



HAL
open science

Membrane binding of zebrafish actinoporin-like protein: AF domains, a novel superfamily of cell membrane binding domains

Ion Gutiérrez-Aguirre, Peter Trontelj, Peter Macek, Jeremy H. Lakey, Gregor
Anderluh

► **To cite this version:**

Ion Gutiérrez-Aguirre, Peter Trontelj, Peter Macek, Jeremy H. Lakey, Gregor Anderluh. Membrane binding of zebrafish actinoporin-like protein: AF domains, a novel superfamily of cell membrane binding domains. *Biochemical Journal*, 2006, 398 (3), pp.381-392. 10.1042/BJ20060206 . hal-00478529

HAL Id: hal-00478529

<https://hal.science/hal-00478529>

Submitted on 30 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Membrane binding of zebrafish actinoporin-like protein: AF DOMAINS, A NOVEL SUPERFAMILY OF CELL MEMBRANE BINDING DOMAINS

Ion Gutiérrez-Aguirre^{*}, Peter Trontelj^{*}, Peter Maček^{*}, Jeremy H. Lakey[†], Gregor Anderluh^{*,‡}

^{*} Department of Biology, Biotechnical faculty, University of Ljubljana, Večna pot 111, 1000 Ljubljana, Slovenia

[†] Institute for Cell and Molecular Biosciences, University of Newcastle upon Tyne, Framlington place, Newcastle upon Tyne, NE2 4HH, UK

[‡] Corresponding author: E-mail, gregor.anderluh@bf.uni-lj.si

SYNOPSIS

Actinoporins are potent eukaryotic pore-forming toxins specific for sphingomyelin-containing membranes. They are structurally similar to members of the fungal fruit-body lectin family that bind cell-surface exposed Thomson-Friedenreich antigen. In this study we found a number of sequences in public databases similar to actinoporins. They originate from three animal and two plant phyla and can be classified in three families according to phylogenetic analysis. The sequence similarity is confined to a region from the C-terminal half of the actinoporin molecule and comprises the membrane binding site with a highly conserved P-[WYF]-D pattern. A member of this novel actinoporin-like protein family from zebrafish was cloned and expressed in *E. coli*. It displays membrane-binding behaviour, but does not have permeabilising activity nor sphingomyelin specificity, two properties typical of actinoporins. We propose that the three families of actinoporin-like proteins and the fungal fruit body lectin family comprise a novel superfamily of membrane binding proteins, tentatively called AF domains (abbreviated from actinoporin-like proteins and fungal fruit-body lectins).

Running title: AF domains, a novel superfamily of membrane binding domains

Keywords:

Protein-membrane domains / AF domains / Actinoporins / Fungal fruit body lectins / Surface plasmon resonance / Lipid monolayer

INTRODUCTION

The protein world is classically divided into soluble and membrane proteins, the latter comprising approximately 30% of all proteins. There is, however, another large group that spans both definitions. Peripheral or amphitrophic proteins are soluble proteins that transiently interact with the lipid membrane [1]. Such interactions are fundamental for many biological processes, as only temporary or permanent binding to cellular membranes elicits certain biological functions. Many examples of protein-membrane interactions have been described at the molecular level. Proteins that are involved in cell signalling events are by far the largest and best-known group of peripheral proteins [2]. Examples include protein kinase C (PKC) conserved 1 (C1) and PKC conserved 2 (C2) domains, pleckstrin homology (PH) domains, Phox domains (PX), Fab1, YOTB, Vac1, and EEA1 domains (VYFE) [2], etc. Examples of other amphitrophic proteins are enzymes that act on or at the membrane [3], intracellular proteins involved in the transport of lipids between cellular compartments [4], annexins [5] and proteins that participate in antimicrobial defence [6].

Furthermore, a large number of protein toxins attach to membrane surfaces during the first step of transmembrane pore formation [7]. Creation of transmembrane pores is a very efficient way of killing cells, as the lipid membrane is the first barrier that needs to be overcome by invading molecules. So-called pore-forming toxins (PFT) are thus one of the biggest and most important groups of natural toxins [8]. PFT are extremely heterogeneous in terms of their lipid or protein receptor specificity. Some of them require a particular lipid, i.e. actinoporins, PFT from sea anemones [9], require sphingomyelin (SM) for optimal activity. Actinoporins are a distinct eukaryotic PFT family. They are basic proteins of about 20 kDa folded in a compact β -sandwich flanked on each side by α -helices [10-12]. The most studied representative is equinatoxin II (EqII) from the sea anemone *Actinia equina*. The unique molecular mechanism of actinoporin action has been unravelled in the last few years. In the first step they attach to SM-containing membranes [12-15], then they transfer an amphipathic N-terminal helix to the lipid-water interface [16,17] and, finally, monomers aggregate on the membrane surface to form a (probably tetrameric) final pore [16,18,19]. The suggested pore is composed of four α -helices, one from each monomer, and very likely also includes membrane lipids [20,21]. Actinoporins thus belong to the α -helical PFT group [22].

For a long time it was believed that actinoporins represent an isolated protein family exclusive to sea anemones and dissimilar to any other protein family. This assumption was

first challenged by the discovery of a hemolytic toxin, from the salivary gland of a marine gastropod, with high sequence similarity to actinoporins [23]. More recently reported structures of fungal lectins from *Xerocomus chrysenteron* (XCL) [24] and *Agaricus bisporus* (ABL) [25] revealed remarkable structural similarity to actinoporins despite less than 15% sequence identity. They both belong to a fungal fruit body lectin protein family (Pfam code PF07367). Members of this family have insecticidal [26] and antiproliferative properties [27]. Both XCL and ABL selectively bind Thomsen-Friedenreich antigen (T-antigen), a Gal β 1-3GalNAc disaccharide present on glycoproteins of some malignant cells [25,28].

In this work we have performed extensive searches of public sequence databases and found a large number of actinoporin-like sequences, which arise from three metazoan and two plant phyla. We found that the conserved domain corresponds to the C-terminal part of EqII including the putative SM-binding site. Actinoporin-like sequences were classified in three families according to phylogenetic analysis. We hypothesize that these sequences are not toxins but intracellular proteins that are able to attach to lipid membranes. To test this hypothesis, we have expressed a representative from zebrafish and found it to bind to lipid membranes *in vitro*. We propose that the three families of actinoporin-like proteins, together with the fungal fruit body lectin family, form a new superfamily of cell membrane binding domains, tentatively called AF domains.

EXPERIMENTAL

Materials - 1,2-Dioleoyl-sn-Glycero-3-Phosphocholine (DOPC), 1,2-Dioleoyl-sn-Glycero-3-[Phospho-rac-(1-glycerol)] (DOPG) and bovine brain sphingomyelin (SM) were from Avanti Polar Lipids (Alabaster, AL, USA). 1-anilinonaphtalene-8-sulfonic acid (ANS) was from Molecular Probes (Eugene, OR). Guanidinium hydrochloride (GdmHCl) was from Merck (Darmstadt, Germany). All other chemicals were from Sigma unless stated otherwise.

Bioinformatics - Public databases were searched by blastp and tblastn algorithms [29]. EqtII sequence was used initially as a probe. When similar sequences were retrieved, databases were also searched with species-specific actinoporin-like protein (ALP) sequences. All available protein and nucleotide databases (as of 5th January 2006) at NCBI were searched, as well as species-specific genomic databases on the Ensembl Genome Browser for Fugu, zebrafish and other vertebrates (human, mouse, Xenopus) (<http://www.ensembl.org>) and Genoscope for *Tetraodon nigroviridis* (<http://www.genoscope.cns.fr>). Expressed sequence tags (EST) were assembled by the CAP EST assembler (<http://bio.ifom-firc.it/ASSEMBLY/assemble.html> [30]). Overlap length was taken as 30 and overlap percent identity was 90. Singletons or contigs were translated in all six frames to retrieve the ALP sequence. All sequences obtained in this way have a low E-value, typically around $1e^{-8}$. Signal peptides were predicted by SignalP [31]. Sequences were also analysed for the presence of other protein domains by using tools from the InterPro database. Secondary structure content was estimated from the amino acid sequence by using PSIPRED [32]. Estimation of secondary structure content from far-UV circular dichroism spectra was done on Dichroweb [33] by using CDSSTR [34].

Phylogenetic Analysis – The final alignment of 98 protein sequences used for the phylogenetic analysis was obtained in three steps. First, an alignment, created by MUSCLE [35], was used to obtain a preliminary neighbour-joining tree. Second, a subset of sequences representing all main branches were aligned taking into account known DSSP-defined secondary structures with the aid of PRALINE [36]. Small corrections were made by hand in order to incorporate known secondary structure information of actinoporins [10-12]. Third, the "profile" option of MUSCLE was used to align the rest of the sequences to the secondary structure based alignment. Alternatively, all sequences were aligned using PRALINE and taking into account pre-defined secondary structures without manual interference. This approach introduced slightly less gaps (alignment lengths were 326 and 315, respectively). Both alignments yielded trees with largely identical topologies indicating that the alignment-

sensitive positions did not bear a substantial phylogenetic signal. An optimal model of amino acid evolution was selected using Modelgenerator (<http://bioinf.nuim.ie/software/modelgenerator>). The model preferred by Akaike Information Criterion two and Bayesian Information Criterion was a Whelan-Goldman [37] model with a gamma-distributed rate heterogeneity among sites. The search for optimal trees was conducted under the maximum likelihood criterion using the hill-climbing algorithm implemented in PHYML [38]. The shape parameter of a discrete (four rates) approximation of the gamma distribution was optimised via maximum likelihood. In the absence of suitable outgroups, midpoint rooting was used to tentatively root the tree.

Cloning and Expression - The coding region for predicted ALP from zebrafish (*Danio rerio*) (Dr1) was amplified by polymerase chain reaction (PCR) from I.M.A.G.E. EST clones 7267252 (GenBank (gb) accession number CN510193) and 7038734 (gb CF998026) by oligonucleotides Dr1+ (5'-CGCGGATCCACTGAGTCTGCCGAGG-3') and Dr1- (5'-GGAATTCACGCGTTAGCAATCCATCTGAGCC-3'). The PCR product was cleaved with *Bam*HI and *Mlu*I and inserted in a precleaved T7-based expression vector. A modified version of pET8c vector was used that leaves H₆S₂LVPR tag with thrombin recognition sequence at the N-terminus for easier purification. The correct construction of the plasmid was verified by nucleotide sequencing. Dr1 was expressed in an *E. coli* Origami (DE3) pLysS strain (Novagen). 20 ml of overnight culture were used to inoculate 0.5 l of M9-LB medium supplemented with 100 µg ml⁻¹ of ampicillin and 25 µg ml⁻¹ of chloramphenicol. Expression of Dr1 was induced at an A₆₀₀ of approximately 0.8 by the addition of isopropyl-1-thio-β-D-galactopyranoside to a 1 mM final concentration. Cells were grown for an additional 3 h, spun down and frozen. Seven grams of frozen cells were thawed into 17.5 ml of 50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 20 mM β-mercaptoethanol, 0.5 mg ml⁻¹ lysozyme, 10 µg ml⁻¹ DNase, 20 µg ml⁻¹ RNase, 0.5 mM phenylmethylsulfonyl fluoride, and 1 mM benzamidine. They were incubated on ice for 45 min with occasional vigorous shaking. Bacteria were sonicated twice during the incubation for 2 min by using a microtip (MSE 150W ultrasonic disintegrator). Broken cells were centrifuged for 30 min at 4 °C at 11,000 g (Sigma). The supernatant was stored at 4 °C, and the pellet was washed with 7.5 ml of the same buffer. Incubation for 20 min on ice was followed by 2-min sonication and centrifugation at the same conditions. The supernatants were merged and applied to a 1-ml Ni-NTA column (Qiagen, Crawley, UK) equilibrated in 50 mM NaH₂PO₄, pH 8.0, 10 mM imidazole, and 20 mM β-mercaptoethanol. Unbound protein was eluted with the same buffer,

and the bound Dr1 was eluted with 300 mM imidazole in the same buffer. The protein was dialyzed three times against 2 l of 5 mM Tris-HCl, pH 8.0 at 4 °C. The histidine tag was cleaved from Dr1 with thrombin according to the supplier's specifications (Novagen). Ion-exchange chromatography on an Äkta™ FPLC™ system (Amersham Biosciences) was used as a final step in purification. The sample was applied to a MonoQ column equilibrated in 20 mM Tris-HCl, pH 8.0 and bound Dr1 was eluted from the column by a gradient of the same buffer with 2 M NaCl. Fractions shown to contain Dr1 by SDS-PAGE gel, were merged, aliquoted and stored at -20°C. Extinction coefficients, $\epsilon^{0.1\%}$, calculated from the sequence at ExPASy Proteomics tools Internet site (<http://us.expasy.org/tools/>), for Dr1 and His-tagged Dr1 used were 1.442 and 1.347, respectively.

Fluorescence Measurements - All fluorescence measurements were performed on a Jasco FP-750 spectrofluorometer (Jasco Corporation, Japan). The sample compartment was equipped with a Peltier thermostated single-cell holder. All experiments were done at 25 °C under constant stirring. Excitation wavelength was fixed at 295 nm, in order to eliminate the contribution of the tyrosines, and the emission spectra were recorded between 310 and 400 nm. Excitation and emission slits were set at 5 nm. Protein was preincubated in 20 mM DTT before spectra acquisition. Protein concentration in the cuvette was 1 μ M in a final volume of 1400 μ l. Buffer was Hepes 10 mM, NaCl 200 mM, DTT 20 mM, pH 7.5, and, where indicated, 6 M GdmHCl was included in the buffer. For the iodide quenching experiments the concentration of Dr1 was 1 μ M in the same buffer. Spectra were recorded in the absence and in the presence of increasing iodide concentrations. The stock of iodide was composed of 2.5 M KI and 1 mM Na₂S₂O₃ 1 mM. All spectra were corrected with the corresponding spectra of the buffer alone and for the dilution caused by the addition of iodide. No further correction for wavelength dependent sensitivity was done. The values of the collisional quenching constant were obtained from the Stern-Volmer equation [39]:

$$F_0/F = K_{SV} * [Q] + 1$$

where F_0 is the fluorescence intensity in the absence of iodide, F is the fluorescence intensity in the presence of iodide, K_{SV} is the collisional Stern-Volmer constant and Q is iodide concentration. The fraction of fluorophores accessible to the soluble quencher (f_a) was calculated by using the Stern-Volmer equation modified for multiple emission centres [39]:

$$F_0/(F_0 - F) = (1/f_a * K_{eff}) * 1/Q + 1/f_a$$

where f_a is the fraction of accessible tryptophans and K_{eff} is the effective quenching constant.

For the ANS binding assay, concentration of Dr1 and ANS in 10 mM Hepes, 200 mM NaCl, 10 mM DTT, pH 7.5 was 15 μ M. Measurements were taken between 12 °C and 90 °C at a heating rate of 1 °C/min. Excitation wavelength was 370 nm and emission was measured at 488 nm. Excitation and emission slits were set at 5 nm. The T_m was calculated as the mean value between the temperature at which fluorescence starts to increase and the temperature at which maximum fluorescence signal is reached.

Circular Dichroism Spectroscopy - Circular dichroism (CD) spectra were measured by Aviv CD spectrometer (Lakewood, N.J., USA). Bandwidth was set to 1 nm and 0.5 nm for far-UV and near-UV CD spectra, respectively. The signal was averaged for 5 s at each wavelength. The temperature was set to 20 °C. The protein concentration in 20 mM NaH₂PO₄, pH 8.0, was 8 μ M in a 1 mm pathlength cuvette and 30 μ M in a 1 cm pathlength cuvette for far-UV and near-UV CD spectra, respectively. Spectra were scanned from 250-190 nm in the far-UV CD region and from 320-240 in near-UV CD region. Dr1 concentration was 6 μ M for unfolding experiments in a 1mm pathlength cuvette. The CD signal at 208 nm was followed at 0.5°C increments. For each data point the signal was averaged over 5 s. Temperature was increased at a rate of 1°C/min.

Surface Pressure Measurements - Surface pressure measurements were carried out with a MicroTrough-S system from Kibron (Helsinki, Finland) at room temperature under constant stirring. The aqueous subphase consisted of 500 μ l of 10 mM Hepes, 200 mM NaCl, pH 7.5. For monolayer insertion experiments the appropriate lipid mixture, dissolved in chloroform:methanol (2:1, v:v), was gently spread over the subphase to create a monolayer. The desired initial surface pressure was attained by changing the amount of lipid applied to the air-water interface. After 10 minutes, for solvent evaporation, the protein, preincubated in the presence of 10 mM DTT was injected through a hole connected to the subphase. The final Dr1 concentration in the Langmuir trough was 2 μ M. The increment in surface pressure vs. time was recorded until a stable signal was obtained.

Surface Plasmon Resonance - Surface plasmon resonance (SPR) measurements were performed on a Biacore X (Biacore, Sweden) apparatus at 25°C. An L1 chip was equilibrated in 20 mM Tris-HCl, 140 mM NaCl, 1 mM EDTA, pH 8.5. Large unilamellar vesicles were prepared by extrusion as described previously [13]. The liposome-coated chip surface was prepared as described in Anderluh et al. [40]. Briefly, liposomes were passed at 0.5 mM lipid concentration across the chip for 10 min at 1 μ l min⁻¹ and were washed with two injections of 100 mM NaOH for 1 min at 30 μ l min⁻¹. Nonspecific binding sites were blocked by one

injection of 0.1 mg ml^{-1} bovine serum albumin for 1 min at $30 \text{ } \mu\text{l min}^{-1}$. For the binding experiment Dr1 was injected at the desired concentration for 90 s at $30 \text{ } \mu\text{l min}^{-1}$. Four different concentrations of Dr1 were used in each experiment. After the sensorgrams were corrected for the contribution of blank injections of buffer, association (k_a) and dissociation (k_d) rate constants were globally fitted to a 1:1 Langmuir model by using a BIAevaluation 3.2 software (Biacore AB). The first ten seconds of the injection and of the dissociation were not included in the fit due to refractive index effects. The dissociation constant (K_D) was calculated as follows:

$$K_D = k_d/k_a.$$

RESULTS

Identification of Actinoporin-Like Protein (ALP) Families – We were able to retrieve 72 sequences by using EqtII sequence as a probe in exhaustive searches of public databases (Figure 1, Table 1). All retrieved sequences had a significant E-value that was typically around 10^{-8} - 10^{-9} . Sequences were from three animal (chordates, cnidarians and molluscs) and two plant (mosses and ferns) phyla. The majority of the sequences were EST clones from teleost fishes (36 sequences) (Table 1). Twelve sequences were actinoporins, twelve further non-actinoporin sequences were found in Cnidaria and the rest were from mosses (3), ferns (1), *Xenopus* (4), chicken (2), skate (1) and the marine gastropod *Monoplex echo* (1). No hits with significant E-value were found in other organisms, i.e. bacteria, *Caenorhabditis elegans*, *Drosophila*, rat, mouse, human or other plants. The ALP sequences are not expressed in large copy number, as the number of EST found was for all organisms very low (Table 1), i.e. typically from 0.002% (15 EST ALP clones from 689581 total in zebrafish; dbEST release 111105, 11th November 2005) to 0.045% (80 EST ALP clones from 179615 total in *Hydra*).

The similarity to EqtII is confined to the C-terminal EqtII region 83-179 (Figure 1 and 2A). The conserved region can be described as a fingerprint of three separate regions with high local similarity. All fish sequences have an insertion of five amino acids between first two parts (Figure 1). This insertion corresponds to the loop between strands $\beta 6$ and $\beta 7$ of the EqtII structure. The second part is the most conserved of all. It possesses an invariant hydrophobic residue (Val¹⁰² in EqtII) and a pattern P-[WYF]-D (Figure 1) located on the broad loop between strands $\beta 7$ and $\beta 8$ (Figure 2A). The most variable part of the domain by length is the insertion between the second and third parts. It is 13-53 residues long and in EqtII comprises the loop between helix B, strand $\beta 10$ (including short strand $\beta 9$). The last motif of the domain encompasses EqtII strands $\beta 11$ and $\beta 12$. It is the least conserved of all three and possesses only two residues that are conserved in more than 90% of all ALP sequences. The conserved region is thus structurally well defined, encompassing almost half of the EqtII molecule (Figure 2A) and is arranged in two layers which flank helix B. Some of the most conserved residues of the domain help maintain its 3D structure. Thus, the space between the two layers is filled with side chains of the hydrophobic amino acids, i.e. Phe¹⁰⁴, Val¹⁰², Val¹¹⁹, Leu¹⁰⁰ and Ile¹²¹ from the second part are in close contact with Ile¹⁷⁴ and Met¹⁶⁴ from the third part of the domain (Figure 2A). Similarly, Leu¹⁰³ from $\beta 5$ and Leu¹³⁶ from helix B are in close contact and position helix B on one side of the layer (Figure 2A).

Other conserved residues together with P-[WYF]-D pattern overlap well with the actinoporin phosphocholine (POC) binding site (Figure 1 and 2A) [12].

No other known protein domains were found in ALP sequences according to searches in the InterPro database (Figure 2B). Furthermore, only a fraction of sequences possess a predicted signal peptide. Those that do include all the actinoporins, seven sequences from Cnidaria, the hemolysin from *Monoplex echo* and only one fish sequence, Hc1.

The similarity with fungal lectins was first revealed by comparison of 3D structures [24]. We did not retrieve sequences of the fungal fruit body lectin family from databases by using EqtlI or other ALP sequences as a probe, since sequence identity to actinoporins is only 11-14%. Nevertheless, there is a high degree of structural conservation with RMSD value of 1.38 Å over 68 C^α atoms of equivalent residues [24] (Figure 2A). Crucially, a binding site for cell-surface T-antigen in ABL is located on the same side of the β-sandwich as the POC site of actinoporins. Structurally equivalent amino acids form the binding site in both cases (Figure 2C) [24,25]. The fungal fruit body lectin family is highly conserved. XCL and other family members have 19-65% identical residues. Apart from sequences presented in the alignment of Figure 2 in Birck et al. [24], we were able to find members of this family in 6 other fungal species and also plants *Marchantia polymorpha* and *Tortula ruralis*, by using XCL as a probe (Table 1). Interestingly, *Tortula ruralis* is at the moment the only organism to possess both ALP and fungal fruit-body lectin family sequences.

The retrieved actinoporins and actinoporin-like proteins, along with fungal fruit body lectins, constitute a highly divergent group of hypothesized homologues. Particularly the lectins, sharing only 6 to 17% of identical amino acid position with other sequences, make a purely sequence-based alignment practically impossible. However, the highly conserved, nearly identical secondary structure (Figure 2C) [24] enabled a clear homology assessment between different regions of the proteins.

Phylogenetic analysis revealed four distinct groups of proteins (Figure 3) Between-group sequence similarity (12-17%) was substantially lower than similarity within groups (45-69%). Based on these results, we tentatively assign these groups to four protein families: (1) ALPs found predominantly in vertebrates, (2) hydrozoan ALPs, (3) actinoporins, and (4) fungal fruit body lectins. The first three groups are referred to as ALP families. The gap in average sequence similarity between the three ALP groups (27%) and lectins (14%) favour a separate phylogenetic position of the latter. The relationships between the ALP families are less obvious, although an outermost position of hydrozoan ALPs seems to be indicated.

Strikingly, a small group of proteins from mosses and ferns are phylogenetically so closely associated with actinoporins that they can be considered as members of the same family.

Membrane-binding behaviour of ALPs - On the basis of the sequence similarity and conservation of residues that participate in POC binding, we hypothesized that ALP will all have membrane-binding behaviour. To check for this, we chose one of the ALP from fishes, zebrafish Dr1, and checked its membrane binding ability with various *in vitro* systems.

Preparation and Characterization of Dr1 – Dr1 was expressed in *Escherichia coli* as a His-tagged protein, purified from the bacterial cytoplasm and cleaved with thrombin to remove His-tag. This procedure yielded pure protein of the expected size 21 kDa (Figure 4). First we needed to confirm that Dr1 is folded. As the ALP domain represents more than half of this protein we assumed that Dr1 structural characteristics are similar to EqtII. PSIPRED estimation of secondary structure content from sequence is H11/S39/T+R50 (percentage of helix/strand/turn + random) (Figure 4). The far-UV CD spectrum of Dr1 confirmed its folded nature and was strikingly similar to EqtII [17]. It has a negative maximum at 218 nm, indicating a high content of β -structure (Figure 4). Indeed, estimation of the secondary structure content by CDSSTR gave H6/S37/T24/R32 in agreement with the estimation from sequence. Thus, Dr1 is a mainly β -sheet protein, very similar to EqtII in secondary structure content. The spectral features of Dr1 are lost when the protein was denatured in 6 M GdmHCl (Figure 4). The spectrum in the presence of GdmHCl is typical of denatured proteins, indicating that Dr1 in normal buffer is folded. Furthermore, Dr1 also has a well defined, and complex, near UV-CD spectrum, indicating a fixed chiral environment of aromatic residues (Figure 4).

The fluorescence spectra of Dr1 in buffer and in the presence of 6M GdmHCl are shown in Figure 5. The emission maximum of Dr1 in solution is 342 nm, typical of a protein, which excludes some of its fluorophores from the solvent. The emission maximum of the spectra in the presence of GdmHCl shows a red-shift to 354 nm, indicative of unfolding and exposure of its three tryptophans to the solvent. We further quenched tryptophan fluorescence of Dr1 by iodide. Stern-Volmer plots for the quenching in the absence and presence of 6M GdmHCl are shown on Figure 5. Quenching constant, K_{sv} , obtained from the slope, is 3.71 M^{-1} for the unfolded sample. The data for Dr1 in buffer deviated from linearity and gave K_{sv} of 1.72 M^{-1} . This indicates that in folded Dr1 a fraction of the tryptophans are hidden within the protein fold, being inaccessible to the iodide. When using a modified Stern-Volmer plot for multiple emitting centres (Figure 5, inset) we obtained the fraction of tryptophans

accessible to the iodide (f_a) of approximately 43%, which means that at least one of the tryptophans is fully exposed to the solvent while the other two are buried or partially buried in the hydrophobic core of the protein.

Finally we characterized the thermostability of Dr1 by ANS binding assay and CD spectroscopy. ANS increases quantum yield when it binds to newly exposed hydrophobic patches of temperature-denatured proteins (Figure 5). The denaturation temperature was calculated as the mean value between the temperature at which fluorescence starts to increase and the temperature at which maximum fluorescence signal was reached, which in the case of Dr1 was 46.3°C. A similar transition was observed with CD when the signal was followed at 208 nm. In this case a T_m obtained was 47.6°C, in close agreement with ANS data.

Based on all of the above results we concluded that Dr1 is folded with at least some of the tryptophans buried in the interior of the protein. Although the melting temperature of Dr1 is lower than EqtII [17,41], the other spectral characteristics are similar. We, therefore, continued to characterize the properties of Dr1 and initially checked whether Dr1 possesses permeabilising activity similar to EqtII. As predicted Dr1, which lacks a permeabilizing amphipatic helix, could not release fluorescent probe calcein from DOPC/SM liposomes, typically used to monitor Eqt activity (data not shown). We then checked whether it can insert into lipid monolayers or bind to immobilized liposomes.

Surface Pressure Measurements – First, we measured the ability of Dr1 to partition to the air-water interface. Dr1 readily interacts with this lipid-free surface in a concentration dependant manner (Figure 6). The maximum increment in the surface pressure was approximately 6 mN/m. Saturation was obtained at 2 μ M protein, so this was the concentration we used in the assays in the presence of lipids. Dr1 is also able to insert into lipid monolayers of different compositions. Figure 6 shows the kinetics of insertion of Dr1 into lipid monolayers composed of DOPC, DOPC:SM (1:1 mol:mol), and DOPC:DOPG (1:1 mol:mol), at an initial pressure of 15 mN/m. Dr1 inserts readily in all three monolayers reaching a final value of approximately 25 mN/m. We then performed similar kinetics at different initial pressures and by plotting the increment in the surface pressure vs the initial pressure we calculated the critical pressure, i.e., the initial pressure at which no protein can insert in the monolayer (Figure 6). The values of the critical pressures were similar for all three compositions: 27.9 mN/m, 27.2 mN/m and 27.0 mN/m for DOPC, DOPC:SM and DOPC:DOPG monolayers, respectively.

Surface Plasmon Resonance Measurements – The binding of Dr1 was tested on supported liposomes by using an L1 chip on a Biacore X. Dr1 readily binds to liposomes composed of either DOPC, DOPC:SM (1:1 mol:mol), and DOPC:DOPG (1:1 mol:mol) (Figure 7). Blanks of the buffer without protein were subtracted from the raw sensorgrams and curves were then fitted with 1:1 Langmuir model. Reasonable fits with low χ^2 values typically around 1 were obtained (Figure 7). Again, in these surface plasmon resonance experiments Dr1 showed a similar level of binding for all lipid compositions tested, in agreement with monolayer insertion data. Kinetic parameters were quite similar for DOPC and DOPC:SM membranes (Table 2). For DOPC:DOPG the association rate was slightly lower and the dissociation rate was higher resulting in a K_D twice as big as in the case of DOPC liposomes.

DISCUSSION

Actinoporin-like Protein Families – A large number of sequences with a considerable degree of similarity to EqII have been retrieved from databases. Phylogenetically, they can be attributed to three distinct families. ALP is obviously an ancient domain, as it spans the metazoan kingdom from diploblasts to vertebrates, and is present in the genomes of some representatives of mosses and ferns. To further discuss phylogenetic and evolutionary scenarios of ALP sequences we need to keep in mind the taxonomic bias of the databases from which most sequences were retrieved. This bias notwithstanding, it is remarkable that despite the wide range of taxa sharing ALP sequences, many large groups with extensive available genomic information lack these domains altogether. Particularly noteworthy is their absence from protozoans, arthropods, nematodes, mammals and flowering plants. The phylogenetic pattern emerging on the maximum likelihood tree (Figure 3) seems to be consistent with a pre-metazoan origin followed by substantial subsequent differentiation in cnidarians, giving rise to the actinoporins and hydrozoan ALPs. The third monophyletic group and ALP family includes largely triploblastic animals, with invertebrates heavily under represented. This under representation seems to be real, not just a consequence of taxonomic bias.

ALP sequences are most abundantly expressed in teleost fishes (Table 1), where up to seven isoforms could be observed in a single species, i.e. in Medaka (*Oryzias latipes*), while some of them, common carp, zebrafish or salmon, have only one or two copies. The medaka fish paralogs are scattered among all five distinct teleost clades indicating that a high genomic diversity might be ancestral to this group. This finding is in line with recently discovered large-scale genome level duplications within the stem lineage of the Actinopterygii (ray-finned fishes) [42]. There are two possible explanations for the lack of diversity in some teleost species and the generally low number of EST clones: first, secondary loss following reduced selection pressure, and second, under representation due to methodological bias. The fact that some of them are expressed in specific developmental stages (see below) supports the latter explanation.

As a general observation it can be concluded that few groups of metazoans display a diverse assemblage of ALPs, particularly anthozoans, hydrozoans, teleosts, and possibly frogs and birds. In some other groups ALPs seem to present as genomic relics (e. g. in snails and cartilaginous fishes), whereas in the majority of well-studied metazoan taxa these genes were lost. Such an unbalanced taxonomic distribution suggests that the biological functions

conducted by ALPs are both diverse and specialized. This is underscored by actinoporins found in mosses and ferns, where their function might be fundamentally different from how they act in sea anemones. Furthermore, in the moss *Tortula ruralis* ALP domains are quite abundantly expressed, present in up to 0.6% of total EST clones available, indicating important functional role. At the present stage of taxonomic sampling it is difficult to offer a plausible explanation for the isolated occurrence of plant ALPs amidst the large metazoan assemblage.

All ALP sequences, apart from actinoporins and the previously characterized echotoxin from the marine snail *Monoplex echo* [23], represent as yet undefined, novel proteins. They are mostly intracellular proteins, since only a few possess a predicted signal peptide. Of those that do, actinoporins and echotoxin, are secreted toxins that also possess an additional amphipathic α -helix at their N-terminus which is used for efficient pore-formation (Figure 1F) [14,43]. In addition, a large fraction of the sequences from *Hydra* also have predicted signal peptide, and it was proposed that they act as allomones for preying or defence using a mechanism similar to actinoporins [44].

The fact that there are so many different fish ALP makes it likely that at least some of them are functional copies. In fact, the expression of two of them was confirmed in Medaka fish by a large *in situ* hybridisation screen of stages 18, 24 and 32 [45]. O13 (MEPD clone 631-136-06-O) was abundantly and homogeneously expressed in all stages, while O14 (MEPD 631-135-09-M) was differentially expressed in the enveloping layer in stages 18 and 24. The exact biological roles of the majority of ALP still need to be determined. It has not escaped our notice that ALP sequences are presented mostly in aquatic organisms. This fact, their abundance and preservation through evolution in widely separated taxonomic groups indicate a functional importance.

Properties of Dr1 – On the basis of EqtII similarity we wanted to check whether these proteins use the conserved ALP domain for membrane attachment. To test this, we cloned a zebrafish ALP, Dr1, and checked its membrane binding ability by biophysical methods, and checked its distribution pattern when expressed in cells. According to the structural characterization, recombinant Dr1 is a folded protein with a measured secondary structure content close to predicted and very similar to EqtII. It, however, is not hemolytic and does not permeabilize liposomes. It also does not show the SM specificity typical of actinoporins, as the kinetic parameters for the binding to immobilized liposomes or insertion into monolayers composed of DOPC or DOPC/SM membranes were very similar (Table 2).

Electrostatic interactions do not contribute significantly to the binding, as the presence of excess negative charge in DOPC/DOPG membranes does not improve the binding. Estimated equilibrium dissociation constants are two orders of magnitude larger than for EqtII binding to the tethered lipid membrane composed of DOPC/SM 1/1 [14]. This can be explained in part by the contribution of the N-terminal amphipathic helix that is not present in Dr1 and, furthermore, Dr1 also lacks the equivalent of EqtII Trp¹¹², which is replaced by Leu. This exposed tryptophan in EqtII contributes significantly to the binding [14] and was the only one that showed changes in ¹⁹F-NMR chemical shift upon addition of SM to PC micelles in a recent study, where EqtII tryptophans were labelled with ¹⁹F [15]. In lipid monolayer experiments Dr1 is close to, but below the threshold for deeper insertion into natural membranes, taking into account that the lateral pressure of the lipids in a membrane bilayer is approximately 30 mN/m [46]. In monolayers of DOPC, EqtII shows critical pressure values around 25 mN/m [47], similar to the ones that we observed with Dr1, and reversible association with DOPC membranes [47], i.e. by using a sub-optimal lipid composition EqtII behaves in a similar way to Dr1.

To summarize, Dr1 can insert into lipid monolayers at low initial surface pressures and can reversibly associate with immobilized liposomes, which suggest that Dr1 can associate weakly with lipid membranes and can, therefore, be assigned as a peripheral membrane protein. Such behaviour is not uncommon in peripheral proteins, a large fraction of yeast PX and PH domains exhibit low affinity and were proposed to require additional domains within the same protein or other protein partners for efficient membrane binding [48].

AF Domains, a Novel Superfamily of Membrane Binding Domains - On the basis of sequence, structural and functional similarities described above we propose that three ALP families and the fruit body fungal lectins are the tentative basis for a novel superfamily, the AF membrane binding domains (abbreviated from Actinoporin-like proteins, Fruit body fungal lectins). So far, members of this family were shown to bind various molecules of cell membranes, specifically the lipids SM [18,19] and GM₁ ganglioside [43] or carbohydrates [25,28]. The common link between ALP and fungal fruit body lectins is a highly conserved structure at almost the full-length extent with weak sequence similarity, suggesting a homeomorphic protein superfamily. On the other hand the absence of structural information for hydrozoan ALP and ALP found predominately in vertebrates precludes any firm conclusions. If high-similarity local regions in ALP are considered as separate domains, we can describe the AF domain as a homology-domain superfamily. The observed widespread

distribution may be explained by this alternative. However, the number and order of highly similar local regions within sequences of ALP is strictly conserved, thus favouring the first possibility. It is quite possible that sequences presented in this paper are only a subset of the bigger family and that many members exist that do not have sufficient sequence similarity to be recognised. Since the AF domain superfamily is ancient and highly divergent, it will be difficult to recognise novel members and trace the evolutionary history of ALP families unless more structural information is available. Currently, the AF superfamily is characterized by a structure rich in β -sheets organized as a stable β -sandwich core flanked by an α -helix and binding determinants for different ligands located on one of the broad loops of the sandwich and α -helix (Figure 2).

ABBREVIATIONS

Abbreviations used: ABL, *Agaricus bisporus* lectin; ALP, actinoporin-like protein; ANS, 1-anilinonaphthalene-8-sulfonic acid; C1, PKC conserved domain 1; C2, PKC conserved domain 2; CD, circular dichroism; DOPC, 1,2-Dioleoyl-*sn*-Glycero-3-Phosphocholine; DOPG, 1,2-Dioleoyl-*sn*-Glycero-3-[Phospho-*rac*-(1-glycerol)]; Dr1, *Danio rerio* actinoporin-like sequence; DTT, dithiothreitol; EqtII, equinatoxin II; EST, expressed sequence tag; GdmHCl, Guanidinium hydrochloride; PFT, pore-forming toxins; PH, pleckstrin homology; PKC, protein kinase C; PCR, polymerase chain reaction; POC, phosphocholine; PX, Phox; SM, sphingomyelin; SPR, surface plasmon resonance; XCL, *Xerocomus chrysenteron* lectin.

ACKNOWLEDGEMENTS

This work was supported by the grants from Slovenian Ministry for Higher Education, Science and Technology. G.A. and J.H.L. are recipients of a Wellcome Trust International Research Development Award. I.G.A. was recipient of a postdoctoral fellowship form the Basque Government. The authors would like to thank to Dr. Eva Žerovnik for the help with CD spectrometer, Simona Strgulc Krajšek for discussions and Mauro Dalla Serra and Gabriella Viero for critical reading of the manuscript.

REFERENCES

- 1 Goni, F. M. (2002) Non-permanent proteins in membranes: when proteins come as visitors (review). *Mol. Membr. Biol.* **19**, 237-245
- 2 Cho, W. and Stahelin, R. V. (2005) Membrane-protein interactions in cell signaling and membrane trafficking. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 119-151
- 3 Stachowiak, O., Schlattner, U., Dolder, M. and Wallimann, T. (1998) Oligomeric state and membrane binding behaviour of creatine kinase isoenzymes: implications for cellular function and mitochondrial structure. *Mol. Cell Biochem.* **184**, 141-151
- 4 Holthuis, J. C. and Levine, T. P. (2005) Lipid traffic: floppy drives and a superhighway. *Nat. Rev. Mol. Cell Biol.* **6**, 209-220
- 5 Gerke, V., Creutz, C. E. and Moss, S. E. (2005) Annexins: linking Ca²⁺ signalling to membrane dynamics. *Nat. Rev. Mol. Cell Biol.* **6**, 449-461
- 6 Leippe, M., Bruhn, H., Hecht, O. and Grotzinger, J. (2005) Ancient weapons: the three-dimensional structure of amoebapore A. *Trends Parasitol.* **21**, 5-7
- 7 Gouaux, E. (1997) Channel-forming toxins: Tales of transformation. *Curr. Opin. Struct. Biol.* **7**, 566-573
- 8 Menestrina, G., Dalla Serra, M. and Lazarovici, P., eds. (2003) Pore-forming Peptides and Protein Toxins, Taylor&Francis, London
- 9 Anderluh, G. and Maček, P. (2002) Cytolytic peptide and protein toxins from sea anemones (Anthozoa : Actiniaria). *Toxicon* **40**, 111-124
- 10 Athanasiadis, A., Anderluh, G., Maček, P. and Turk, D. (2001) Crystal structure of the soluble form of equinatoxin II, a pore-forming toxin from the sea anemone *Actinia equina*. *Structure* **9**, 341-346
- 11 Hinds, M. G., Zhang, W., Anderluh, G., Hansen, P. E. and Norton, R. S. (2002) Solution structure of the eukaryotic pore-forming cytolytic equinatoxin II: Implications for pore formation. *J. Mol. Biol.* **315**, 1219-1229
- 12 Mancheño, J. M., Martín-Benito, J., Martínez-Ripoll, M., Gavilanes, J. G., and Hermoso, J. A. (2003) Crystal and electron microscopy structures of sticholysin II actinoporin reveal insights into the mechanism of membrane pore formation. *Structure* **11**, 1319-1328
- 13 Anderluh, G., Barlič, A., Podlesek, Z., Maček, P., Pungerčar, J., Gubenšek, F., Zecchini, M. L., Dalla Serra, M. and Menestrina, G. (1999) Cysteine-scanning

- mutagenesis of an eukaryotic pore-forming toxin from sea anemone - Topology in lipid membranes. *Eur. J. Biochem.* **263**, 128-136
- 14 Hong, Q., Gutiérrez-Aguirre, I., Barlič, A., Malovrh, P., Kristan, K., Podlesek, Z., Maček, P., Turk, D., Gonzalez-Mañas, J. M., Lakey, J. H. and Anderluh, G. (2002) Two-step membrane binding by equinatoxin II, a pore-forming toxin from the sea anemone, involves an exposed aromatic cluster and a flexible helix. *J. Biol. Chem.* **277**, 41916-41924
- 15 Anderluh, G., Razpotnik, A., Podlesek, Z., Maček, P., Separovic, F. and Norton, R. S. (2005) Interaction of the Eukaryotic Pore-forming Cytolysin Equinatoxin II with Model Membranes: (19)F NMR Studies. *J. Mol. Biol.* **347**, 27-39
- 16 Malovrh, P., Viero, G., Dalla Serra, M., Podlesek, Z., Lakey, J. H., Maček, P., Menestrina, G. and Anderluh, G. (2003) A novel mechanism of pore formation - Membrane penetration by the N-terminal amphipathic region of equinatoxin. *J. Biol. Chem.* **278**, 22678-22685
- 17 Kristan, K., Podlesek, Z., Hojnik, V., Gutiérrez-Aguirre, I., Gunčar, G., Turk, D., Gonzalez-Mañas, J. M., Lakey, J. H., Maček, P., and Anderluh, G. (2004) Pore formation by equinatoxin, a eukaryotic pore-forming toxin, requires a flexible N-terminal region and a stable beta-sandwich. *J. Biol. Chem.* **279**, 46509-46517
- 18 Belmonte, G., Pederzoli, C., Maček, P. and Menestrina, G. (1993) Pore formation by the sea anemone cytolysin equinatoxin II in red blood cells and model lipid membranes. *J. Membr. Biol.* **131**, 11-22
- 19 Tejuca, M., Dalla Serra, M., Ferreras, M., Lanio, M. E. and Menestrina, G. (1996) Mechanism of Membrane Permeabilisation by Sticholysin I, a Cytolysin Isolated from the Venom of the Sea Anemone *Stichodactyla helianthus*. *Biochemistry* **35**, 14947-14957
- 20 Valcarcel, C. A., Dalla Serra, M., Potrich, C., Bernhart, I., Tejuca, M., Martinez, D., Pazos, F., Lanio, M. E. and Menestrina, G. (2001) Effects of lipid composition on membrane permeabilization by sticholysin I and II, two cytolysins of the sea anemone *Stichodactyla helianthus*. *Biophys. J.* **80**, 2761-2774
- 21 Anderluh, G., Dalla Serra, M., Viero, G., Guella, G., Maček, P. and Menestrina, G. (2003) Pore formation by equinatoxin II, a eukaryotic protein toxin, occurs by induction of nonlamellar lipid structures. *J. Biol. Chem.* **278**, 45216-45223

- 22 Anderluh, G. and Lakey, J. H. (2005) Lipid Interactions of α -Helical Protein Toxins. In Protein-Lipid Interactions. From Membrane Domains to Cellular Networks (Tamm, L. K., ed.), pp. 141-162, Wiley-VCH, Weinheim, Germany
- 23 Kawashima, Y., Nagai, H., Ishida, M., Nagashima, Y. and Shiomi, K. (2003) Primary structure of echotoxin 2, an actinoporin-like hemolytic toxin from the salivary gland of the marine gastropod *Monoplex echo*. *Toxicon* **42**, 491-497
- 24 Birck, C., Damian, L., Marty-Detraves, C., Lougarre, A., Schulze-Briese, C., Koehl, P., Fournier, D., Paquereau, L. and Samama, J. P. (2004) A new lectin family with structure similarity to actinoporins revealed by the crystal structure of *Xerocomus chrysenteron* lectin XCL. *J. Mol. Biol.* **344**, 1409-1420
- 25 Carrizo, M. E., Capaldi, S., Perduca, M., Irazoqui, F. J., Nores, G. A. and Monaco, H. L. (2005) The antineoplastic lectin of the common edible mushroom (*Agaricus bisporus*) has two binding sites, each specific for a different configuration at a single epimeric hydroxyl. *J. Biol. Chem.* **280**, 10614-10623
- 26 Trigueros, V., Lougarre, A., Ali-Ahmed, D., Rahbe, Y., Guillot, J., Chavant, L., Fournier, D. and Paquereau, L. (2003) *Xerocomus chrysenteron* lectin: identification of a new pesticidal protein. *Biochim. Biophys. Acta* **1621**, 292-298
- 27 Yu, L., Fernig, D. G., Smith, J. A., Milton, J. D. and Rhodes, J. M. (1993) Reversible inhibition of proliferation of epithelial cell lines by *Agaricus bisporus* (edible mushroom) lectin. *Cancer Res.* **53**, 4627-4632
- 28 Damian, L., Fournier, D., Winterhalter, M. and Paquereau, L. (2005) Determination of thermodynamic parameters of *Xerocomus chrysenteron* lectin interactions with N-acetylgalactosamine and Thomsen-Friedenreich antigen by isothermal titration calorimetry. *BMC. Biochem.* **6**, 11
- 29 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389-3402
- 30 Huang, X. (1992) A contig assembly program based on sensitive detection of fragment overlaps. *Genomics* **14**, 18-25
- 31 Bendtsen, J. D., Nielsen, H., Von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783-795
- 32 McGuffin, L. J., Bryson, K. and Jones, D. T. (2000) The PSIPRED protein structure prediction server. *Bioinformatics.* **16**, 404-405

- 33 Whitmore, L. and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668-W673
- 34 Johnson, W. C. (1999) Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* **35**, 307-312
- 35 Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792-1797
- 36 Simossis, V. A., Kleinjung, J. and Heringa, J. (2005) Homology-extended sequence alignment. *Nucleic Acids Res.* **33**, 816-824
- 37 Whelan, S. and Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691-699
- 38 Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696-704
- 39 Eftink, M. R. and Ghiron, C. A. (1981) Review. Fluorescence quenching studies with proteins. *Anal. Biochem.* **114**, 199-227
- 40 Anderluh, G., Beseničar, M., Kladnik, A., Lakey, J. H., and Maček, P. (2005) Properties of nonfused liposomes immobilized on an L1 Biacore chip and their permeabilization by a eukaryotic pore-forming toxin. *Anal. Biochem.* **344**, 43-52
- 41 Poklar, N., Lah, J., Salobir, M., Maček, P. and Vesnaver, G. (1997) pH and Temperature-Induced Molten Globule-Like Denatured States of Equinatoxin II: A Study by UV-Melting, DSC, Far- and Near-UV CD Spectroscopy, and ANS Fluorescence. *Biochemistry* **36**, 14345-14352
- 42 Christoffels, A., Koh, E. G., Chia, J. M., Brenner, S., Aparicio, S. and Venkatesh, B. (2004) Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. *Mol. Biol. Evol.* **21**, 1146-1151
- 43 Shiomi, K., Kawashima, Y., Mizukami, M. and Nagashima, Y. (2002) Properties of proteinaceous toxins in the salivary gland of the marine gastropod (*Monoplex echo*). *Toxicon* **40**, 563-571
- 44 Sher, D., Knebel, A., Bsoor, T., Nesher, N., Tal, T., Morgenstern, D., Cohen, E., Fishman, Y. and Zlotkin, E. (2005) Toxic polypeptides of the hydra--a bioinformatic approach to cnidarian allomones. *Toxicon* **45**, 865-879
- 45 Henrich, T., Ramialison, M., Wittbrodt, B., Assouline, B., Bourrat, F., Berger, A., Himmelbauer, H., Sasaki, T., Shimizu, N., Westerfield, M., Kondoh, H. and Wittbrodt,

- J. (2005) MEPD: a resource for medaka gene expression patterns. *Bioinformatics*. **21**, 3195-3197
- 46 Brockman, H. (1999) Lipid monolayers: why use half a membrane to characterize protein-membrane interactions? *Curr. Opin. Struct. Biol.* **9**, 438-443
- 47 Caaveiro, J. M. M., Echabe, I., Gutiérrez-Aguirre, I., Nieva, J. L., Arrondo, J. L. R., and Gonzalez-Mañas, J. M. (2001) Differential interaction of equinatoxin II with model membranes in response to lipid composition. *Biophys. J.* **80**, 1343-1353
- 48 Yu, J. W., Mendrola, J. M., Audhya, A., Singh, S., Keleti, D., DeWald, D. B., Murray, D., Emr, S. D. and Lemmon, M. A. (2004) Genome-wide analysis of membrane targeting by *S. cerevisiae* pleckstrin homology domains. *Mol. Cell* **13**, 677-688
- 49 Anderluh, G., Podlesek, Z. and Maček, P. (2000) A common motif in parts of Cnidarian toxins and nematocyst collagens and its putative role. *Biochim. Biophys. Acta* **1476**, 372-376

Tables

Table 1

The distribution of AF superfamily members in public databases.

Family / Organism**1. Actinoporin-like proteins found predominately in vertebrates****Cnidaria***Hydractinia echinata* (1; 0.01%)**Mollusca***Monoplex echo*^a (1)**Vertebrata****Chondrichthyes***Leucoraja erinacea* (1; 0.009%)**Teleostei***Cyprinus carpio* (1; 0.009%), *Danio rerio* (2; 0.002%), *Fugu rubripes*^b (5), *Gasterosteus aculeatus* (2; 0.002%), *Haplochromis chilotes* (5; 0.06%), *Misgurnus anguillicaudatus* (1; 0.009%), *Oncorhynchus mykiss* (5; 0.02%), *Oryzias latipes* (7; 0.03%), *Paralichthys olivaceus* (1; 0.026%), *Salmo Salar* (2; 0.008%), *Tetraodon nigroviridis*^b (5)**Amphibia***Xenopus tropicalis* and *X. laevis* (4; 0.0003%)**Aves***Gallus gallus* (2; 0.01%)**2. Hydrozoan actinoporin-like proteins***Hydra magnipapillata* (10; 0.025%), *Hydra vulgaris* (1; 0.66%)**3. Actinoporins****Cnidaria***Actinaria villosa*^a (1), *Actinia equina*^a (3), *Heteractis crispa*^a (1), *Heteractis magnifica*^a (1), *Phyllo-discus semoni*^a (1), *Oulactis orientalis*^a (2), *Sagartia rosea*^a (1), *Stichodactyla helianthus*^a (2)**Plants****Bryophyta***Physcomitrella patens* (1; 0.014%), *Tortula ruralis* (2; 0.62%)**Pteridophyta***Selaginella lepidophylla* (1; 0.191%)**4. Fungal fruit body lectins****Fungi***Agaricus bisporus*^a (1), *Arthrotrichum oligospora*^a (1), *Aspergillus fumigatus*^b (1), *Aspergillus terreus*^b (2), *Coccidioides immitis*^b (1), *Gibberella moniliformis*^b (1), *Gibberella zeae* (1; 0.012%), *Neosartoria fischeri*^b (1), *Neurospora crassa* (1; 0.004%), *Paxillus involutus* (3; 0.1%), *Phanerochaete chrysosporium*^b (2), *Pleurotus cornucopiae*^a (2), *Podospora anserina*^b (1), *Xerocomus chrysenteron*^a (1)**Plants***Marchantia polymorpha* (3; 0.85), *Tortula ruralis* (1; 0.02)

Number of different sequences in a single organism and percentage of EST clones, when available, is reported in parentheses.

^a protein sequence was obtained by protein sequencing or was deduced from mRNA.^b sequences were retrieved from the genome data.

Table 2

Kinetic parameters for the binding of Dr1 to supported liposomes as determined by SPR. The buffer was 140 mM NaCl, 20 mM Tris-HCl, 1 mM EDTA, pH 8.5. n=2 or 3; average \pm S.D.

Liposomes	k_a ($M^{-1}s^{-1}$) $\times 10^3$	k_d (s^{-1}) $\times 10^{-3}$	K_D (M) $\times 10^{-6}$
DOPC	1.7 ± 0.2	3.2 ± 1.5	1.9 ± 1.0
DOPC:SM (1:1, mol:mol)	1.3 ± 0.2	2.7 ± 0.3	2.1 ± 0.6
DOPC:DOPG (1:1, mol:mol)	1.2 ± 0.3	3.8 ± 0.5	3.7 ± 1.2

FIGURE LEGENDS

Figure 1. Alignment of actinoporin-like proteins

Selected examples from three families of ALP are presented. An empty line separates each family from top to bottom: hydrozoan ALP, actinoporins, and ALP found predominately in vertebrates. Selected examples of fungal fruit body lectin family are shown for the comparison below the line of asterisks, which highlights highly conserved residues in ALP. Alignment is shaded according to conservation. Black, amino acids conserved in more than 90% of sequences; dark gray with light letters, conserved in more than 60% of sequences; light gray with black letters, conserved in more than 40% of sequences. Alignment was prepared and colored using GeneDoc (<http://www.psc.edu/biomed/genedoc/>). The numbers in parentheses on the left denote the number of residues before the conserved domain. The size of insertion between helix and fifth beta strand is shown in parentheses within the alignment. The number of residues after the domain is shown in parentheses on the right of the alignment. The length of each protein in amino acids is written on the far right. The numbering above the sequence is according to the sequence of mature EqtII. Abbreviations of protein and species names are as follows: AGAbis, *Agaricus bisporus*; Avt1, actinoporin from *Actinaria villosa*; Dr1, *Danio rerio*; Eqt2, equinatoxin II from *Actinia equina*; Fr, *Fugu rubripes*; Ga, *Gasterosteus aculeatus*; Gg, *Gallus gallus*; Hc, *Haplochromis chilotes*; He, *Hydractinia echinata*; HlyME; hemolysin from *Monoplex echo*; Hm, *Hydra magnipapillata*; Hv, *Hydra vulgaris*; Le, *Leucoraja erinacea*; Ma, *Misgurnus anguillicaudatus*; MARpol, *Marchantia polymorpha*; NEUcra, *Neurospora crassa*; Ol, *Oryzias latipes*; Om, *Oncorhynchus mykiss*; OrA, actinoporin from *Oulactis orientalis*; PHAchr, *Phanerochaete chrysosporium*; PLEcor, *Pleurotus cornucopiae*; Pp, *Physcomitrella patens*; Ss, *Salmo salar*; StII, sticholysin II from *Stichodactyla helianthus*; TORrur, *Tortula ruralis*; Tn, *Tetraodon nigroviridis*; Tr, *Tortula ruralis*; XERchr, *Xerocomus chrysenteron*; Xl, *Xenopus laevis*; Xt, *Xenopus tropicalis*.

Figure 2. Structural properties of actinoporin-like proteins

(A) Structural features of the conserved domain using the known structure of EqtII. The structure of the whole molecule of EqtII (PDB code 1IAZ) is presented on the left. The conserved region is colored blue. The side chains of residues that participate in POC binding and are highlighted in the alignment in C are shown as red sticks. Residues Y and D from the conserved motif P-[WYF]-D are shown in green as sticks. POC is colored orange. The next

image to the right shows T-antigen binding site in ABL (PDB 1Y2X and 1Y2W). Amino acids that participate in T-antigen binding, and are highlighted in the alignment in C, are presented as sticks. T-antigen is colored green. Two right images show the CPK presentation of side chains of some of the residues that are more than 90% conserved in all ALP. (B) Structural organization of ALP. Red, predicted signal peptide; green, cnidocyst sorting signal [49]; yellow, amphipatic helix needed for permeabilising activity [14,16,17]; blue, the conserved ALP region. (C) Structural alignment of actinoporins and fungal fruit body lectins. Elements of secondary structure as defined by DSSP are presented above the sequences. Residues that participate in POC binding (POC) in StII and residues that participate in Thomsen-Friedenreich (T-)antigen binding (Tant) in ABL are shown below the sequences. PDB codes are written beside the names of proteins. The numbering is according to mature proteins. Color scheme; green with white characters: hydrophobic amino acids (V, L, I); green with black characters: aromatic amino acids (W, Y, F); green with red characters: small hydrophobic amino acids (A, G); yellow: polar amino acids (N, Q, M, T, S, C); blue: positively charged amino acids (K, R, H); red: negatively charged amino acids (D, E); white: Pro.

Figure 3. Phylogenetic relationships of actinoporins, actinoporin-like proteins and fungal fruit body lectins inferred by maximum likelihood under a Whelan-Goldman gamma model of amino acid evolution. Numbers at nodes are bootstrap values in per cent (200 replicates); n.s. indicates non-supported nodes with a bootstrap value below 50 per cent. Abbreviations of the protein and species names are as follows: ARToli, *Arthrotrypis oligospora*; ASPfum, *Aspergillus fumigatus*; ASPter, *Aspergillus terreus*; Cc, *Cyprinus carpio*; COCimm, *Coccidioides immitis*; GIBmon, *Gibberella moniliformis*; GIBzea, *Gibberella zeae*; HMgIII, an actinoporin from *Heteractis magnifica*; NEOfis, *Neosartoria fischeri*; PAXinv, *Paxillus involutus*; Po, *Paralichthys olivaceus*; PODans, *Podospira anserine*; PsTX20A, an actinoporin from *Phyllodiscus semoni*; RTXA, an actinoporin from *Heteractis crispa*; Sagros, an actinoporin from *Sagartia rosea*; Sl, *Selaginella lepidophylla*; SORMac, *Sordaria macrospora*. Other names are as in Figure 1.

Figure 4. Isolation and secondary structure of Dr1

Top: One μg of protein was resolved on 12% SDS-PAGE gel and stained by Coomassie Blue. M, molecular weight marker; 1, EqII; 2, His-tagged Dr1; 3, thrombin cleaved Dr1.

Middle: Prediction of secondary structure elements of Dr1 by PSIPRED [32]. Bold italics, helix; bold, strand; normal, other (turns, irregular). Lowercase letters, confidence of prediction 0-4 as reported by program; uppercase letters, confidence 5-9. ALP domain is shaded. Elements of EqtII secondary structure determined from 3D structure are denoted by underlining with a full line for β -strands and dashed line for α helix. Bottom: Far-UV CD spectra of purified Dr1. The concentration of protein in 20 mM NaH₂PO₄, pH 8.0 was 8 μ M (solid line). Protein was denatured by 90 min incubation in 6M GdmHCl (dashed line). The spectra of EqtII from [17] is shown for a comparison as a dashed line. The inset shows Near-UV CD spectra of 30 μ M Dr1 in 20 mM NaH₂PO₄, pH 8.0. All spectra were taken at 20°C.

Figure 5. Tryptophan fluorescence and unfolding of Dr1

Top: Tryptophan fluorescence of Dr1 in 10 mM Hepes, 200 mM NaCl, 20 mM DTT, pH 7.5 (solid line) and in 6M GdmHCl (dashed line). Protein concentration was 1 μ M. Excitation wavelength was 295 nm. Spectra were measured at 25 °C and under constant stirring. Middle: Stern-Volmer plots for the quenching with iodide in 10 mM Hepes, 200 mM NaCl, 20 mM DTT, pH 7.5 (solid circles) and in 6M GdmHCl (open circles). The modified Stern-Volmer plot for multiple emitting centres is shown in the inset for the protein in the buffer without denaturant. Other conditions are as in top panel. Bottom: Denaturation of Dr1 followed by ANS fluorescence (dashed line) or CD spectroscopy (dots and solid line). Protein and ANS concentrations were 15 μ M. Buffer was 10 mM Hepes, 200 mM NaCl, 20 mM DTT, pH 7.5. Excitation and emission wavelengths were fixed at 370 and 488 nm, respectively and slits for both were 5 nm. Experiment was performed in a temperature-controlled holder with constant stirring. For unfolding followed by CD spectroscopy Dr1 concentration was 6 μ M in 20 mM NaH₂PO₄, pH 8.0. The CD signal at 208 nm was measured at 0.5°C increments. For each data point the signal was averaged 5 s. The bandwidth was 1 nm. Temperature was increased in both cases at a ratio of 1°C/min. The data from CD spectroscopy was fitted to a sigmoidal curve by Origin software, yielding a midpoint of transition at 47.6 \pm 0.6°C.

Figure 6. Interaction of Dr1 with lipid monolayers

Top: Accumulation at the air-water interface. Different concentrations of Dr1 were injected into a subphase composed of 10 mM Hepes, 200 mM NaCl, pH 7.5 and changes in the surface pressure were registered with a surface pressure microbalance. Dr1 was always

preincubated with 10 mM DTT. Experiment was carried out under constant stirring and at room temperature. Middle: Insertions into lipid monolayers at an initial pressure of approximately 15 mN/m. Concentration of Dr1 in the subphase was 2 μ M. Other experimental conditions are as in A. Bottom: Critical pressure plots for different lipid monolayers. Squares, DOPC monolayers; circles, DOPC/SM monolayers; triangles, DOPC/DOPG monolayers.

Figure 7. Interaction of Dr1 with supported liposomes

Liposomes were immobilized on the surface of L1 chip as described [40]. The flow-rate was 30 μ l/min. The running buffer was 140 mM NaCl, 20 mM Tris-HCl, 1 mM EDTA, pH 8.5. The injection lasted 90 s and the dissociation was followed for 5 min. Each panel shows sensorgrams for 2.5, 5, 10 and 20 μ M Dr1. Sensorgrams were fitted to a 1:1 Langmuir binding model, fits are shown as thick lines superimposed on sensorgrams.

Figure 1

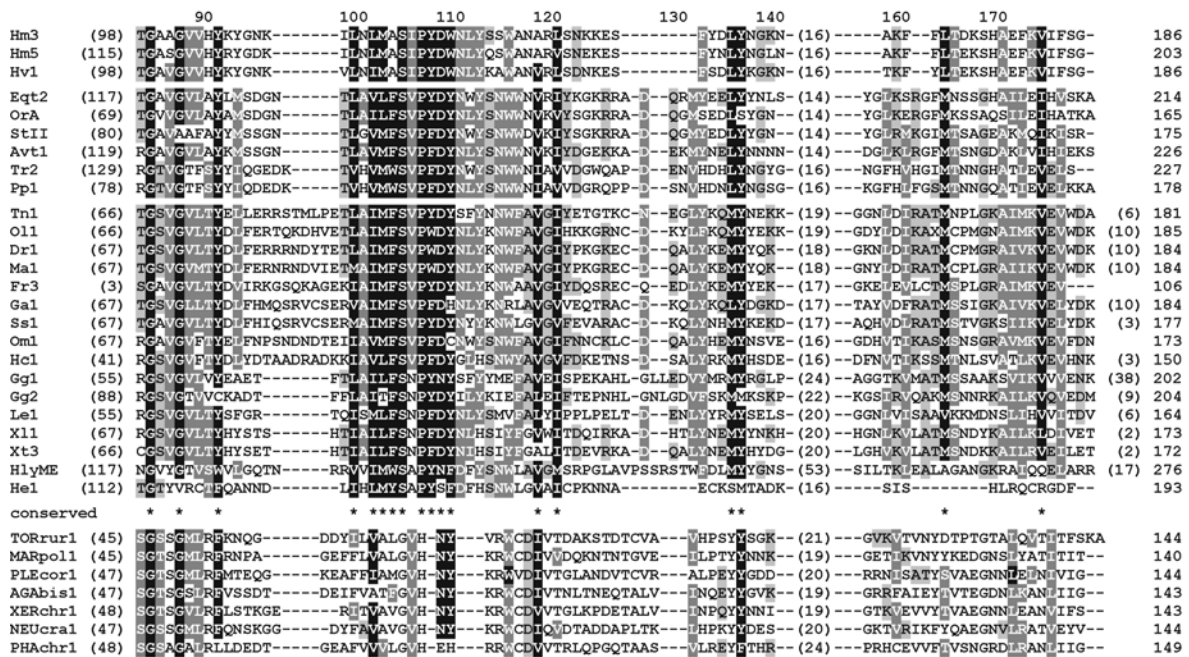


Figure 2

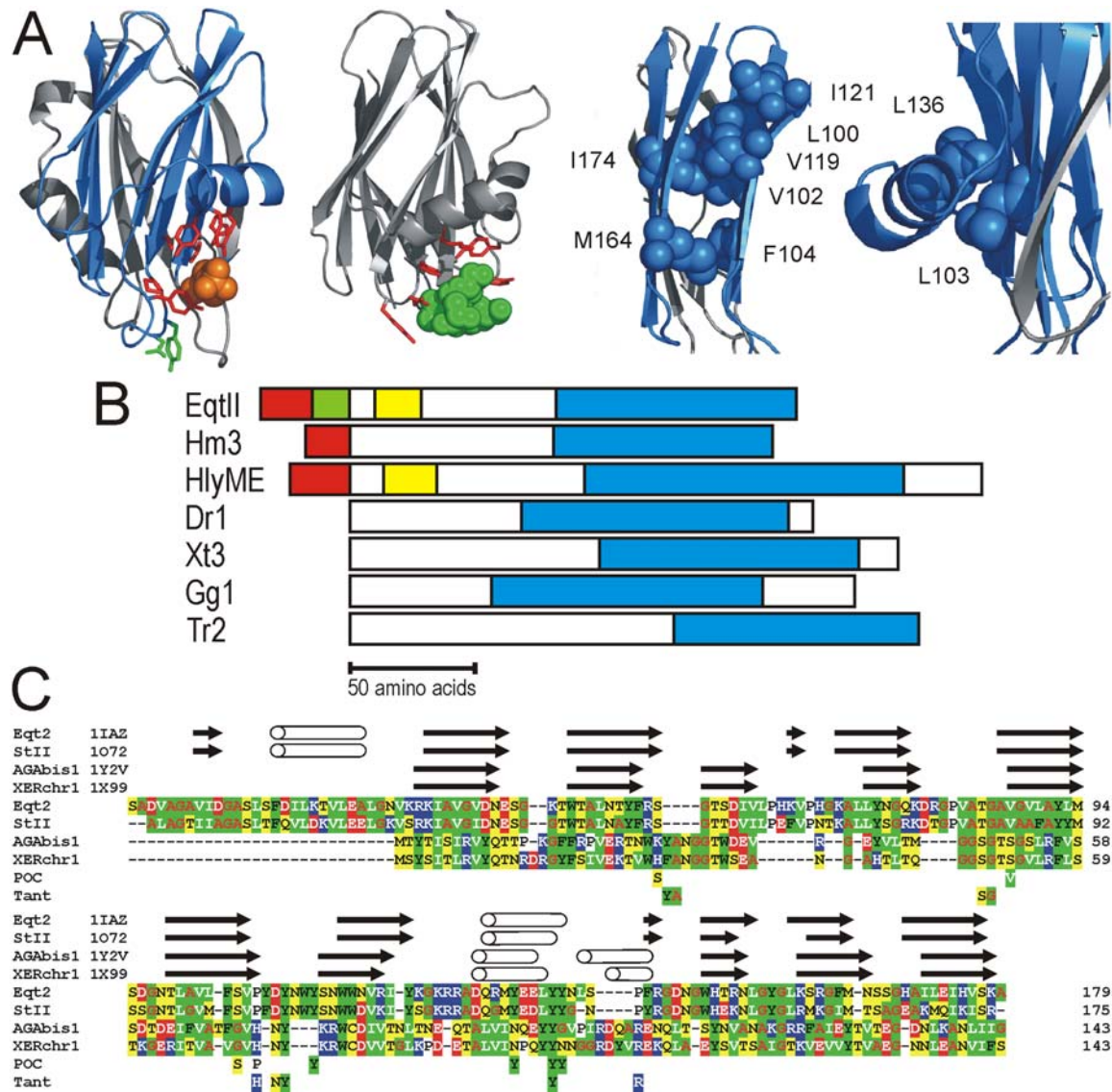


Figure 3

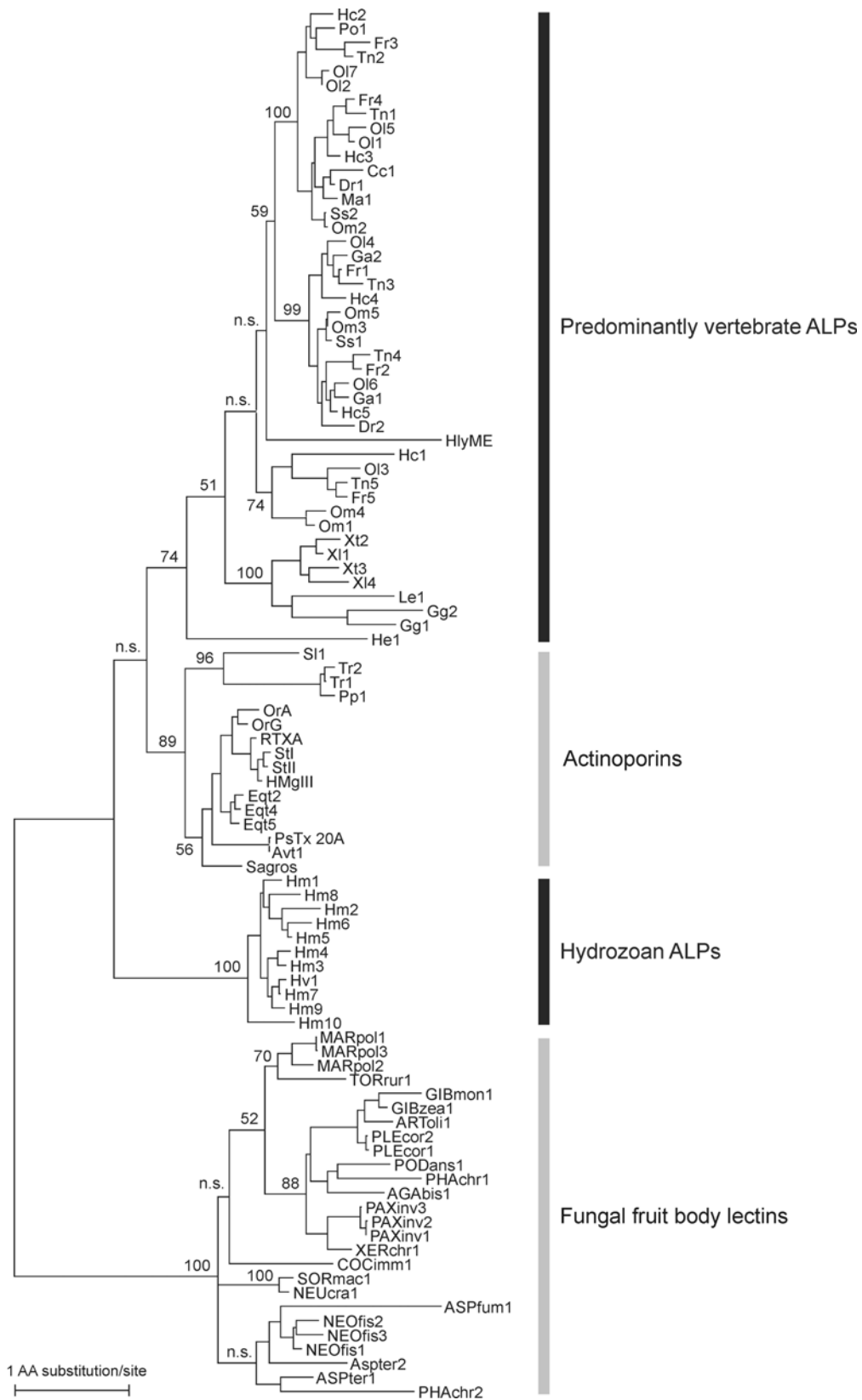
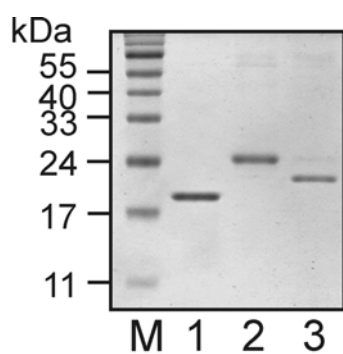


Figure 4



GsteSaeAvaaNVssRRHATVEITnLTNNY
 CFlnpkvyleNGEtSNPPQptvRPLktevC
 TfskSAAHatgsvqVLTYdlferrrNDYTE
 TLAIMFSVPwDynlyknWFAVGIIYpKGKEc
 DqALYKEMYyQKNQHGfvreEANGsgInfe
 GknLDIRATMCPmgraIVKVEVWdKLLSPm
 aqmDC

Helix Strand Other

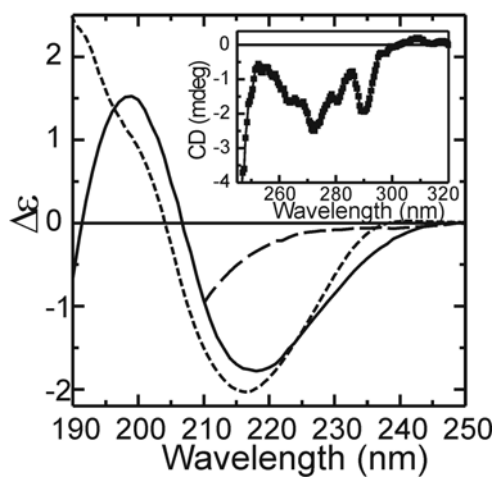


Figure 5

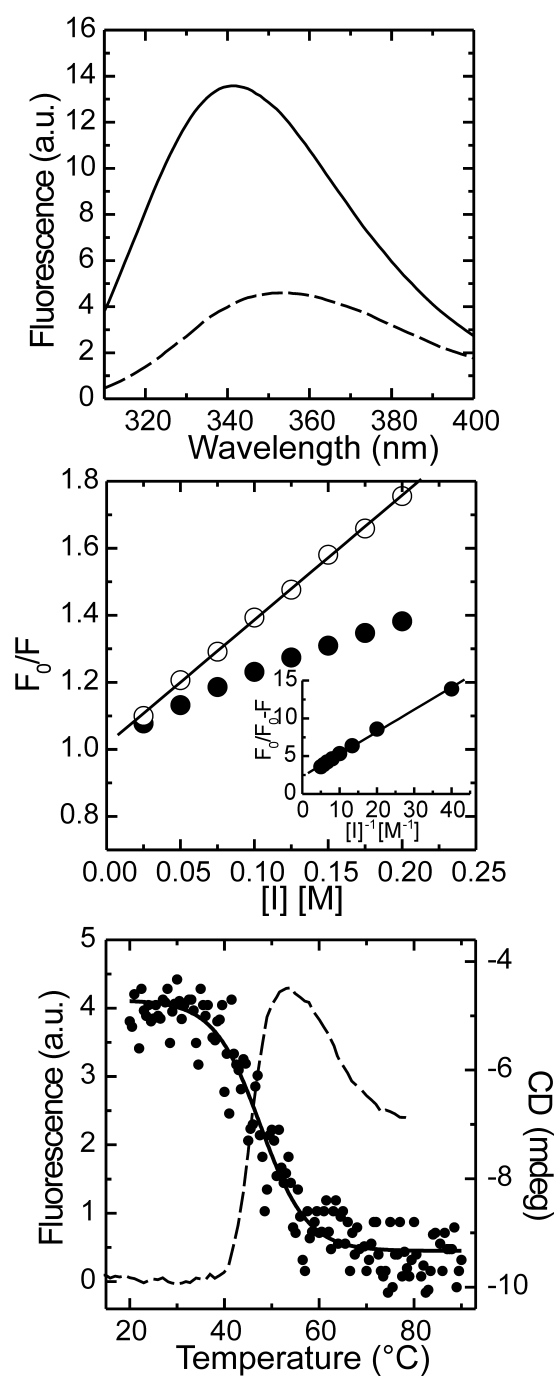


Figure 6

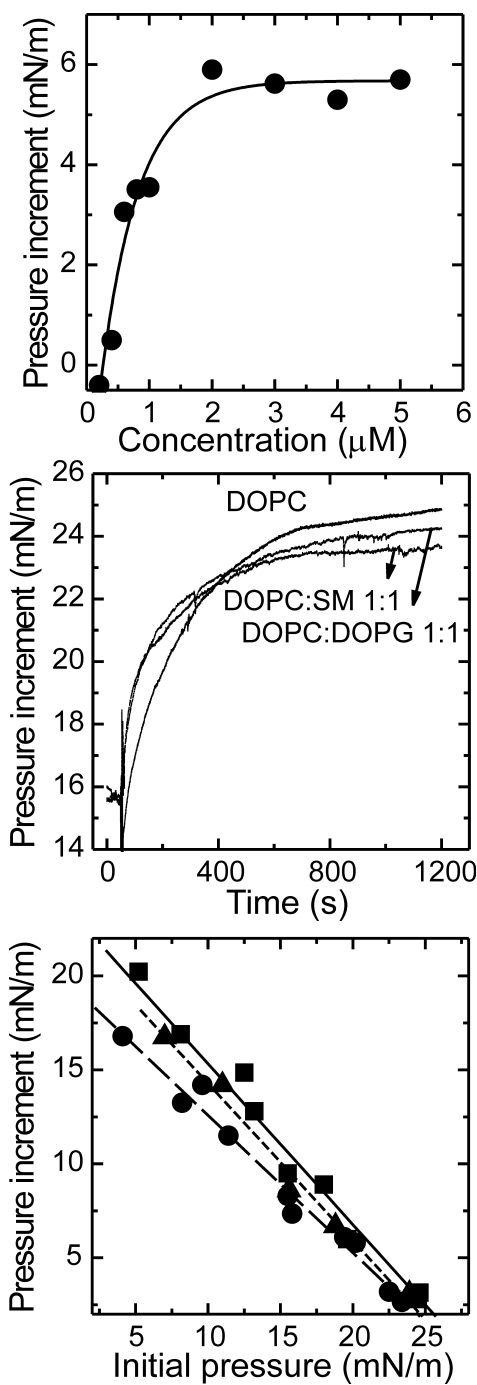


Figure 7

