

# Evolution of an antifreeze glycoprotein

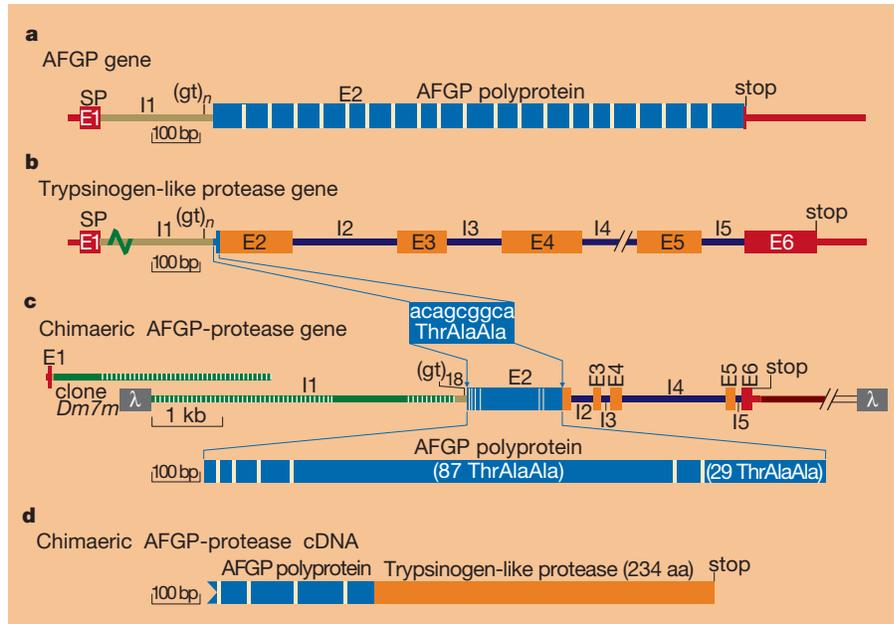
A blood protein that keeps Antarctic fish from freezing arose from a digestive enzyme.

The ice-binding antifreeze glycoprotein (AFGP) that circulates in the blood of Antarctic notothenioid fishes enables them to avoid freezing in their perpetually icy environment<sup>1</sup>. This crucial survival protein probably arose from a functionally unrelated pancreatic trypsinogen-like protease<sup>2</sup>. We have now discovered an important intermediate in this evolutionary process — transcriptionally active chimaeric genes that encode both an AFGP polyprotein and the protease, confirming the protease origin of AFGP and indicating how it was created.

AFGP binds to and arrests the growth of ice crystals that enter the fish, thereby preventing the fish from freezing. There are at least eight forms of the protein of different sizes (AFGP 1–8), all composed of repeats of a simple glycotriptide monomer (Thr-Ala/Pro-Ala-) with a disaccharide attached to each threonine residue<sup>1</sup>. These isoforms are encoded as distinct molecules linked in a series by conserved three-residue spacers (mostly Leu/Phe-Ile/Asn-Phe) within a large polyprotein; post-translational removal of the spacers produces many copies of mature AFGP per gene<sup>2,3</sup>. A typical AFGP gene comprises two exons and a single intron, with the small exon (E1) encoding the signal peptide and the large exon (E2) encoding the AFGP polyprotein (Fig. 1a).

All notothenioid AFGP polyprotein genes characterized so far encode only the small isoforms (AFGPs 6, 7 and 8)<sup>2,4</sup>. To determine whether the large isoforms (AFGPs 1–5) are also individually encoded, we screened a genomic library obtained from *Dissostichus mawsoni* (the giant Antarctic toothfish) and isolated a clone, *Dm7M*, that contains an unusual gene encoding not only the large AFGP isoforms but also a trypsinogen-like protease in tandem (Fig. 1c, middle).

The chimaeric gene in clone *Dm7M* spans about 9.6 kilobase pairs (kb), which we sequenced completely. It is a partial gene lacking its 5' end, which we subsequently obtained by amplification of genomic DNA using *Dm7M*-specific primers in the polymerase chain reaction (PCR) (Fig. 1c, top left). The complete reconstructed chimaeric gene (~11 kb) consists of six exons and five introns, and its arrangement is identical to that of independent trypsinogen-like protease genes<sup>2</sup> (Fig. 1b), with the notable exception of a hybrid exon 2 (Fig. 1c, middle). Hybrid E2 has a large 5' AFGP polyprotein-coding segment that encodes seven AFGP molecules (two large and five



**Figure 1** Key molecular components in the evolution of the AFGP gene from the protease gene in Antarctic notothenioids. **a**, An independent AFGP polyprotein gene; blue, AFGP coding sequences; pink, spacer sequences. SP, signal peptide. **b**, An independent trypsinogen-like protease gene. **c**, Chimaeric AFGP–protease gene. PCR amplification of genomic DNA produced the 5' sequence (top left) of the partial chimaeric gene in clone *Dm7M* (middle), which is expanded at the hybrid exon E2 to show the AFGP polyprotein (bottom). **d**, Chimaeric AFGP–protease cDNA from a chimaeric gene distinct from *Dm7M*. Homologous sequence segments have identical colours. The notothenioid AFGP gene arose from recruitment of the front (exon E1 and intron I1) and the tail (E6) of an ancestral protease gene, *de novo* creation of the AFGP coding region by repeated duplications of the Thr-Ala-Ala-coding element that straddles the I1–E2 junction of the protease gene (enlarged between **b** and **c**), and deletion of the bulk of the protease sequence (E2–I5). This transcriptionally active chimaeric AFGP–protease gene confirms the protease origin of the AFGP gene and indicates how it evolved.

small isoforms) (Fig. 1c, bottom), and a small 3' segment that is similar in sequence to the trypsinogen-like protease exon 2 reported previously<sup>2</sup>.

The location of the AFGP-polyprotein-coding segment in the chimaeric gene corresponds to that of the single Thr-Ala-Ala-coding element in the protease gene (this element straddles the junction of E2 with the first intron (I1))<sup>2</sup> (Fig. 1b), which strongly indicates that the repetitive AFGP-polyprotein-coding sequence arose from expansion of the element (together with a spacer whose origin is unknown) through iterative duplications. The chimaeric-gene intron-1 sequence, apart from two insertions, is very similar to that of the distinct AFGP genes, including the (gt)<sub>n</sub>-minisatellite DNA-bearing I1 sequence that was apparently inherited from the ancestral trypsinogen-like protease (this (gt)<sub>n</sub> sequence persists in the present-day trypsinogen-like gene; Fig. 1b,c). This (gt)<sub>n</sub> sequence presumably facilitated the first duplication of the ancestral Thr-Ala-Ala-coding element through an accidental replication slippage<sup>5,6</sup>. Two of the tripeptide repeats may have partial ice-binding activity (the smallest functional isoform, AFGP8,

has only four repeats) and as the Antarctic water chilled towards freezing, they perhaps became selected upon for *de novo* expansion through further replication slippage<sup>5,6</sup> or unequal crossing over<sup>7</sup>.

The *Dm7M* chimaeric AFGP–protease gene is probably not a pseudogene, as it has all the expected elements of a functional gene, and similar chimaeric genes in the toothfish are found to be transcribed. We obtained partial AFGP–protease chimaeric complementary DNAs by using reverse transcription with PCR of toothfish pancreatic messenger RNA (Fig. 1d). The encoded partial AFGP polyprotein contains four AFGP molecules, followed without interruption by the entire trypsinogen-like protease. This AFGP polyprotein differs from that encoded in clone *Dm7M* (Fig. 1a), indicating that there is more than one chimaeric AFGP–protease gene in the toothfish genome. We found by Southern-blot analysis of genomic DNA that chimaeric AFGP–protease genes also exist in other notothenioid species (results not shown).

The simultaneous presence of a present-day protease gene carrying the incipient Thr-Ala-Ala-coding element, a transcriptionally

active chimaeric AFGP–protease gene, and independent AFGP genes in the toothfish genome confirms that the AFGP sequence started out as a small integral part of the ancestral protease gene, underwent expansion, and acquired independence by shedding the bulk of the protease sequence (E2 to I5)<sup>7</sup>. By capturing the chimaeric evolutionary intermediates, we have a rare view of the genesis of a new protein that ultimately enabled the Notothenioidae suborder to rise to dominance in the freezing Southern Ocean<sup>8,9</sup>.

Chi-Hing C. Cheng, Liangbiao Chen

Department of Ecology, Ethology and Evolution, University of Illinois,

Urbana, Illinois 61801, USA

e-mail: c-cheng@uiuc.edu

- DeVries, A. L. *Comp. Biochem. Physiol. A* **7**, 627–640 (1982).
- Chen, L., DeVries, A. L. & Cheng, C.-H. C. *Proc. Natl Acad. Sci. USA* **94**, 3811–3816 (1997).
- Hsiao, K. C. *et al. Proc. Natl Acad. Sci. USA* **87**, 9265–9269 (1990).
- Chen, L., DeVries, A. L. & Cheng, C.-H. C. *Proc. Natl Acad. Sci. USA* **94**, 3817–3822 (1997).
- Cheng, C.-H. C. in *Fishes of Antarctica* (eds di Prisco, G., Pisano, E. & Clarke, A.) 311–328 (Springer, Milan, 1998).
- Levinson, G. & Gutman, G. A. *Mol. Biol. Evol.* **4**, 203–221 (1987).
- Lewin, B. *Genes*, Vol. V (Cell, Cambridge, Massachusetts, 1994).
- Eastman, J. T. *Antarctic Fish Biology* (Academic, San Diego, California, 1993).
- Clarke, A. & Johnston, I. A. *Trends Ecol. Evol.* **11**, 187–228 (1996).

Animal behaviour

## Aquatic sex pheromone from a male tree frog

Many creatures use chemical signals (pheromones) as sources of information about the world around them<sup>1–3</sup>. For example, a sex pheromone produced by one sex (usually the female) of a particular species induces an immediate behavioural response in the opposite sex of the same species<sup>2,3</sup>. However, very little is known about amphibian pheromones<sup>4</sup>. We have now discovered and characterized an aquatic, female-attracting pheromone from the parotoid and rostral glands of a male frog, the magnificent tree frog, *Litoria splendida*. To our knowledge, this pheromone, which we have named splendipherin, is the first

pheromone from an anuran (frog or toad) to be identified.

The skin glands of anurans secrete defence compounds, including many types of biologically active peptides<sup>5,6</sup>. For example, *L. splendida* (Fig. 1) exudes several biologically active peptides from the parotoid and rostral glands (situated at the rear and front of the head, respectively), including the broad-spectrum antimicrobial caerin-1 peptides (Fig. 2)<sup>7,8</sup>. However, no anuran pheromones have been reported, apart from a possible alarm pheromone in tadpoles of the toad *Bufo bufo*<sup>9</sup>.

Because anurans often breed in aquatic conditions and deposit their eggs in water, it is possible that they produce peptides as water-soluble aquatic pheromones. This proposal is supported by the identification of the sex pheromone sodefrin (a decapep-



Figure 1 The magnificent tree frog, *Litoria splendida*.

tide, SIPSKDALK-OH) from the cloacal gland of the male newt *Cynops pyrrhogaster*<sup>10</sup>.

We monitored the peptide content of secretions from the parotoid and rostral glands of male and female *L. splendida* every month for three years by using high-performance liquid chromatography and electrospray mass spectrometry. No animals were killed during this procedure<sup>11</sup>. A minor component that is found only in male secretions can be seen in the partial HPLC profile shown in Fig. 2. This peptide, which we call splendipherin, comprises 25 amino-acid residues with the sequence GLVSSIGKALGGLLADVVKSKGQPA-OH.

The concentration of splendipherin peaks during the breeding season (January to March) to about 1% of the total peptide content of the glandular secretion, and decreases tenfold during June to November. Unlike the other peptides in Fig. 2, splendipherin has no antimicrobial activity. The natural peptide contains L-amino acids, as confirmed by the preparation of synthetic L-splendipherin, which has the same pheromone activity as the natural material.

The pheromone behavioural tests were carried out in a glass tank (0.65 × 2.0 × 0.75 metres) containing water 2 cm deep. Female *L. splendida* frogs were placed in the centre of the tank, where they remained for up to ten minutes without moving. Adding 40 nanograms of splendipherin in water (a final concentration of 10<sup>-13</sup> M) to a gauze pad 1 m from the frog elicited a distinct change in posture and increased alertness in the frog within 20 seconds. The female then walked slowly towards the pad, sat on it, and remained in a sitting position until she was removed. This experiment has been repeated several times and in different directions, using eight different females, with a 100% success rate.

The average time between introducing the pheromone and the female adopting a sitting position on the gauze pad is 7 minutes. The speed of the initial recognition indicates that splendipherin may be acting as a surfactant, but this has yet to be confirmed experimentally. When more than 4,000 nanograms of pheromone is added to the pad, the female becomes confused and

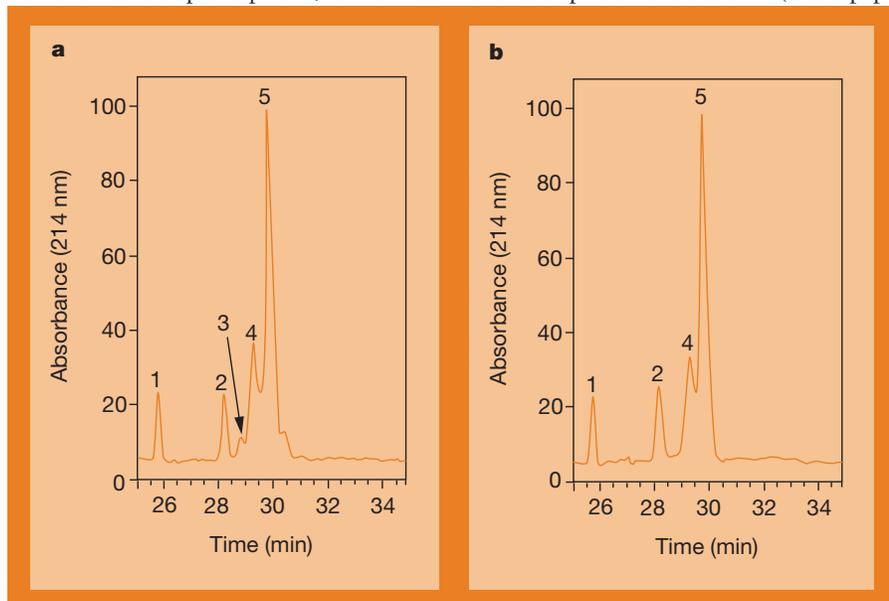


Figure 2 Partial high-performance liquid chromatography (HPLC) traces of the glandular secretions of *Litoria splendida*. a, Male; b, female. For experimental conditions, see refs 8,12. The components are as follows: 1, GLWQKIKDKASELVGIVEGVK-NH<sub>2</sub> (caerin 3.1; antibiotic)<sup>12</sup>; 2, GLVSSIGKALGGLLADVVKSKGQPA-OH (caerin 2.1; weak antibiotic)<sup>12</sup>; 3, GLVSSIGKALGGLLADVVKSKGQPA-OH (splendipherin); 4, GLFSLVLAGAVAKHVLPHWVPVIAEKL-NH<sub>2</sub> (caerin 1.6; antibiotic)<sup>8</sup>; 5, GLLSVLGSAKHLVPHWVPVIAEHL-NH<sub>2</sub> (caerin 1.1; antibiotic)<sup>7,8</sup>. For the full HPLC trace, see ref. 12. Caerin 2.1 differs from splendipherin in that an arginine residue replaces a lysine residue at position 8. Caerin 2.1 has no pheromone activity towards either male or female *L. splendida*.