



appearing in mRNA are in uppercase font whereas introns and promoter sequences are in lowercase font. The deduced protein sequence is shown with the signal peptide in lowercase font. The presumed TATA box and polyadenylation signals are shown in bold font. The transcription start site at base pair 33 was predicted using the BDGP neural network prediction program ([http://www.fruitfly.org/seq\\_tools/promoter](http://www.fruitfly.org/seq_tools/promoter)). The positions of partially degenerate primers used to amplify the bulk of the coding sequence from cDNA are underlined with arrows showing orientation, whereas the primer used to obtain the 3' portion of the cDNA by anchor PCR is double underlined. The majority of the sequence was obtained from these two overlapping cDNA sequences. Genomic DNA was the template used to amplify 230 bases at the 5' end of the gene by inverse PCR (5' primer shown highlighted in grey) as well as a fragment of the gene containing two introns by using the anchor PCR primer (double underlined) with a 3' primer (grey highlighting).

**Supporting Table 1:** Comparison of observed and predicted tryptic peptide masses for the large snow flea AFP.

| Observed Mass (Da) <sup>1</sup> | Predicted Mass (Da) <sup>2</sup> | Residue Range | Missed Cleavages |
|---------------------------------|----------------------------------|---------------|------------------|
| 1046±1                          | 1045.5                           | 1-10          | 1                |
| 4474±4                          | 4472.9                           | 3-57          | 1                |
| 3738±3                          | 3734.6                           | 11-57         | 0                |
| 3487±3                          | 3485.6                           | 58-103        | 2                |
| 3245±3                          | 3243.5                           | 58-100        | 1                |
| 3016±3                          | 3015.4                           | 64-103        | 1                |
| 2774±3                          | 2773.2                           | 64-100        | 0                |
| 4735±5                          | 4735.2                           | 101-162       | 1                |
| 4492±4                          | 4493.1                           | 104-162       | 0                |
| 1424±1                          | 1422.6                           | 163-179       | 0                |
| n.d.                            | 1733.9                           | 180-203       | 0                |

<sup>1</sup> Uncertainty was estimated as half of the width of the ion peak at 50% of maximum intensity, which averaged 0.1%.

<sup>2</sup> The predicted monoisotopic masses were calculated using PeptideMass (Wilkins, M.R., Lindskog, I., Gasteiger, E., Bairoch, A., Sanchez, J.-C., Hochstrasser, D.F., and Appel, R.D. *Detailed peptide characterisation using PEPTIDEMASS - a World-Wide Web accessible tool*; Electrophoresis 18(3-4), 403-408 (1997).