protein samples were diluted to approximately 1 pmol μl^{-1} in HPLC running buffer A, where buffer A was 95:5 water:acetonitrile and 0.05% trifluoroacetic acid and buffer B was 5:95 water:acetonitrile and 0.05% trifluoracetic acid. Approximately 20 pmol was injected onto a home made protein desalting trap and was washed with excess buffer A. Bound components were passed to a homemade capillary HPLC column packed with Jupiter C_{18} reversed phase sorbent (3.5 μm bead size, 150 Å pore size, column i.d. 180 μm, length ∼10 cm) and were eluted with an increasing gradient of buffer B. The eluent was passed directly to the continuous flow nanospray source of the mass spectrometer, operating in positive ion mode. The instrument acquired full scan mass spectra (m/z 500–2100) every 4.5 s.

3. Results

We have studied the effects of calcium binding on a triplet of cbEGFs derived from the microneme protein EtMIC4. A triplet was chosen for study to ensure that the cbEGF domains were in as near-native protein context as possible and therefore exhibit the calcium binding properties of the intact protein [12]. The number was limited to three because of the difficulty of refolding, in vitro, a structure rich in disulphide bonds, but also to allow comparisons to be made with other triplet EGFs that are already well characterised [13,14]. Since a large amount of protein was required, we used a bacterial expression system that has proven to be a powerful tool to express and refold fragments of fibrillin-1, an extracellular matrix protein that contains 43 cbEGFs [9]. $cbEGF_{7–9}$ of EtMIC4 was selected because this triplet contains three well conserved consensus sequences for calcium binding and also possesses several cleavage recognition sites for the serine proteases used during protection assays in this study.

3.1. Purification and refolding of $cbEGF_{7–9}$

Time course studies were carried out to determine the optimum time for harvesting bacterial cells expressing $cbEGF_{7–9}$. At 3 h post-induction, recombinant protein expression was at a maximum with no apparent degradation of the product (data not shown). After denaturation in 6 M guanidine–HCl, the protein was partially purified by metal ion chromatography. Peak fractions were pooled, made 0.1 M with respect to DTT in 0.1 M Tris–HCl pH 8 to reduce the protein, then acidified and further purified by RP-HPLC chromatography (Fig. 2A). The sample was concentrated and refolded in vitro using an oxido-shuffling system (see materials and methods). Since there are no previously published data about the refolding of these particular cbEGFs, we carried out in vitro refolding trials in the presence or absence of calcium. After refolding in the presence of 5 mM calcium, the resulting RP-HPLC chromatogram of the refolded oxidised protein (Fig. 2B) was dominated by a single well-resolved peak, which eluted ear-

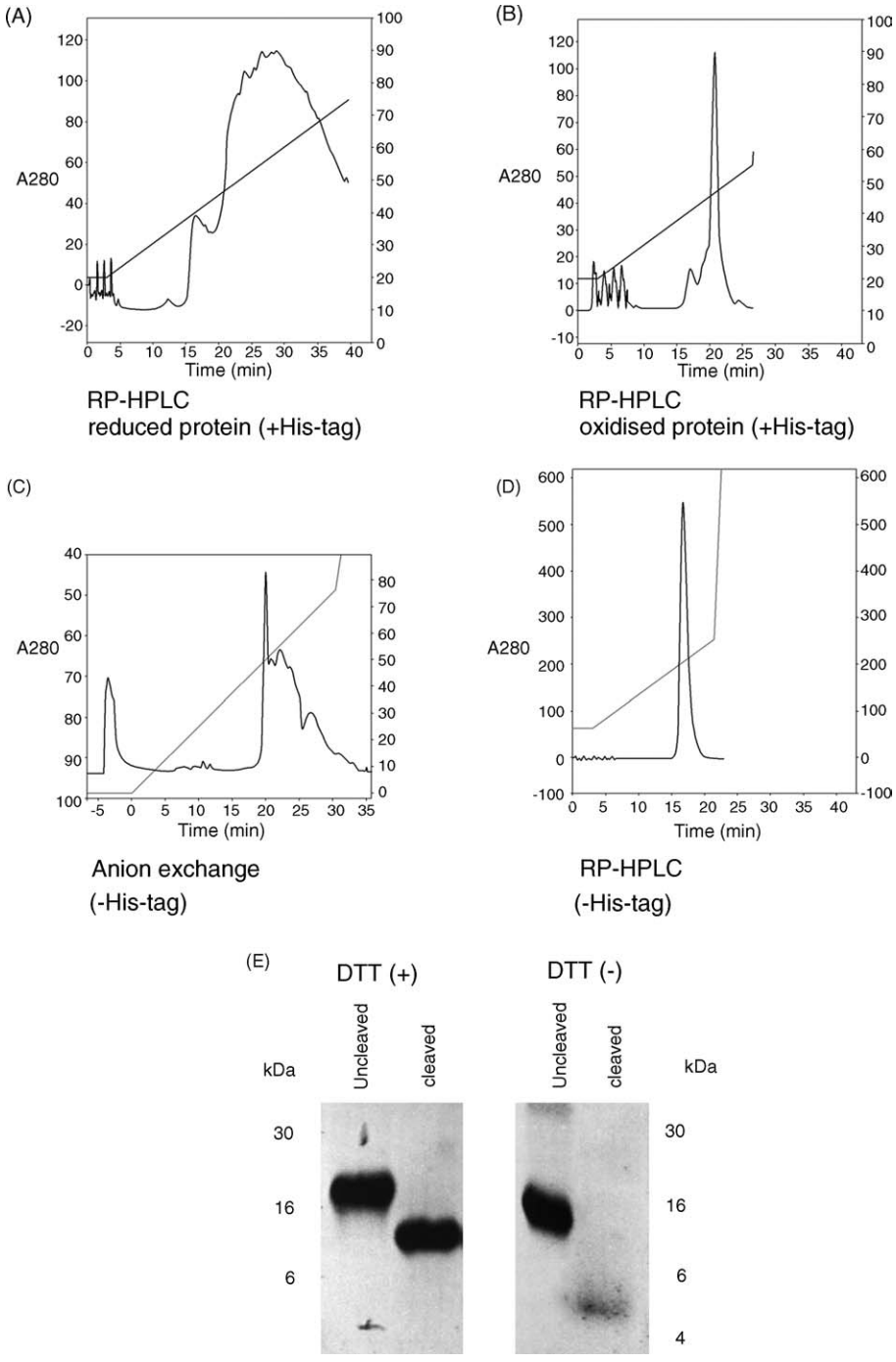

Fig. 2. Purification and refolding of cbEGF$_{7–9}$. After dialysis and acidification, reduced cbEGF$_{7–9}$ was purified by one round of reverse phase chromatography (A) then subjected to in vitro refolding, followed by a second round of reverse phase chromatography (B). Comparison of the elution profiles in A and B clearly shows the dramatic tightening of the elution peak and earlier elution time that occurred after oxidation, which indicates successful refolding. After digestion with Factor X, cbEGF$_{7–9}$ was separated from the cleaved $(His)_6$ tag by anion exchange (C), then subjected to one further round of reverse phase chromatography (D). Samples of uncleaved and cleaved cbEGF$_{7–9}$ polypeptides were examined by SDS-PAGE (E), to confirm the complete removal of the $(His)_6$ tag.

lier than the broad peak obtained with the reduced-denatured protein (Fig. 2A). This is similar behaviour to that observed for cbEGF fragments from other proteins [13,14] and strongly suggests that the oxidised protein has a more hydrophilic and homogenous conformation, consistent with what would be expected of native cbEGFs.

The $(His)_6$ tag was removed by Factor Xa cleavage and the cleaved protein separated from the tag by anion exchange chromatography (Fig. 2C). Release of the tag was confirmed by an observed decrease in size of the recombinant protein, as monitored by SDS-PAGE (Fig. 2E). Recombinant cbEGF$_{7–9}$ was finally concentrated by RP-HPLC (Fig. 2D)

and lyophilised. A final yield of ∼9 mg was obtained and the molecular mass of the oxidised protein was measured to be 13,359.5 ± 0.18 Da by electrospray mass spectrometry (data not shown). This agrees well with the theoretical mass of the polypeptide, assuming all disulphide bonds are formed, of 13,360.38 Da. Overall, during the refolding process cbEGF$_{7–9}$ behaved in a very similar manner to other recombinant cbEGF triplets [15,16]. Taken together, the measured molecular mass, the requirement for calcium in the oxido-shuffling process and the observed shifts in chromatogram profile and SDS-PAGE migration of cbEGF$_{7–9}$ under reduced and oxidised conditions suggested that the triplet was refolded into its native form.

3.2. *Calcium protects cbEGF$_{7–9}$ from proteolytic digestion*

To evaluate whether the putative calcium binding EGF repeats indeed bound calcium ions, we investigated the susceptibility of refolded cbEGF$_{7–9}$ to several classes of serine proteinase that are present in the intestinal lumen or potentially associated with inflammatory processes during *E. tenella* infection. Previous studies have demonstrated that calcium binding to cbEGFs domains affords protection against proteolysis [17,18]. In the presence of 1 mM calcium, cbEGF$_{7–9}$ was partially protected against digestion by trypsin, chymotrypsin and elastase, whereas in the presence of 10 mM EGTA, substantial degradation of cbEGF$_{7–9}$ occurred within 20 min (Fig. 3A–C). Resistance to proteolysis was attributable to calcium binding to the protein since the inclusion of 1 mM $MgCl_2$ in the tryptic digestion buffers gave very little protection (compare Fig. 3D with 3A). This is consistent with the behaviour of other well characterised cbEGF modules, such as those from fibrillin [19]. Protection against trypsin digestion is particularly significant since the activity of this enzyme is enhanced by calcium [20].

Protected cleavage sites within cbEGF$_{7–9}$ were investigated in detail by mass spectrometry to determine whether specific areas of the polypeptide were particularly protected against digestion. Tryptic cleavages sites can be clearly defined, since this enzyme has a well characterised specificity (trypsin cleaves C-terminal to lysine and arginine residues unless they are followed by a proline residue). Our triplet contains five susceptible tryptic cleavage sites: Lys9, Arg18, Lys34, Lys90 and Lys128. Fig. 4 shows the mass spectrometric analysis of the cbEGF prior to reduction. Under such conditions, the molecule remains intact despite evidence of protease activity, since it is held together by disulphide bonds. In the absence of calcium ions, at least three cleavage events can be detected, as evidenced by sequential increases of ∼18 Da, caused by addition of H and OH at either side of the cleaved bond (Fig. 4C). This produces species of measured molecular masses 13,377, 13,393 and 13,414 Da. The major species corresponds to protein in which two bonds have been cleaved.

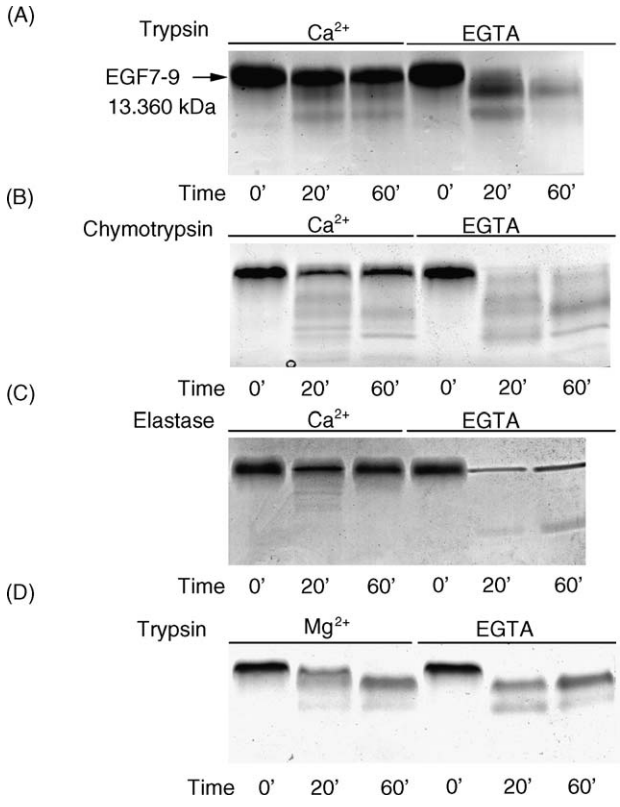

Fig. 3. Protease protection assays. Refolded cbEGF$_{7–9}$ was incubated with the serine proteinases trypsin, chymotrypsin or elastase in the presence of calcium (1 mM, panels A–C), magnesium (1 mM, panel D) or with EGTA (10 mM, all panels). Even after 60 min of incubation, in the presence of calcium, intact cbEGF$_{7–9}$ could be seen for all enzyme treatments indicating that calcium protects against proteolysis. The protective effect of calcium is specific since substitution with magnesium was not protective.

To investigate which sites were cleaved predominately by trypsin in the absence of calcium, we reduced the disulphide bonds in the digested cbEGF triplet and reanalysed the sample by mass spectrometry. We found peptides liberated by cleavages C-terminal to residues Lys9, Arg18 and Lys34 (Fig. 5). We found no evidence for cleavage after Lys90 or Lys128, despite the fact that these represent trypsin consensus sites. These data indicate that these sites are intrinsically protected from protease digestion even in the absence of calcium.

In the presence of calcium, mass spectrometric analysis of non-reduced protein indicated that the majority of molecules remained uncleaved. Analysis of digested samples, treated with trypsin in the presence of calcium, also indicated partial cleavage in the N-terminal region of the triplet. Similar analyses of chymotrypsin and elastase treated samples yielded data that were more problematic to interpret. Chymotrypsin cleaves preferentially C-terminal to large hydrophobic residues, but has a number of secondary cleavage sites. Elastase cleaves adjacent to small, uncharged amino acids, such as alanine and valine. The cbEGF triplet has multiple susceptible sites to both chymotrypsin and elastase cleavage and, coupled with the complexity of disulphide bonds,

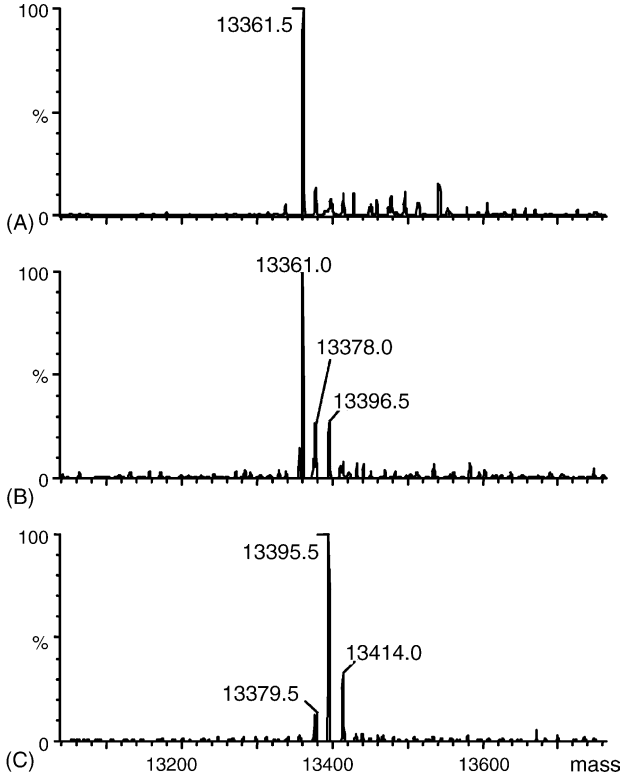

Fig. 4. Deconvoluted electrospray mass spectra of cbEGF$_{7–9}$. Electrospray mass spectra were acquired for protein in the absence of protease (A), in the presence of protease and calcium (B) and in the presence of protease and EGTA (C). All mass spectra were deconvoluted to convert *m/z* information to a mass scale. The measured mass of the oxidised refolded polypeptide (A) agrees well with its theoretical mass of 13360 Da, assuming all disulphide bonds are made. After incubating with trypsin in the presence of 1 mM calcium (B), the majority of the polypeptide remains undigested although some increased signals are evident to the right of the main peak corresponding to sequential additions of 18 Da indicating up to two cleavages in some molecules. After incubating with trypsin in the presence of EGTA (C), the mass of the most abundant species is 13393.5 Da, which corresponds to peptides bearing two cleavages. Also present are signals corresponding to 1 and 3 cleavages.

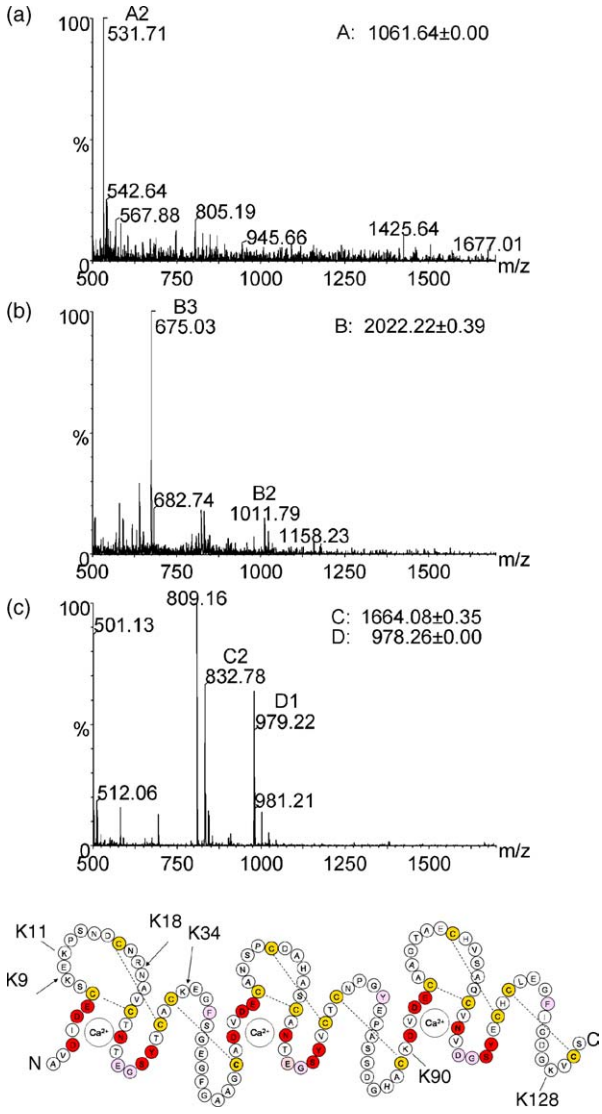

Fig. 5. Electrospray mass spectra of peptides liberated after tryptic digestion of cbEGF$_{7–9}$ in the presence of EGTA and subsequent reduction of disulphide bonds. Peptides were separated by online capillary HPLC and analysed by electrospray mass spectrometry. Some peptides were of low abundance or co-eluted with other peptides, accounting for increased numbers of peaks. In each spectrum, relevant peaks are annotated with the peptide label and the number of charges associated with the ion. (a) Mass spectrum of the peptide spanning residues 10–18 (measured mass 1061.64 Da, calculated mass 1061.46 Da); (b) mass spectrum of peptide spanning residues 1–18 (measured mass 2022.22 Da, calculated mass 2021.88 Da); (c) mass spectrum of peptides spanning residues 1–9 (measured mass 978.26 Da, calculated mass 978.43 Da) and residues 19–34 (measured mass 1664.08 Da, calculated mass 1663.66 Da). Other peaks within spectrum (c) correspond to fragments of the peptide 1–9 occurring in the mass spectrometer and confirm the identity of this peptide. The cartoon indicates tryptic cleavage sites.

the location of protected cleavage sites could not be determined. However, we performed mass spectrometric analysis of non-reduced samples, in the presence and absence of calcium, after treatment with either chymotrypsin or elastase. These data demonstrated conclusively that, in the presence of calcium, the cbEGF triplet is significantly protected from proteolytic digestion. In the absence of calcium, smaller proteolytic fragments were generated (data not shown).

3.3. Calcium binding causes a change in conformation of cbEGF$_{7–9}$

Our results from protease protection assays suggest that calcium confers protection against proteolysis but that certain areas of the cbEGF triplet are inherently partially protease resistant, even in the absence of calcium. The data also suggest that calcium binding imparts significant protection to proteases, not just in the direct region of the calcium binding site but also in regions that are distant from the binding site. NMR studies have shown significant conformational change in other cbEGF domains on binding calcium [21] and we wanted to assess whether our cbEGF triplet undergoes a conformational rearrangement on calcium binding that can account for the increased resistance to proteases.

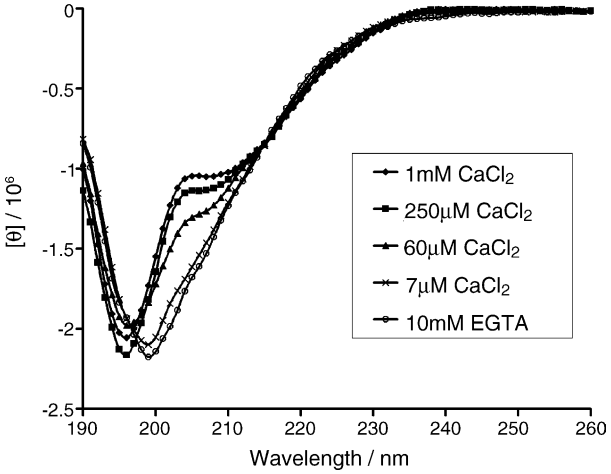

Fig. 6. Far UV circular dichroism spectra of refolded cbEGF repeats in the presence of (upwards at 200 nm) 10 mM EGTA, 7 μM $CaCl_2$, 60 μM $CaCl_2$, 250 μM $CaCl_2$ and 1 mM $CaCl_2$ all in 20 mM Tris pH 7.5 All CD spectra were corrected by subtraction of spectra acquired for buffer alone and were smoothed.

The effects of calcium ions on the global folding of $cbEGF_{7–9}$ were investigated by comparing far UV CD spectra of the triplet in the presence of different concentrations of $CaCl_2$ at pH 7.5. CD spectra are shown in Fig. 6. In the presence of EGTA, to chelate endogenous calcium ions, the CD spectrum has a minimum at ∼198 nm and is characteristic of a protein in predominately random coil structure. This is consistent with NMR determinations of other EGF domains in the absence of calcium, which possess little secondary structure. In the presence of increasing concentrations of calcium ions the CD spectrum gradually changes; there is a blue shift to the minimum to ∼195 nm, and an increased ellipiticity at around 205 nm. A loose isodichroic point is evident at around 197 nm. The final CD spectrum, in the presence of 1 mM $CaCl_2$ is difficult to interpret in terms of particular secondary structural components, however, it is indicative of a significantly more ordered structure. In addition, the spectral changes show clearly that chelation of calcium ions alter the conformation of the cbEGF triplet. The isodichroic point suggests that as the levels of calcium are increased, the structure of the protein shifts from one conformation to the other without intermediates.

4. Discussion

We have shown that cbEGFs from the EtMIC4 protein of *E. tenella* are able to bind calcium and that the presence of calcium in the environment determines their conformation. Both CD and proteolysis analyses suggest that $cbEGF_{7–9}$ and, by inference, the extracellular region of EtMIC4 that contains 31 tandem EGFs, behaves in a manner similar to other proteins containing multiple cbEGFs [22]. In the presence of calcium, all the cbEGFs in EtMIC4 could potentially adopt a more extended, rigid conformation and be protected from proteolysis and even in the absence of calcium the organisation of the EGFs may confer significant structural protection against proteases. Since the invasive stages (sporozoites and merozoites) of *E. tenella* have to survive within the protease-rich environment of the chicken intestine, the protection of a secreted/surface-bound microneme protein against proteolysis may be crucial to its function.

Measurements of cbEGFs in the presence of calcium from the solution structure of a cbEGF dimer [7] and from rotary shadowing electron microscopy measurements of purified fibrillin [20] agree that a single cbEGF extends ∼2.9 nm in length. In the absence of calcium, the molecule adopts a wider conformation that is ∼30% shorter. Thus, in the presence of extracellular levels of calcium (∼2 mM) the 31 EGFs of EtMIC4 could theoretically adopt a stable, extended structure of some 100 nm in length, with only limited degrees of flexibility in putative 'hinge' regions between cbEGF-EGF junctions and putative thrombospondin type I domains at the N- and C-termini of the extracellular domain. This would offer a distinct advantage to EtMIC4 in interacting with host cell ligands once it reaches the parasite surface. Some cbEGFs have been shown directly to exhibit low-affinity binding activity [23], so the multiple copies of cbEGFs in EtMIC4 could serve to increase the avidity of binding of these sequences to host ligands during invasion. Alternatively, the rigidified structure of the tandem EGFs could serve as a scaffold from which the four N-terminal TSP-1 domains of EtMIC4 are propelled closer towards their target ligands. Experiments are ongoing to determine the binding properties of different regions of the EtMIC4 protein. In contrast to the extracellular milieu, levels of calcium within the late secretory pathway of apicomplexan parasites are likely to be in the micromolar concentration range, although these have not been directly determined. Under these conditions, the 31 EGFs of EtMIC4 may adopt a shorter, flexible, and more compacted structure that could favour the packaging of the protein into the constricted space of the microneme organelle.

Acknowledgements

We thank Karen Billington for technical assistance and Mick Gill for help with preparation of the figures. This research was supported by an IAH studentship to J.P. and BBSRC for funding to A.C.G. and F.M.T.

References

[1] Dubremetz JF, Garcia-Reguet N, Conseil V, Fourmaux MN. Apical organelles and host-cell invasion by Apicomplexa. Int J Parasitol 1998;28:1007–13.
[2] Carruthers VB, Giddings OK, Sibley LD. Secretion of micronemal proteins is associated with toxoplasma invasion of host cells. Cell Microbiol 1999;1:225–35.
[3] Jewett TJ, Sibley LD. Aldolase forms a bridge between cell surface adhesins and the actin cytoskeleton in apicomplexan parasites. Mol Cell 2003;11:885–94.

[4] Buscaglia CA, Coppens I, Hol WG, Nussenzweig V. Sites of interaction between aldolase and thrombospondin-related anonymous protein in plasmodium. Mol Biol Cell 2003;14:4947–57.
[5] Campbell ID, Bork P. Epidermal growth factor-like modules. Curr Opin Struct Biol 1993;3:385–92.
[6] Stacey M, Chang GW, Davies JQ, et al. The epidermal growth factor-like domains of the human EMR2 receptor mediate cell attachment through chondroitin sulfate glycosaminoglycans. Blood 2003;102:2916–24.
[7] Downing AK, Knott V, Werner JM, Cardy CM, Campbell ID, Handford PA. Solution structure of a pair of calcium-binding epidermal growth factor-like domains: implications for the Marfan syndrome and other genetic disorders. Cell 1996;85:597–605.
[8] Tomley FM, Billington KJ, Bumstead JM, Clark JD, Monaghan P. EtMIC4: a microneme protein from *Eimeria tenella* that contains tandem arrays of epidermal growth factor-like repeats and thrombospondin type-I repeats. Int J Parasitol 2001;31:1303–10.
[9] Knott V, Downing AK, Cardy CM, Handford P. Calcium binding properties of an epidermal growth factor-like domain pair from human fibrillin-1. J Mol Biol 1996;255:22–7.
[10] Raleigh EA, Murray NE, Revel H, et al. McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. Nucleic Acids Res 1988;16:1563–75.
[11] Handford PA, Baron M, Mayhew M, et al. The first EGF-like domain from human factor IX contains a high-affinity calcium binding site. Embo J 1990;9:475–80.
[12] Smallridge RS, Whiteman P, Doering K, Handford PA, Downing AK. EGF-like domain calcium affinity modulated by N-terminal domain linkage in human fibrillin-1. J Mol Biol 1999;286:661–8.
[13] Whiteman P, Smallridge RS, Knott V, Cordle JJ, Downing AK, Handford PA. A G1127S change in calcium-binding epidermal growth factor-like domain 13 of human fibrillin-1 causes short range conformational effects. J Biol Chem 2001;276:17156–62.
[14] Suk JY, Jensen S, McGettrick A, et al. Structural consequences of cysteine substitutions C1997Y and C1997R in calcium-binding epidermal growth factor-like domain 30 of human fibrillin-1. J Biol Chem 2004;279:51258–65.
[15] Mayhew M, Handford P, Baron M, Tse AG, Campbell ID, Brownlee GG. Ligand requirements for Ca^{2+} binding to EGF-like domains. Protein Eng 1992;5(6):489–94.
[16] Stacey M, Lin HH, Hilyard KL, Gordon S, McKnight AJ. Human epidermal growth factor (EGF) module-containing mucin-like hormone receptor 3 is a new member of the EGF-TM7 family that recognizes a ligand on human macrophages and activated neutrophils. J Biol Chem 2001;276(22):18863–70.
[17] Reinhardt DP, Ono RN, Sakai LY. Calcium stabilizes fibrillin-1 against proteolytic degradation. J Biol Chem 1997;272(2):1231–6.
[18] McGettrick AJ, Knott V, Willis A, Handford PA. Molecular effects of calcium binding mutations in Marfan syndrome depend on domain context. Hum Mol Genet 2000;9(13):1987–94.
[19] Cardy CM, Handford PA. Metal ion dependency of microfibrils supports a rod-like conformation for fibrillin-1 calcium-binding epidermal growth factor-like domains. J Mol Biol 1998;276(5):855–60.
[20] Reinhardt DP, Mechling DE, Boswell BA, Keene DR, Sakai LY, Bachinger HP. Calcium determines the shape of fibrillin. J Biol Chem 1997;272:7368–73.
[21] Werner JM, Knott V, Handford PA, Campbell ID, Downing AK. Backbone dynamics of a cbEGF domain pair in the presence of calcium. J Mol Biol 2000;296:1065–78.
[22] Handford PA, Mayhew M, Baron M, Winship PR, Campbell ID, Brownlee GG. Key residues involved in calcium-binding motifs in EGF-like domains. Nature 1991;351:164–7.
[23] Davis SJ, Ikemizu S, Wild MK, van der Merwe PA. CD2 and the nature of protein interactions mediating cell–cell recognition. Immunol Rev 1998;163:217–36.