Structural Prediction and *In Silico* Physicochemical Characterization for Mouse Caltrin I and Bovine Caltrin Proteins

Ernesto J. Grasso, Adolfo E. Sottile and Carlos E. Coronel

Instituto de Investigaciones Biológicas y Tecnológicas (IIByT) CONICET–Universidad Nacional de Córdoba, and Instituto de Ciencia y Tecnología de los Alimentos (ICTA), Cátedra de Química Biológica, Departamento de Química Industrial y Aplicada, Facultad de Ciencias Exactas, Físicas y Naturales, Córdoba, Argentina.

ABSTRACT: It is known that caltrin (calcium transport inhibitor) protein binds to sperm cells during ejaculation and inhibits extracellular Ca^{2+} uptake. Although the sequence and some biological features of mouse caltrin I and bovine caltrin are known, their physicochemical properties and tertiary structure are mainly unknown. We predicted the 3D structures of mouse caltrin I and bovine caltrin by molecular homology modeling and threading. Surface electrostatic potentials and electric fields were calculated using the Poisson–Boltzmann equation. Several different bioinformatics tools and available web servers were used to thoroughly analyze the physicochemical characteristics of both proteins, such as their Kyte and Doolittle hydropathy scores and helical wheel projections. The results presented in this work significantly aid further understanding of the molecular mechanisms of caltrin proteins modulating physiological processes associated with fertilization.

KEYWORDS: homology modeling, threading, mouse caltrin I, bovine caltrin, hydropathy plots, electrostatic surface potentials

CITATION: Grasso et al. Structural Prediction and *In Silico* Physicochemical Characterization for Mouse Caltrin I and Bovine Caltrin Proteins. *Bioinformatics and Biology Insights* 2016:10 225–236 doi: 10.4137/BBI.S38191.

TYPE: Original Research

RECEIVED: June 06, 2016. **RESUBMITTED:** September 27, 2016. **ACCEPTED FOR PUBLICATION:** October 03, 2016.

ACADEMIC EDITOR: J. T. Efird, Associate Editor

PEER REVIEW: Six peer reviewers contributed to the peer review report. Reviewers' reports totaled 2457 words, excluding any confidential comments to the academic editor.

FUNDING: This work was supported by grants from Secretaría de Ciencia y Tecnología (SECyT), Universidad Nacional de Córdoba, and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) PIP No. 11220120100617CO, Argentina, to CEC. The authors confirm that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: ejgrasso@conicet.gov.ar

Paper subject to independent expert blind peer review. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE). Provenance: the authors were invited to submit this paper.

Published by Libertas Academica. Learn more about this journal.

Introduction

On ejaculation, sperm cells are suspended in the seminal plasma and are consequently exposed to proteins and other molecules mainly secreted by the seminal vesicles. Among these, caltrin (calcium transport inhibitor), a secretory protein detected in several species, inhibits extracellular Ca^{2+} uptake[1–4] by binding to specific regions of the sperm plasma membrane over the acrosome (mouse caltrin I and bovine caltrin) and the tail (bovine caltrin), but does not adhere to the distal portion of the sperm head or to the midpiece.[5]

The biological properties of mouse caltrin I and bovine caltrin have been thoroughly studied.[1,4,6] Mouse caltrin I and bovine caltrin are small, basic proteins (Mr 5411 and 6126, pI 8.3 and 9.5, respectively).[1,6,7] Mouse caltrin I was identified as a kazal-type trypsin inhibitor also known as SPINK3/P12,[1,8] which contains six cysteine residues that are not reactive with thiol reagents until the protein has been treated with reducing agents such as dithiothreitol.[3] Cysteines form -S-S- bridges in native mouse caltrin I, which appear to play an important role in the biological and immunological activity of this protein.[9] In fact, reduction and alkylation abolished the Ca^{2+} transport inhibitory activity.[9] The specific binding of mouse caltrin I to the sperm surface on the acrosomal region[7] suggests the existence of caltrin receptors, or precise protein–phospholipid interactions. On the other hand, it was demonstrated that the inhibitory effect of bovine caltrin on sperm Ca^{2+} uptake is promoted by interactions with phosphatidylserine (PS) and becomes an enhancer when the protein is separated from this anionic phospholipid.[10] When the anions are removed, the protein acquires a new conformation, detectable by fluorescence measurements, and becomes an enhancer of sperm Ca^{2+} uptake.[10]

At first sight, the amino acid sequence of mouse caltrin I[2] is too different from that of bovine caltrin[10] to assume a similar structure of these two proteins.[6] Nevertheless, both proteins inhibit Ca^{2+} transport in the sperm cells, sustaining the hypothetical role of modulators of acrosomal exocytosis without disturbing the capacitation process during the sperm's journey along the female reproductive tract. Thus, spermatozoa can maintain the acrosomal integrity required for interaction with the oviductal epithelium and their fertilizing potential.[11] Our rationale is that two highly different proteins from different species have a similar biological behavior and this behavior does not depend directly on the primary structure but on

properties arising from their supramolecular features defined by their physicochemical characteristics.

We used two automated systems, namely, SWISS-MODEL (http://swissmodel.expasy.org/) and I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER/), for modeling the 3D structures of mouse caltrin I and bovine caltrin proteins from their amino acid sequences.[12] Homology modeling or comparative modeling (SWISS-MODEL Workspace) relies on evolutionarily related structures (templates) to generate a structural model of the protein of interest (target). The process typically comprises the following steps: (1) template identification, (2) template selection, (3) model building, and (4) model quality estimation.[13] However, I-TASSER Suite predicts the 3D structures of proteins by threading[14] (see "Materials and methods" section for further description). Several different bioinformatics tools and available web servers were used to thoroughly analyze the physicochemical characteristics of mouse caltrin I and bovine caltrin. We tried to associate the known biological behavior of mouse caltrin I and bovine caltrin with their 3D structures and physicochemical properties in order to further understand the molecular mechanisms of caltrin proteins in modulating physiological processes associated with fertilization.

Materials and Methods

Physicochemical characterization. Amino acid sequences of mouse caltrin I and bovine caltrin were obtained from the NCBI protein database (http://www.ncbi.nlm.nih.gov/) in FASTA format and used for further analyses. Proteins were subjected to physicochemical characterization by using ExPASy ProtParam tools[15] for theoretical measurements, such as molecular weight, isoelectric point (pI), extinction coefficient, instability index, and aliphatic index. Kyte and Doolittle hydropathy scores (KD scores) and residues accessibility ratios were obtained using the ExPASy Protscale tool.[16,17]

Protein structure prediction by homology modeling and threading. We submitted the amino acid sequences of mouse caltrin I and bovine caltrin to the SWISS-MODEL Workspace (http://swissmodel.expasy.org/) and I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) servers. The SWISS-MODEL Template Library was searched in parallel with both BLAST[18] and HHblits[19] to identify templates and to obtain target-template alignments. The combined use of these two methods guarantees good alignments at high and low sequence identity levels. Models were built based on the target-template alignment using Promod-II.[20] Coordinates that are conserved between the target and the template were copied from the template to the model. Insertions and deletions were remodeled using a fragment library. Side chains were then rebuilt and the geometry of the resulting model was regularized by using a force field.[12] Model quality was assessed with the local composite scoring function QMEAN, which uses several statistical descriptors expressed as potentials of mean force: geometrical features of the model (pairwise atomic distances, torsion angles, and solvent accessibility) are compared to statistical distributions obtained from experimental structures and scores.[12]

I-TASSER searches the query by threading, using LOMETS through a nonredundant structure library to identify structural templates. LOMETS is a meta-threading method containing eight-fold-recognition programs (PPAS, EnvPPAS, wPPAS, dPPAS, dPPAS2, wdPPAS, MUSTER, and wMUSTER).[14] For each target, simulations generate a large ensemble of structural conformations, called decoys. To select the final models, I-TASSER uses the SPICKER program to cluster all the decoys based on the pairwise structure similarity and reports up to five models that correspond to the five largest structure clusters. The confidence of each model is quantitatively measured by a C-score that is calculated based on the significance of threading template alignments and the convergence parameters of the structure assembly simulations. C-score is typically in the range of [–5, 2], where a C-score of a higher value signifies a model with a higher confidence and vice versa. TM-score and Root Mean Square Deviation of atomic position (RMSD) are estimated based on C-score and protein length following the correlation observed between these qualities.[14] Finally, the predicted models for the two caltrin proteins were subjected to energy minimization using ModRefiner.[21] The energy-minimized structures were assessed using PROCHECK Ramachandran plots and ProSA-web.[22] All the modeled proteins were superimposed with the template using UCSF Chimera 1.10.[23] Secondary structures were predicted from the templates using ESPript 3.0 and from the sequences using PSIPRED v 3.3.

Structure analysis. Molecular surfaces were estimated by UCSF Chimera© 1.10.[23] The software shows solvent-excluded molecular surfaces, composed of probe (1.4 Å radius) contact, toroidal, and reentrant surface components.[24] These differ from solvent-accessible surfaces that are traced out by the probe center. Protein theoretical volume and area were also estimated by UCSF Chimera© 1.10. Molecular radius of gyration, nuclear radius, and van der Waals radius were calculated by Yasara© software.

We used DeepView (the Swiss-Pdb Viewer v 3.7) to estimate the electrostatic surface potential and electric fields of mouse caltrin I and bovine caltrin. Dielectric constants were adjusted at D80 and D4 (the most commonly used dielectric constant value, which is believed to account for electronic polarization and small backbone fluctuations)[25] for solvent and protein interior, respectively, with cutoff values of kT/e selected at 1.8 and –1.8 mV and ionic strength at 145 mM. We used two charge models: currently, the protein is assumed to be at pH 7.0 with default protonation state for all residues. In one case, only charged residues (Arg, Lys, Glu, and Asp) are taken into account and the charges are assumed to be located at the corresponding (non-H) atom positions; in the other case, atomic partial charges were used.

Results and Discussion

Mouse caltrin I and bovine caltrin modeling. By using the ProtParam tools, we computed several physicochemical properties for mouse caltrin I, bovine caltrin, and human insulin, as standard (see Table 1). The instability index computed was 51.99 and 24.86 for mouse caltrin I and bovine caltrin, respectively, and 43.05 for human insulin. A protein with an instability index lower than 40 is predicted as stable, while a value greater than 40 predicts that the protein may be unstable. The instability index is useful for storing proteins in the correct solvent. For instance, it is well known that insulin monomer (instability index = 43.05) is unstable and tends to macroscopically aggregate in aqueous solution during storage, causing loss of hormone biological activity, which is a major obstacle for developing long-term delivery formulations.[26]

Table 1. Theoretical physicochemical properties for mouse caltrin I, bovine caltrin and human insulin.

	MOUSE CALTRIN I	BOVINE CALTRIN	INSULIN (HUMAN)
Residues	57	48	74
Mr	6126	5411	8372.5
pI	9.5	8.3	15
Ins. Index (II)	51.99	24.86	43.05
Stability (t.t.)	Unstable	Stable	Unstable
Aliph. Index	75.26	71.25	88.24
GRAVY	−0.161	−0.994	−0.165
Ex. Coeff.*	3355	6990	6335
Ex. Coeff. Red.*	2980	–	5960

Notes: *Extinction coefficients are in units of M^{-1} cm^{-1}, at 280 nm measured in water.

Thus, only bovine caltrin would be suitable for long-term storage in aqueous solution.

The grand averages of hydropathy (GRAVY), which are indications of the hydrophilic and soluble behavior of proteins, were −0.161 and −0.994 for mouse caltrin I and bovine caltrin, respectively. The theoretically calculated extinction coefficient, which is in direct correlation with the cysteine, tryptophan, and tyrosine content, was 3355 M^{-1} cm^{-1} at 280 nm for mouse caltrin I, assuming that all pairs of cysteine residues form cystine, or 2980 M^{-1} cm^{-1} if all cysteine residues are reduced. For bovine caltrin, which has no cysteine but contains a tryptophan residue, it was 6990 M^{-1} cm^{-1}. The aliphatic index of both caltrin proteins showed high values of 75.26 and 71.25 for mouse caltrin I and bovine caltrin, respectively, indicating stability over a wide temperature range.[27]

The half-life is a prediction of the time it takes for half the amount of protein in a cell to disappear after its synthesis. ProtParam relies on the "N-end rule" that relates the half-life of a protein to the identity of its N-terminal residue; the prediction is given for three model organisms (human, yeast, and *Escherichia coli*).[15] Estimated half-life values were the same for both caltrin proteins: >20 hours in yeast (*in vivo*) and >10 hours in *E. coli* (*in vivo*) with the exception in mammalian reticulocytes (*in vivo*), where they were 4.4 and 1.9 hours for mouse caltrin I and bovine caltrin, respectively.

Based on the sequence previously reported,[1,6] the predicted 3D models for mouse caltrin I and bovine caltrin are shown in Figure 1. The accuracy of a comparative model is related to the percentage of sequence identity (the number of characters that match exactly between two different sequences)[28] on which it is based, correlating with the relationship between the structural and sequence similarity (the degree of resemblance between two sequences when they are compared)[28] of two proteins.

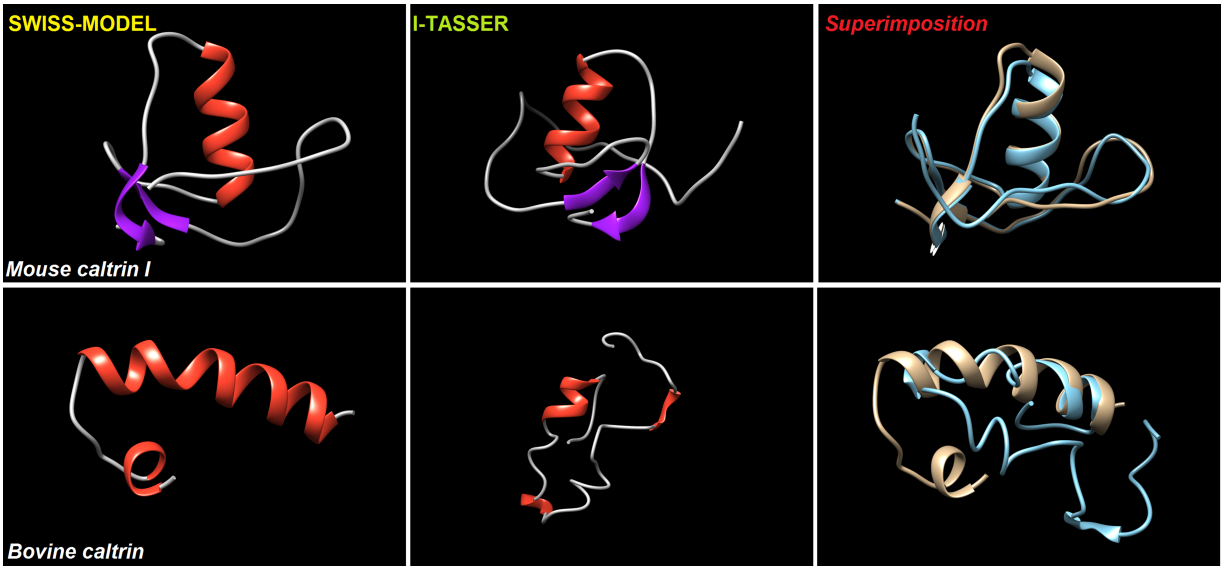


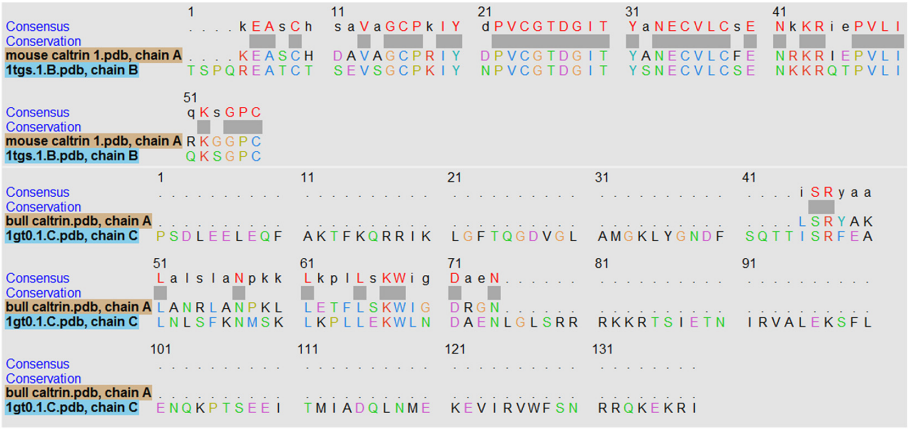
Figure 1. Homology models predicted for mouse caltrin I and bovine caltrin using SWISS-MODEL Workspace and I-TASSER Suite. Both models for each protein were superimposed using UCSF Chimera 1.10.

High-accuracy comparative models are based on more than 50% sequence identity with their templates.[29] For the mouse caltrin I model, the best template, identified by the SWISS-MODEL workspace and obtained with Blast, was the *pancreatic secretory trypsin inhibitor*, a kazal-type protease inhibitor (PDB: 1tgs.1.B), with a sequence identity of 71.15%, sequence similarity of 0.55, and coverage (the fraction of the query sequence structure that can be predicted from the template, and the plausibility of the resulting model) of 0.91. QMEAN4 score was −0.38 and the normalized QMEAN4 of the model was within the limits of Z-scores <1 (Supplementary Fig. 1). Using I-TASSER, the best identified template was also a *recombinant variant of human pancreatic secretory trypsin inhibitor*, kazal type (PDB: 1hptA), with a sequence identity of 61%, a coverage of 0.98, a C-score of 0.77, and a TM-score of 0.82 ± 0.09. Thus, the quality of both models for mouse caltrin I is good.

For the bovine caltrin model, the best template, identified by the SWISS-Model workspace and obtained with HHblits, was the *octamer-binding transcription factor 1* (PDB: lgt0.l.C) with a sequence identity of 33.33%, sequence similarity of 0.36, and a coverage of 0.63. QMEAN4 score was −2.25 and the normalized QMEAN4 of the model was within the limits of Z-scores >2. In contrast, using I-TASSER, the best identified template was the *STHK carboxy-terminal region in complex with CGMP* (PDB: 4d7sA) with a sequence identity of 27% and a coverage of 0.85. Figure 2 shows the sequence alignments obtained by using both servers.

We did not observe significant differences between the two mouse caltrin I models (SWISS-MODEL and I-TASSER), but in the case of the bovine caltrin model, a stronger template homology was achieved with the SWISS-MODEL (sequence identity of 36% vs 27%). Therefore, for further analysis, the models developed by the SWISS-MODEL Workspace were used. The models were submitted to energy minimization using ModRefiner, and the structures were assessed for both geometric and energy aspects. The ProSA-web[22] assesses the

A



B

Rank	PDB Hit	Iden1	Iden2	Cov	Norm. Z-score	Download Align.
1	1hptA	0.61	0.60	0.98	4.59	Download
2	1cgiA	0.61	0.60	0.98	1.25	Download
3	2ovoA	0.25	0.25	0.96	4.64	Download
4	1tgsI	0.66	0.65	0.98	2.16	Download
5	1cgiI	0.61	0.60	0.98	1.66	Download
6	1hptA	0.61	0.60	0.98	4.91	Download
7	1cgiI	0.60	0.60	0.96	2.38	Download
8	1hptA	0.61	0.60	0.98	1.10	Download
9	1tgsI	0.66	0.65	0.98	3.62	Download
10	1tgsI	0.66	0.65	0.98	3.91	Download

C

Rank	PDB Hit	Iden1	Iden2	Cov	Norm. Z-score	Download Align.
1	4d7sA	0.27	0.25	0.85	0.86	Download
2	3lmlA	0.12	0.08	0.83	0.53	Download
3	4ga0A	0.12	0.17	0.90	0.83	Download
4	1xp4A	0.19	0.12	1.00	0.73	Download
5	5aizA	0.25	0.23	0.83	0.59	Download
6	1hn3A	0.29	0.23	0.79	0.72	Download
7	4xrmA	0.16	0.17	0.77	0.88	Download
8	3r64A	0.00	0.23	0.10	0.64	Download
9	4ed9A	0.40	0.27	0.62	0.25	Download
10	2z5hA	0.19	0.19	0.98	1.05	Download

Figure 2. (**A**) Alignment of mouse caltrin I and bovine caltrin with their templates (pancreatic secretory trypsin inhibitor and octamer-binding transcription factor 1, respectively) using SWISS-MODEL Workspace. (**B**, **C**) Alignment of mouse caltrin I and bovine caltrin with the best 10 templates obtained using I-TASSER Suite.

protein structures by Z-scores, which are indicative of overall model quality. The Z-scores were −4.77 and −5.87 for the templates and −4.85 and −1.49 for the homology modeled proteins for mouse caltrin I and bovine caltrin, respectively (Fig. 3). The great difference observed in the Z-score value for bovine caltrin model may be due to the low sequence identity and the similarity between the template and the model. Nevertheless, when the template and bovine caltrin model were superimposed (Supplementary Fig. 2), a good correlation was observed (RMSD = 0.104 Å of C-α atoms only, calculated using the Needleman–Wunsch algorithm and BLOSUM-62 matrix), see Supplementary Table 1. This was also observed when the mouse caltrin I model and its template were superimposed (RMSD = 0.386 Å) with minimal differences in loop regions. Thus, the models passed this quality test because a good model-template superimposition tends to have less than 1 Å root mean square error for the main-chain atoms, which is comparable to the accuracy of a medium-resolution nuclear magnetic resonance (NMR) structure, or a low-resolution X-ray structure.[29] Simultaneously, the PROCHECK Ramachandran plot analysis[30] indicated a good quality for both homology modeled proteins: mouse caltrin I – 96.49% and bovine caltrin – 95.83% of residues in the most favored regions (Supplementary Fig. 3).

Secondary structures were predicted based on the templates by using ESPript 3.0 and based on the sequence of caltrin proteins by using PSIPRED v 3.3 (Fig. 4). Mouse caltrin I showed a good correlation of predicted structures between the model prediction and sequence prediction, but

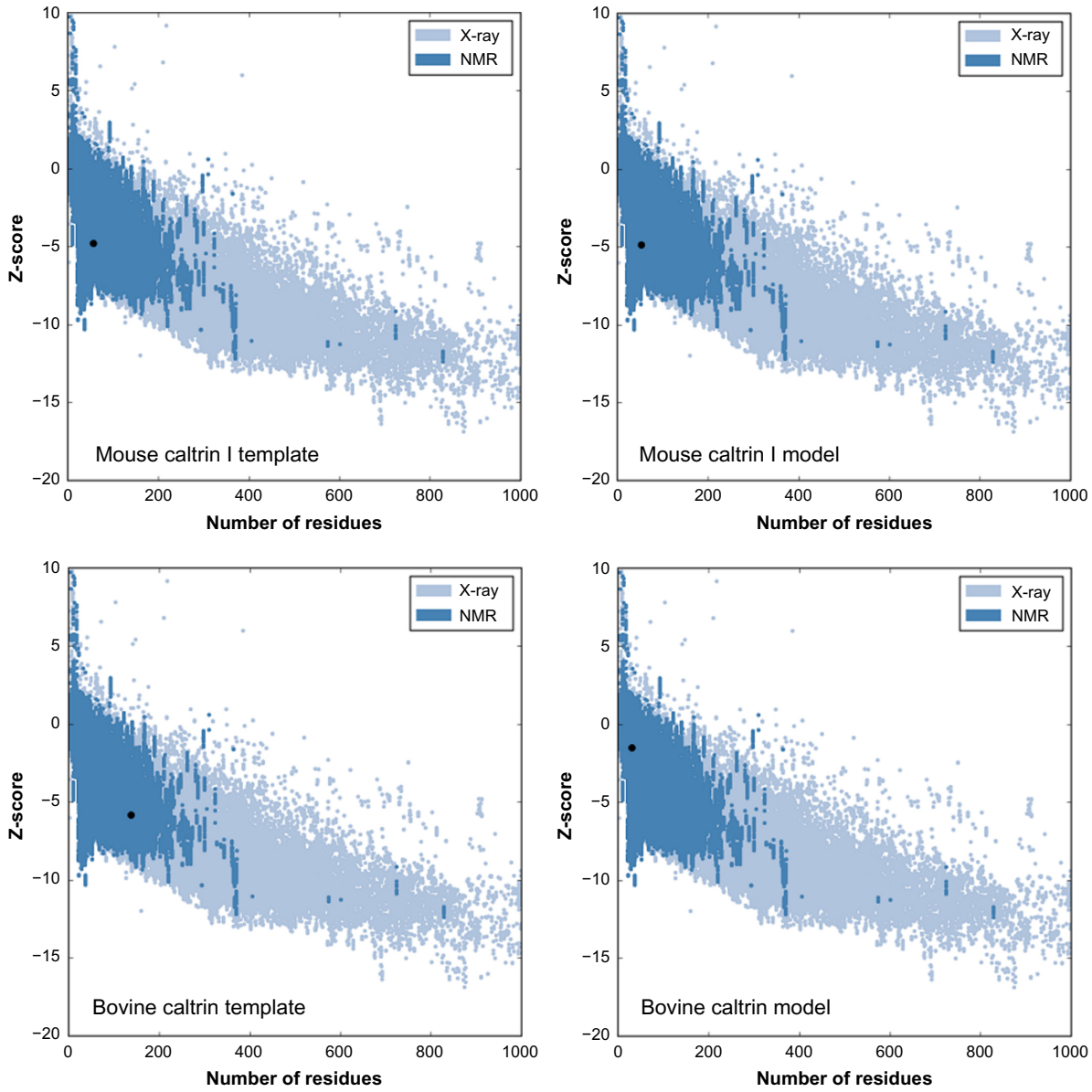


Figure 3. Z-score (highlighted as a black dot) is displayed in a plot that contains the Z-scores of all experimentally determined protein chains currently available in the Protein Data Bank. Groups of structures from different sources (X-ray and NMR) are distinguished by different colors (light blue and dark blue, respectively).

two β-sheets were not predicted using only the sequence. The group of kazal-type serine protease inhibitors is characterized by a well-preserved amino acid sequence containing three disulfide bridges.[31] Thus, the mouse caltrin I model shows three disulfide bridges (C10, C17, and C25 cross-linked to C39, C36, and C57, respectively), as also shown by Luo et al.[32] Based on the model prediction around 42% of the total amino acid residues formed secondary structures that included one α-helix of 10 residues of 17.3 Å length (ECVLCFENRK residues from positions 35 to 44, Fig. 5A), three β-sheets that involve 14 residues (PVCGT, ITY, and LIRKGG residues from positions 23 to 27, 30 to 32, and 50 to 55, respectively), a β-turn at D28G29 and random structures, including loops (33 residues, representing 58% of total structure). It was proposed that D22 and/or Y21 but not R19 were responsible for the sperm-binding site, while R19 but neither D22 nor Y21 was indispensable for trypsin inhibition.[32] The 3D model for mouse caltrin I predicted in this work has experimental support. The circular dichroism spectra of this protein in the wavelengths of 200–250 nm showed two negative bands with magnitudes of -8.1×10^3 and -1.31×10^4 deg cm^2 dmol^{-1} around 220 nm (band I) and 205 nm (band II), respectively, which suggest a considerable amount of ordered structures including a helix and a mixture of β-forms and β-turns.[32]

The predicted secondary structures of bovine caltrin, based on the template and the sequence, were also similar, but a longer α-helix was predicted based on the protein sequence (Fig. 4). Bovine caltrin, which lacks trypsin inhibitory activity,[7] has two α-helix (based on the model) stretched from S16 to A19 (9.0 Å length) and from A26 to G43 residues (28.8 Å length), and a β-turn from L21 to N23 residues (Figs. 4 and 5B). A fragment of 13 residues of bovine caltrin was studied, using circular dichroism, showing two minima at ~205 and ~220 nm and a crossover at ~200 nm.[33] Such a spectrum is characteristic of a predominantly α-helical conformation.[33] Thus, the secondary structure shown in the molecular model for bovine caltrin accounts for the CD spectra.[33]

Figure 5C–F shows the solvent-excluded and solvent-accessible molecular surfaces that were determined for the two caltrin proteins. Mouse caltrin I showed a theoretical volume of 5870 Å^3/molecule, with solvent-excluded and solvent-accessible surface areas of 3097.8 and 3889.94 Å^2, respectively. Bovine caltrin displayed a theoretical volume of 3489 Å^3/molecule, with solvent-excluded and solvent-accessible surface areas of 2315.09 and 3136.92 Å^2, respectively. Mouse caltrin I showed a gyration radius of 10.23 Å, a nuclear radius of 19.163 Å, and a van der Waals radius of 20.987 Å, which were measured using Yasara© software. Bovine caltrin showed a gyration radius of 10.291 Å, a nuclear radius of 19.107 Å, and a van der Waals radius of 20.904 Å. Although showing similar geometric values, mouse caltrin I appears to have a spherical shape, while bovine caltrin seems to be a rod-like protein.

Theoretical electrostatic surface potentials of mouse caltrin I and bovine caltrin. Electrostatic interactions in macromolecular systems arise from the presence of local charges, the polarization stemming from the non-spherical distribution of electron density around atoms, the redistribution of electrons caused by local electrical fields, and the reorientation of polar groups in the solute and

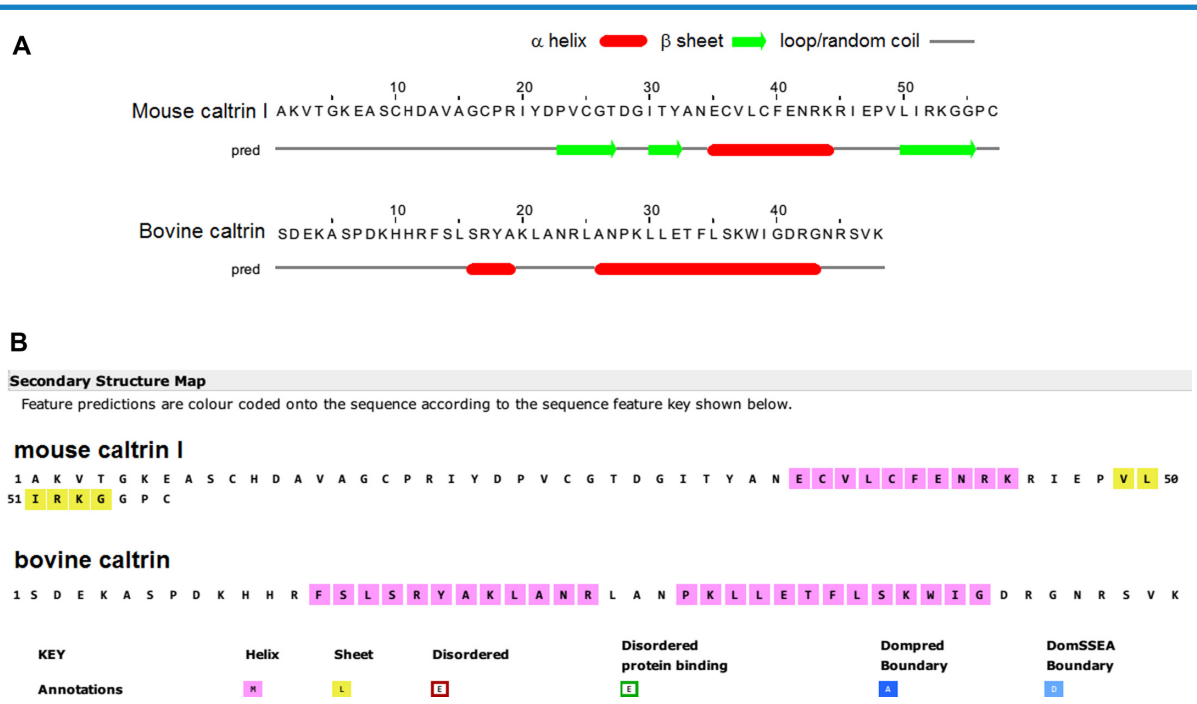


Figure 4. (**A**) Secondary structure predicted based on the templates (PDB ID: 1tgs. 1.B and 1 gt0.1.C for mouse caltrin I and bovine caltrin, respectively) by using ESPript 3.0. (**B**) Secondary structure predicted based on the sequence of caltrin proteins by using PSIPRED v 3.3.

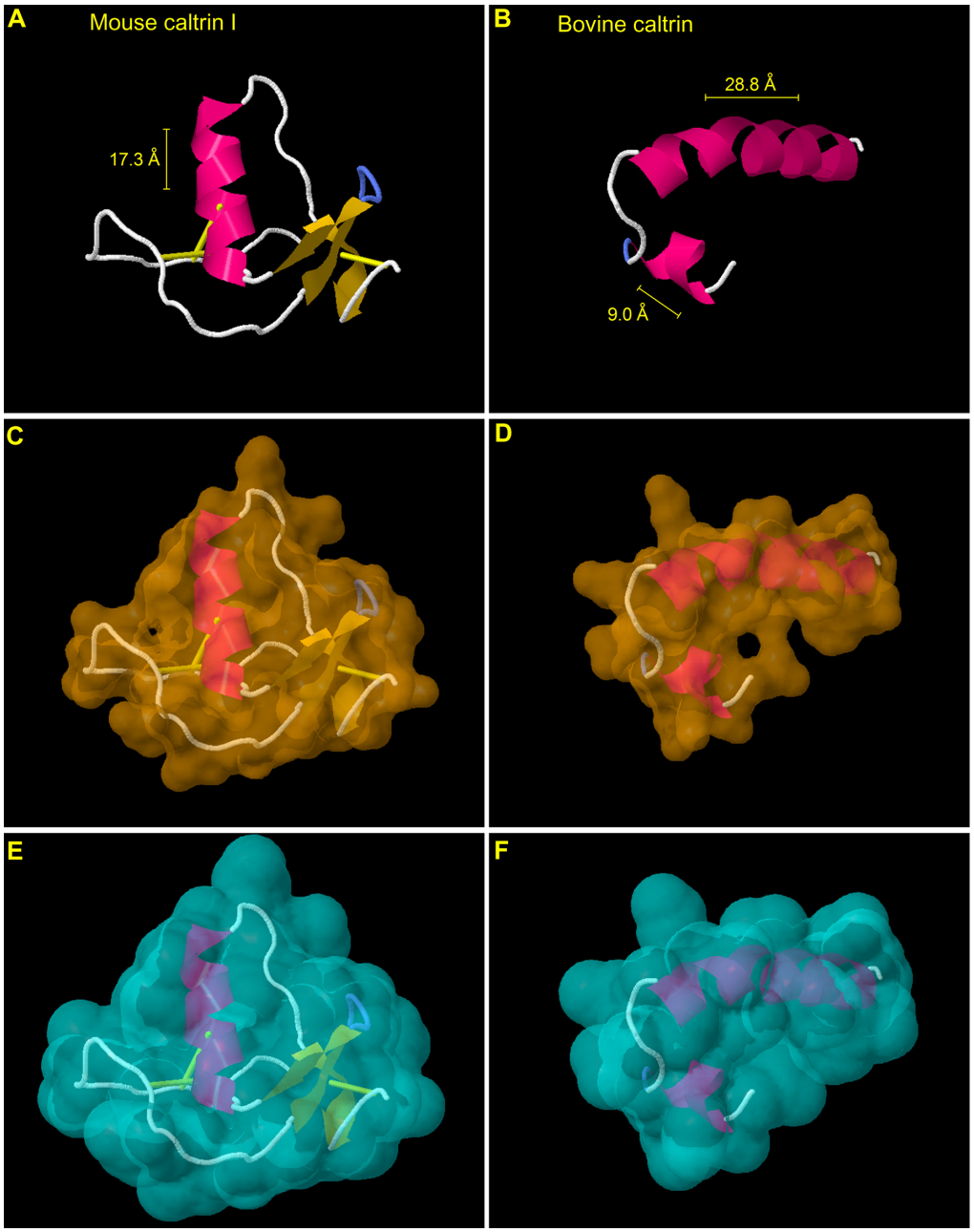


Figure 5. (**A**, **B**) Predicted models. (**C**, **D**) Molecular surface. (**E**, **F**) Accessible surface area (ASA) or Lee–Richards surfaces. All surfaces were calculated using UCSF Chimera 1.10.

solvent molecules in response to the electric field.[34] In order to model electrostatic interactions in proteins, one might assume that charges on a protein interact through a medium characterized by a single dielectric constant (D) and that all interactions can be described by Coulomb's law.[35] However, this approach fails since the protein and the solvent have very different dielectric properties. A more realistic approach considers that the protein and the solvent region have different D and, consequently, the interactions cannot be computed using Coulomb's law. To solve the different dielectric properties of the system, the Poisson–Boltzmann equation was used.[35]

Displaying the distribution of the electric charge at the molecular surface enables protein–protein or protein–substrate interactions to be studied.[36] Because bovine caltrin has a pI of 8.3, it is expected to develop a positive electrostatic surface potential using both charge models (Fig. 6C and D). Figure 7 shows the theoretical electric field spreading out into the solvent, which was calculated with the same setting as for the electrostatic surface potential. By computing the electric field, we can define its relative ability to attract or repel other molecules.[36] Note that bovine caltrin generates mainly positive electric fields (Fig. 7C and D), and hence, it may have a net positive charge at the pH of the seminal plasma.[10]

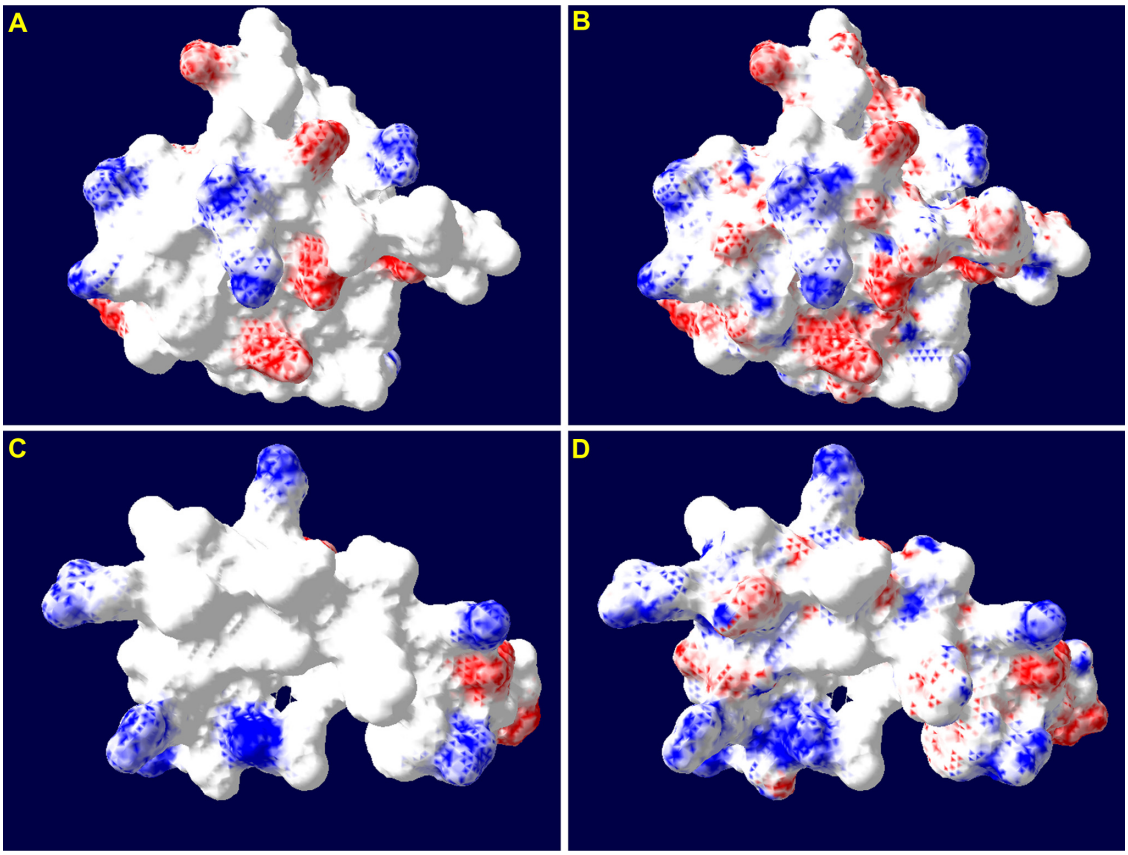


Figure 6. Distribution of the electric charge at the molecular surface. (**A**, **B**) Charged residues and partial charges for mouse caltrin I, respectively. The molecular surface is colored with a red (–1.8 mV as cutoff), to white (neutral points), to blue (1.8 mV as cutoff) color gradient. (**C**, **D**) Idem for bovine caltrin. Ionic strength = 145 mM.

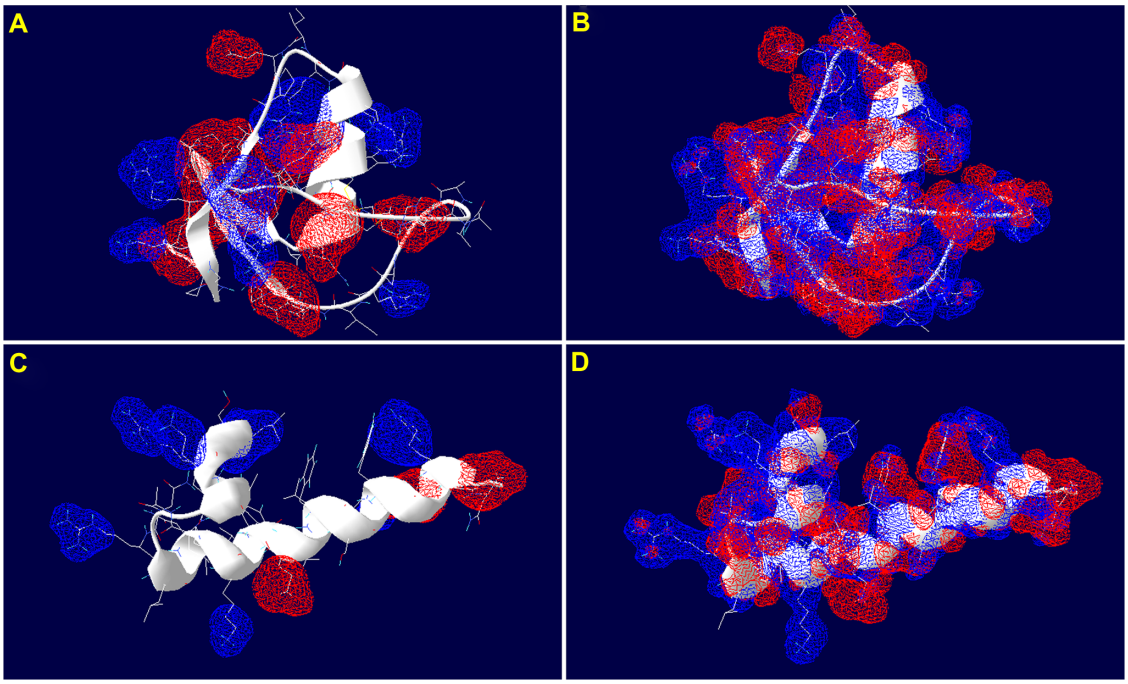


Figure 7. Theoretical electric fields for mouse caltrin I and bovine caltrin. (**A**, **B**) Charged residues and partial charges for mouse caltrin I, respectively. (**C**, **D**) Charged residues and partial charges for bovine caltrin, respectively. A positive value of 1.87 mV is used as a cutoff to delimit a blue contour of those grid points whose value is higher than the given cutoff. Similarly, a negative cutoff of –1.87 mV is used to delimit a red contour lower than the given cutoff. Ionic strength = 145 mM.

Bovine caltrin has a bimodal behavior: Ca^{2+} transport inhibitor and Ca^{2+} transport enhancer in the sperm cells.[10] The anionic cofactors of the seminal plasma stabilize the inhibitory form of bovine caltrin.[37] High-performance liquid chromatography of the ether extract of seminal plasma enabled citric and L-lactic acids to be identified in the samples, and thin-layer chromatography in silica gel detected the presence of the anionic phospholipids, PS, phosphatidylglycerol, and cardiolipin.[10] Of these anionic cofactors identified in the ether extract of acidified bovine seminal plasma, only PS converted the enhancer form of caltrin into the inhibitory molecule at pH 7.4.[10] This event was associated with conformational changes in the protein due to the binding of the phospholipid, forming protein–phospholipid complexes.[10] This interaction is noncovalent, and thus, the generated positive net electric field accounts for the binding of PS to bovine caltrin. Loss of the bovine caltrin–PS interaction stabilizes the enhancer Ca^{2+} uptake form of bovine caltrin[10] with the consequent spontaneous acrosome reaction during capacitation, diminishing the fertilizing potential of sperm cells.

As far as we know, mouse caltrin I does not bind to PS present in the seminal plasma. Nevertheless, mouse caltrin I has a high homology with rat caltrin (see Supplementary Fig. 4). Rat caltrin penetration (adsorption) into model membranes was dependent on phospholipid phase states and polar headgroups (data not shown). The highest rat caltrin adsorption was observed in negatively charged surfaces and expanded lateral phase states (data not shown). Based on our predictions, mouse caltrin I could generate positive surface potentials (Fig. 6A and B) and positive electric fields (Fig. 7A and B) in a similar way as bovine caltrin. Thus, we could assume that mouse caltrin I binding to sperm cell membranes, during ejaculation, probably occurs in lateral expanded and negatively charged regions of such structures, as observed with rat caltrin.

Mouse caltrin I and bovine caltrin hydropathy score. KD scores[16] and residues accessibility to solvent[17] for mouse caltrin I and bovine caltrin were obtained using ExPASy ProtScale tools (Fig. 8). The individual values in kcal mol^{-1} used for KD score calculations and for accessible residues ratio calculations for the 20 amino acids are shown in Supplementary Tables 2 and 3, respectively. The increase (more positive) of KD values above 0 kcal mol^{-1} indicates that the amino acids located in that region of the protein are more hydrophobic.[38] The major energetic factors favoring the partitioning of an amino acid chain from aqueous solution into a membrane are hydrophobic interactions; the factors that maintain them in the aqueous phase are interactions of polar and charged side chains with water.[38] It is assumed that amino acid sequences that are sufficiently hydrophobic and sufficiently long (≥20 residues) imply the existence of a transmembrane α-helix.[38] We confirmed this by determining the KD scores for the dimeric transmembrane

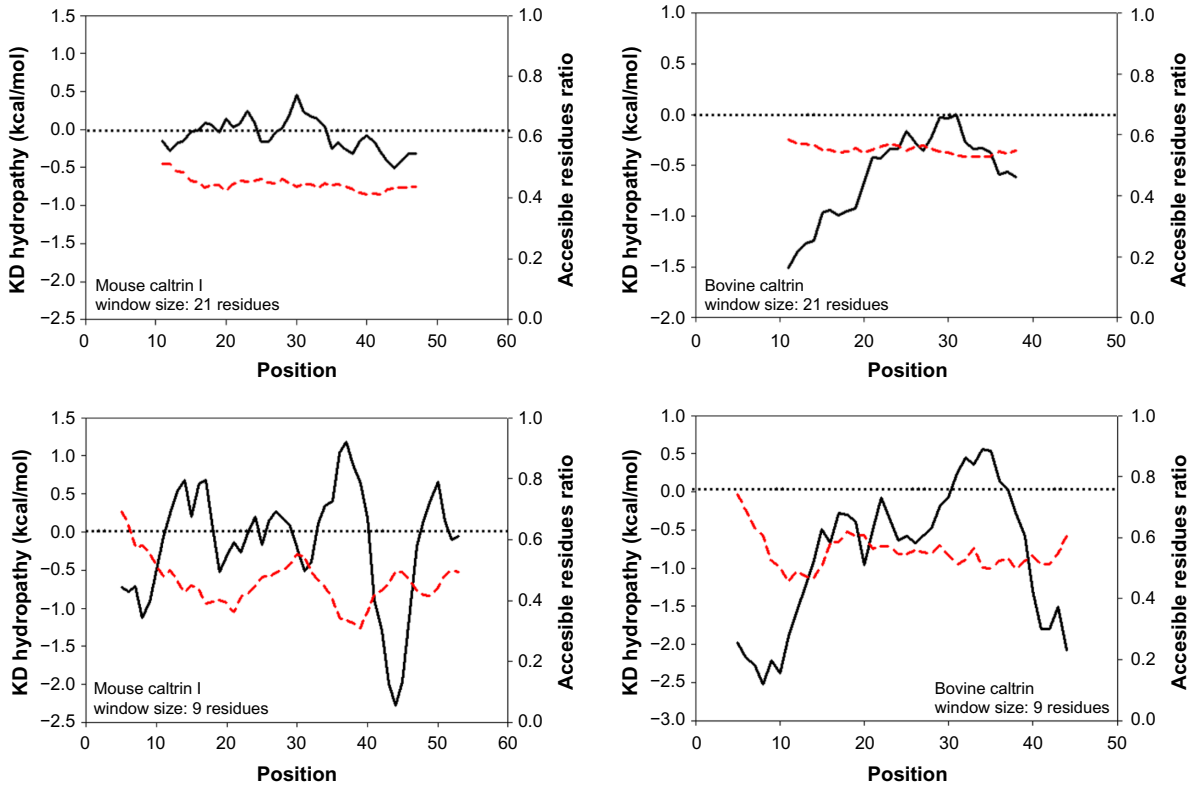


Figure 8. KD scores for mouse caltrin I and bovine caltrin at 21 and 9 residues of window sizes (black continuous lines). Free energies are represented in the vertical axis in kcal mol^{-1} for the transfer of each side chain from a nonaqueous to an aqueous environment. Red dashed lines indicate accessible residue ratios for mouse caltrin I and bovine caltrin at 21 and 9 residues of window sizes.

domain of human glycophorin A (PDB: 1 AFO, structure resolved by NMR) as control (Supplementary Fig. 5). When using a window length of 21 residues (this refers to the number of amino acids scored at the same time; the window moves down one amino acid each time, and each set receives a score that is plotted),[16] two transmembrane α-helixes (continuous bars) were predicted. This was not the case for mouse caltrin I or bovine caltrin, which showed lower values of KD scores than the glycophorin A molecule. This is consistent with the GRAVY score measured, indicating that both proteins are mainly hydrophilic (see Table 1). However, when reducing the window length to 9 residues, it displayed a different hydropathy profile, with maximum values of 1.81 and 0.565 kcal mol^{-1} for mouse caltrin I and bovine caltrin, respectively, indicating hydrophobic regions[16] of these proteins. The region of maximum hydrophobicity of mouse caltrin I (positions 33–41) corresponds to ANECVLCFE residues, where the last 7 are involved in the only α-helix predicted for this protein. Bovine caltrin has a maximum hydrophobicity at positions 30–38, corresponding to LLETFLSKW residues, all forming the α-helix structure. Previously, Sitaram and Nagaraj[33] showed a longer hydropathy profile for the same protein from residues 28–40. In our interpretation of KD scores, we define all values greater than 0 kcal mol^{-1} as hydrophobic (see Ref. 16) and the hydrophobic domains of bovine caltrin shown by Sitaram and Nagaraj[33] had cutoff values below 0 kcal mol^{-1}. Consequently, we defined a narrower, but hydrophobic, segment (residues 30–38) for bovine caltrin than that reported by Sitaram and Nagaraj.[33]

It was mentioned in the previous section that bovine caltrin binds to anionic phospholipids as reported by San Agustin and Lardy.[10] In addition, Sitaram et al.[39] showed that bovine caltrin has bactericide activity by permeabilizing the bacterial membranes. Two 13-residue segments located between residues 14–26 and 28–40 in the bovine caltrin sequence (SLSRYAKLANRLA and PKLLETFLSKWIG) were proposed to interact with lipids.[39] These authors demonstrated that a synthetic peptide corresponding to this 27-residue segment has antimicrobial activity comparable to that of the whole molecule of bovine caltrin. Additionally, bovine caltrin at 37 μM (200 μg/mL) was able to release about 30% of the total hyaluronidases of the acrosome and 50% of the cytosolic lactate dehydrogenase from epididymal sperm by membrane permeabilization. This event was prevented by PS, presumably through caltrin–phospholipid complex formation while PC was ineffective.[40] Based on the KD scores, we support the hypothesis that the segment with the maximum hydrophobicity of bovine caltrin molecule (Supplementary Fig. 6) has the ability to spontaneously partition into the lipid bilayer of membranes, inducing the reorganization of the fatty acid acyl chain arrangement, thereby altering the physical properties of the membrane.

The hydrophobicity analysis of the two caltrin proteins was completed by helical wheel projections (a type of plot or visual representation used to illustrate the properties of α-helices in proteins).[41] The wheels are projections of the amino acid side chains onto a plane perpendicular to the axis of the helix. The perimeter of each wheel corresponds to the backbone of the polypeptide chain and the external spokes to the side chains. For an α-helix with 3.6 residues per turn, adjacent side chains in the sequence are separated by 100° of arc on the wheel.[41] Figure 9 shows the α-helices of mouse caltrin I, bovine caltrin, and glycophorin A as control. As expected, glycophorin A showed an amphipathic profile and mouse caltrin I showed a hydrophobic stabilization arc (at least three hydrophobic residues on the same side of the wheel at n = residue number, $n \pm 3$, $n \pm 4$ distributions of other residues related to n)[41] at residues C2, C5, and F6; and the α-helix was also stabilized by C2, V3, L4, C5, and F6 turn (more hydrophobic region). Hydrophobic stabilization arcs are clusters of hydrophobic residues that stabilize an α-helix.[42] Within each helical segment, the side chains of residues numbered $n \pm 3$, n, $n \pm 4$ are in the most favorable positions to interact (interhelical interactions).[41,42] Only the larger helical segment predicted for bovine caltrin showed hydrophobic stabilization arcs (residues L5, L6, F9, L10, and W13), in a similar manner to glycophorin A, suggesting a clear amphiphilic structure. Thus, as mentioned above, it is possible that the segment between residues 30 and 38 of bovine caltrin, in which we observed the maximum hydrophobicity (Fig. 8), could account for the alteration of the physical properties of membranes (bactericide activity).[33]

Mouse caltrin I binds to the head on the acrosome region of sperm cells and also inhibits trypsin activity.[7] It was proposed that D22 and/or Y21 residues but not R19 were responsible for the sperm-binding site of mouse caltrin I, while R19 but neither D22 nor Y21 was indispensable for trypsin inhibition.[32] We could assume that such residues are in a polar environment. By calculating the hydropathy profile and solvent accessibility of mouse caltrin residues, it was observed that D22, Y21, and R19 have KD scores of −0.271, −0.135, and −0.523 kcal mol^{-1}, indicating hydrophilicity, and a moderate solvent accessibility (0.411, 0.362, and 0.401, respectively, using a window length of 9 residues; and 4, 2, and 7, respectively, as predicted by the I-TASSER suite, see Supplementary Fig. 7). Thus, residue R19 is one of the most exposed to solvent and, although residues D22 and Y21 are slightly hydrophilic, they are poorly exposed to the solvent. Thus, the binding of mouse caltrin I to its receptor in the acrosome seems to be mainly due to electrostatic interactions (hydrophilic KD scores) but the residues are buried (Supplementary Fig. 7). This suggests that conformational changes could be needed during mouse caltrin I–membrane receptor interaction.

Conclusions

In the present work, we predicted the 3D structures for mouse caltrin I and bovine caltrin by various bioinformatics tools and servers. Although they have similar geometric

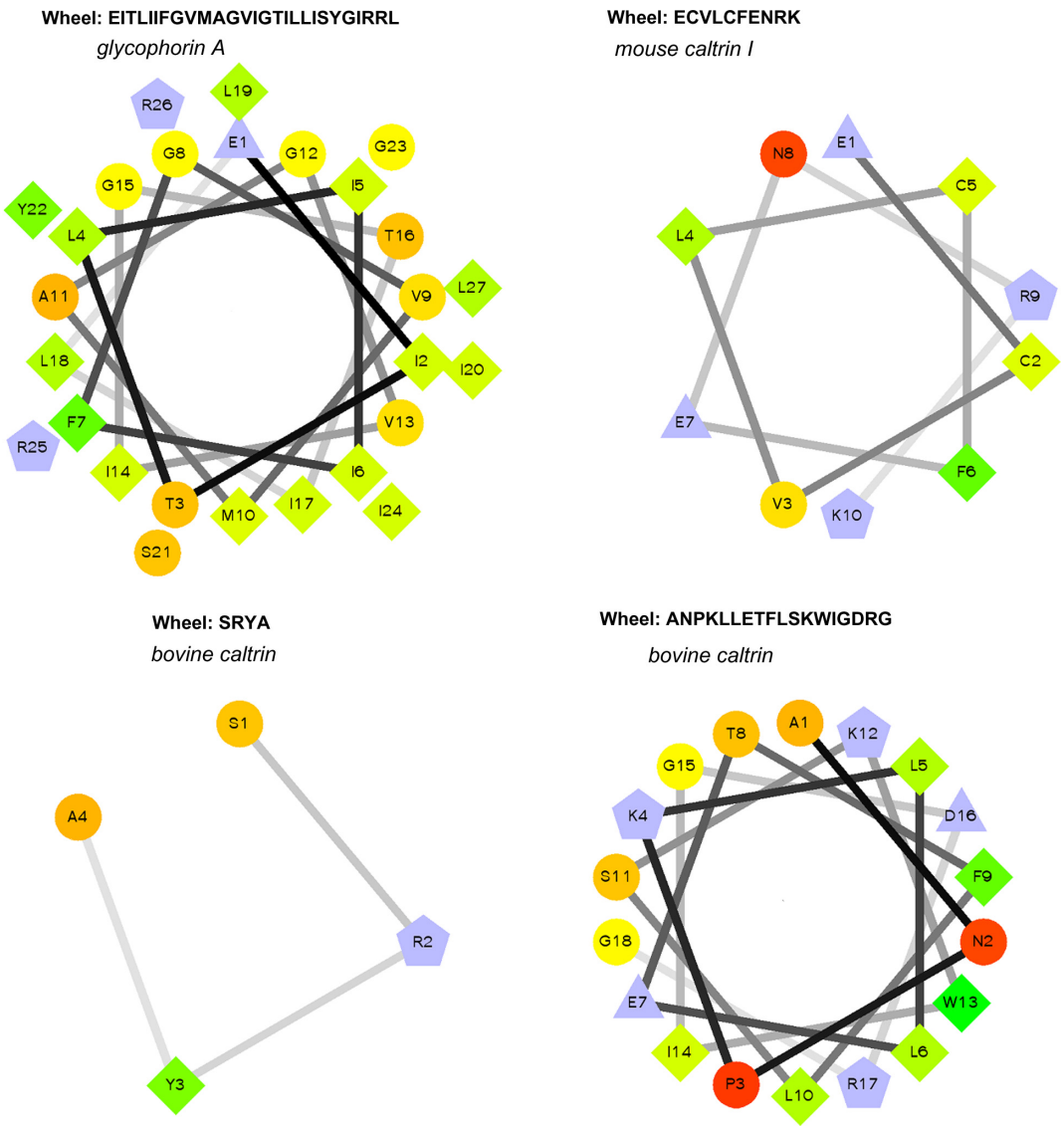


Figure 9. Helical segments of glycophorin A, mouse caltrin I, and bovine caltrin. Wheels presents the hydrophilic residues as circles, hydrophobic residues as diamonds, potentially negatively charged as triangles, and potentially positively charged as pentagons. Hydrophobicity is color coded as well: the most hydrophobic residue is green, and the amount of green decreases proportionally to the hydrophobicity, with zero hydrophobicity coded as yellow. Hydrophilic residues are coded red with pure red being the most hydrophilic (uncharged) residue, and the amount of red decreasing proportionally to the hydrophilicity. The potentially charged residues are light blue.

values, mouse caltrin I seems to have a spherical shape, while bovine caltrin looks like a rod. Mouse caltrin I structure is stabilized by three disulfide bridges and the predicted α-helix for the two proteins by hydrophobic stabilization arcs. Both proteins generate mainly positive electric fields in correlation with their pIs, which probably facilitates the binding of seminal fluid anions only to bovine caltrin. Knowing the structure and physicochemical properties of caltrin proteins is of great importance for understanding the molecular mechanisms of these proteins for protecting acrosomal integrity during sperm capacitation.[4,11]

Acknowledgment

This work was previously the subject of a poster presentation at the Sociedad Argentina de Biofísica Reunión Anual 2016.

Author Contributions

Conceived and designed the experiments: EJG, AES, CEC. Analyzed the data: EJG, AES. Wrote the first draft of the manuscript: EJG. Contributed to the writing of the manuscript: AES. Agreed with manuscript results and conclusions: EJG, AES, CEC. Made critical revisions and approved the final version: CEC. All the authors reviewed and approved the final manuscript.

Supplementary Material

Supplementary Figure 1.
Supplementary Figure 2.
Supplementary Figure 3.
Supplementary Figure 4.
Supplementary Figure 5.

Supplementary Figure 6.
Supplementary Figure 7.
Supplementary Table 1.
Supplementary Table 2.
Supplementary Table 3.

REFERENCES

1. Chen LY, Lin YH, Lai ML, Chen YH. Developmental profile of a caltrin-like protease inhibitor, P12, in mouse seminal vesicle and characterization of its binding sites on sperm surface. *Biol Reprod*. 1998;59(6):1498–505.
2. Coronel CE, San Agustin JT, Lardy HA. Purification and structure of caltrin-like proteins from seminal vesicle of the guinea pig. *J Biol Chem*. 1990;265(12):6854–9.
3. Coronel CE, Lardy HA. Functional properties of caltrin proteins from seminal vesicle of the guinea pig. *Mol Reprod Dev*. 1992;33(1):74–80.
4. Clark EN, Corron ME, Florman HM. Caltrin, the calcium transport regulatory peptide of spermatozoa, modulates acrosomal exocytosis in response to the egg's zona pellucida. *J Biol Chem*. 1993;7(268):5309–16.
5. San Agustin JT, Hughes P, Lardy HA. Properties and function of caltrin, the calcium-transport inhibitor of bull seminal plasma. *FASEB J*. 1987;1(1):60–6.
6. Lardy HA. Happily at work. *J Biol Chem*. 2003;278(6):3499–509.
7. Winnica DE, Novella ML, Dematteis A, Coronel CE. Trypsin/acrosin inhibitor activity of rat and guinea pig caltrin proteins. Structural and functional studies. *Biol Reprod*. 2000;63(1):42–8.
8. Zalazar L, Saez Lancellotti TE, Clementi M, et al. SPINK3 modulates mouse sperm physiology through the reduction of nitric oxide level independently of its trypsin inhibitory activity. *Reproduction*. 2012;143(3):281–95.
9. Coronel CE, Winnica DE, Novella ML, Lardy HA. Purification, structure, and characterization of caltrin proteins from seminal vesicle of the rat and mouse. *J Biol Chem*. 1992;267(29):20909–15.
10. San Agustin JT, Lardy H. Bovine seminal plasma constituents modulate the activity of caltrin, the calcium-transport regulating protein of bovine spermatozoa. *J Biol Chem*. 1990;265(12):6360–7.
11. Dematteis A, Miranda SD, Novella ML, et al. Rat caltrin protein modulates the acrosomal exocytosis during sperm capacitation. *Biol Reprod*. 2008;79(3):493–500.
12. Biasini M, Bienert S, Waterhouse A, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res*. 2014;42(Web Server issue):W252–8.
13. Schwede T, Sali A, Eswar N, Peitsch MC. Protein structure modelling. In: Schwede T, Peitsch MC, eds. *Computational Structural Biology*. Singapore: World Scientific Publishing; 2008;3–36.
14. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. 2015;1(12):7–8.
15. Wilkins MR, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol*. 1999;112:531–52.
16. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982;157(1):105–32.
17. Janin J. Surface and inside volumes in globular proteins. *Nature*. 1979;277(5696):491–2.
18. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;3(215):403–10.
19. Remmert M, Biegert A, Hauser A, Söding J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods*. 2011;2(9):173–5.
20. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 1997;18(15):2714–23.
21. Hasan MA, Alauddin SM, Al Amin M, Nur SM, Mannan A. In silico molecular characterization of cysteine protease YopT from *Yersinia pestis* by homology modeling and binding site identification. *Drug Target Insights*. 2014;8:1–9.
22. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 2007;35(Web Server issue):W407–10.
23. Pettersen EF, Goddard TD, Huang CC, et al. UCSF Chimera – a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605–12.
24. Bystroff C. MASKER: improved solvent-excluded molecular surface area estimations using Boolean masks. *Protein Eng*. 2002;15(12):959–65.
25. Li L, Li C, Zhang Z, Alexov E. On the dielectric "constant" of proteins: smooth dielectric function for macromolecular modeling and its implementation in DelPhi. *J Chem Theory Comput*. 2013;4(9):2126–36.
26. Owens DR. New horizons–alternative routes for insulin therapy. *Nat Rev Drug Discovery*. 2002;7(1):529–40.
27. Ikai J. Thermostability and aliphatic index of globular proteins. *J Biochem*. 1988;88(6):1895–8.
28. Pearson WR. *An introduction to sequence similarity ("homology") searching*. Curr Protoc Bioinformatics. 2013 Jun; Chapter 3:Unit 3.1. John Wiley & Sons. Hoboken, New Jersey.
29. Baker D, Sali A. Protein structure prediction and structural genomics. *Science*. 2001;5540(294):93–6.
30. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR*. 1996;8(4):477–86.
31. Laskowski MJ, Kato I. Protein inhibitors of proteinases. *Annu Rev Biochem*. 1980;49:593–626.
32. Luo CW, Lin HJ, Gopinath SC, Chen YH. Distinction of sperm-binding site and reactive site for trypsin inhibition on p12 secreted from the accessory sex glands of male mice. *Biol Reprod*. 2004;70(4):965–71.
33. Sitaram N, Nagaraj R. A synthetic 13-residue peptide corresponding to the hydrophobic region of bovine seminalplasmin has antibacterial activity and also causes lysis of red blood cells. *J Biol Chem*. 1990;18(265):10438–42.
34. Moult J. Electrostatics. *Curr Opin Struct Biol*. 1992;2(2):223–9.
35. Neves-Petersen MT, Petersen SB. Protein electrostatics: a review of the equations and methods used to model electrostatic equations in biomolecules – applications in biotechnology. *Biotechnol Annu Rev*. 2003;9:315–95.
36. Klapper I, Hagstrom R, Fine R, Sharp K, Honig B. Focusing of electric fields in the active site of Cu-Zn superoxide dismutase: effects of ionic strength and amino-acid modification. *Proteins*. 1986;1(1):47–59.
37. Babcock DF, Singh JP, Lardy HA. Alteration of membrane permeability to calcium ions during maturation of bovine spermatozoa. *Dev Biol*. 1979;69(1):85–93.
38. Engelman DM, Steitz TA, Goldman A. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. *Annu Rev Biophys Biophys Chem*. 1986;15:321–53.
39. Sitaram N, Subbalakshmi C, Krishnakumari V, Nagaraj R. Identification of the region that plays an important role in determining antibacterial activity of bovine seminalplasmin. *FEBS Lett*. 1997;400(3):289–92.
40. San Agustin JT, Lardy HA. Lysogenic activity of enhancer caltrin and the influence of phospholipids on its expression. *Biol Reprod*. 1993;49(4):723–9.
41. Schiffer M, Edmundson AB. Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. *Biophys J*. 1967;2(7):121–35.
42. Scheregaga HA, Nemethy G, Steinberg IZ. The contribution of hydrophobic bonds to the thermal stability of protein conformations. *J Biol Chem*. 1962;237:2506–8.