Biophysical Chemistry 189 (2014) 25–32

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Characterization of an aphid-specific, cysteine-rich protein enriched in salivary glands

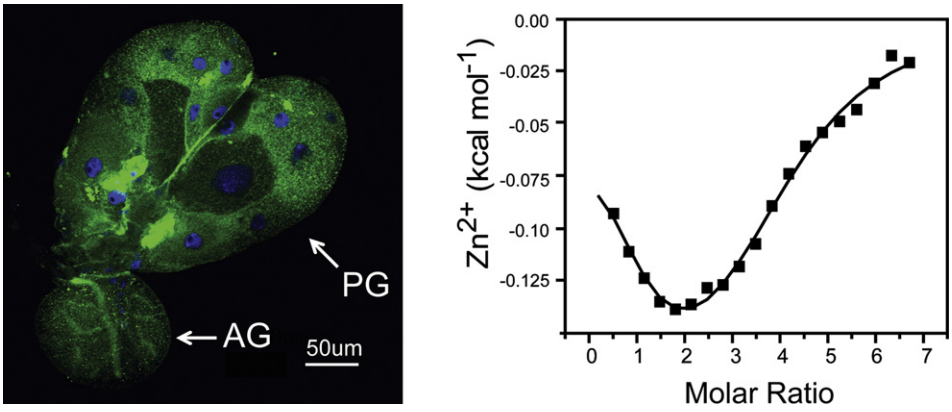
Kun Guo [a,b], Wei Wang [b], Lan Luo [b], Jun Chen [a], Ya Guo [a], Feng Cui [b,*]

[a] Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100193, China
[b] State Key Laboratory of Integrated Management of Pest Insects and Rodents, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

HIGHLIGHTS

- An aphid specific cysteine-rich protein is expressed highly in salivary glands.
- This protein is a monomer globular molecule with a high extent of beta strand.
- This protein is able to bind zinc ions at two binding sites.
- Aphids require more of this protein when feeding on plants.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 5 February 2014
Received in revised form 22 March 2014
Accepted 25 March 2014
Available online 1 April 2014

Keywords:
Aphid
Salivary gland
Cysteine-rich protein
Zinc ion binding
Gene interference

ABSTRACT

Aphids secrete saliva into the phloem during their infestation of plants. Previous studies have identified numerous saliva proteins, but little is known about the characteristics (physical and chemical) and functions of these proteins in aphid–plant interactions. This study characterized an unknown protein (ACYPI39568) that was predicted to be enriched in the salivary glands of pea aphid. This protein belongs to an aphid-specific, cysteine-rich protein family that contains 14 conserved cysteines. ACYPI39568 is a monomeric globular protein with a high beta strand extent. The binding stoichiometric ratios for Zn^{2+} and ACYPI39568 were approximately 3:1 and 1:1 at two binding sites. ACYPI39568 was predominantly expressed in the first instar stage and in the salivary glands. Aphids required more ACYPI39568 when feeding on plants than when feeding on an artificial diet. However, the interference of ACYPI39568 expression did not affect the survival rate of aphids on plants.

1. Introduction

Aphids comprise more than 4700 species worldwide [1]. Polyphagous species, such as *Myzus persicae* (Sulzer) and *Aphis gossypii* Glover, are notorious pests in agriculture, forestry, and horticulture. Unlike other insects with chewing or licking mouthparts, aphids have caused great economic losses by inflicting direct damage and by acting as vectors of phytopathogenic viruses, during which their saliva functions essentially [2,3]. Salivation continues throughout probing and feeding on phloem sap [4]. Many studies have focused on aphid saliva in recent years; however, the function of aphid saliva in the insect's feeding and colonization behavior remains almost unknown.

* Corresponding author. Tel.: +86 10 64807218; fax: +86 10 64807099.
E-mail address: cuif@ioz.ac.cn (F. Cui).

http://dx.doi.org/10.1016/j.bpc.2014.03.006
0301-4622/

Miles reviewed previous enzymatic studies on aphid saliva [2]. Abundant genomic and proteomic data, including the sequenced genome of the pea aphid *Acyrthosiphon pisum*, were used to identify the molecular composition of saliva. The expressed sequence tags (ESTs) from the salivary glands of *Ac. pisum* and *M. persicae* were employed to identify candidate saliva proteins and their abundance [5,6]. Numerous proteins have been identified from the saliva of various aphid species by using mass spectrometry-based proteomic profiling [7,8]. Glucose oxidase, glucose dehydrogenase, NADH dehydrogenase, α-glucosidase, and α-amylase were discovered in the saliva of *M. persicae* [9]. Nine proteins have been identified in the saliva of *Ac. pisum*, of which four were homologs of known proteins: an angiotensin-converting enzyme, an M1 zinc-dependent metalloprotease, a glucose–methanol–choline oxidoreductase, and a regucalcin homolog [10]. Twelve and seven proteins have been identified from the saliva of *Sitobion avenae* and *Metopolophium dirhodum*, respectively [11]. A previous study focused on the Russian wheat aphid *Diuraphis noxia* and found that aphids fed with different substrates secrete watery saliva with qualitative and quantitative differences in soluble proteins [12]. The composition of salivary proteins may be an important determinant of the plant range of different aphid species [13].

The functions of individual saliva proteins in aphid–plant interaction have not been explored adequately. Knockdown of the well-known saliva protein c002 alters the foraging and feeding behavior as well as reduces the survival of pea aphid on fava beans [14,15]. The fecundity of *M. persicae* is enhanced by c002 overexpression in *Nicotiana benthamiana* [6]. Mp10, another candidate saliva protein from *M. persicae*, specifically induces chlorosis and local cell death in *N. benthamiana*; this protein reportedly suppresses pathogen-associated molecular pattern-triggered immunity [6]. However, the physical and chemical characteristics of individual saliva proteins must be elucidated to understand the interaction between saliva proteins and plant counterpart molecules.

Carolan et al. used transcriptomic and proteomic approaches to predict 42 candidate effector proteins enriched in the salivary glands of pea aphid, 33 of which possess unknown functions [5]. Among these unknown proteins, four belong to an undescribed protein family, which is characterized by having 14 conserved cysteine residues [16]. This family contains 13 genes but encodes for 14 proteins because one gene (ACYPI008667) has two isoform transcripts [16]. In this study, this family is designated as aphid cysteine-rich protein family (AphidCRP). The present study aims to explore the physical and chemical characteristics, temporal and spatial expression profiles, and possible involvement in aphid feeding on host plants of the ACYPI39568 gene belonging to the AphidCRP family.

2. Materials and methods

2.1. Insect

The pea aphid *Ac. pisum* was collected in 2010 from peas (*Pisum sativum*) at Yuxi, Yunnan Province, China, and reared on fava beans (*Vicia fabae*) in climate chambers under the following conditions: 21 ± 1 °C, 60% ± 5% relative humidity, and 16 h: 8 h (L:D) photoperiod.

2.2. Protein sequence analysis

A secretory signal peptide was analyzed with the SignalP server (http://www.cbs.dtu.dk/services/SignalP) [17], and the protein molecular weight was calculated at http://web.expasy.org/protparam [18]. Potential O- and N-glycosylation sites were predicted at http://www.cbs.dtu.dk/services/NetOGlyc [19] and http://www.cbs.dtu.dk/services/NetNGlyc [20], respectively. The secondary structure was predicted at http://scratch.proteomics.ics.uci.edu. Protein function was predicted at http://zhanglab.ccmb.med.umich.edu/I-TASSER [21–23]. Confidence score, i.e., EC-score, was used to estimate the quality and functional similarities of predicted models. EC-score > 1.1 indicates functional similarity between the query and template proteins. Both homologous proteins and EST sequences from other species were tracked down with BlastP/tBlastn software at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and then aligned with Bioedit software [24]. A phylogenetic tree was constructed with the neighbor-joining method (complete deletion and Poisson correction model) using MEGA 5 software [25]. Bootstrap analysis (1000 replicates) was applied to evaluate the internal support of the tree topology.

2.3. RNA isolation and cDNA synthesis

Total RNA was extracted from whole bodies or various tissues of pea aphid using the RNeasy Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol and then treated with DNase to eliminate DNA contamination. cDNA was reverse-transcribed from 2 μg of total RNA using MMLV reverse transcriptase (Promega, Madison, WI, USA) according to the manufacturer's protocol.

2.4. Protein expression, purification, and antibody preparation

The full length of ACYPI39568 was cloned from the cDNA library of whole bodies using primers ACYPI39568-F1 and ACYPI39568-R1 (Table 1). ACYPI39568 was inserted into the vector pFastbacI between the EcoRI and XhoI sites using primers ACYPI39568-EcorI and ACYPI39568-XhoI with a hexa histidine-tag at the 3′ terminal (Table 1). The resulting plasmids after sequencing verification were used to generate recombinant baculoviruses with the Bac-to-Bac baculovirus system (Invitrogen, Carlsbad, CA, USA). *Spodoptera frugiperda* Sf9 cells were infested with the recombinant baculovirus at a multiplicity of infection of 3 and then incubated at 28 °C for 72 h. The cell-free medium was loaded on a Ni-NTA agarose column (Qiagen) that was equilibrated with a binding buffer (20 mM Tris–HCl, 50 mM NaCl, 10 mM imidazole, pH 8.0). After washing with a binding buffer containing 20 mM imidazole, the protein was eluted from the column with a binding buffer containing 100 mM imidazole and then dialyzed against a dialysis buffer (20 mM Tris–HCl, pH 8.0) at 4 °C overnight. The dialyzed protein was

Table 1
Primer sequence.

Primer name	Primer (5′ → 3′)	PCR product (bp)
ACYPI39568-F	AATCACAGCAATAATACACATAC	829
ACYPI39568-R	ATAATTTGCATACCACTGTCAC	
ACYPI39568-EcorI	GCGCGAATTCATGTTTATGT	437
ACYPI39568-XhoI	GGCCCTCGAGTTAGTGGTGGTGGTGGTGGTGCTGGCAAGTT	
ACYPI39568-qPCR-F	AGACTCAGGCATTCCCGAAAC	104
ACYPI39568-qPCR-R	AGCCACTGTTGCATTCGTCAC	
L27-qPCR-F	TCGTTACCCTCGGAAAGTC	108
L27-qPCR-R	GTTGGCATAAGGTGGTTGT	
dsACYPI39568-F	AAGGCCGTCTTTGCTTAC	246
dsACYPI39568-R	GAAATTGATTTGTGGGTTC	
dsGFP-F	CACAAGTTCAGCGTGTCCG	420
dsGFP-R	GTTCACCTTGATGCCGTTC	

applied to a Q-sepharose Fast Flow column (GE Healthcare, Piscataway, NJ, USA). After the column was washed with a dialysis buffer containing 50 mM NaCl, the ACYPI39568 protein was eluted with a dialysis buffer containing 400 mM NaCl and then concentrated with Amicon ultra-4 (Millipore, Billerica, MA, USA). Protein concentration was determined using the Pierce™ BCA protein assay kit (Pierce Biotechnology, Rockford, IL, USA) according to the manufacturer's protocol. Peptide fragments including residues 1–10, 27–36, 43–52, 63–72, and 83–92 of the mature protein were synthesized as antigens to immunize rabbits and to obtain an anti-ACYPI39568 polyclonal antibody with Abmart Inc. (Shanghai, China).

2.5. Molecular weight analysis by MALDI-TOF mass spectrometer

The purified recombinant ACYPI39568 (2 μg, 10 μL) was desalted with ZipTip μ-C18 (Millipore). The sample was eluted and spotted onto a target plate using an elution buffer (50% acetonitrile/0.1% TFA) containing sinapic acid (5 mg/mL) matrix. Mass spectrometry was performed using an AXIMA-CFR Plus MALDI-TOF mass spectrometer from KRATOS Analytical (Shimadzu, Manchester, UK). Positive ion mass spectra were acquired in the m/z range of 9000 to 20,000 through a linear mode.

2.6. Measurement of circular dichroic (CD) spectra

The CD spectra of ACYPI39568 (0.47 mg/mL) were recorded from 190 nm to 260 nm at 25 °C on a Pistar n-180 (Applied Photophysics Ltd. Leatherhead, UK) using a quartz cell with a path length of 0.1 cm. The spectra were composed of an average of three consecutive scans at a scan rate of 50 nm/min. The protein was in a buffer containing 10 mM Tris–HCl at pH 8.0. The spectra were corrected by subtracting the signal from the protein-free solutions recorded under the same conditions. The CD data were reported as the mean residue ellipticity [θ] in the unit of degrees$\cdot$cm$^2\cdot$dmol^{-1} and submitted to the online K2D2 server (http://k2d2.ogic.ca) to estimate the protein secondary structure.

2.7. Analytical ultracentrifugation

Sedimentation velocity experiments were conducted on the recombinant ACYPI39568 using an Optima XL-I ultracentrifuge (Beckman Coulter, Fullerton, CA, USA). Sedimentation was monitored at 60,000 rpm and 20 °C by absorbance at 280 nm with a 30 s interval. Data were analyzed using Sedfit 14.0 software (http://www.analyticalultracentrifugation.com/sedfit.htm).

2.8. Isothermal titration calorimetry

The binding affinity and thermodynamic properties associated with the complexation of Zn^{2+}, Mn^{2+}, or Ca^{2+} with ACYPI39568 were determined using isothermal titration calorimetry in an iTC_{200} microcalorimeter (GE Healthcare). Traces of metal ions were stripped from the recombinant protein samples by adding 1 mM EDTA before dialysis against 20 mM Tris–Hcl and 50 mM NaCl at pH 8.0 overnight. The dialysate buffer was recovered for constructing blank ITC controls and metal ion solutions of varying concentrations: $ZnCl_2$ (0.5, 1.5, 1.7, and 2.0 mM), $MnCl_2$ (0.5 and 1.5 mM), and $CaCl_2$ (0.5 and 1.5 mM) (Sigma-Aldrich, Louis MO, USA). Before each experiment, the sample cell was filled with 250 μL of 0.05 mM protein ACYPI39568. The titrant (different metal ion solutions) was loaded in the syringe. All measurements were performed under constant stirring (1000 rpm) at a constant temperature (25 °C). Small aliquots of the titrant (typically 2 μL) were successively injected into the solution of the working cell with an injection duration of 4 s at 180 s intervals. As an exception, the first injection volume and duration were 0.5 μL and 1 s, respectively, to remove any irregularity caused by capillary action. A blank control acquired by adding the titrant into the buffer without protein was subtracted from the titration curve. The result was analyzed using the Origin software for ITC supplied with the iTC_{200} microcalorimeter.

2.9. Activity assay of phospholipase A2 (PLA2) and porphobilingen synthase

PLA2 activity was measured according to the manufacturer's protocol (EnzChek Phospholipase A2 kit; Life Technologies, Gaithersburg, MD, USA). Briefly, the purified protein ACYPI39568 was diluted with a reaction buffer to 5, 50, and 500 μg/mL. Approximately 50 μL of standard samples (0 U/mL to 10 U/mL PLA2 from honey bee venom for producing a standard curve), PLA2 reaction buffer as control, or ACYPI39568 was mixed with 50 μL of the substrate–liposome mixture in individual microplate wells. PLA2 activity was monitored at 515 nm/575 nm in a fluorescence microplate reader (Synergy 4™, BioTek Instruments, Winooski, VT, USA). The activity assay of porphobilingen synthase was carried out as previously described [26].

2.10. Immunohistochemistry

Immunohistochemistry was carried out to localize the expression of ACYPI39568 in vivo in aphid salivary glands using purified anti-ACYPI39568 polyclonal antibody. Detailed experimental steps are provided in the study of Mutti et al. [15]. Images were taken using a Zeiss LSM 710 confocal microscope (Zeiss, Oberkochen, Germany).

2.11. Quantification of ACYPI39568 gene expression in tissues and developmental stages

Tissues from the brain, salivary gland, ovary, and digestive gut were dissected from 7 to 13 adult aphids for RNA isolation, and six replicates per tissue were prepared. RNA was also extracted from first-instar, second-instar, third-instar, fourth-instar, and adult aphids. Six replicates and five individuals per replicate were prepared for each developmental stage. Two pairs of primers, ACYPI35968-qPCR-F and ACYPI35968-qPCR-R for ACYPI39568 and L27-qPCR-F and L27-qPCR-R (Table 1) for *Ac. pisum* ribosomal protein L27 transcript (accession number in GenBank CN584974) as internal control, were designed to amplify 104- and 108-bp fragments, respectively, in qPCR. qPCR was conducted on a Roche Light cycler 480 (Roche, Mannheim, Germany) initiated with a 2 min incubation period at 95 °C, followed by 40 cycles of 20 s at 95 °C, 20 s at 58 °C, and 20 s at 68 °C. The transcript level of ACYPI39568 was normalized to that of L27. Melting curve analysis was performed to confirm the specificity of amplification. Difference in transcript levels was analyzed by one-way ANOVA, followed by Tukey's test for multiple comparisons using SPSS 17.0 software. Values are presented as the mean ± SEM.

2.12. Comparison of ACYPI39568 transcription in diet-fed and plant-fed pea aphids

One-day-old adult aphids were transferred from fava beans to an artificial diet [27] for feeding for 24 h. The diet was sealed between two layers of stretched parafilm that covered a sterile plastic cap (Falcon, Primaria, NJ, USA) where an aphid stayed. Six replicates and 10 individuals per replicate were sampled for the diet-fed group and the fava bean-fed group. The transcript levels of ACYPI39568 in the heads (containing the salivary glands) of the two groups were compared through qPCR. Differences were statistically evaluated by *t*-test using SPSS 17.0. Values are presented as mean ± SEM.

2.13. Double-stranded RNA (dsRNA) synthesis and delivery

dsRNA fragments of green fluorescent protein (GFP) and ACYPI39568 were prepared using the T7 RiboMAX Express RNAi system (Promega) and then purified with Wizard SV Gel and PCR Clean-Up System (Promega) following the manufacturer's protocol. PCR primers

dsACYPI39568-F, dsACYPI39568-R and dsGFP-F, dsGFP-R (Table 1) were used to prepare 246- and 420-bp dsRNA fragments of ACYPI39568 and GFP in their open reading frames, respectively. The dorsal abdomens of third-instar aphids were injected with 23 nL of dsRNA at 6 μg/μL through a glass needle using Nanoliter 2000 (World Precision Instruments, Sarasota, FL, USA). Six groups of third-instar aphids and 15 individuals in each group were injected and then reared on fava bean plants. Dead aphids within several hours after injection were removed; their death was attributed to mechanical damage. The interference efficiency on the ACYPI39568 expression in the whole body was examined at 3 d after dsRNA injection using qPCR. The survival curves of the dsACYPI39568-RNA and dsGFP-RNA injection groups were statistically compared with the Kaplan–Meier method using Log Rank (Mantel–Cox) test in SPSS 17.0.

3. Results

3.1. Phylogenetic analysis of the AphidCRP family

The AphidCRP family does not have any known conserved domain; in addition, members of this family are not homologous to other known proteins. This family is characterized by 14 conserved cysteine residues, which form three CXXC motifs, one CXC motif, and one CXXXC motif (Fig. S1). The amino acid sequence identities of the 13 proteins of the pea aphid ranged from 30% to 89%. In addition to pea aphids, other Aphididae species also possess AphidCRP members. One to 14 transcripts of the AphidCRP family have been identified in EST libraries of *Ac. kondoi*, *Rhopalosiphum padi*, *Ap. gossypii*, *M. persicae*, *M. ascalonicus*, and *Toxoptera citricida*; however, judging the gene number in these aphid species is difficult because of insufficient genome information. Some but not all AphidCRP members are secretory proteins (Fig. S1). However, this family does not exist in non-aphid organisms. Phylogenetic analysis showed that AphidCRP genes diverged after the divergence of aphid species, except in *Ac. pisum* and *Ac. kondoi*, where orthologous AphidCRP genes were detected (Fig. 1).

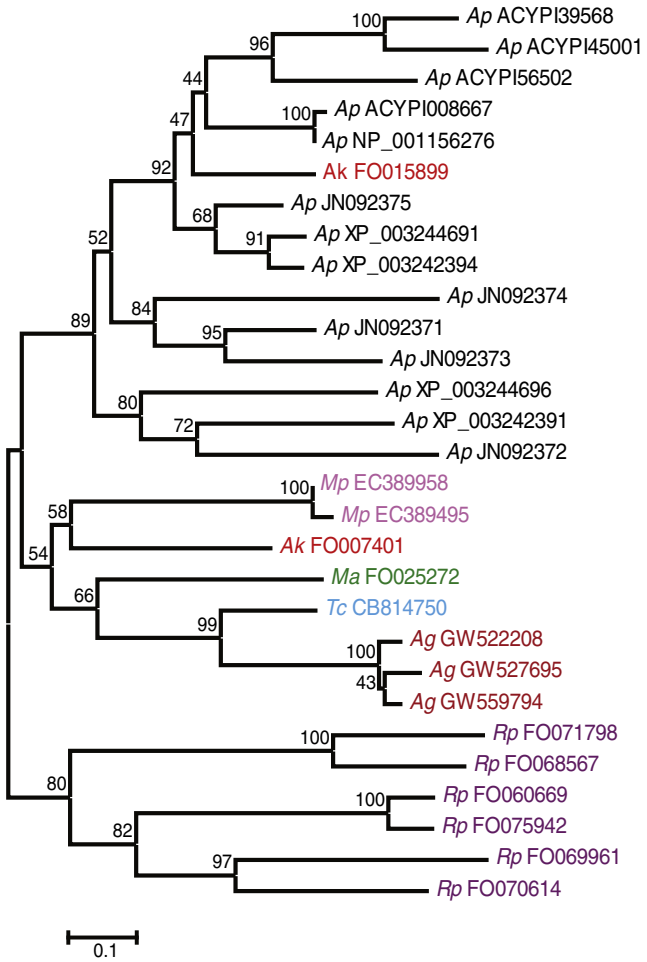


Fig. 1. Neighbor-joining phylogenetic tree of the AphidCRP family from various aphids. Bootstrap supporting values (1000 replicates) are shown at the branched nodes. Different colors indicate protein sequences from different aphid species. *Ap*: *Acyrthosiphon pisum*. *Ak*: *A. kondoi*. *Mp*: *Myzus persicae*. *Ma*: *M. ascalonicus*. *Tc*: *Toxoptera citricida*. *Ag*: *Aphis gossypii*. *Rp*: *Rhopalosiphum padi*.

3.2. Sequence character of ACYPI39568

For ACYPI39568, a secretory signal peptide was predicted at the N-terminus with a cleavage site between residues Ser28 and Ser29. A potential O-glycosylation site was predicted at Thr130 of the protein precursor. Secondary structure prediction showed that the protein consisted of several beta strands that may form a beta sheet structure (Fig. S2). The enzyme homolog of ACYPI39568 in the PDB library was highly similar to the PLA2 (EC 3.1.1.4) of *Crotalus durissus terrificus*, which requires calcium as a cofactor. However, the EC-score for the enzyme homology prediction (0.52) was very low, indicating low prediction accuracy (Table S1). Other members of the AphidCRP family were predictably similar to chloroumconate cydoisomerase, porphobilinogen synthase, ferredoxin hydrogenase, PLA2, proteasome endopeptidase complex, RNA-directed RNA polymerase, ferredoxin-NAD reductase, dimethylallytranstransferase, or fatty-acyl-CoA synthase, with calcium, manganese, zinc, or nickel as a cofactor (Table S1).

3.3. Physical and chemical properties of ACYPI39568

The full length of ACYPI39568 plus a His-tag at the C-terminus was recombinantly expressed in Sf9 cells and then purified from the cell culture medium. The molecular weight of mature ACYPI39568 plus a His-tag at the C-terminus was measured by MALDI-TOF to be 12,176 Da (Fig. 2C), which was close to the predicted 12,189 Da. However, the molecular weight viewed on SDS-PAGE and western blot was considerably larger than 12,176 Da (Fig. 2A and B). Strong disulfides or mismatching of disulfides among protein molecules may account for this abnormal mobility on SDS-PAGE.

The circular dichroic spectrum of ACYPI39568 exhibited one minima at 208 nm and was estimated to contain 3.6% of alpha helix and 44.5% of beta strand (Fig. 3A). Analysis of sedimentation velocity revealed that ACYPI39568 is a monomer in the solution, with a sedimentation coefficient of 1.65 S. The frictional coefficient ratio ($f/f_0 = 1.1$) fitted a globular protein structure (Fig. 3B). An ITC titration curve was observed when zinc ion was titrated in the protein solution (Fig. 3C). The titration data best fitted into a two-binding-site model. For one binding site, the binding stoichiometric ratio for Zn^{2+} and ACYPI39568 was approximately 3:1, with a dissociation constant (Kd) of 23.9 μM. The ratio for the other binding sites was 1:1, with a Kd of 2.4 μM. The complexation of Zn^{2+} with ACYPI39568 was primarily driven by a change in enthalpy ($\Delta H1^\circ = -2.07$ kcal/mol, $\Delta H2^\circ = -0.47$ kcal/mol) rather than entropy ($-T\Delta S1^\circ = -4.23$ kcal/mol, $-T\Delta S2^\circ = -7.18$ kcal/mol). When Mn^{2+} or Ca^{2+} was titrated in the protein solution, no titration curve was observed, indicating that ACYPI39568 did not bind these two ions. Although ACYPI39568 was most similar to PLA2 in functional prediction, enzymatic analysis showed that ACYPI39568 did not have PLA2 activity. ACYPI39568 did not show porphobilingen synthase activity, although it bound zinc ions.

3.4. Temporal and spatial expression of ACYPI39568

Real-time quantitative PCR (qPCR) showed that the expression of ACYPI39568 peaked in the first instar stage and then dropped during development (Fig. 4A). The salivary gland was the predominant organ

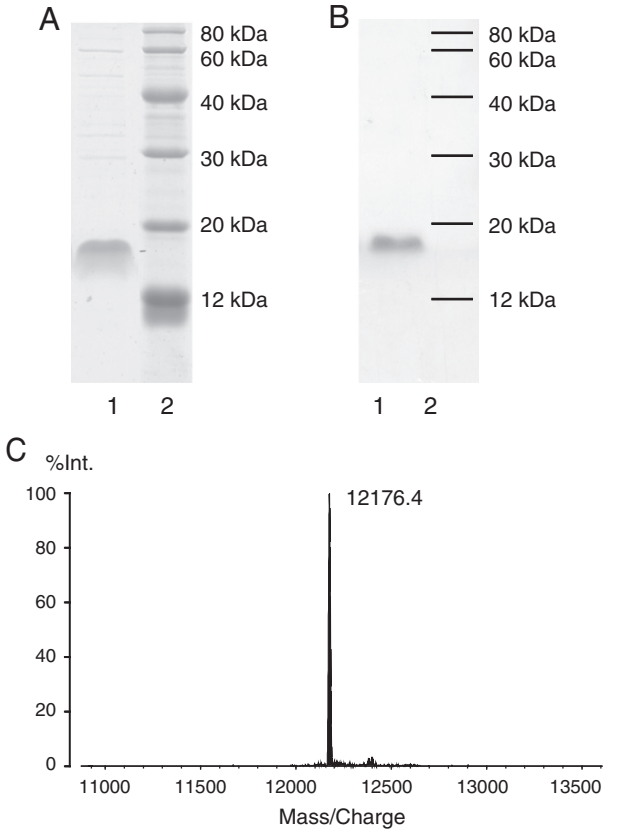


Fig. 2. ACYPI39568 recombinantly expressed in Sf9 cells. SDS/PAGE (A) and western blot (B) of purified ACYPI39568. 1: ACYPI39568. 2: Marker. (C) Molecular weight of ACYPI39568 measured by MALDI-TOF.

for the expression of ACYPI39568 (Fig. 4B). Immunohistochemistry at the salivary glands revealed that ACYPI39568 was unevenly expressed in both primary and accessory glands. ACYPI39568 was absent in a pair of large secretory cells in the middle of the primary salivary gland (Fig. 4C and 4D).

3.5. Effect of ACYPI39568 on aphid survival on plants

The expression level of ACYPI39568 was examined in the heads (containing the salivary glands) of aphids fed with fava beans or artificial diet. The expression of ACYPI39568 was significantly reduced by 31% in the diet-fed aphids compared with the fava bean-fed aphids ($P = 0.02$) (Fig. 5A). This result indicated that aphids required more ACYPI39568 protein when feeding on plants than when feeding on an artificial diet. When dsACYPI39568-RNA was injected to the third-instar aphids, the expression of ACYPI39568 was reduced by 32.4% at 3 d after injection (Fig. 5B). This interference did not affect the survival rate of the aphids living on host plants ($P = 0.62$) (Fig. 5C).

4. Discussion

Similar to most piercing-sucking insects, aphids secrete saliva into the phloem to digest nutrition and to oppose the defense reactions of plants. Salivary proteins have been thought to have crucial functions in aphid–plant interaction. Numerous proteins in aphid saliva have been found or predicted by proteomics or transcriptomics [5,9–11]. However, very few proteins have known functions in aphid–plant interaction. In this study, we characterized one member (ACYPI39568) of the AphidCRP family, a putative saliva protein family that exclusively exists in aphids with 14 conserved cysteines without known functions. ACYPI39568 can bind zinc ions and is highly expressed in the salivary gland. Although aphids require a high expression of this protein when feeding on plants, decreasing its expression did not influence aphid survival.

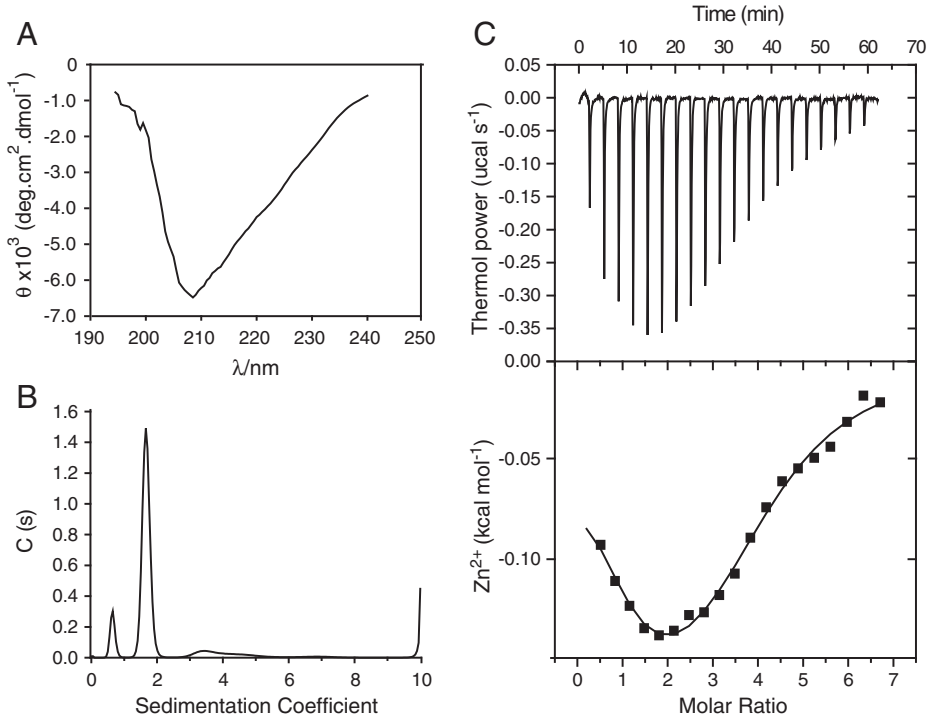


Fig. 3. Physical and chemical properties of ACYPI39568. (A) Circular dichroism spectra of ACYPI39568 in 10 mM Tris–HCl (pH 8.0). (B) $c(s)$ analysis of analytical ultracentrifugation for ACYPI39568. (C) Isothermal titration calorimetry of ACYPI39568 with Zn^{2+}. Top panel: Heat release with incremental addition of $ZnCl_2$ to the ACYPI39568 solution. Bottom panel: Data fitted with best-fit two-site binding model.

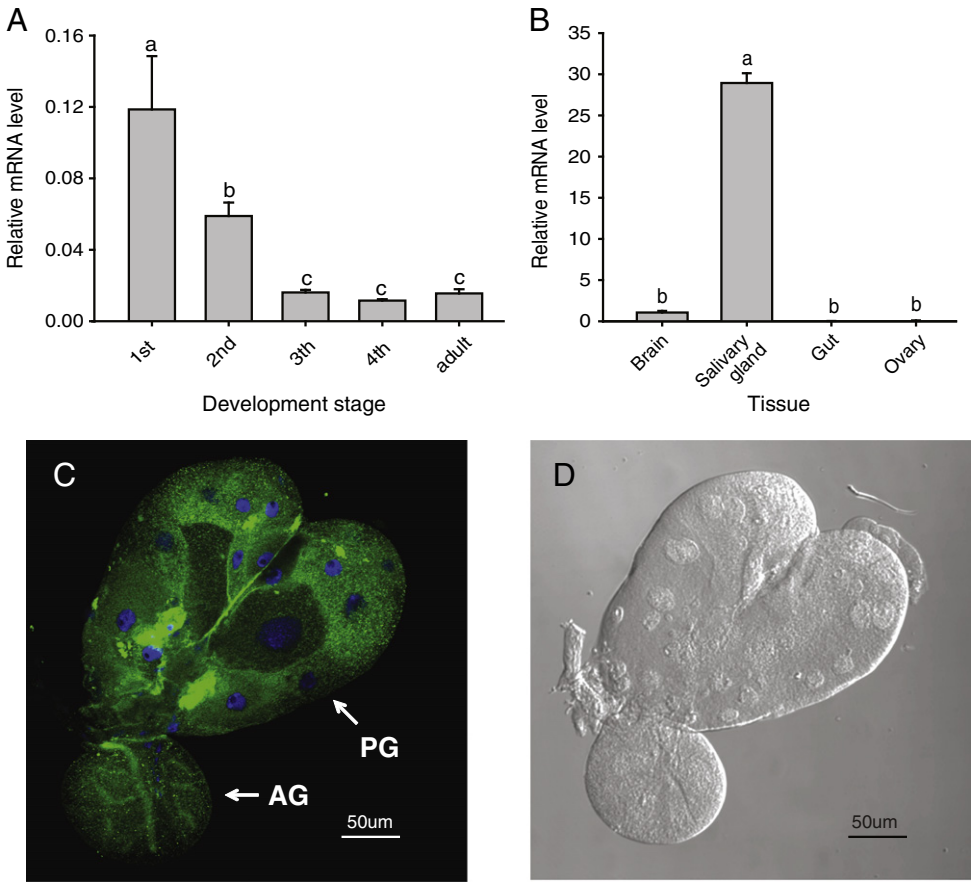


Fig. 4. Temporal and spatial expression of ACYPI39568 in pea aphid. ACYPI39568 transcript levels in four tissues (A) and different developmental stages (B) measured with real-time quantitative PCR. Values are represented as mean ± SEM. Different letters indicate significant differences at $P < 0.05$. (C) Immunohistochemical localization of ACYPI39568 in the salivary glands of pea aphid. Green indicates the positive signal of anti-ACYPI39568 polyclonal antibody staining. Nuclei are shown in blue. PG, principal gland. AG, accessory gland. The negative control (without anti-ACYPI39568 antibody) does not show any positive staining. Scale bar = 50 μm. (D) Differential interference contrast image.

Although both AphidCRP and c002 are involved in aphid–plant interaction, some differences still exist between the two. ACYPI39568 is expressed in both primary and accessory glands, whereas c002 is expressed uniquely in the primary gland [14,15]. Knockdown of c002 leads to the mortality of aphids, whereas interference of ACYPI39568 does not influence aphid survival. The reason for such a difference in phenotypes is possibly because c002 is a single gene whereas ACYPI39568 belongs to one gene family where other members can enact similar or redundant functions. Other members of this family should be investigated to clarify the functions of the whole family.

Members of the AphidCRP family have experienced diversity in evolution. Five members have 15 cysteines, which could form homopolymers or heteropolymers in aphids. Alternative splicing occurs in ACYPI008667 [16]. The translation of JN092375 stops earlier to possibly lose its original function. The 13 genes are scattered on nine scaffolds of the genome and transcribed in both directions [16]. Phylogenetic analysis revealed no correlation between their genome location and sequence similarity. One possible reason for this gene dispersion is to escape both the adaptive and maladaptive effects of recombination [5]. The diversity of AphidCRPs may render aphids successfully adaptive to their living environment.

AphidCRP does not resemble other cysteine-rich proteins. The well-known CRP family is characterized by the presence of two LIM domains that are linked to short glycine-rich repeats [28]. In vertebrates, members of the CRP family mediate protein–protein interactions and are functional in cell differentiation, cytoskeletal remodeling, and transcriptional regulation [29]. Several CRP proteins have been identified in insects. The Mlp60A and Mlp84B of *Drosophila melanogaster* and the BmMlp of *Bombyx mori* are associated with muscle formation [30,31]. The EsMlp of *Epiblema scudderiana* is involved in the cold hardiness of diapause [32]. Cysteine-rich secretory proteins (CRISPs) contain 16 strictly conserved cysteines, 10 of which are clustered in the C-terminal of proteins [33]. In mammals, CRISPs mainly exist in the reproductive tract and function in sperm maturation, sperm ovum fusion, and immune systems [34–36]. In non-mammals, CRISPs exist mainly in the venom and can block Na^+, K^+, Ca^{2+}, and cyclic nucleotide-gated channels [33]. Several CRPs in plants and microbes are involved in different biological functions [37–39]. Basing on the sequence similarity and physical and chemical characteristics, we hypothesized that AphidCRP does not belong to the previously identified CRPs.

CXXC, CXC, and CXXXC motifs exist in AphidCRPs. The CXXC motif performs different biological functions. CXXC is the Zn^{2+} binding domain of DNA methyltransferase I and functions in distinguishing between unmethylated and hemimethylated DNA in vitro [40,41]. The CXXC domain is also the active site of thioredoxin-related proteins [42]. The trithorax gene encodes a protein that contains 27 CXXC motifs with Zn^{2+} binding ability and regulates the bithorax complex for segment determination in *Drosophila* [43]. CXXC has also been detected in the mixed-lineage leukemia protein involved in chromosomal translocation [44] in the Gal/GalNAc lectin of *Entamoeba histolytica* and variant-specific surface proteins of *Giardia lamblia* [45,46]. AphidCRPs have three CXXC motifs, which are the probable binding sites of zinc

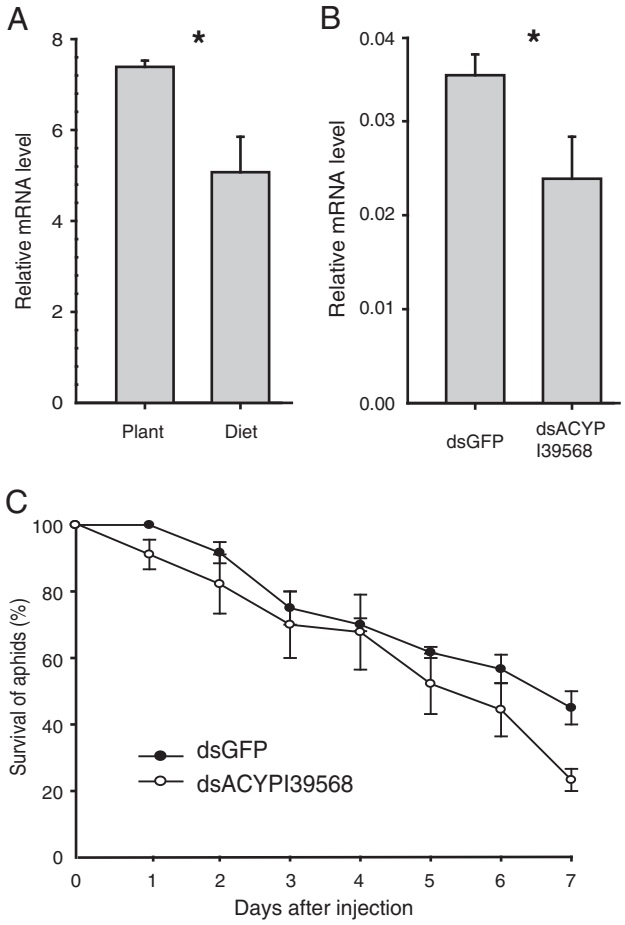


Fig. 5. Effect of ACYPI39568 on pea aphid survival on plants. (A) Transcript levels of ACYPI39568 in the heads (containing the salivary gland) of diet-fed and plant-fed aphids determined by real-time quantitative PCR. *, $P < 0.05$. (B) Transcript levels of ACYPI39568 in aphids at 3 d after dsGFP or dsACYPI39568 injection. *, $P < 0.05$. (C) Survival curves of aphids on fava beans after dsGFP or dsACYPI39568 injection.

ions. The CXC motif is common in cysteine-rich peptides [47]. CXC motifs have disulfide isomerase activity in numerous proteins, such as in the nod factor perception protein of *Medicago truncatula* [48], yphP protein of *Bacillus subtilis* [49], and purine nucleoside phosphorylases of *Sulfolobus solfataricus* and *Pyrococcus furiosus* [50]. The CXXXC motif in the rough endoplasmic reticulum protein complex prolyl 3-hydroxylase 1 and cartilage-associated protein cyclophilin B exhibits disulfide isomerase activity in vitro [51].

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bpc.2014.03.006.

Acknowledgments

The authors would like to thank Yuanyuan Chen for the assistance in the ITC assay at the Institute of Biophysics, Chinese Academy of Sciences. This work was supported by grants from the Major State Basic Research Development Program of China (973 Program) (No. 2012CB114102) and the National Natural Science Foundation of China (No. 31272364 and No. 81102745).

References

[1] G. Remaudière, M. Remaudière, Catalogue of the World's Aphididae, Institut National de la Recherche Agronomique, Paris, 1997.
[2] P.W. Miles, Aphid saliva, Biol. Rev. 77 (1999) 71–85.
[3] T. Will, W.F. Tjallingii, A. Thönnessen, A.J. van Bel, Molecular sabotage of plant defense by aphid saliva, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 10536–10541.
[4] W.F. Tjallingii, Salivary secretions by aphids interacting with proteins of phloem wound responses, J. Exp. Bot. 57 (2006) 739–745.
[5] J.C. Carolan, D. Caragea, K.T. Reardon, N.S. Mutti, N. Dittmer, K. Pappan, F. Cui, M. Castaneto, J. Poulain, C. Dossat, Predicted effector molecules in the salivary secretome of the pea aphid (*Acyrthosiphon pisum*): a dual transcriptomic/proteomic approach, J. Proteome Res. 10 (2011) 1505–1518.
[6] J.I. Bos, D. Prince, M. Pitino, M.E. Maffei, J. Win, S.A. Hogenhout, A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid), PLoS Genet. 6 (2010) e1001216.
[7] F. Cui, C.M. Smith, J. Reese, O. Edwards, G. Reeck, Polymorphisms in salivary-gland transcripts of Russian wheat aphid biotypes 1 and 2, Insect Sci. 19 (2012) 429–440.
[8] D.A. Elzing, G. Jander, The role of protein effectors in plant–aphid interactions, Curr. Opin. Plant Biol. 16 (2013) 451–456.
[9] N. Harmel, E. Létocart, A. Cherqui, P. Giordanengo, G. Mazzucchelli, F. Guillonneau, E. De Pauw, E. Haubruge, F. Francis, Identification of aphid salivary proteins: a proteomic investigation of *Myzus persicae*, Insect Mol. Biol. 17 (2008) 165–174.
[10] J.C. Carolan, C.I.J. Fitzroy, P.D. Ashton, A.E. Douglas, T.L. Wilkinson, The secreted salivary proteome of the pea aphid *Acyrthosiphon pisum* characterised by mass spectrometry, Proteomics 9 (2009) 2457–2467.
[11] S.A. Rao, J.C. Carolan, T.L. Wilkinson, Proteomic profiling of cereal aphid saliva reveals both ubiquitous and adaptive secreted proteins, PLoS One 8 (2013) e57413.
[12] W.R. Cooper, J.W. Dillwith, G.J. Puterka, Salivary proteins of Russian wheat aphid (Hemiptera: Aphididae), Environ. Entomol. 39 (2010) 223–231.
[13] W.R. Cooper, J.W. Dillwith, G.J. Puterka, Comparisons of salivary proteins from five aphid (Hemiptera: Aphididae) species, Environ. Entomol. 40 (2011) 151–156.
[14] N.S. Mutti, Y. Park, J.C. Reese, G.R. Reeck, RNAi knockdown of a salivary transcript leading to lethality in the pea aphid, *Acyrthosiphon pisum*, J. Insect Sci. 6 (2006) 1–7.
[15] N.S. Mutti, J. Louis, L.K. Pappan, K. Pappan, K. Begum, M.S. Chen, Y. Park, N. Dittmer, J. Marshall, J.C. Reese, G.R. Reeck, A protein from the salivary glands of the pea aphid, *Acyrthosiphon pisum*, is essential in feeding on a host plant, Proc. Natl. Acad. Sci. U. S. A. (2008) 9965–9969.
[16] Y. Zhang, G.X. Wu, K. Guo, W. Wang, X.P. Ding, S.M. Song, Y.Y. Xu, F. Cui, Cloning and expression profiling of a saliva protein family at different developmental stages in *Acyrthosiphon pisum* (Hemiptera: Aphididae), Acta Entomol. Sin. 54 (2011) 1445–1451.
[17] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods 8 (2011) 785–786.
[18] E. Gasteiger, C. Hoogland, A. Gattike, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, The proteomics protocols handbook, in: J.M. Walker (Ed.), Protein Identification and Analysis Tools on the ExPASy Server, Humana Press, New York, 2005, pp. 571–607.
[19] C. Steentoft, S. Vakhrushev, H. Joshi, Y. Kong, M. Vester-Christensen, K. Schjoldage, K. Lavrsen, S. Dabelsteen, N. Pedersen, L. Marcos-Silva, R. Gupta, E. Bennett, U. Mandel, S. Brunak, H. Wandall, S. Levery, H. Clausen, Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology, EMBO J. 32 (2013) 1478–1488.
[20] R. Gupta, E. Jung, S. Brunak, Prediction of N-Glycosylation Sites in Human Proteins, 10, NetNGlyc, 2004. (http://www.cbs.dtu.dk/services/NetGlyc/).
[21] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinforma. 9 (2008) 40.
[22] A. Roy, A. Kucukural, Y. Zhang, I-TASSER: a unified platform for automated protein structure and function prediction, Nat. Protoc. 5 (2010) 725–738.
[23] A. Roy, J. Yang, Y. Zhang, COFACTOR: an accurate comparative algorithm for structure-based protein function annotation, Nucleic Acids Res. 40 (2012) W471–W477.
[24] T. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95 / 98 / NT, Nucleic Acids Symp. Ser. 41 (1999) 95–98.
[25] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Mol. Biol. Evol. 28 (2011) 2731–2739.
[26] A. Cheh, J. Neilands, Struct bond, in: J.D. Dunitz (Ed.), The δ-Aminolevulinate Dehydratases: Molecular and Environmental Properties, Springer, Berlin, 1976, pp. 211–216.
[27] G. Febvay, B. Delobel, Y. Rahbe, Influence of the amino acid balance on the improvement of an artificial diet for a biotype of *Acyrthosiphon pisum* (Homoptera: Aphididae), Can. J. Zool. 66 (1988) 2449–2453.
[28] R. Weiskirchen, J.D. Pino, T. Macalma, K. Bister, M.C. Beckerle, The cysteine-rich protein family of highly related LIM domain proteins, J. Biol. Chem. 270 (1995) 28946–28954.
[29] R. Weiskirchen, K. Günther, The CRP/MLP/TLP family of LIM domain proteins: acting by connecting, Bioessays 25 (2003) 152–162.
[30] B.E. Stronach, S.E. Siegrist, M.C. Beckerle, Two muscle-specific LIM proteins in *Drosophila*, J. Cell Biol. 134 (1996) 1179–1195.
[31] J.S. Hwang, H.J. Go, T.W. Goo, E.Y. Yun, K.H. Choi, S.I. Seong, J.S. Chang, S.M. Lee, B.H. Lee, I.S. Kim, cDNA cloning and mRNA expression of LIM protein gene homologue from the silkworm, *Bombyx mori*, Biotechnol. Lett. 26 (2004) 1469–1473.
[32] T. Bilgen, T.E. English, D.C. McMullen, K.B. Storey, EsMlp, a muscle-LIM protein gene, is up-regulated during cold exposure in the freeze-avoiding larvae of *Epiblema scudderiana*, Cryobiology 43 (2001) 11–20.
[33] Y. Yamazaki, T. Morita, Structure and function of snake venom cysteine-rich secretory proteins, Toxicon 44 (2004) 227–231.
[34] D.J. Cohen, D. Busso, V. Da Ros, D.A. Ellerman, J.A. Maldera, N. Goldweic, P.S. Cuasnicu, Participation of cysteine-rich secretory proteins (CRISP) in mammalian sperm–egg interaction, Int. J. Dev. Biol. 52 (2008) 737–742.
[35] D.J. Cohen, J.A. Maldera, G. Vasen, J.I. Ernesto, M.W. Muñoz, M.A. Battistone, P.S. Cuasnicú, Epididymal protein CRISP1 plays different roles during the fertilization process, J. Androl. 32 (2011) 672–678.

[36] D.J. Cohen, V.G. Da Ros, D. Busso, D.A. Ellerman, J.A. Maldera, N. Goldweic, P.S. Cuasnicú, Participation of epididymal cysteine-rich secretory proteins in sperm–egg fusion and their potential use for male fertility regulation, Asian J. Androl. 9 (2007) 528–532.
[37] S. Busby, R. Ebright, Transcription activation by catabolite activator protein (CAP), J. Mol. Biol. 293 (1999) 199–213.
[38] J. Darlix, M. Lapadat-Tapolsky, H. de Rocquigny, B.P. roques, First glimpses at structure function relationships of the nucleocapsid protein of retrovirused, J. Mol. Biol. 254 (1995) 523–537.
[39] I. Petty, R. donald, A. Jackson, Multiple genetic determinants of barley stripe mosaic virus influence lesion phenotype on *Chenopodium amaranticolor*, Virology 198 (1994) 218–226.
[40] T.H. Bestor, Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain, EMBO J. 11 (1992) 2611–2617.
[41] M. Pradhan, P.O. Estève, H.G. Chin, M. Samaranayke, G.D. Kim, S. Pradhan, CXXC domain of human DNMT1 is essential for enzymatic activity, Biochemistry 47 (2008) 10000–10009.
[42] S. Quan, I. Schneider, J. Pan, A. Von Hacht, J.C. Bardwell, The CXXC motif is more than a redox rheostat, J. Biol. Chem. 282 (2007) 28823–28833.
[43] A.M. Mazo, D.H. Huang, B.A. Mozer, I.B. Dawid, The trithorax gene, a trans-acting regulator of the bithorax complex in *Drosophila*, encodes a protein with zinc-binding domains, Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 2112–2116.
[44] Z.B. Xia, MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 8342–8347.
[45] B. Loftus, I. Anderson, R. Davies, U.C.M. Alsmark, J. Samuelson, P. Amedeo, P. Roncaglia, M. Berriman, R.P. Hirt, B.J. Mann, The genome of the protist parasite *Entamoeba histolytica*, Nature 433 (2005) 865–868.
[46] R.D. Adam, Y.M. Yang, T.E. Nash, The cysteine-rich protein gene family of *Giardia lamblia*, Mol. Cell. Biol. 12 (1992) 1194–1201.
[47] K.J. Woycechowsky, R.T. Raines, The CXC motif: a functional mimic of protein disulfide isomerase, Biochemistry 42 (2003) 5387–5394.
[48] B. Lefebvre, D. Klaus-Heisen, A. Pietraszewska-Bogiel, C. Hervé, S. Camut, M.C. Auriac, V. Gasciolli, A. Nurisso, T.W. Gadella, J. Cullimore, Role of N-glycosylation sites and CXC motifs in trafficking of *Medicago truncatula* Nod factor perception protein to plasma membrane, J. Biol. Chem. 287 (2012) 10812–10823.
[49] U. Derewenda, T. Boczek, K.L. Gorres, M. Yu, L.W. Hung, D. Cooper, A. Joachimiak, R.T. Raines, Z.S. Derewenda, Structure and function of *Bacillus subtilis* YphP, a prokaryotic disulfide isomerase with a CXC catalytic motif, Biochemistry 48 (2009) 8664–8671.
[50] G. Cacciapuoti, I. Peluso, F. Fuccio, M. Porcelli, Purine nucleoside phosphorylases from hyperthermophilic *Archaea require* a CXC motif for stability and folding, FEBS J. 276 (2009) 5799–5805.
[51] Y. Ishikawa, H.P. Bächinger, An additional function of the rough endoplasmic reticulum protein complex prolyl 3-hydroxylase 1 cartilage-associated protein cyclophilin B the CXXXC motif reveals disulfide isomerase activity in vitro, J. Biol. Chem. 288 (2013) 31437–31446