ResearchOnline@JCU

This file is part of the following work:

Brinkman, Diane L. (2008) ***The molecular and biochemical characterisation of venom proteins from the box jellyfish, Chironex fleckeri.*** **PhD Thesis, James Cook University.**

Access to this file is available from:

https://doi.org/10.25903/58rd%2Dsq03

Copyright © 2008 Diane L. Brinkman

The author has certified to JCU that they have made a reasonable effort to gain permission and acknowledge the owners of any third party copyright material included in this document. If you believe that this is not the case, please email researchonline@jcu.edu.au

The Molecular and Biochemical Characterisation of Venom Proteins from the Box Jellyfish, *Chironex fleckeri*

Thesis submitted by

Diane Louise BRINKMAN

BSc *UQ*, PGDipSc *JCU*

in December 2008

for the degree of Doctor of Philosophy

in the School of Pharmacy & Molecular Sciences
James Cook University
Australia

Statement of Access

I, the undersigned, author of this work, understand that James Cook University will make this thesis available for use within the University Library and, via the Australian Digital Theses network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act.

Diane Brinkman Date

Statement of Sources

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Diane Brinkman Date

ELECTRONIC COPY

I, the undersigned, the author of this work, declare that the electronic copy of this thesis provided to the James Cook University Library, is an accurate copy of the print thesis submitted, within the limits of the technology available.

_______________________________ _______________

Signature Date

Declaration on Ethics

The research presented and reported in this thesis was conducted within the guidelines for research ethics outlined in the National Statement on Ethics Conduct in Research Involving Human (1999), the Joint NHMRC/AVCC Statement and Guidelines on Research Practice (1997), the James Cook University Policy on Experimentation Ethics. Standard Practices and Guidelines (2001), and the James Cook University Statement and Guidelines on Research Practice (2001). The proposed research methodology received clearance from the James Cook University Experimentation Ethics Review Committee (Approval Number A901).

Diane Brinkman Date

Statement on the Contribution of Others

Scientific Collaborations

I gratefully thank the following people and organisations for providing resources and scientific/technical assistance to me during my candidature:

- Dr Jamie Seymour and Ms Griselda Ávila-Soria (JCU), Reef HQ (Great Barrier Reef Marine Park Authority) and Queensland Surf Life Saving Association for generously providing jellyfish samples
- Dr Jamie Seymour and Ms Teresa Carrette (JCU) for performing lethality assays on crayfish (Animal Ethics Approval No. A901), as described in Chapter 4
- Dr Heinrich Körner, Mrs Julie Webster and staff of the Immunogenetics Research Facility (JCU) for assisting in the preparation of mouse antibodies against *C. fleckeri* venom
- Dr Lisa-Anne Gershwin (JCU) for providing information and advice relating to box jellyfish taxonomy
- Dr Andrew Greenhill and Professor Phillip Summers (JCU) for kindly providing heparinised sheep blood, sheep blood agar plates and technical advice for haemolytic assays
- Analysis of *C. fleckeri* venom proteins using mass spectrometry has been facilitated by access to the Australian Proteome Analysis Facility (APAF) established under the Australian Government's Major National Research Facilities Program

Financial Assistance

This research project was funded by James Cook University, the Australian Lions Foundation and the Cooperative Research Centre (CRC) for Reef Research, as detailed below.

- JCU Prestige Postgraduate Research Scholarship (2005–2007)
- JCU Graduate Research Scheme Grant (2005)
- Australian Lions Foundation Grant (2005)
- CRC Reef Research Grant (2005–2006)

Publications Arising from Thesis

At the time of thesis submission, two manuscripts describing the research findings of Chapters 5 and 6 were already published and an invited review based on Chapters 1 and 8 was submitted. Three additional manuscripts are currently in preparation. Details of each manuscript are provided below.

- **Brinkman, D.**, Burnell, J., 2007. Identification, cloning and sequencing of two major venom proteins from the box jellyfish, *Chironex fleckeri*. *Toxicon* 50, 850-860. (Chapter 5)

- **Brinkman, D.**, Burnell, J., 2008. Partial purification of cytolytic venom proteins from the box jellyfish, *Chironex fleckeri*. *Toxicon* 51, 853-863. (Chapter 6)

- **Brinkman, D.**, Burnell, J., *submitted*. Biochemical and molecular characterisation of cubozoan protein toxins. *Toxicon.* (Chapters 1 and 8)

- Ávila-Soria, G., **Brinkman, D.**, Burnell, J., *in prep*. Molecular aspects of box jellyfish cDNA expression libraries and screening by antibody probes. *FEBS Letters.* (Chapter 3)

- **Brinkman, D**., Burnell, J., *in prep*. Biochemical characterisation of nematocyst-derived venom proteins from the box jellyfish, *Chironex fleckeri*. *Toxicon*. (Chapter 4)

- **Brinkman, D.**, Burnell, J., *in prep*. Expression of box jellyfish (*Chironex fleckeri*) toxins in *Escherichia coli*. *Protein Expression and Purification.* (Chapter 7)

Acknowledgements

Firstly, I wish to thank my supervisor Professor Jim Burnell for his continued support and encouragement during my PhD candidature. I have greatly benefited from his in-depth knowledge and practical experience in molecular biology and biochemistry, during my transition from chemist to molecular biochemist.

I gratefully thank Dr Jamie Seymour for his enthusiastic support of my project, his willingness to collaborate and for providing information and advice relating to his expertise in the biology and ecology of Australian jellyfish.

I also thank the research staff and students in the Department of Biochemistry and Molecular Biology at JCU for providing technical advice, training when necessary, and access to instrumentation and research facilities. I particularly thank my laboratory colleagues, Ursula Tems, Griselda Ávila-Soria, Ronelle Welton and Stacey Cole for their helpfulness, good company and personal support during my research project.

I also express my gratitude to the Australian Institute of Marine Science for supporting my applications for study leave during my postgraduate studies at JCU. I especially thank Dr Lyndon Llewellyn and Dr Cédric Robillot for encouraging me to undertake doctoral studies in the field of marine toxinology. Special thanks also to Dr Kathryn Burns for her encouragement and personal support during my postgraduate studies.

Finally, I am very grateful for the unwavering support and encouragement of my husband, Richard, and the welcome distraction of our endearing children, Lindsay and Elliot.

Abstract

Chironex fleckeri is a dangerous Australasian box jellyfish that inflicts painful, debilitating and potentially life-threatening stings in humans. The venom of *C. fleckeri* contains a variety of bioactive proteins that are cytolytic, cytotoxic, inflammatory or lethal, however, few individual venom proteins have been thoroughly characterised and their mechanism(s) of action remain unclear. Hence, the primary objectives of this thesis were to identify and characterise the major protein components in *C. fleckeri* venom, provide insight into their possible structures, functions and mechanisms of action, and explore the potential to express recombinant venom proteins in bacteria.

Two of the most abundant proteins contained in the nematocysts of *C. fleckeri* were identified in this study as *C. fleckeri* toxin-1 (CfTX-1) and toxin-2 (CfTX-2). The two proteins also represent the first *C. fleckeri* venom proteins to be successfully cloned and sequenced. The molecular masses of CfTX-1 and CfTX-2 (~43 and 45 kDa, respectively) were determined by SDS-PAGE, and both proteins were strongly antigenic to commercially available box jellyfish antivenom (CSL Ltd) and rabbit polyclonal antibodies raised against nematocyst-derived *C. fleckeri* venom. A combination of N-terminal amino acid sequencing, peptide mass fingerprinting, RT-PCR and cDNA library screening was used to isolate and clone cDNA encoding CfTX-1 and -2 (1789 and 1624 bp, respectively). Searches of non-redundant protein databases revealed that the deduced amino acid sequences of mature CfTX-1 and CfTX-2 (436 and 445 residues, respectively) were similar to three lethal, haemolytic box jellyfish toxins: CqTX-A from *Chiropsalmus quadrigatus*, CrTXs from *Carybdea rastoni* and CaTX-A from *Carybdea alata*. The protein and cDNA sequences of the five box jellyfish toxins were not similar to any other sequence in protein or nucleotide databases, supporting a hypothesis that the toxins may have evolved as highly specialised cubozoan toxins. Following a multiple sequence alignment of the five protein sequences, several short, but highly conserved regions of amino acids coincided with a predicted transmembrane spanning region, which could be involved in a pore-forming mechanism of action. Furthermore, remote protein homology predictions for the family of box jellyfish toxins suggested weak structural similarities and, hence, inferred function to pore-forming insecticidal δ-endotoxin proteins.

CfTX-1 and -2 were difficult to separate using electrophoretic or chromatographic methods, however, the two proteins were significantly co-purified from *C. fleckeri* venom using size exclusion chromatography or cation exchange chromatography. The native molecular mass of the co-purified CfTX proteins was 370 kDa, suggesting the formation of oligomeric

quaternary structures. The co-purified CfTX proteins were potently haemolytic to sheep erythrocytes (HU_{50} = 14 ng/mL) and caused the formation of large cleared zones of haemolysis in 5% sheep blood agar, thus confirming predictions of a pore-forming mechanism of action. Due to the significant sequence similarity of CfTX-1 and -2 to CrTX-A, CaTX-A and CqTX-A, the CfTX proteins could also be lethal, painful, inflammatory and dermonecrotic, and as such, may be the primary cause of life-threatening effects in envenomed humans.

During size exclusion and ion exchange chromatography experiments, a second major cytolysin (145 kDa) and a minor cytolysin (70 kDa) were also partially purified from *C. fleckeri* venom. The 145 kDa cytolysin, comprised of two major proteins (~39 and 41 kDa), was twice as haemolytic to sheep erythrocytes (HU_{50} = 7 ng/mL) as co-purified CfTX-1 and -2. Notably, the 39 and 41 kDa proteins were not significantly antigenic to CSL box jellyfish antivenom or rabbit antibodies raised against nematocyst-derived venom and the proteins were not always present in different batches of nematocysts. Due to the relatively high abundance of the 39 and 41 kDa proteins in some batches of *C. fleckeri* nematocysts compared to others, the variable concentrations of the proteins could have a major impact on the potency and variety of biological activities elicited by *C. fleckeri* venom. Furthermore, the 39 and 41 kDa proteins may represent a novel class of cytolytic proteins that are produced by cubozoan jellyfish.

Another strategy used to identify putative *C. fleckeri* venom proteins in this study involved the construction and screening of a *C. fleckeri* tentacle cDNA expression library with antibodies raised against *C. fleckeri* venom. Although no putative venom clones were isolated using this approach, cDNA clones encoding 44 independent non-venom proteins were isolated, sequenced and characterised, thus providing the first preliminary survey of the transcriptome of *C. fleckeri*.

Novel studies were also undertaken to assess whether CfTX-1 and -2 could be heterogeneously expressed in *Escherichia coli* for further functional and structural characterisation and/or for potential use in future antivenom production or other therapeutic applications. Although results of the studies indicated that CfTX-1 and -2 can be expressed in a bacterial host, expression levels were too low (ng protein/g cells range) to permit further characterisation. Furthermore, a large proportion of the expressed CfTX proteins formed insoluble inclusion bodies that required solubilisation and refolding prior to purification as soluble, potentially active recombinant proteins.

The research presented in this thesis, typified by the isolation, identification and characterisation of two major protein toxins in *C. fleckeri* venom, will benefit future research investigating the mechanisms of action of box jellyfish venom proteins in envenomed humans and potentially assist in the development of improved clinical treatments for box jellyfish stings.

Table of Contents

Statement of Access ii

Statement of Sources iii

Declaration on Ethics iv

Statement on the Contribution of Others v

Publications Arising from Thesis vi

Acknowledgements vii

Abstract viii

List of Tables xv

List of Figures xvi

CHAPTER 1 1

Current Understanding of *Chironex fleckeri* and its Venom –Background and Research Aims 1

1.1 Introduction 1
1.2 Brief Overview of Taxonomy and Life Cycle 2
1.3 General Anatomy and Behaviour of *Chironex fleckeri* 4
1.4 Morphology and Function of *Chironex fleckeri* Nematocysts 6
1.4.1 Microbasic Mastigophores 8
1.4.2 Trirhopaloids 9
1.4.3 Atrichous Isorhizas (without spines) 9
1.4.4 Holotrichous Isorhizas (complete coverage with spines) 9
1.4.5 Variation in Nematocyst Type Ratios during Development 10
1.5 Development of Cnidarian Nematocysts 10
1.6 The Physiology of Nematocyst Discharge (Exocytosis) 11
1.7 Isolation of Jellyfish Venom 13
1.8 Bioactivity of *C. fleckeri* Venom 14
1.8.1 Lethal Activity and Potential Mechanism(s) of Action 15
1.8.2 Cytolytic Activity 20
1.8.3 Severe Pain and Damage to Skin 21
1.9 Effects on Venom Bioactivity 23
1.9.1 Buffer Composition 23
1.9.2 Temperature and Storage 24
1.9.3 Ionic Strength 25
1.9.4 pH 25
1.9.5 Lipid Membrane Components and Osmotic Protectants 26
1.10 Fractionation of *C. fleckeri* Venom Components 26
1.10.1 Size Exclusion Chromatography 28

1.10.2 Ion Exchange Chromatography ... 29
1.10.3 Immunoaffinity Chromatography ... 30
1.10.4 Hydrophobic Interaction Chromatography (HIC) ... 31
1.10.5 Electrophoretic Separation Methods ... 31
1.11 Characterisation of Toxic Proteins from Other Jellyfish ... 33
1.12 Clinical Effects of *C. fleckeri* Envenomation ... 36
1.12.1 Symptoms ... 36
1.12.2 Treatment ... 37
1.13 Project Aims and Research Outcomes ... 40

CHAPTER 2 ... 43
General Methods and Materials ... 43
2.1 Sample Collection and Storage ... 43
2.2 Preparation of *C. fleckeri* Venom ... 43
2.2.1 Bead Mill Homogenisation ... 43
2.2.2 Lyophilised Milked Venom ... 44
2.2.3 Chemical Induction of Nematocyst Discharge ... 44
2.3 General Protein Methods ... 44
2.3.1 Protein Concentration ... 44
2.3.2 Polyclonal Antibodies against *C. fleckeri* Venom ... 44
2.3.3 SDS-PAGE Analysis ... 45
2.3.4 Western Blot Analysis ... 46
2.3.5 Two Dimensional Gel Electrophoresis (2DGE) ... 47
2.3.6 N-Terminal Amino Acid Sequencing ... 47
2.3.7 Peptide Mass Fingerprinting ... 48
2.3.8 Haemolytic Activity ... 48
2.4 General Nucleic Acid Methods ... 49
2.4.1 Quantification and Purity of Nucleic Acids ... 49
2.4.2 Isolation of Total RNA ... 49
2.4.3 Isolation of Poly(A)$^+$ RNA ... 50
2.4.4 Isolation of Plasmid DNA ... 51
2.4.5 Restriction Enzyme Digestion ... 51
2.4.6 DNA Ligation ... 51
2.4.7 Polymerase Chain Reaction (PCR) ... 51
2.4.8 Reverse Transcription Polymerase Chain Reaction (RT-PCR) ... 52
2.4.9 Agarose Gel Electrophoresis ... 53
2.4.10 Purification of DNA Fragments ... 53
2.4.11 Preparation of Radiolabelled DNA Probes ... 53
2.4.12 *C. fleckeri* cDNA Expression Library Construction and Titration ... 54
2.4.13 Immunoscreening of Expressed cDNA Library Clones ... 54
2.4.14 DNA Screening of cDNA Library Phage ... 55

2.4.15 *In vivo* Excision of pBluescript SK- Clones ... 56
2.4.16 DNA Screening Bacterial Colonies ... 56
2.4.17 DNA Sequencing ... 56
2.5 Bacterial Methods ... 57
2.5.1 Preparation of Competent Cells ... 57
2.5.2 Transformation of Competent Cells ... 57
2.5.3 Bacterial Glycerol Stocks ... 57

CHAPTER 3 ... 58

Analysis of a *C. fleckeri* Tentacle cDNA Expression Library – the Search for Putative Venom Proteins ... 58

3.1 Introduction ... 58
3.2 Methods ... 59
3.3 Results and Discussion ... 60
3.3.1 Immunodetection of cDNA Clones using CSL box jellyfish antivenom ... 60
3.3.2 Immunodetection of cDNA Clones using Mouse and Rabbit Antibodies Raised Against *C. fleckeri* Nematocyst Venom ... 71
3.3.3 Failure to Detect Putative Venom Clones ... 72
3.3.4 Elongation Factor-1α ... 74
3.3.5 General Features of Other Commonly Detected Full-length cDNA Clones ... 77
3.3.6 Thypedin – a Thymosin-like Protein or a Pedin Precursor? ... 79
3.4 Commentary ... 81

CHAPTER 4 ... 83

Biochemical Characterisation of *Chironex fleckeri* Venom Proteins ... 83

4.1 Introduction ... 83
4.2 Methods and Materials ... 84
4.2.1 Preparation of Nematocyst Venom ... 84
4.2.2 SDS-PAGE and Western Blot Analysis ... 84
4.2.3 Two Dimensional Gel Electrophoresis (2DGE) ... 84
4.2.4 Examination of Bead Mill Extraction Efficiency ... 84
4.2.5 Effects of NaCl Concentration on Nematocyst Protein Extraction ... 85
4.2.6 Lethal Activity of Whole Nematocysts Extracts ... 85
4.2.7 Comparison of *C. fleckeri* Venom Proteins from Different Extract Sources ... 86
4.2.8 Antibody Specificity towards *C. fleckeri* Venom Proteins ... 87
4.2.9 Interspecies Comparison of Nematocyst Extracts ... 87
4.3 Results and Discussion ... 88
4.3.1 SDS-PAGE and Western Blot Analysis ... 88
4.3.2 Two Dimensional Gel Electrophoresis (2DGE) ... 91
4.3.3 Examination of Bead Mill Extraction Efficiency ... 93
4.3.4 Effects of NaCl Concentration on Nematocyst Protein Extraction ... 94

4.3.5 Lethal Activity of Whole Nematocysts Extracts ... 96
4.3.6 Comparison of *C. fleckeri* Venom Proteins from Different Extract Sources ... 99
4.3.7 Interspecies Comparison of Nematocyst Extracts ... 103
4.3.8 Antibody Specificity towards *C. fleckeri* Venom Proteins ... 106

CHAPTER 5 ... 109

Identification and Molecular Characterisation of Two Major Venom Proteins from *Chironex fleckeri* ... 109

5.1 Introduction ... 109
5.2 Methods and Materials ... 110
5.2.1 SDS-PAGE and Western Blot Analysis ... 110
5.2.2 N-Terminal and Internal Amino Acid Sequencing ... 110
5.2.3 Isolation of Full-Length cDNA Clones Encoding Two Major Venom Proteins . 110
5.3 Results ... 112
5.3.1 SDS-PAGE and Western Blot Analysis ... 112
5.3.2 N-Terminal and Internal Amino Acid Sequencing ... 113
5.3.3 Isolation of Full-Length cDNA Clones Encoding Two Major Venom Proteins . 113
5.3.4 Phylogenetic Analysis ... 119
5.3.5 Secondary Structure Analysis ... 119
5.3.6 Tertiary Structure Analysis ... 120
5.4 Discussion ... 123

CHAPTER 6 ... 127

Fractionation of *Chironex fleckeri* Venom Proteins ... 127

6.1 Introduction ... 127
6.2 Methods and Materials ... 128
6.2.1 Preparation of Nematocyst Extracts ... 128
6.2.2 SDS-PAGE and Western Blot Analysis ... 128
6.2.3 Immunoaffinity Chromatography ... 128
6.2.4 Size Exclusion Chromatography ... 129
6.2.5 Cation Exchange Chromatography ... 130
6.2.6 Haemolytic Activity ... 130
6.3 Results ... 131
6.3.1 Immunoaffinity Chromatography ... 131
6.3.2 Size Exclusion Chromatography ... 134
6.3.3 Cation Exchange Chromatography ... 139
6.4 Discussion ... 142

CHAPTER 7 .... **148**

Expression of Major *C. fleckeri* Venom Proteins in *E. coli* **148**

7.1 Introduction 148
7.2 Methods 149
7.2.1 Subcloning CfTX-1 into pProEX HTc 149
7.2.2 Subcloning CfTX-2 into pProEX HTc 149
7.2.3 Small-Scale Expression of CfTX-1 and -2 150
7.2.4 Large-Scale Expression of CfTX-1 and -2 151
7.2.5 Purification of Recombinant CfTX-1 and -2 151
7.2.6 SDS-PAGE and Western Blot Analysis 154
7.3 Results 155
7.3.1 Subcloning CfTX-1 and -2 from pBSK- into pProEX HTc 155
7.3.2 Small-Scale Expression of CfTX-1 and -2 155
7.3.3 Large-Scale Expression of CfTX-1 and -2 157
7.4 Discussion 163

CHAPTER 8 .... **166**

General Conclusions and Future Research **166**

8.1 Research Overview 166
8.2 Is There an Easy Way to Isolate and Identify *C. fleckeri* Venom Proteins? 166
8.3 Biochemical Characterisation of *C. fleckeri* Venom Proteins 167
8.4 Molecular and Biochemical Characterisation of CfTX-1 and -2 171
8.5 Additional Cytolytic Proteins in *C. fleckeri* Nematocyst Venom 172
8.6 The Emergence of a Novel Family of Bioactive Cubozoan Proteins 172
8.7 Future Research 175

REFERENCES **178**

APPENDICES

Appendix 1 Peptide Mass Fingerprinting A1
Appendix 2 Functional Classification of Antigenic *C. fleckeri* Proteins A2
Appendix 3 *C. fleckeri* Nucleotide Sequence Clusters Obtained Using CSL Box Jellyfish Antivenom Immunodetection A4
Appendix 4 Nucleotide and Deduced Amino Acid Sequences of Expressed *C. fleckeri* Clones Obtained by Immunoscreening a Tentacle cDNA Library A6
Appendix 5 Eukaryote EF1α Protein Sequences A46
Appendix 6 CfTX-1 Mass Fingerprinting Data – APAF A52
Appendix 7 Calculated Tryptic Peptide Monoisotopic Ion Masses of CfTX-1 and -2 .... A57
Appendix 8 Nucleotide and Deduced Amino Acid Sequences of *C. fleckeri* Venom Protein Expression Constructs A63

List of Tables

Table 1.1 Summary of partially purified bioactive proteins from the box jellyfish, *C. fleckeri* 27

Table 2.1 Isoelectric focusing programme for the separation of *C. fleckeri* nematocyst proteins. 47

Table 2.2 Components of PCR reaction mixtures for amplification of DNA with Taq DNA polymerase, based on a final reaction volume of 20μL........ 52

Table 3.1 Forward (F) and reverse (R) primers designed according to the nucleotide sequences of pBSK clones partially encoding thypedin, ABP or HSP70. 59

Table 3.2 Summary of antigenic *C. fleckeri* cDNA clones obtained by immunodetection with CSL box jellyfish antivenom........ 62

Table 5.1 Degenerate forward (F) and reverse (R) primers designed according to the internal peptide sequences, FIAMVVQR and NDDLEKFNFVPVK, respectively, present in a 43 kDa major *C. fleckeri* nematocyst protein. 111

Table 6.1 Purification of two *C. fleckeri* haemolysins (370 and 145 kDa) using Superdex 200 size exclusion chromatography. 137

Table 7.1 Program for the elution of recombinant CfTX-1 extracted under native conditions from a nickel-affinity column. 152

Table 7.2 Program for the elution of recombinant CfTX-2 extracted under native conditions from a nickel-affinity column. 152

Table 7.3 Program for the elution of recombinant CfTX-1 extracted under denaturing conditions from a nickel-affinity column. 153

Table 7.4 Program for the elution of recombinant CfTX-2 extracted under denaturing conditions from a nickel-affinity column. 154

Table 8.1 Novel family of bioactive box jellyfish proteins........ 173

List of Figures

Figure 1.1 Spatial distribution of *C. fleckeri* populations in Australian coastal waters 1

Figure 1.2 Taxonomic scheme for *C. fleckeri* and selected cnidarians 3

Figure 1.3 Life cycle of *C. fleckeri* 4

Figure 1.4 Basic anatomy of *C. fleckeri* 5

Figure 1.5 Radial arrangement of nematocysts on *C. fleckeri* tentacles 6

Figure 1.6 The cnidome of *C. fleckeri*. 7

Figure 1.7 Ultrastructure of the microbasic mastigophore 8

Figure 1.8 Schematic diagram of nematocyst discharge 12

Figure 1.9 Mediators and signalling pathways involved in the regulation of intracellular Ca^{2+} in mammalian cardiocytes. 17

Figure 1.10 The main signalling pathways and mediators involved in the control of vascular smooth muscle 19

Figure 1.11 Mediators and signalling pathways involved in inflammation and pain 22

Figure 1.12 Flowchart of biochemical and molecular biology methods used to characterise the proteins contained in *C. fleckeri* venom 42

Figure 3.1 Secondary screen of clone 1-2 using CSL box jellyfish antivenom immunodetection 60

Figure 3.2 Expression profile of *C. fleckeri* tentacles categorised according to general protein function 61

Figure 3.3 Full-length cDNA (clone 4-5-1) encoding EF1α detected by immunoscreening the *C. fleckeri* cDNA library with CSL box jellyfish antivenom 63

Figure 3.4 Full-length cDNA (clone 3-9-1) encoding 3-PGDH obtained by immunoscreening the *C. fleckeri* cDNA library with CSL box jellyfish antivenom 64

Figure 3.5 Full-length cDNA (clone 7-1) encoding HSP70-1, obtained by DNA screening the *C. fleckeri* cDNA library with a gene-specific radiolabelled oligonucleotide probe 66

Figure 3.6 Full-length cDNA (clone 13-5) encoding ABP (coronin 1B), obtained by DNA screening the *C. fleckeri* cDNA library with a gene-specific radiolabelled oligonucleotide probe 68

Figure 3.7 Incomplete cDNA (clone 3-9) encoding thypedin obtained by DNA screening the *C. fleckeri* cDNA library with a gene-specific radiolabelled oligonucleotide probe.. ... 70

Figure 3.8 Molecular phylogenetic tree of eukaryotes based on Neighbour Joining analysis of EF1α amino acid sequences. 75

Figure 3.9 Multiple sequence alignment of highly basic N-terminal signature motifs in short coronins 78

Figure 3.10 (a) Alignment of two distinct β-thymosin-like repeat sequences (types 1 and 2) in the partial sequence of thypedin (*C. fleckeri*). Multiple sequence alignment of *C. fleckeri* β-thymosin-like repeats (b) type 1 and (c) type 2 with their respective best matches obtained by BLAST analysis, including thypedin homologues and different members of the β-thymosin family 80

Figure 3.11 Multiple alignment of the repeat sequence for pedin in thypedin from *H. vulgaris* and corresponding sequences in thypedin homologues from *N. vectensis* and *C. fleckeri* 81

Figure 4.1 Typical profiles of non-lyophilised nematocyst venom proteins following SDS-PAGE and western blot analysis using 15% polyacrylamide gels (electrophoresis time = 45min) 89

Figure 4.2 Typical profiles of non-lyophilised nematocyst venom proteins following SDS-PAGE and western blot analysis using 15% polyacrylamide gels (electrophoresis time = 2h) 89

Figure 4.3 Reducing (R) and non-reducing (NR) SDS-PAGE analysis of non-lyophilised nematocyst venom proteins 90

Figure 4.4 Effects of cysteine alkylation on the molecular masses of non-lyophilised nematocyst venom proteins 90

Figure 4.5 Separation of nematocyst venom proteins by 2DGE 92

Figure 4.6 Western blot of nematocyst venom proteins separated by 2DGE 92

Figure 4.7 Comparison of nematocyst proteins from non-lyophilised nematocysts obtained using bead mill (BM) and direct extraction methods 94

Figure 4.8 Effect of NaCl concentration on the extraction of non-lyophilised nematocyst proteins using bead mill homogenisation. 95

Figure 4.9 Comparison of protein profiles of nematocyst extracts 1 and 2 (E1 and E2), used in lethality studies on *C. quadricarinatus*. 98

Figure 4.10 Comparison of *C. fleckeri* bead mill venom proteins from non-lyophilised (NL) and lyophilised (L) nematocysts 99

Figure 4.11 Comparison of nematocyst extracts (Nem.) and milked venom (Milked) 101

Figure 4.12 Further comparison of bead mill nematocyst extracts (Nem.) from non-lyophilised nematocysts (NL) and milked venom (Milked) 101

Figure 4.13 Comparison of nematocyst venom proteins from *C. fleckeri* (*C.f.*) and *C. bronzie* (*C.b.*) 105

Figure 4.14 Haemolytic activity of *C. fleckeri* and *C. bronzie* venoms on 5% sheep blood agar 106

Figure 4.15 Binding of rabbit, mouse and sheep antibodies to *C. fleckeri* bead mill nematocyst proteins 107

Figure 4.16 Binding of rabbit, mouse and sheep antibodies to *C. fleckeri* bead mill nematocyst proteins. 107

Figure 5.1 SDS-PAGE and western blot analysis of *C. fleckeri* nematocyst venom 112

Figure 5.2 Nucleotide and deduced amino acid sequences of CfTX-1 116

Figure 5.3 Nucleotide and deduced amino acid sequences of CfTX-2 117
Figure 5.4 Comparison of five homologous box jellyfish proteins using ClustalW multiple amino acid sequence alignment 118
Figure 5.5 Phylogenetic relationships of five homologous box jellyfish proteins 119
Figure 5.6 Crude 3-D model predictions for (a) CfTX-1 and (b) CfTX-2 122
Figure 5.7 Peptide maps of (a) CfTX-1 and (b) CfTX-2 124
Figure 6.1 Separation of nematocyst extract proteins by CSL antivenom immunoaffinity chromatography 132
Figure 6.2 Separation of nematocyst venom proteins using CfTX-specific immunoaffinity chromatography 133
Figure 6.3 Separation of nematocyst venom proteins by CfTX-specific immunoaffinity chromatography 134
Figure 6.4 Fractionation of crude nematocyst extracts by size exclusion chromatography..... 136
Figure 6.5 Dose-response curves of sheep erythrocyte lysis by crude venom and partially purified *C. fleckeri* haemolysins obtained using size-exclusion chromatography 138
Figure 6.6 Kinetics of purified CfTX-1 and -2 haemolytic activities at different temperatures.. 139
Figure 6.7 Fractionation of crude nematocyst extracts by cation exchange chromatography . 141
Figure 7.1 Small-scale expression of CfTX-1 and -2 in *E. coli* NM522................................ 156
Figure 7.2 Bacterial growth following induction of protein expression. 156
Figure 7.3 Purification of soluble CfTX-1 under native conditions.. 158
Figure 7.4 Purification of CfTX-1 from inclusion bodies under denaturing conditions 159
Figure 7.5 A comparison of proteins in the pellet before and after solubilisation of CfTX-1 inclusion bodies under denaturing conditions 160
Figure 7.6 Purification of soluble CfTX-2 under native conditions....................................... 161
Figure 7.7 Purification of CfTX-2 from inclusion bodies under denaturing conditions 162
Figure 7.8 A comparison of two cDNA library screening strategies for the detection of CfTX-1 and -2 clones 165

CHAPTER 1

Current Understanding of *Chironex fleckeri* and its Venom – Background and Research Aims

1.1 Introduction

The Australasian box jellyfish, *Chironex fleckeri*, is considered to be the most dangerous jellyfish to humans in the world (Wiltshire et al., 2000) and in Australia alone, approximately 70 fatal envenomations (~55 of which were children) have been reported (Fenner and Harrison, 2000; Ramasamy et al., 2004). The most recent fatalities in Australia include a seven year old girl near Bamaga (northern Queensland, January, 2006; ABC News, 2006) and a 6-year-old boy at the Tiwi Islands (Northern Territory, November, 2007; ABC News, 2007).

The spatial distribution of *C. fleckeri* populations includes the tropical, coastal waters of Queensland, the Northern Territory and Western Australia (Hartwick, 1987); a region extending from north of Agnes Water to Exmouth (Fenner and Harrison, 2000) (Figure 1.1).

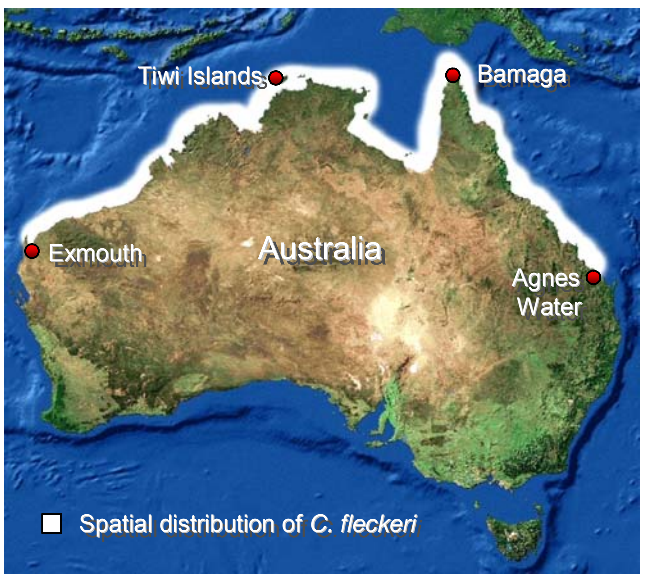


Figure 1.1 Spatial distribution of *C. fleckeri* populations in Australian coastal waters.

Adult *Chironex* populations reach a seasonal maximum during warmer months (November to May), however, their abundance is generally sporadic and unpredictable (Hartwick, 1991). Although this so-called “stinger season” represents the most dangerous time period for humans, envenomations by *Chironex* have been reported in every month of the year in the Northern Territory and every month except July in Queensland (Fenner and Harrison, 2000).

1.2 Brief Overview of Taxonomy and Life Cycle

C. fleckeri is a cubozoan jellyfish that belongs to phylum Cnidaria (Figure 1.2). Cubozoans, scyphozoans and hydrozoans are broadly classed as “jellyfish” because they all develop medusae during their life cycle. In contrast, anthozoans lack the medusa stage during their life cycle (Rifkin et al., 1996). The life cycle of *C. fleckeri* is partitioned between estuarine and coastal marine environments (Figure 1.3). In autumn, adult medusae enter coastal estuaries, undergo sexual reproduction via paired spawning to produce large numbers of fertilised eggs in the water and subsequently die (Hartwick, 1987). Resulting larvae rapidly develop into microscopic free-swimming planulae that search for suitable substrates for temporary settlement (Hartwick, 1991). Following settlement, the planulae transform into primary creeping (“roving”) polyps that further explore their substrate for a site on which to permanently attach themselves (Hartwick, 1987; Burnett et al., 1996a).

The sessile polyp matures and becomes fully self-sufficient, feeding on plankton and reproducing asexually by budding new creeping polyps from its body wall. The small, harmless polyps grow during winter and add tentacles, but never exceed 1mm in diameter (Hartwick, 1991). In spring, the polyps undergo metamorphosis in 10–14 days to form small, juvenile medusae that break away from their rocky substrates and migrate downstream towards the sea (Hartwick, 1991). The juvenile jellyfish initially feed on plankton, but as they grow larger and develop more tentacles and greater numbers of nematocysts, they are able to capture larger prey such as crustaceans and fish (Hartwick, 1991). On average, *C. fleckeri* reach a size that is dangerous to humans within 2-3 months (Hartwick, 1991).

Phylum Cnidaria

- **Class Cubozoa** (box jellyfish)
  - **Order Carybdeida** (single-tentacled pedadium)
    - Genus *Carybdea*
      - *C. rastoni*, *C. alata, C. marsupialis, C. xaymacana*
    - Genus *Carukia*
      - *C. barnesi* (Irukandji), *C. shinju*
    - Genus *Alatina*
      - *A. mordens*, *A. moseri* (Hawaiian box jellyfish)
    - Genus *Malo*
      - *M. maxima*, *M. kingi*
  - **Order Chirodropida** (multi-tentacled pedalium)
    - **Genus *Chironex***
      - ***C. fleckeri***
    - Genus *Chiropsalmus*
      - *C. quadrigatus*, *C. quadrumanus*
    - Genus *Chiropsella*
      - *C. bronzie*
- Class Scyphozoa ('true' jellyfish)
  - Order Semaeostomeae (scalloped medusae)
    - Genus *Chrysaora* (sea nettle)
      - *C. quinquecirrha*
    - Genus *Cyanea* (hair jellyfish)
      - *C. capillata* (grows up to 1m in diameter)
  - Order Coronatae (grooved medusae)
  - Order Rhizostomeae (lack tentacles)
    - Genus *Catostylus*
      - *C. mosaicus* (blubber jellyfish)
    - Genus *Cassiopea* (upside-down jellyfish)
      - *C. xamachana*, *C. ndrosia*
- Class Hydrozoa (hydroids)
  - Subclass Siphonophora, Order Cystonectae
    - Genus *Physalia*
      - *P. physalis* (Portuguese man-of-war)
  - Subclass Hydroida, Order Milleporida
    - Genus *Millepora*
      - *M. platyphylla* (fire coral)
- Class Anthozoa (sea anemones and true corals)
  - Order Actiniaria (solitary sea anemomes)
  - Order Zoanthinaria
  - Order Corallimorpharia
  - Order Scleratinia (hard corals)

Figure 1.2 Taxonomic scheme for *C. fleckeri* and selected cnidarians. (Adapted from Rifkin, 1996a; modified according to Gershwin, 2005, 2006.)

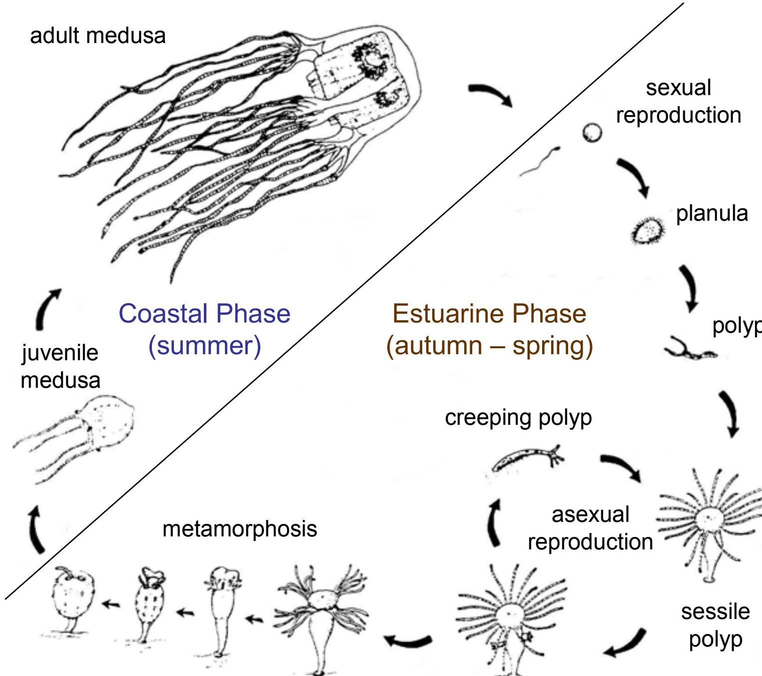


Figure 1.3 Life cycle of *C. fleckeri.* (Adapted from Hartwick, 1987.)

1.3 General Anatomy and Behaviour of *Chironex fleckeri*

C. fleckeri, like other cubozoans, is generally characterised by its cuboid or box-shaped bell (Figure 1.4). A mature *Chironex* can measure up to 325 mm in bell height (Kinsey, 1986) and weigh more than 6 kg (Marsh and Slack-Smith, 1986). At the lower corners of the bell, there are four complex and fleshy lobes, called pedalia, to which clusters of tentacles are attached (Williamson, 1985). This tentacle arrangement provides a morphological distinction between the cubozoan families Chirodropida, to which *C. fleckeri* belongs, and Carybdeida, which includes species such as *Carukia barnesi* (Irukandji) (Figure 1.2). Whereas chirodropids have multiple tentacles extending from each pedalium, carybdeids have a less complex arrangement, with only one tentacle attached to each pedalium.

Juvenile *C. fleckeri* may have only a few tentacles extending from each pedalium, but mature specimens can have up to 15 tentacles attached to each pedalium; 60 in total. Fully extended, each tentacle can measure more than 3 metres in length (Burnett et al., 1996a) and when fully retracted, tentacle length can be dramatically reduced to ~8 cm (Williamson, 1985).

A sense organ (rhopalium) is located on each side of the bell. The 4 rhopalia each contain a balance organ (statocyst) and a complex eye (ocellus), including lens and retina (Marsh and

Slack-Smith, 1986; Rifkin, 1996b). The ocelli are light sensitive (Rifkin, 1996b) and generally enable the jellyfish to sense and avoid large obstructions in their environment (Hartwick, 1987).

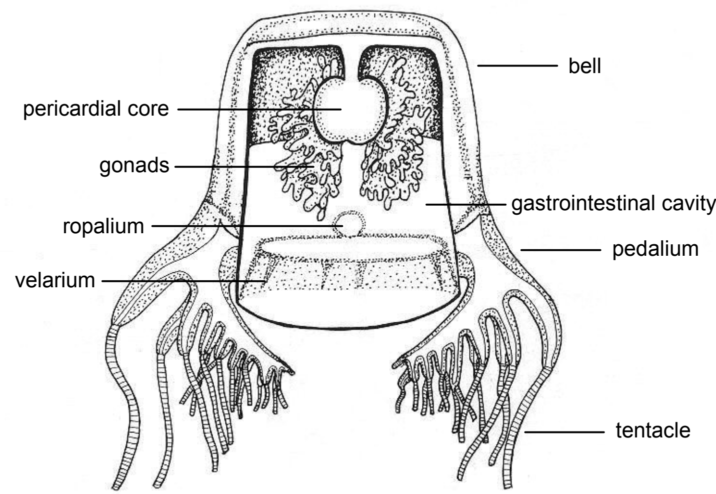


Figure 1.4 Basic anatomy of *C. fleckeri.* (Drawing based on photographs and preserved specimens, by Jan Hynes. Adapted from Rifkin, 1996b.)

Unlike most jellyfish, cubozoans are powerful swimmers with good control of speed and direction (Marsh and Slack-Smith, 1986). Chirodropid jellyfish, including *C. fleckeri* and *Chiropsella bronzie*[*], propel themselves through the water, bell-first with tentacles trailing behind, using a powerful and rhythmic "pulse and coast" style of jet propulsion (Shorten et al., 2005; Burnett et al., 1996a). Based on laboratory and field studies, adult *C. fleckeri* can achieve maximal velocities ≥ 400 m/h (Shorten et al., 2005; Seymour et al., 2004), although extremely high velocities of 4–5 knots/h (7.4–9.3 km/h) in short bursts have also been reported (Burnett et al., 1996a). Cubozoan jellyfish can contract or extend their tentacles voluntarily, and while hunting, the tentacles trail behind the swimming animal, extended to their full length to maximise prey capture. The natural prey of *C. fleckeri* are fish, crustaceans and other invertebrates, while natural predators include turtles and large fish (Hartwick, 1987; Williamson, 1985).

Human envenoming almost invariably results from a victim swimming, diving or walking into the tentacles of the jellyfish (Hartwick, 1987). As box jellyfish are often colourless, they

[*] Reported as *Chiropsalmus sp.* in Shorten et al. (2005); reclassified as *Chiropsella bronzie* in Gershwin (2006).

can be extremely difficult to see in the water, particularly in areas of high turbidity where the cubozoans concentrate their feeding activities (Marsh and Slack-Smith, 1986; Hartwick, 1987).

1.4 Morphology and Function of *Chironex fleckeri* Nematocysts

Nematocysts are small non-living capsules that proliferate the external cell layers of jellyfish tentacles. In *C. fleckeri*, the nematocysts are arranged radially in discrete rings along each tentacle in a characteristic "ladder-rung" pattern (Figure 1.5). Conservatively, it has been estimated that an average adult *Chironex* carries a total of 4–5 billion nematocysts on its tentacles (Hartwick, 1987; Williamson, 1985).

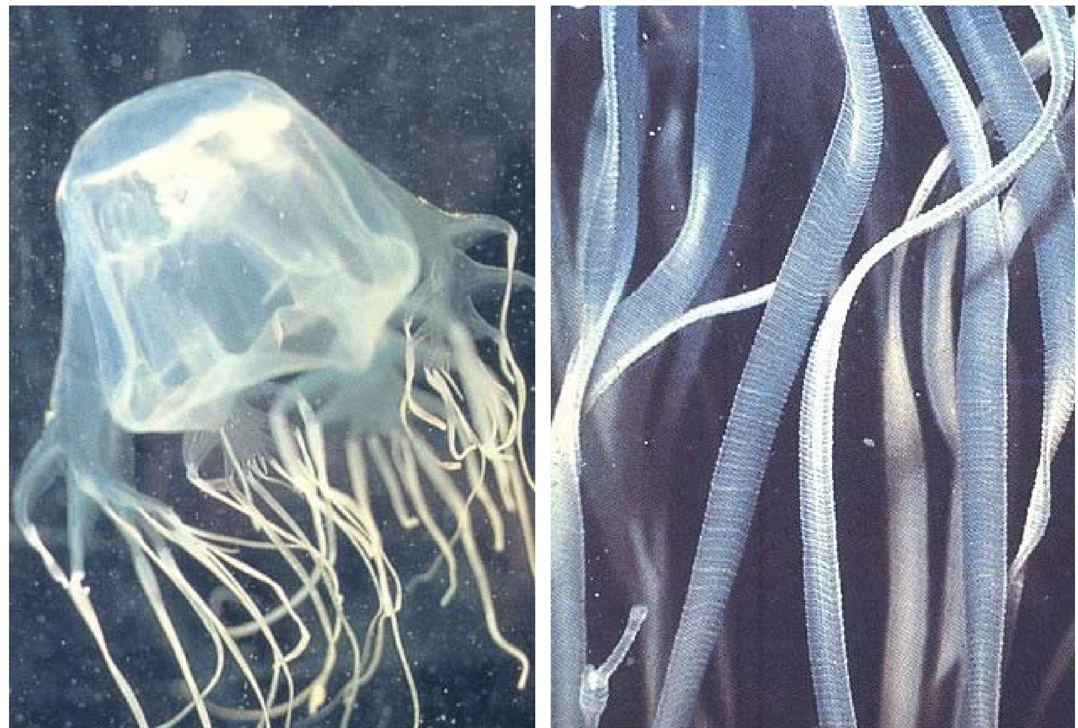


Figure 1.5 Radial arrangement of nematocysts on *C. fleckeri* tentacles. (Photographs courtesy of Dr Bob Hartwick. Reproduced from Williamson et al. (1996) with the permission of the University of New South Wales Press Ltd.)

The basic structure of an undischarged nematocyst consists of a capsule, containing a matrix and inverted tubule, and a lid-like structure called an operculum (Rifkin, 1996b). In nature, nematocysts function to capture prey and provide defence (Marsh and Slack-Smith, 1986; Rifkin, 1996b). Once killed or paralysed, captured prey are conveyed by the tentacles to the mouth and gastrointestinal cavity of the jellyfish (Rifkin, 1996b).

Different Cnidaria have characteristic types of nematocysts, and their shape, size and internal structures/arrangements can be used to identify individual species (Marsh and Slack-Smith, 1986). The cnidome of *C. fleckeri* (Figure 1.6) comprises four different types of nematocysts, two of which have distinct size classes: microbasic mastigophores (large and small), atrichous isorhizas, holotrichous isorhizas and trirhopaloids (large and small)† (Endean et al., 1969; Rifkin and Endean, 1983; Rifkin, 1996b).

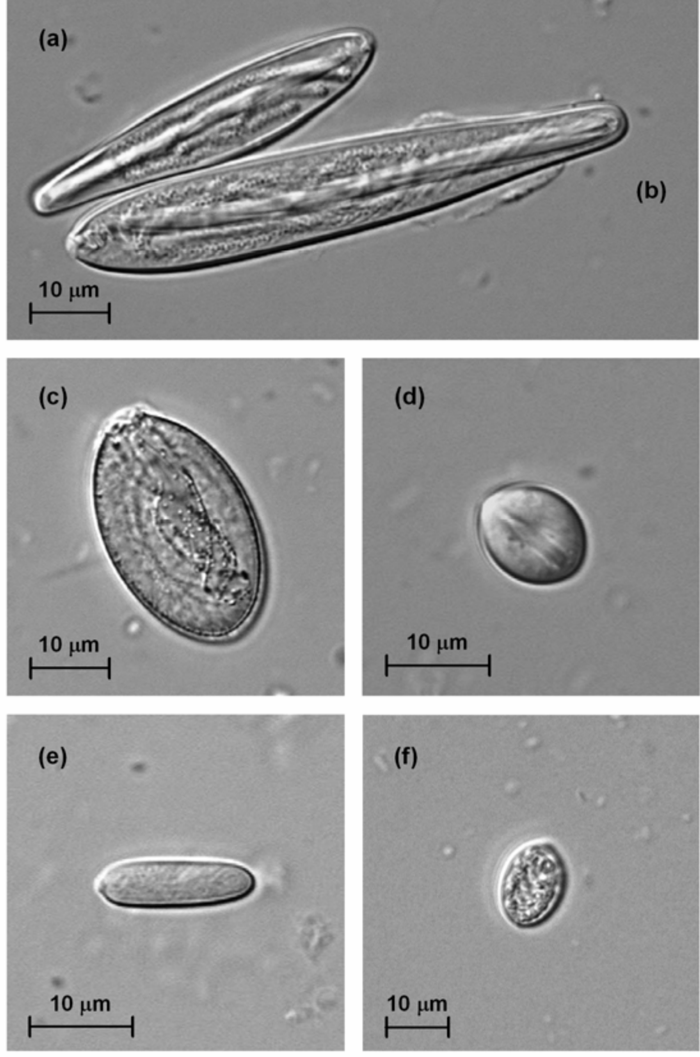


Figure 1.6 The cnidome of *C. fleckeri*. (a) Small and (b) large microbasic mastigophores (length 22–90µm; penetrative); (c) large trirhopaloid (average length ~34µm; penetrative); (d) small trirhopaloid (average length ~12µm); (e) atrichous isorhiza (length 5–17µm; agglutinant); (f) holotrichous isorhiza (average length ~15µm; volvent). (Photographs by Diane Brinkman.)

† formerly classified as heterotrichous microbasic euryteles

1.4.1 Microbasic Mastigophores

Microbasic mastigophores constitute approximately 43% of the mature *C. fleckeri* cnidome (Endean et al., 1969; Carrette et al., 2002a) and exhibit both penetrative and injecting functions (Rifkin and Endean, 1983). In an undischarged state, the microbasic mastigophore is a smooth, elongated, "cigar-shaped" capsule, which is slightly tapered at one end and ranges in length from 22–90µm (Rifkin and Endean, 1983) (Figure 1.7a). A clover-leaf shaped operculum is situated at the apical end of the capsule. Inside the nematocyst, a long inverted tubule with a stiffened spine-covered base (or butt) arises from the broader, apical end of the capsule. The base, approximately 2µm in diameter, is loosely coiled to form a helix and extends almost the length of capsule, where it merges into a finer tubule approximately 1µm in diameter. The inverted tubule, usually about 6 times the length of the capsule, is pleated and tightly coiled in a helical formation. The lumen of the tubule is covered in rows of fine internalised spines. A granular matrix is present in the capsule and the lumen of the inverted tube. During discharge of the mastigophore, the tubule rapidly everts through the operculum, penetrates the cuticle of the victim/prey and injects venom from the capsule (Figure 1.7b).

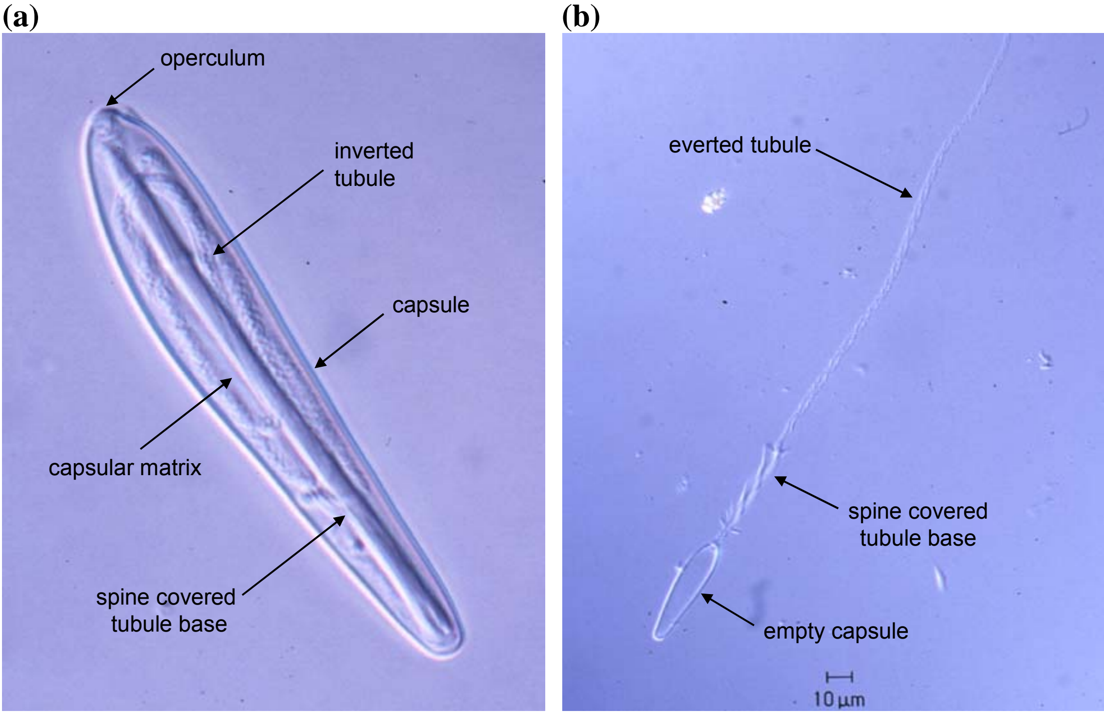


Figure 1.7 Ultrastructure of the microbasic mastigophore. (a) undischarged nematocyst; (b) discharged nematocyst. (Reproduced from Mastro et al. (1998) with permission from Dr Lyndon Llewellyn.)

1.4.2 Trirhopaloids

Trirhopaloids constitute approximately 13% of the mature *C. fleckeri* cnidome (Endean et al., 1969). They are present as two distinct size classes, with the large, ellipsoid trirhopaloids averaging ~34μm in length (Figure 1.6c) and the small, more spherical trirhopaloids averaging ~12μm (Figure 1.6d) (Rifkin and Endean, 1983). The undischarged trirhopaloid capsule contains a deeply pleated butt with large spines and a long thread, measuring approximately 20 times the length of the capsule. The structure of trirhopaloids indicates both penetrative and injector functions (Rifkin and Endean, 1983). However, unlike extracts from isolated microbasic mastigophores, trirhopaloid extracts do not elicit powerful responses in animal musculature (Endean and Rifkin, 1975). Instead, trirhopaloids are suggested to act as penetrating grappling hooks, capable of securing and dragging the integument of prey into closer contact with the jellyfish tentacle (Rifkin and Endean, 1983).

1.4.3 Atrichous Isorhizas (without spines)

Atrichous isorhizas are small cylindrical nematocysts (5–17μm in length) that represent approximately 42% of the mature *C. fleckeri* cnidome (Rifkin and Endean, 1983; Endean et al., 1969) (Figure 1.6e). They contain a long, butt-less thread devoid of spines, which is usually 12–14 times the length of the capsule (Rifkin and Endean, 1983). Absence of spines on the thread infers lack of penetrative function. However, the nematocysts contain a sticky fluid, which appears to facilitate adhesion of the tentacle to the integument of the prey/victim and thus suggests an agglutinant function (Rifkin, 1996b).

1.4.4 Holotrichous Isorhizas (complete coverage with spines)

Holotrichous isorhizas are the least prevalent nematocysts in the *C. fleckeri* cnidome and represent only ~2% of total nematocysts (Endean et al., 1969). The ellipsoid capsule of the holotrichous isorhiza averages 15μm in length and contains a butt-less thread covered entirely with helically arranged spines (Endean and Rifkin, 1975) (Figure 1.6f). When discharged, the thread generally forms a spiral, which is thought to provide a method of small prey entanglement (Rifkin, 1996b).

1.4.5 Variation in Nematocyst Type Ratios during Development

Studies of nematocyst type ratios in *C. fleckeri* during development have shown a dramatic increase in the proportion of microbasic mastigophores as the jellyfish matures from juvenile to adult (Carrette et al., 2002a). This ontogenetic shift in nematocyst ratios has been linked to the changing diet of *C. fleckeri* during development. Whereas juvenile *C. fleckeri* feed predominantly on small crustaceans such as the prawn, *Acetes australis*, during maturation, the jellyfish increasingly include fish in their diet. An increase in the proportion of microbasic mastigophores in mature *C. fleckeri* effectively increases the total amount of venom that is available to kill or paralyse larger prey and thus may facilitate the dietary shift to piscivory (Carrette et al., 2002a; Kintner et al., 2005). Furthermore, the proportion of trirhopaloids increases with jellyfish size/maturity, supporting an hypothesis that the "grappling hook"-like function of trirhopaloids may better serve to anchor fish rather than crustaceans, which are typically smaller, with more irregular surfaces. In contrast, a decrease in the proportion of isorhizas occurs during maturation, suggesting that the entanglement function of these nematocysts may be limited during the capture of relatively smooth animals, such as fish, compared to crustaceans (Carrette et al., 2002a).

1.5 Development of Cnidarian Nematocysts

Although little is known specifically about the development of cubozoan nematocysts, a model for the complex process of cnidarian nematocyst morphogenesis has been proposed based on studies in *Hydra* (Engel et al., 2002). The model involves assembly of proteins within a large post-Golgi vesicle to form a double-layered capsule containing a matrix and an inverted tubule. Nematocyst development begins within specialised cells called nematoblasts, which subsequently undergo terminal division to form nematocytes. During an early growth phase, the capsule primordium forms and grows by addition of new protein-filled vesicles to the original vesicle harbouring the capsule. A high molecular weight glycoprotein synthesised in the endoplasmic reticulum (ER), Nowa (nematocyst outer wall antigen), is transported within vesicles to the nematocyst vesicle and enriches the outer wall layer of the capsule. Minicollagens synthesised in the ER are also transported within separate vesicles to the nematocyst vesicle and accumulate in the capsular matrix (Engel et al., 2002).

Following outer wall construction, fusion of additional protein-containing vesicles at the apex of the capsule leads to the formation of an external tubule, with the wall of the capsule and

the tubule eventually forming a continuous structure. The outer tubule subsequently invaginates into the nematocyst. Formation of tubule spines occurs by condensation of the protein, spinalin, within the lumen of the inverted tubule. During maturation of the capsule, the minicollagens in the capsular matrix polymerise to form the insoluble inner wall of the capsule. Nowa molecules become cross-linked (via covalent disulphide bonds) to minicollagen polymers at the interface between the inner and outer capsule layers, which leads to compaction of the wall structure and formation of a highly stable composite biomaterial (Engel et al., 2002; Meier et al., 2007).

A final late stage of maturation involves the synthesis of poly-γ-glutamate (pG) in the capsular matrix, which is accompanied by swelling of the nematocyst due to increasing internal pressure (Szczepanek et al., 2002). This process serves to generate the extremely high pressures required to drive discharge of the mature nematocyst. The minicollagens of the capsule inner wall provide the necessary tensile strength to withstand the high internal pressures of the matrix (~150 atmospheres) when the nematocyst is in its undischarged state (Holstein et al., 1994).

Enzymes required for spinalin, pG and other protein syntheses are imported into the capsular matrix during nematocyst formation when the tubule and capsule walls are still sufficiently permeable to allow the passage of proteins (Szczepanek et al., 2002). Dynamic microtubule scaffolding in the nematocyte is proposed to play two important roles during nematocyst morphogenesis. Firstly, microtubules provide "tracks" to guide the transport of additional protein-filled vesicles to the site of the growing nematocyst vesicle and secondly, cage-like microtubule scaffolding appears to facilitate stabilisation of nematocyst shape during construction (Engel et al., 2002).

In *Hydra*, nematocysts are formed in clusters of differentiating nematocytes within the body column (Szczepanek et al., 2002). Upon completion of differentiation, the clusters of nematocytes break up into single cells that migrate to the epithelium of the tentacles.

1.6 The Physiology of Nematocyst Discharge (Exocytosis)

Within the surface layers of the jellyfish tentacle, each mature nematocyst-containing nematocyte is surrounded by several epidermal cells called accessory cells (Rifkin, 1996b; Endean et al., 1991). Each nematocyte has a cnidocil (triggering apparatus) attached to its apical end, which is linked by fibres to the operculum of its associated nematocyst; immature nematocytes are easily recognisable as they lack a cnidocil (Endean et al., 1991). When mechanically and/or chemically stimulated, the cnidocil of the nematocyte initiates nematocyst discharge (Rifkin, 1996b; Kass-Simon and Scappaticci, 2002). Chemoreceptors on the

epidermis of the tentacles are possibly activated by chemicals released from potential prey, thus sensitising the nematocytes for nematocyst discharge upon mechanical stimulation by struggling prey (Rifkin, 1996b).

In *C. fleckeri*, the nematocyte assumes two different positions within the tentacle, referred to as "resting" and "fire-ready" states (Rifkin, 1996b). In the "resting" state, the nematocyte is retracted below the surface of the tentacle by a fibrillar network and covered by the apical regions of adjacent accessory cells to conceal the cnidocil. This inactive state is believed to occur when a jellyfish is well fed and thus avoids the unnecessary discharge of nematocysts. In the "fire-ready" (active) state, the apical ends of the neighbouring accessory cells are retracted, allowing projection of the nematocyte to the surface of the epidermis and exposure of the cnidocil to the external environment (Endean, 1988a; Rifkin, 1996b). The severity of prey/victim envenoming is therefore likely to depend on several factors, including the number of mastigophores positioned in an active position at the time of contact (Endean, 1988a), the level of maturity of the nematocysts (Endean et al., 1991) and whether the jellyfish tentacles are fully extended or retracted (Burnett et al., 1996a).

Nematocyst discharge is one of the fastest cellular processes in nature (Holstein and Tardent, 1984). During discharge of the microbasic mastigophore in *C. fleckeri*, the events that occur are similar to those described for the stenotele nematocyst in *Hydra attenuata* (see Figure 1.8). When discharge is initiated, the operculum of the nematocyst opens rapidly and the spine-covered butt of the tubule bursts explosively through the hole.

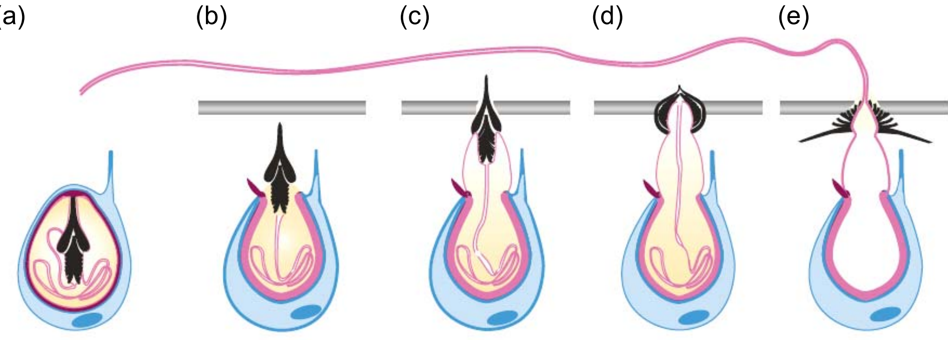


Figure 1.8 Schematic diagram of nematocyst discharge. The nematocyte (blue) harbours a single nematocyst (pink walls, tubule and operculum; black stylets (large spines)) (a) undischarged stenotele nematocyst from *Hydra attenuata* (b) operculum opens ejecting stylets (c) stylets puncture cuticle of prey/victim; (d) stylets expand; (e) tubule everts into subsurface tissue and venom is discharged. (Reproduced from Nüchter et al. (2006) with permission from Elsevier.)

As the butt emerges, its spines temporarily form a piercing structure, which facilitates penetration of the butt into the tissue of the prey/victim (Rifkin and Endean, 1983). In *Hydra*, this first step in nematocyst discharge is incredibly fast (0.7–1.4μs), with an average velocity of 9.3–18.6 m/s, and the generated acceleration ($1.34–5.41 \times 10^6$ *g*) is sufficiently forceful for the large stenotele spines (stylets) to penetrate the cuticle of crustacean prey (Nüchter et al., 2006). The spines of the butt fan outwards, presumably to anchor the butt securely in place. As the tightly coiled thread everts, it rotates such that successive thread spines cut a path through the tissue of the prey/victim (Rifkin and Endean, 1983). In *Hydra*, this second step is slower and characterised by contraction of the capsule (~18 μs), a brief arrest phase and tubule eversion (~2.7 m/s) (Nüchter et al., 2006).

The mechanism of venom delivery to the prey/victim by *C. fleckeri* is likely to be similar to that described for the Mediterranean jellyfish, *Rhopilema nomadica*. During tubule eversion, venom is translocated from the capsular matrix to the external environment via hollow, outward-facing spines or barbs located along the surface of the everted tubule (Lotan et al., 1996). The progressive release of venom along the tubule facilitates venom delivery to both vascular and non-vascular tissue, which may explain the rapid onset of venom-induced systemic effects, as well as the severe pain, damage and inflammation of the skin at the site of envenomation (Rifkin, 1996b). Eversion also facilitates the release of the contents of the tubule lumen, which coat the now outward-facing spines. In some cases, upon complete eversion, the capsular matrix also bursts through the end of the tubule (Rifkin, 1996b).

1.7 Isolation of Jellyfish Venom

Historically, *C. fleckeri* venom has been extracted most commonly from either whole tentacles or isolated nematocysts that were fresh, frozen or lyophilised (e.g. Crone and Keen, 1969; Othman and Burnett, 1990; Endean et al., 1993; Bloom et al., 1998; Carrette and Seymour, 2004). A variety of techniques have been used to extract active venom from *C. fleckeri* tentacles, including mortar-and-pestle homogenisation and sonication (e.g. Crone and Keen, 1969; Comis et al., 1989; Othman and Burnett, 1990). However, the use of whole tentacle extracts has been criticised because envenomed animals are never injected with tentacle material other than the contents of the nematocysts (Endean et al., 1993; Bloom et al., 1998). Furthermore, nematocyst-free tentacle tissues have also been shown to have pharmacological activity of their own (Endean and Noble, 1971; Endean et al., 1993; Ramasamy et al., 2005a).

Isolation of *C. fleckeri* nematocysts was first achieved by lightly grinding thawed tentacles in a mortar-and-pestle to liberate crude nematocysts, then allowing the nematocyst mixture to autolyse in seawater at 4°C for 5 days to remove tentacle debris (Endean et al., 1969). A subsequent modification of the method, involving the autolysis of excised whole *C. fleckeri* tentacles in seawater at 4°C over several days and lyophilisation of the isolated nematocyst sediments, provided a sound method for preserving intact nematocysts containing bioactive venom for extended periods of time (Bloom et al., 1998). In recent years, this method of nematocyst isolation has been routinely used by several jellyfish researchers (e.g. Carrette and Seymour, 2002b; 2004; Ramasamy et al., 2003, 2004; Winter et al., 2007a, b; Bailey et al., 2005; Kim et al, 2006).

As an additional step, some researchers have used density centrifugation to separate the different classes of isolated jellyfish nematocysts. For example, two types of *Chrysaora quinquecirrha* nematocysts, large oval isorhizas (specific gravity 1.616) and small oval euryteles (specific gravity 1.563) were successfully separated using caesium chloride density gradient centrifugation (Burnett and Calton, 1973). Similarly, *C. fleckeri* nematocysts were separated into four classes (microbasic mastigophores, trirhopaloids, atrichous isorhizas and holotrichous isorhizas) using a 76% Urografin™ density gradient (Endean and Rifkin, 1975).

Active venom has been successfully extracted from isolated nematocysts by electric all-glass micro-grinding (Endean et al., 1969), sonication (Bloom et al., 1998), mortar-and-pestle homogenisation (Wiltshire et al., 2000) and bead-mill homogenisation (Carrette and Seymour, 2004). Alternatively, an innovative method for the extraction of nematocyst venom from intact *C. fleckeri* tentacles involves the electrical stimulation of freshly excised tentacles to discharge their nematocyst venom through human amniotic membrane (Barnes, 1967). The nematocyst discharge, referred to as “milked venom”, is collected from the underside of the amnion and is currently used by CSL Ltd (Australia) to produce commercially available box jellyfish antivenom.

1.8 Bioactivity of *C. fleckeri* Venom

Early studies of extracts from frozen tentacle and isolated nematocysts have reported several biological activities associated with *C. fleckeri* venom, including lethal, dermonecrotic, myotoxic, cardiotoxic, haemodynamic and haemolytic effects (Baxter and Marr, 1969; Endean et al., 1969; Endean and Henderson, 1969; Crone and Keen, 1969; Freeman and Turner, 1969,

1971; Naguib et al., 1988). Venom from isolated nematocysts elicited powerful contractions in several muscle types, including striated musculature of barnacles, and skeletal, respiratory and extravascular smooth musculature of rats (Endean and Henderson, 1969). Rat hearts exposed to nematocyst venom displayed bradycardia (decreased heart rate) associated with progressive decrease in ventricular relaxation (Endean and Henderson, 1969). In more recent experiments, *C. fleckeri* nematocyst venom injected intravenously into rats produced a transient hypertensive response followed by hypotension and cardiovascular collapse within minutes (Ramasamay et al., 2004). Exposure of chick biventer cervicis nerve-muscle preparations to *C. fleckeri* and *C. bronzie* venoms significantly decreased indirect twitches of the tissue and inhibited contractile responses to exogenous acetylcholine; effects indicative of venom-induced neurotoxicity (Ramasamy et al., 2003). Direct twitches of the nerve-muscle preparations were also significantly decreased following exposure to *C. fleckeri* and *C. bronzie* venoms (suggesting venom-induced myotoxicity) and subsequent histological examination revealed non-selective damage to the skeletal muscle cells (Ramasamy et al., 2003). In another study, cultured rat cardiocytes exposed to *C. fleckeri* nematocyst venom ceased spontaneous contractions and failed to recover (Bailey et al., 2005).

Most (if not all) of the biological activity associated with *C. fleckeri* venom is attributed to microbasic mastigophore nematocysts (Endean and Rifkin, 1975). Extracts from ruptured mastigophores, isolated by density gradient centrifugation, exhibited activity against barnacle muscle and rat diaphragm musculature comparable to mixed nematocyst extracts. In contrast, extracts from ruptured trirhopaloids or atrichous isorhizas produced no effect on the same tissues (Endean and Rifkin, 1975).

1.8.1 Lethal Activity and Potential Mechanism(s) of Action

The lethality of *C. fleckeri* venom has been established in a wide variety of experimental animals including fish, prawns, crayfish, mice, rats, guinea pigs, sheep, rabbits and monkeys (Endean et al., 1969; Kintner et al., 2005; Carrette and Seymour, 2004; Baxter and Marr, 1969; Baxter and Marr, 1975; Comis et al. 1989; Endean et al. 1993; Wiltshire et al., 2000). However, differences in the source of jellyfish venom and the units of reported lethal doses have made comparison of results difficult.

In both experimentally injected mammals and envenomed humans, death caused by *C. fleckeri* venom appears to be due to pulmonary and/or myocardial arrest (Freeman and Turner, 1969, 1971; Lumley et al., 1988). However, recent research using nematocyst-derived

venom suggests that rapid death from *C. fleckeri* venom is more likely to be caused by severe cardiovascular effects rather than respiratory depression (Ramasamay et al., 2004). An intravenous route of venom entry also appears necessary for *C. fleckeri* and other cubozoan jellyfish to exert their lethal effects in mammals. Mice and rats injected with *C. fleckeri* venom via subcutaneous, intradermal or intraperitoneal routes have survived doses that would otherwise be fatal to intravenously injected animals within seconds (Endean et al., 1969; Baxter et al., 1968). Similar findings were reported for rats and rabbits injected with *Chiropsalmus quadrigatus* nematocyst venom (Noguchi et al., 2005; Koyama et al., 2003).

Despite several decades of research, the mechanism(s) underlying the lethal action of *C. fleckeri* venom remains unclear. Early studies suggested that venom-induced calcium ion (Ca^{2+}) transport across muscle fibre membranes plays an important role in muscle contracture and paralysis (Endean and Henderson, 1969). Crude *C. fleckeri* nematocyst venom caused an influx of Ca^{2+} into barnacle muscle fibres (Endean and Henderson, 1969) and reduced Ca^{2+} uptake by the sarcoplasmic reticulum (SR) of mouse striated muscles (Endean and Henderson, 1974). Verapamil (a phenylalkylamine L-type Ca^{2+} channel blocker) corrected arrhythmias in mice elicited by *C. fleckeri* tentacle extracts (Burnett and Calton, 1983) and countered the effects of a 150 kDa myotoxic nematocyst venom protein on smooth muscle and diaphragm musculature (Endean, 1988b), thus suggesting that venom-induced Ca^{2+} influx occurs via L-type Ca^{2+} channels in the sarcolemma (plasma membrane) of muscle fibres (Figure 1.9).

However, several reports have suggested that venom-induced intracellular Ca^{2+} entry is not mediated by L-type Ca^{2+} channel activation. In one study, researchers found that a lethal, cardiotoxic 20 kDa protein purified from milked venom did not facilitate Ca^{2+} transport through membranes, but instead caused the formation monovalent cation channels in the sarcolemma (Olson et al., 1984). More recently, nematocyst venoms from three cubozoan jellyfish, *C. fleckeri, C. bronzie* and *Carybdea xaymacana*, applied to rat myocytes caused significant irreversible increases in intracellular Ca^{2+} that were not blocked by verapamil, but were inhibited by lanthanum (La^{3+}), a non-specific cation channel and pore blocker (Bailey et al., 2005). Therefore, the rapid influx of Ca^{2+} into cardiac cells could be attributed to the indiscriminate entry of ions via large pores in the cell membrane formed after exposure to *C. fleckeri* venom rather than via Ca^{2+} channels (Bailey et al., 2005). Furthermore, verapamil was ineffective in preventing or reversing the *in vivo* cardiovascular effects of *C. fleckeri* tentacle and nematocyst venoms in anaesthetised piglets (Tibballs, 1988) and rats (Ramasamy et

al., 2004), although the conclusions of these studies remain contentious (Burnett, 1998; Burnett and Calton, 2004).

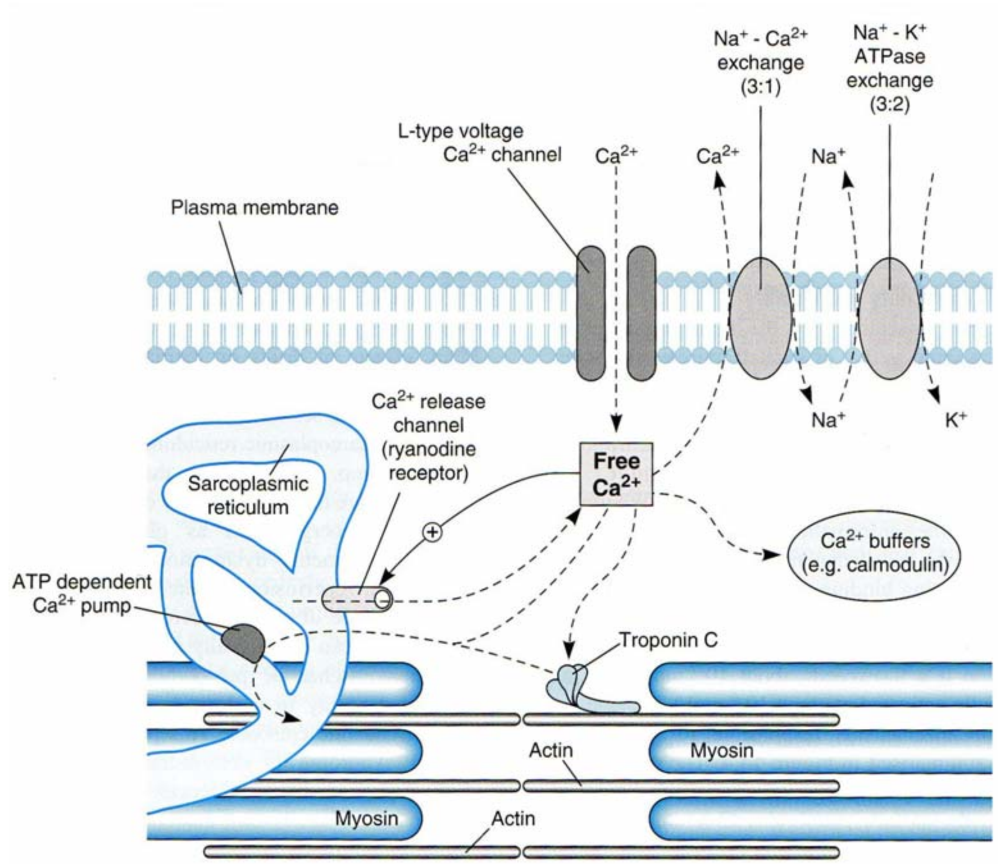


Figure 1.9 Mediators and signalling pathways involved in the regulation of intracellular Ca^{2+} in mammalian cardiocytes. (Reproduced from Rang et al. (1999).)

Following an extensive study of the effects of *C. fleckeri* tentacle extracts on ferret heart papillary muscles and isolated rat and ferret myocytes, a different mechanism underlying the cardiotoxic effects of *C. fleckeri* venom was proposed which emphasised the importance of intracellular sodium ions (Na^{+}) as well as Ca^{2+} (Mustafa et al., 1995). Examination of the antagonistic effects of nifedipine (a dihydropyridine L-type Ca^{2+} channel blocker) indicated that elevated intracellular Ca^{2+} concentrations caused by exposure to *C. fleckeri* tentacle extracts were unlikely to occur via the influx of extracellular Ca^{2+} through L-type Ca^{2+} channels. Furthermore, studies involving ryanodine (an agonist of Ca^{2+} release channels in the SR) and caffeine (a SR inhibitor), suggested that venom-induced elevation of intracellular Ca^{2+} concentrations was unlikely to involve Ca^{2+} release from the SR. However, the effects of the

venom were eliminated in the absence of extracellular Na^+ and by blockade of Na^+/Ca^{2+} exchangers with nickel ions (Ni^{2+}), indicating that Na^+ influx plays a primary role in the mechanism of action. To investigate the possible route(s) of Na^+ influx, studies involving tetrodotoxin (a Na^+ voltage-gated channel blocker), strophanthidin (a Na^+/K^+ ATPase pump inhibitor) and amiloride (a Na^+/H^+ exchange blocker), suggested that none of these mechanisms were related to venom-induced increases in intracellular Na^+ concentration. Instead, the group proposed that *C. fleckeri* venom induces an initial Na^+ influx via non-specific cation channels, thus increasing intracellular Na^+ concentration and depolarising the cell membrane. Secondly, Ca^{2+} efflux is decreased and/or Ca^{2+} influx is enhanced via Na^+/Ca^{2+} exchangers, which leads to increased intracellular Ca^{2+} and, in turn, activates Ca^{2+}-activated non-specific cation channels. The cell membrane is depolarised further, increasing Na^+ influx, thus enhancing further Ca^{2+} influx, until calcium overload is reached. This consequential overload of Ca^{2+} is likely to explain the rapid and irregular cardiac rhythm and conduction block, which typically characterises the cardiac response of intact animals to *C. fleckeri* venom (Mustafa et al., 1995). Although the nature of the non-specific ion channel involved in the initial Na^+ influx is uncertain, the cardiotoxic component(s) of *C. fleckeri* venom may act by opening existing ion channels in myocyte membranes or act as an ionophore by inserting itself into the membranes of the muscle cells (Mustafa et al., 1995).

However, more recent investigations of the *in vivo* cardiovascular and lethal effects of chirodropid venoms suggest that unlike L-type Ca^{2+} channels, Na^+ channels do not significantly contribute to the mechanism of action of cubozoan venoms (Koyama et al., 2003; Noguchi et al., 2005; Winter et al., 2007a). In a study of *C. quadrigatus* nematocyst venom in rabbits, venom-induced effects were significantly reduced by continuous infusion of diltiazem (a benzothiazepine L-type Ca^{2+} channel blocker), but were unaffected by Na^+ channel blockers (lidocaine and flecainide), thus suggesting that Ca^{2+} channels are more important to the mechanism(s) of action (Koyama et al., 2003). The effects of *C. quadrigatus* venom were also significantly attenuated in anaesthetised rats by prophylactic administration and continuous infusion of nicardipine (a dihydropyridine L-type Ca^{2+} channel blocker) and to a lesser extent by verapamil (Noguchi et al., 2005). In studies on rat aorta, felodipine (another dihydropyridine L-type Ca^{2+} channel blocker) significantly reduced the contractile effects of *C. fleckeri* nematocyst venom, but not the contractile effects of *C. bronzie* nematocyst venom (Winter et al., 2007a).

Researchers have also investigated other potential mechanisms of action that may explain the cardiovascular effects of chirodropid nematocyst venoms. In studies on pre-contracted

endothelium-intact rat aorta, *C. fleckeri* and *C. bronzie* nematocyst venoms elicited a two-phase response (initial relaxation followed by sustained contraction) (Winter et al., 2007a). However, in endothelium-denuded rat aorta this response was significantly inhibited, suggesting that the cardiovascular effects of the chirodropid venoms are partially mediated by the endothelium due to the release of relaxing or contracting factors. Further investigations revealed that NOLA (a nitric oxide synthase inhibitor) significantly reduced the initial phase of venom-induced relaxation in endothelium-intact aorta but not the secondary phase of sustained contraction, thus indicating that the chirodropid venoms promote the release of the endothelium-derived relaxing factor, nitric oxide (NO) (Winter et al., 2007a) (Figure 1.10).

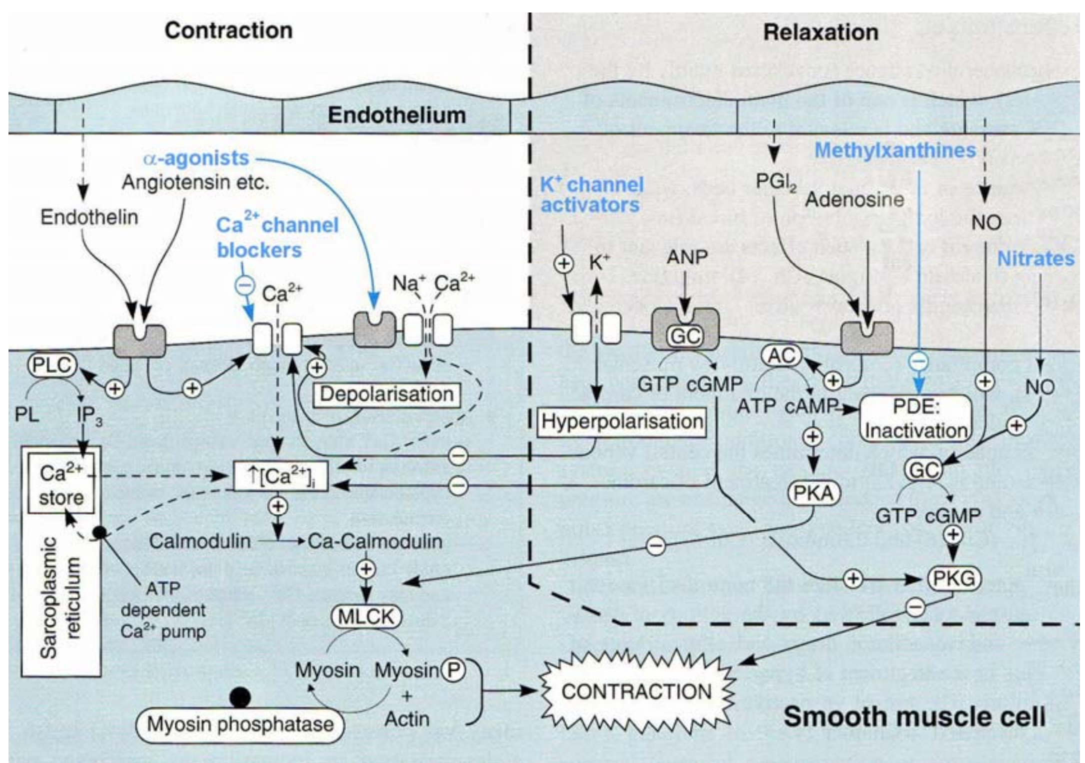


Figure 1.10 The main signalling pathways and mediators involved in the control of vascular smooth muscle. (Adapted from Rang et al. (1999).)

Using atropine (a muscarinic receptor antagonist), BQ 123 (an ET_A receptor antagonist) or HOE 140 (a BK_2 receptor antagonist), no significant inhibitory effects were observed in endothelium-intact aorta exposed to the chirodropid venoms suggesting that endogenous acetylcholine, endothelin or bradykinin, respectively, are not involved in promoting NO release (Winter et al., 2007a). Similarly, the responses of anaesthetised rats exposed to *C. quadrigatus*

nematocyst venom were unaffected by treatment with atropine or BQ 123 (Noguchi et al., 2005). In another study, ketanserin (a 5-HT_{2A} receptor antagonist) had no significant effect on the cardiovascular responses in anaesthetized rats exposed to *C. fleckeri* nematocyst venom, indicating that 5-HT_{2A} receptors, which play an important role in cardiac regulation, are not involved in the mechanism of venom action (Ramasamy et al., 2005a). Likewise, pretreatment with prazosin (an α_1-adrenoceptor antagonist) did not affect *in vitro* and *in vivo* cardiovascular responses to the nematocyst venoms of *C. fleckeri*, *C. bronzie* or *C. quadrigatus*, suggesting that α_1-adrenoceptors do not play a major role in the cardiovascular effects of chirodropid venoms (Winter et al., 2007a; Noguchi et al., 2005; Ramasamy et al., 2005a).

1.8.2 Cytolytic Activity

In general, cytolysins can be divided into two distinct classes: catalytic or stoichiometric (Hessinger, 1988). Catalytic cytolysins are enzymes that alter cell membrane permeability and/or structure by chemically changing membrane components (e.g. phospholipases which hydrolyse phospholipid substrates). In contrast, stoichiometric cytolysins interact with membranes to physically alter the membrane structure and permeability, which results in the formation of size- and/or charge-specific channels, or larger, less specific membrane pores (e.g. bacterial endotoxins). Stoichiometric cytolysins can also disrupt membrane integrity by miscellising membrane components (Hessinger, 1988).

In earlier studies of *C. fleckeri*, no phospholipase A or C activity (indicative of catalytic cytolysins) was detected in nematocyst venom, milked venom or tentacle extracts (Endean and Henderson, 1969; Baxter and Marr, 1969; Keen and Crone, 1969a; Crone, 1976). However, relatively high levels of phospholipase A_2 (PLA2) activity were detected in tentacle extracts of *C. fleckeri*, using a more sensitive assay than those used in previous studies (Nevalainen et al., 2004). The researchers postulated that PLA2 could play a role in the localised and systemic effects caused by jellyfish stings in humans (Nevalainen et al., 2004). However, because tentacle tissues were used in the study, there is no direct evidence as yet to specifically link PLA2 activities with nematocyst-derived toxins.

A cytolytic 20 kDa cardiotoxin purified from *C. fleckeri* milked venom has also been reported (Olson et al., 1984). The toxin was lethal to mice, lysed cultured chick cardiocytes and affected the ionic permeability of lipid bilayer membranes (composed of cholesterol and phosphatidylcholine) by producing monovalent cation channels. More recently, exposure of intact cultured rat cardiocytes to *C. fleckeri* nematocyst venom caused the formation of

numerous circular lesions with internal diameters of ~50–80nm (Bailey et al., 2005), thus indicating the presence of at least one stoichiometric cytolysin.

Haemolytic activity has also been detected in whole and partially purified tentacle and nematocyst extracts from *C. fleckeri*, demonstrating the cytolytic nature of one or more venom components (e.g. Endean and Henderson, 1969; Crone and Keen, 1969; Keen and Crone. 1969a; Baxter and Marr, 1969; Bailey et al., 2005). Yet, despite several reports of *C. fleckeri* venom causing haemolysis *in vitro* and in experimental animals, haemolysis has not been extensively observed in envenomed humans (Bailey et al., 2005; Tibballs, 2006). Some evidence also suggests that haemolytic activity of cubozoan venoms is unrelated to lethal activity (Bailey et al., 2005; Collins et al., 1993). However, these suggestions are contradicted by several studies in which proteins with both lethal and haemolytic properties have been isolated from the venoms of box jellyfish, including *C. rastoni*, *C. alata* and *C. quadrigatus* (Nagai et al., 2002, 2000a, b), as well as *C. fleckeri* (see Section 1.10). Moreover, general lytic activity is a common property of lethal, cytotoxic peptides or proteins produced by other venomous organisms, including hydra (e.g. Tamkun and Hessinger, 1981), snakes (e.g. Dufton and Hider, 1988; Kini and Evans, 1989) and stonefish (e.g. Garnier et al., 1995; Ghadessy et al., 1996).

1.8.3 Severe Pain and Damage to Skin

During human envenoming, millions of *C. fleckeri* nematocysts are discharged into the skin causing immediate, sharp and often excruciating pain and the formation of closely-spaced transverse lesions in a characteristic "ladder-rung" pattern at the site of the sting (Williamson, 1985). The local acute inflammation reaction that follows is swift and vigorous (Burnett et al., 1996b) and full thickness dermonecrosis and permanent scarring of the skin may ensue (Fenner, 1991). Early investigations of *C. fleckeri* venom in experimental animals showed that intradermal injections of tentacle, milked and nematocyst venoms all caused the development of localised haemorrhagic areas and dermonecrosis at the sites of injection (Keen and Crone, 1969b; Baxter and Marr, 1969).

Several researchers have attempted to elucidate the mechanism(s) underlying the severe pain and inflammation caused by *C. fleckeri* venom. Inflammatory pain is mediated by several chemical components, which act on pain receptors (nociceptors) directly or indirectly via complex signalling pathways (Figure 1.11). In experiments using tentacle-derived venom, pre-treatment of rats with anti-histamine and anti-serotonin compounds (diphenhydramine

hydrochloride and D-lysergic acid diethylamide tartrate, respectively) was ineffective against the venom-induced formation of lesions, suggesting that *C. fleckeri* venom has a direct affect on the epidermis and dermis rather than via promotion of histamine or serotonin release (Keen and Crone, 1969b). There was also little evidence to suggest that *C. fleckeri* venom caused increased capillary permeability, resulting from the release of vasoactive components, such as bradykinin (Keen and Crone, 1969b; Baxter and Marr, 1969). Moreover, pain-producing substances including acetylcholine, histamine, 5-hydroxytryptamine (serotonin) were not detected in tentacle extracts (Freeman and Turner, 1969). Therefore, initially, the pain associated with envenomation was attributed to direct tissue damage or the release of potassium from damaged cells (Freeman and Turner, 1969; Keen and Crone, 1969b).

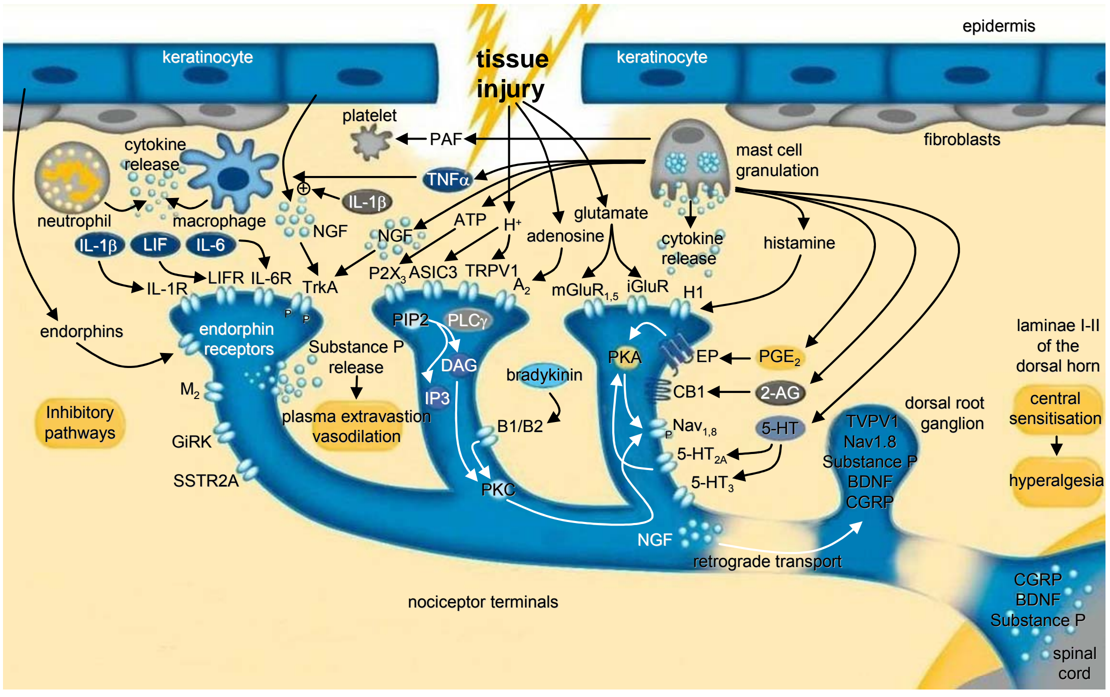


Figure 1.11 Mediators and signalling pathways involved in inflammation and pain. (Adapted from www.sapphirebioscience.com.)

However, in a later study, localised capillary leakage was observed in rats injected intradermally with *C. fleckeri* milked venom (Burnett and Calton, 1986), contradicting previous reports. Reduction of capillary leakage was achieved by pre-treatment of the rats with a variety of anti-inflammatory agents, including methysergide (an anti-serotonin agent), methacin (a prostaglandin inhibitor) and piripost (a leukotriene inhibitor).

In a more recent study involving *C. fleckeri* tentacle extracts, the burning pain sensation associated with cnidarian envenomation was linked to the activation of transient receptor potential vanilloid-I (TRPV1) channels in nociceptive neurons (Cuypers et al., 2006). Pre-treatment of rats with BCTC[‡], a TRVP1 channel antagonist, greatly reduced the induced pain effects caused by *Cyanea capillata* venom, thus demonstrating the potential therapeutic value of BCTC in the treatment of cnidarian stings. However, the efficacy of BCTC was not tested in rats injected with *C. fleckeri* venom, and unfortunately, the study was undertaken using cnidarian venoms extracted from acid-treated tentacles rather than isolated nematocysts.

1.9 Effects on Venom Bioactivity

Several factors affect the biological activity of *C. fleckeri* venom, including buffer composition, storage, temperature, ionic strength and pH.

1.9.1 Buffer Composition

The type of buffers and additives used for extraction of *C. fleckeri* venom can have a dramatic impact on the stability of the bioactive venom components. For example, the lethal and bioactive effects of fractionated venom were rapidly inactivated in Tris-HCl (pH 7.4) buffer compared to using 0.3M phosphate buffer (Endean, 1987). Nematocyst venom activity was also more stable in 0.3M phosphate buffer (pH 7.4) than in 0.9% saline (Endean and Henderson, 1969). However, addition of peptone or gelatine to 0.9% saline dramatically enhanced the stability over time of lethal and haemolytic activities in nematocyst and milked venoms (Baxter and Marr, 1969).

The potential benefits of adding EDTA, a metalloprotease inhibitor, remain unclear. In one study, addition of EDTA inactivated both the lethal and haemolytic activities of the nematocyst and milked venoms (Baxter and Marr, 1969). However in other studies, the addition of EDTA to extraction buffers had little effect on haemolytic activity (Crone and Keen, 1971) or typical activity against barnacle musculature (Endean and Henderson, 1969).

In a later study, the stability of tentacle extracts, determined by monitoring lethal and haemolytic activities over time, was greatest when there was no addition of buffer during tentacle homogenisation (Comis et al., 1989). Comparison of various buffers revealed that extract stability was significantly decreased in phosphate-buffered saline (PBS), PBS fortified

[‡] *N*-(4-tertiarybutylphenyl)-4-(3-chloropyridin-2-yl)tetrahydropyrazine-1(2*H*)-carboxamide

with divalent cations (Ca^{2+}, Mg^{2+} or Zn^{2+}), water or 10mM EDTA, and particularly in Tris-HCl buffer. In contrast, extract stability was less adversely affected using buffers such as 0.1M ammonium acetate, 1mM phenylmethylsulphonylfluoride (PMSF) and 0.3 M phosphate. Interestingly, the group also found that filtration of the crude tentacle extracts through Sep-pak C_{18} cartridges significantly improved the stability of extracts over time, however, other researchers found no significant improvement in stability using this procedure (Othman and Burnett, 1990).

1.9.2 Temperature and Storage

Several researchers have reported the thermolability of *C. fleckeri* venom (e.g. Baxter and Marr, 1969; Endean and Henderson, 1969; Endean et al., 1969; Endean, 1987; Bloom et al., 1998; Carrette and Seymour, 2004). For example, one group found that 0.9% saline extracts from *C. fleckeri* nematocysts stored at 4°C lost bioactivity against barnacle muscle within 24–48 hours, whereas extracts stored at ambient temperature (23–25°C) lost activity within 3 hours (Endean and Henderson, 1969). This study also revealed that the thermolability of extracts is clearly dependent on buffer composition because nematocyst venom in 0.3M phosphate (pH 7.4) retained substantial or full bioactivity against barnacle muscle after 24h incubation at ambient temperature. More recently, researchers have reported that the activity of nematocyst venom can remain stable following exposure to relative high temperatures (up to 30ºC) for short periods of time (≤ 1.5 h) (Carrette and Seymour, 2004; Winter et al., 2007b). In both cases, the nematocyst venoms were prepared by homogenisation of lyophilised nematocysts in distilled water. In contrast, all activity of *C. fleckeri* venom is lost when it is exposed to excessive heat (≥ 42°C), strongly indicating that the bioactive components of the venom are proteins (Endean and Henderson, 1969; Endean et al., 1969; Carrette and Seymour, 2004; Winter et al., 2007b).

Some studies have shown that frozen or freeze-dried *C. fleckeri* venom can retain most of its biological activity for extended periods of time (e.g. Baxter and Marr, 1969; Endean and Henderson, 1969). Following storage of nematocyst venom in 0.9% saline at -10ºC for 2 months, no appreciable loss in activity occurred, however, after 12 months, a 40% reduction in activity was observed (Endean et al., 1969). In contrast, the contents of intact nematocysts stored at -10ºC retained full activity over one year (Endean et al., 1969). Studies have also revealed that repeated freeze-thawing or lyophilisation-reconstitution of *C. fleckeri* venom also results in a significant reduction in bioactivity (Baxter and Marr, 1969; Bloom et al., 1998; Winter et al., 2007b).

1.9.3 Ionic Strength

Few researchers have investigated the effects of ionic strength on the bioactivity of *C. fleckeri* venom, however, one group observed that nematocyst venom in 0.9% saline (154mM NaCl) that was lyophilised and stored for 6 months at –10°C retained full activity when reconstituted in distilled water (Endean and Henderson, 1969). In contrast, nematocyst venom prepared using 2.5% and 3.5% saline (~430 and 600mM NaCl, respectively) prior to lyophilisation were found to be inactive upon reconstitution in distilled water. Furthermore, venom that was prepared using distilled water prior to lyophilisation retained its activity following reconstitution, but the latency period prior to observing a venom-induced response in muscle was significantly increased (Endean and Henderson, 1969). Consequently, the results of this study suggest that an ionic strength similar to physiological saline is more suitable for the extraction of bioactive *C. fleckeri* venom components.

In contrast to the findings for *C. fleckeri* venom, other researchers who purified lethal, haemolytic and dermonecrotic venom components from the box jellyfish, *Carybdea rastoni*, found that the haemolytic activity of the components were significantly stabilised in buffers of high ionic strength (Nagai et al., 2000a). The group recommended that active venom components should be stored under highly saline conditions (0.8M NaCl) and refrigerated (4ºC) rather than frozen or lyophilised. Using this strategy, 90% of the activity of the purified toxins was retained for more than 6 months.

1.9.4 pH

In an earlier study, the effects of pH on the bioactivity of *C. fleckeri* nematocyst venom were assessed on barnacle musculature (Endean and Henderson, 1969). Typical contractile activity was observed between pH 3.5–10, however, musculature responses with the shortest latency periods occurred between pH 6–10. In another study, the haemolytic activity of milked venom and nematocyst venom was optimal between pH 6.3–7 (Baxter and Marr, 1969). More recently, no significant difference in the cardiovascular effects of nematocyst venom in anaesthetised rats was observed between pH 5–9, however, the effects were markedly reduced at pH 3 (Winter et al., 2007b).

1.9.5 Lipid Membrane Components and Osmotic Protectants

The inhibitory effects of various membrane components and osmotic protectants on the haemolytic activity of *C. fleckeri* extracts have also been assessed. In earlier studies, neither egg yolk lecithin (phosphatidylcholine) nor cholesterol inhibited the haemolytic activity of *C. fleckeri* tentacle extracts, however, significant inhibition of haemolysis was achieved by the addition of gangliosides, *N*-acetylneuraminic acid and sucrose (Keen and Crone, 1969a; Crone, 1976). In a more recent study, little or no effect on the haemolytic activity was observed following pre-incubation of erythrocytes with gangliosides, phosphatidylserine and sphingomyelin, and only minor inhibition was observed for phosphatidylethanolamine and phosphatidylcholine (Bloom et al., 2001). Furthermore, cholesterol almost completely abolished the haemolytic activity of the venom, contradicting previous findings.

1.10 Fractionation of *C. fleckeri* Venom Components

Substantial research effort has focussed on the isolation and purification of bioactive proteins in nematocyst venom and tentacle extracts of *C. fleckeri*, however, the results between research groups vary significantly (see Table 1.1 for summary). Historically, the characterisation of bioactive venom proteins from cubozoan jellyfish has been hampered by the instability of the venom proteins and their propensity to aggregate/disaggregate and adhere to preparative surfaces or chromatography media (e.g. Othman and Burnett, 1990; Endean et al., 1993; Bloom et al., 1998). Variability in extract sources and/or the extraction and fractionation techniques used by researchers have also affected results.

Although comparison of the experimental data is difficult, it appears that *C. fleckeri* venom contains multiple lethal proteins varying in molecular mass. Lethal fractions are associated with cardiotoxicity (when tested) and often, but not always, associated with lytic activity. Proteins with relatively low molecular masses (10–30 kDa) are common to the SDS-PAGE protein profiles of lethal chromatography fractions (Olson et al., 1984; Naguib et al., 1988; Othman and Burnett, 1990; Endean et al., 1993), however, higher molecular mass proteins (40–50 kDa) are also frequently observed (Calton and Burnett, 1986; Othman and Burnett, 1990; Endean et al., 1993; Bloom et al., 2001). At least two proteins (120 and 70 kDa) are associated with dermonecrotic activity in addition to haemolytic and/or lethal activities (Naguib et al., 1988).

An overview of the various fractionation techniques used by researchers to fractionate and characterise bioactive *C. fleckeri* venom components are described in the following sections.

Table 1.1

Summary of partially purified bioactive proteins from the box jellyfish, *C. fleckeri*

Reference	Source[a]	Method[b]	Estimated Molecular Mass (kDa)		Biological
			Native	Denatured[c]	Activities[d]
Crone and Keen (1969*, 1971),	T, M	SE	(1) 150		L, C
Keen and Crone (1969b),		(* also CEX, AEX)	(2) 70		L, C, H, D
Freeman and Turner (1971),					
and Freeman (1974)					
Baxter and Marr (1969)	N, M	SE	(1) 350		
			(2) 10-30		L, H
			(3) 1.8		
Olson et al. (1984)	M	IA (α-*Physalia*), SE	65	**20** + several bands	L, C, P+
Calton and Burnett (1986)	T	IA (CSL antivenom)		150, 50	L, C
Naguib et al. (1988)	T	IA (α-C*hironex*)		120	L, H, D
				70	L, D
				14.5	L
				130,115,90,80,	ambiguous
				75,52,29	activities
Othman and Burnett (1990)	T_L	CEX (also AEX)		50, 24	L
	T_L	HIC		24	L
				45-50	L
	T_L	IA (α-C*hironex*)		**24** + several bands	L
Collins et al. (1993)	T_L	mAb (α-C*hironex*)		50	H
Endean (1987) and	N_M, N	SE	(1) 600	59,43,25,**18**	L, M
Endean et al. (1993)			(2) 150	59,43,25,**18**	L, M
			(3) 70	26,10	L, H
			(4) <14		
Bloom et al. (1998)	N_L	SE, NPAGE	438,225,207,45,22		L
	N_L	SDS-CGE		201,**174**,112,	
				71,39,31	
Bloom et al. (2001)	N_L	PEG-CE, SDS-CGE	(1)	201	hepatotoxic
			(2)	**173**,163,112,71	(F4>F1>F2>F3)
			(3)	39,14	
			(4)	<14	L, H, IR

[a] Sources of extracts include tentacle extract (T), milked venom (M) and nematocyst venom (N); subscripts L and M indicate lyophilised sources and venom isolated specifically from microbasic mastigophore nematocysts, respectively.
[b] Purification methods include size-exclusion (SE), cation-exchange (CEX), anion-exchange (AEX), and immunoaffinity (IA) chromatography; mAb indicates monoclonal antibodies; antibodies used for IA and mAb studies are indicated in brackets; electrophoretic methods include native PAGE (NPAGE), PEG-capillary electrophoresis (PEG-CE) and SDS-capillary gel electrophoresis (SDS-CGE).
[c] The molecular mass of denatured proteins was determined by SDS-PAGE, unless SDS-CGE is specified; major proteins are shown in bold.
[d] Biological activities: lethal (L), cardiotoxic (C), haemolytic (H), dermonecrotic (D), pore-forming in lipid membranes (P+), musculotoxic (M), induction of immunological response (IR) and hepatotoxic.

1.10.1 Size Exclusion Chromatography

Separation of *C. fleckeri* venom proteins by size exclusion chromatography has been undertaken by several independent researchers. In an early study, fractionation of tentacle extracts by size exclusion chromatography enabled the partial purification of a lethal, haemolytic and dermonecrotic 70 kDa protein and a 150 kDa protein that was lethal but not significantly haemolytic or dermonecrotic (Crone and Keen, 1969; Keen and Crone, 1969b). In a similar study, separation of tentacle extracts by size exclusion chromatography resulted in the elution of two lethal fractions, composed of several different proteins (Othman and Burnett, 1990). In this case, however, the native molecular masses of the peaks were not determined to allow comparison with the results of other researchers.

Using nematocyst and milked venoms instead of tentacle extracts, size exclusion chromatography yielded two major protein peaks (350 and 1.75 kDa), however, fractions containing proteins in the native molecular mass range 10–30 kDa exhibited the most potent haemolytic and lethal activities (Baxter and Marr, 1969). In contrast, size exclusion fractionation of venom from isolated microbasic mastigophores identified two lethal proteins, with molecular weights of approximately 600 kDa and 150 kDa (Endean, 1987). Both proteins induced contracture in several types of mammalian musculature, but only the 150 kDa protein exhibited activity against crustacean musculature. In a subsequent experiment (Endean et al., 1993), fractionation of nematocyst venom also enabled the isolation of a lethal 70 kDa haemolysin in addition to the 600 and 150 kDa myotoxins obtained in the previously study. Following reducing SDS-PAGE analysis, the haemolysin contained two major bands (10 and 26 kDa), whereas, the myotoxins contained a dominant 18 kDa protein band and three minor proteins bands (25, 43 and 59 kDa). Consequently, both myotoxins were proposed to be aggregates of the 18 kDa subunit. In contrast, chromatographic separation of extracts from tentacles devoid of nematocysts resulted in the isolation of a neurotoxic fraction (150 kDa) containing numerous proteins (including a major 25 kDa protein), and a haemolytic fraction (70 kDa) composed of 52 and 25 kDa subunits (Endean et al., 1993).

More recently, the fractionation of venom from lyophilised *C. fleckeri* nematocysts by size exclusion chromatography yielded only a single, broad protein peak of unreported molecular mass (Bloom et al., 1998). However, native PAGE analysis of concentrated lethal fractions which eluted prior to the large protein peak, revealed the presence of several proteins (438, 225, 207, 45 and 22 kDa).

1.10.2 Ion Exchange Chromatography

Several research groups have also fractionated *C. fleckeri* venom using ion exchange chromatography. Using weak cation exchange-size exclusion chromatography (CM-Sephadex), two bioactive fractions (150 kDa and 70 kDa) were isolated in approximately equal proportions from both *C. fleckeri* tentacle extracts and milked venom (Crone and Keen, 1971), consistent with previous size exclusion chromatography studies (Crone and Keen, 1969). The more strongly retained 70 kDa fraction was shown to have lethal, haemolytic and dermonecrotic properties, whereas the less retained 150 kDa fraction exhibited only lethal activity.

Separation of *C. fleckeri* tentacle extracts by weak anion-exchange chromatography also resulted in the partial purification of two bioactive fractions (Comis et al., 1989), however the nature of the activities are incongruous with the results of the previous study. Several retained protein fractions corresponding to a major protein peak were found to be lethal and haemolytic, whereas, the unretained fractions (containing positively charged or neutral proteins) were found to be lethal but not haemolytic. To be compatible with the previous study, the lethal and haemolytic fractions would have been less retained than the fractions that were lethal but not haemolytic. Furthermore, the molecular masses of the active fractions were not determined for comparison with other studies.

In a subsequent study, fractionation of *C. fleckeri* tentacle extracts by cation exchange chromatography resulted in the elution of a large unretained protein peak followed by a broad, poorly resolved pattern of proteins across the increasing NaCl concentration gradient (Othman and Burnett, 1990). Although the venom proteins were poorly resolved, a lethal factor was partially purified from the retained fractions. Non-reducing SDS-PAGE analysis revealed that the lethal factor exclusively contained a 24 kDa protein. The researchers also described a dense 50 kDa protein band that was present in both lethal and non-lethal retained fractions, however, close examination of the published SDS-PAGE gel results indicates that the two bands are different in molecular mass and, hence, unlikely to be the same protein. Following separation of the tentacle extracts by anion exchange chromatography, lethality was restricted to the unretained fraction (Othman and Burnett, 1990). Haemolytic activity was also detected in the unretained fractions and some of the retained fractions. However, in contrast to the cation exchange chromatography results, the 24 kDa protein band was observed in both retained and unretained fractions.

1.10.3 Immunoaffinity Chromatography

The first attempt to fractionate *C. fleckeri* venom by immunoaffinity chromatography involved the use of monoclonal mouse antibodies raised against the venom of the cnidarian *Physalia physalis* (Portuguese man-of-war) (Olson et al., 1984). Separation of lyophilised *C. fleckeri* milked venom by α-*Physalia* immunochromatography resulted in the partial purification of a lethal cardiotoxin capable of producing monovalent cation channels in lipid bilayers. Analysis of retained and unretained column fractions by non-reducing SDS-PAGE revealed that the specific fraction contained a single 20 kDa protein whereas the non-specific fraction contained two proteins (20 and 50 kDa). Following size exclusion chromatography of the concentrated specific fraction, a single protein peak corresponding to a native molecular mass of 65 kDa eluted from the column, suggesting the formation of a trimer of 20 kDa subunits.

In a subsequent study, an immunoaffinity column was prepared using CSL box jellyfish antivenom (polyclonal sheep antibodies raised against *C. fleckeri* milked venom) to fractionate *C. fleckeri* tentacle extracts (Calton and Burnett, 1986). Compared to the crude venom, the lethal and cardiotoxic activities of the specific fraction were significantly higher, but haemolytic and dermonecrotic activities were significantly lower. In contrast, the non-specific fraction was significantly less lethal, dermonecrotic, haemolytic and cardiotoxic than either the crude venom or specific fraction. Following SDS-PAGE analysis, two major proteins (50 and 150 kDa) were identified in the specific fraction, in contrast to the 20 kDa protein band obtained using anti-*Physalia*-toxin immunoaffinity chromatography (Olson et al., 1984).

In an extensive study utilising monoclonal mouse antibodies raised against *C. fleckeri* tentacle extracts, 13 immunoaffinity columns were prepared for the fractionation of tentacle venom (Naguib et al., 1988). Several proteins were detected in the specific fractions of the immunoaffinity columns by non-reducing SDS-PAGE (130, 120, 115, 90, 80, 75, 70, 52, 29 and 14.5 kDa). Lethal, haemolytic and dermonecrotic activities were assessed in each specific fraction. However, due to the heterogeneous SDS-PAGE protein profiles of most of the specific fractions, the activities of only the 120, 70 and 14.5 kDa proteins could be unambiguously determined. The 120 kDa protein was lethal, haemolytic and dermonecrotic, whereas the 70 kDa protein was lethal and haemolytic but not dermonecrotic. In contrast, the 14.5 kDa protein was lethal but neither haemolytic nor dermonecrotic.

In a repeat of the previous study, three immunoaffinity columns were prepared with the monoclonal mouse antibodies that specifically bound the 120, 70 and 14.5 kDa proteins and the columns were used to fractionate *C. fleckeri* tentacle extracts (Othman and Burnett, 1990). However, on this occasion, the protein fractionation patterns of the three columns were identical and consisted of a large non-specific protein peak and a smaller specific protein peak. Both the specific and non-specific fractions were lethal. Furthermore, in contrast to the previous study, SDS-PAGE analysis of the specific fractions revealed the presence of a major 24 kDa protein, which was absent from the unretained fractions, plus several other proteins (12–150 kDa).

1.10.4 Hydrophobic Interaction Chromatography (HIC)

The use of HIC has only been reported once for the fractionation of *C. fleckeri* tentacle extracts (Othman and Burnett, 1990). Although the majority of crude extract was unretained by the column, three minor protein peaks were observed during elution of the retained components. Only the unretained peak and the most strongly retained peak were lethal. SDS-PAGE analysis indicated that the lethal specific peak contained several proteins including a major 24 kDa protein, whereas the unretained lethal peak was dominated by one or more 45–50 kDa major protein(s).

1.10.5 Electrophoretic Separation Methods

A variety of electrophoretic separation methods have been used to characterise *C. fleckeri* venom proteins. In an interesting study, non-reducing SDS-PAGE was used to compare the protein composition of *C. fleckeri* milked venom with extracts obtained from lyophilised whole tentacles and nematocysts isolated from frozen tentacles (Wiltshire et al., 2000). Milked venom contained a major doublet of 17 kDa proteins and a minor 40 kDa protein. Whole tentacle extract contained a complex mixture of proteins with at least seven major protein bands, including proteins with similar molecular masses to those in the milked venom. The nematocyst venom also contained numerous proteins, including a dominant 40 kDa protein and at least two other major proteins (>106 kDa). Western blot analysis, using *C. fleckeri* box jellyfish antivenom, confirmed that the 17 kDa protein, common to the milked venom and tentacle extracts, and the 40 kDa protein present in all three extracts were detected by the antiserum. In freeze-thaw experiments, the group also observed a sequential decrease in antibody binding to the 40 kDa protein, associated with an increase in binding to the 17 kDa protein. These results

suggested that the larger protein was being degraded to the smaller protein as the freeze-thawing was repeated.

In earlier studies, researcher used flat bed isoelectric focussing to fractionate tentacle extracts (Othman and Burnett, 1990). Separation of two lethal fractions with pI values ranging from 6.00–6.85 and 7.55–8.50 was achieved. The protein profiles of the two fractions varied, with the more acidic fraction containing a 45 kDa protein and several proteins less than 30 kDa and the more basic fraction containing a predominant 24 kDa band. Using an alternative isoelectric focussing method, the group isolated a lethal factor with a native molecular mass of 160 kDa, which under non-reducing SDS-PAGE analysis conditions yielded a 67 kDa protein. However, when the experiment was repeated, SDS-PAGE results varied and two major proteins (30 and 12 kDa) were detected instead of a single 67 kDa protein (Othman and Burnett, 1990).

In one of the most recently reported fractionation studies, *C. fleckeri* nematocyst venom was separated into four fractions using native capillary electrophoresis (Bloom et al., 2001). Although the majority of venom proteins eluted in the first three fractions, haemolytic and lethal activity was confined to the fourth fraction. SDS-capillary gel electrophoresis (SDS-CGE) analysis revealed that fraction 1 contained a single protein (201 kDa), fraction 2 contained multiple proteins (173, 163, 112 and 71 kDa), fraction 3 contained two proteins (39 and 14 kDa) and proteins contained in fraction 4 were all less than 14 kDa. In most cases, the molecular masses of the proteins were comparable to a previous experiment in which reducing SDS-CGE analysis of unfractionated *C. fleckeri* nematocyst venom detected several proteins (201, 174, 112, 71, 39 and 31 kDa) (Bloom et al., 1998).

In another study, three monoclonal mouse antibodies were produced against whole *C. fleckeri* tentacle extracts (Collins et al., 1993). Each antibody had demonstrated anti-haemolytic activity, but no anti-lethal activity. The antibodies were subsequently used to characterise whole tentacle extract proteins separated by SDS-PAGE (non-reducing and reducing) and native PAGE. Western blot analyses revealed that two of the antibodies bound specifically to a 50 kDa suspected haemolysin under non-reducing SDS-PAGE conditions, but no binding of the third antibody was observed. Under reducing conditions, none of the antibodies bound to any of the tentacle proteins, suggesting that the epitope of the antigen may have been destroyed by the reduction of disulphide bonds. Unusually, no binding of any of the antibodies was observed in native PAGE of the tentacle venom, and the group suggested that in the absence of SDS, the 50 kDa protein may have formed aggregates unable to enter the polyacrylamide gel.

1.11 Characterisation of Toxic Proteins from Other Jellyfish

A broad range of bioactive toxins have also been isolated from the venoms of other potentially harmful jellyfish. For example, researchers isolated a 178 kDa haemolysin, two α-chymotrypsin-like serine proteases (40 and 85 kDa) and a PLA2 (14–16 kDa) from the nematocyst venom of the Mediterranean scyphozoan jellyfish, *Rhopilema nomadica* (Gusmani et al., 1997). In a previous study, the N-terminal amino acid sequence of a 17 kDa PLA2 isolated from *R. nomadica* tentacles closely resembled PLA2 toxins isolated from the venoms of a lizard and a honeybee (Lotan et al., 1996). Hence, the researchers suggested that the occurrence of PLA2-like toxins in jellyfish may explain the local and systemic effects observed in envenomed humans.

In addition to research on *C. fleckeri*, a number of studies have investigated the bioactive proteins of other cubozoan jellyfish. Indeed, multiple bioactive proteins have been isolated from the carybdeid, *Carybdea marsupialis*. One research group isolated a single nematocyst-derived cytolysin from an Adriatic specimen of *C. marsupialis*, CARTOX (107 kDa), that was haemolytic to sheep erythrocytes (not human or rabbit erythrocytes), lacked phospholipase C activity and acted as a pore-forming protein (Rottini et al., 1995). In another study, a lethal neurotoxin (120 kDa) and three cytolysins (220, 139 and 36 kDa) were isolated from the nematocyst venom of *C. marsupialis* collected from the Caribbean Sea near Mexico (Sánchez-Rodríguez et al., 2006). The neurotoxin caused paralysis and death in crabs and the three cytolysins were haemolytic to human erythrocytes.

In studies on *C. rastoni*, partially purified venom (pCrTX) and three highly toxic proteins (CrTX-I, CrTX-II and CrTX-III), isolated from tentacle extracts, were lethal and haemolytic, induced platelet aggregation (by increasing indiscriminate cation permeability), acted as calcium-dependent vasoconstrictors and damaged the uptake/storage mechanisms of noradrenaline (Azuma et al., 1986a, b, c). The molecular masses of CrTX-I, -II and -III, determined by SDS-PAGE (49, 100 and 51 kDa, respectively), were unchanged under reducing and non-reducing conditions (Azuma et al., 1986c), suggesting that the proteins do not form intermolecular disulphide bonds. Also, the CrTX proteins were unstained by dansylhydrazine, indicating that the proteins are not glycosylated. In another study, haemolytic and myotoxic factors were also partially purified from tentacle extracts of *C. rastoni* (Othman et al., 1996), however, few details of the study have been published.

In more recent studies, researchers isolated and characterised several major venom proteins from three box jellyfish species including, CrTX-A and -B (43 and 46 kDa) from *C. rastoni*, CaTX-A and -B (43 and 45 kDa) from *C. alata*, and CqTX-A (44 kDa) from *C. quadrigatus* (Nagai et al., 2000a, b, 2002). Subsequent assays revealed that purified CrTX-A, CrTX-B, CaTX-A and CqTX-A were potently haemolytic, CrTX-A, CaTX-A and CqTX-A were lethal to crayfish, CrTX-A and -B were lethal to mice and CrTX-A caused inflammation and necrosis of the skin (Nagai et al., 2002, 2000a, b).

Immunolocalisation studies of the *C. rastonii* toxins (CrTX-A and -B) indicated that CrTX-A is mostly contained within the nematocysts, whereas CrTX-B is confined to the tentacle tissue. Hence, researchers proposed that CrTX-B is synthesised within the tentacle, modified to CrTX-A, then transported to the nematocyst (Nagai et al., 2000a). Similarly, during the purification of CaTX-A and -B from *C. alata* tentacles, CaTX-A (not CaTX-B) was confined to the nematocysts (Nagai et al., 2000b). In a parallel study, a basic 42 kDa haemolytic protein (CAH1) was also isolated from the nematocyst venom of *C. alata* venom (Chung et al., 2001). The partial N-terminal amino acid sequence of CAH1 was very similar (if not identical) to the deduced sequence reported for CaTX-A (Nagai et al., 2000b), suggesting that CAH1 and CaTX-A are the same protein.

CqTX-A was the most abundant protein detected in *C. quadrigatus* nematocyst venom, thus implicating it as the causative toxin in fatal *C. quadrigatus* stings (Nagai, 2003). Although *C. quadrigatus* is considered to be more dangerous to humans than *C. rastonii* and *C. alata*, a comparison of the lethal and haemolytic activities of purified CrTX-A, CaTX-A and CqTX-A revealed that CqTX-A is less potent than CrTX-A and CaTX-A (Nagai et al., 2002). However, this apparent contradiction may be explained by the fact that *C. quadrigatus* has tentacles greater in number and length than those of *C. rastonii* and *C. alata.* Hence, *C. quadrigatus* has the potential to inject much larger doses of toxin than the smaller, four-tentacled carybdeids (Nagai, 2003).

The researchers also successfully cloned and sequenced cDNA encoding CrTX-A/-B, CaTX-A and CqTX-A using a combination of internal peptide sequencing, degenerative RT-PCR and 5'/3'-RACE. Secondary structure analysis of the amino acid sequences predicted the presence of an N-terminal amphiphilic α-helix in each protein, which could be involved in the mechanism of cytolytic activity (Nagai et al., 2002, 2000a, b).

Very little is currently known about the bioactive proteins of the potentially life-threatening carybdeid *C. barnesi*, however, SDS-PAGE protein profiles of tentacle extract included four

major proteins (40, 45, 80 and 106 kDa) and several minor proteins (Wiltshire et al., 2000). Notably, the 106 kDa protein is similar in molecular mass to the cytolysin, CARTOX, from *C. marsupialis* (Rottini et al., 1995) and CrTX-II from *C. rastoni* (Azuma et al., 1986c). In comparison, *C. barnesi* bell extracts contained only three major proteins (50, 80 and >106 kDa) (Wiltshire et al., 2000). In the bell extract of one *C. barnesi* specimen, a 40 kDa protein was cross-reactive to CSL box jellyfish antivenom raised in sheep against *C. fleckeri* milked venom.

Haemolytic activity is a common property of venoms from several species of box jellyfish including *C. alata* (Chung et al., 2001), *C. marsupialis* (Rottini et al., 1995), *C. bronzie*[§], *C. xaymacana*, as well as *C. fleckeri* (Bailey et al., 2005). Haemolytic activity has also been detected in purified toxins from the box jellyfish *C. rastoni* (CrTX-A and -B), *C. alata* (CaTX-A and -B) and *C. quadrigatus* (CqTX-A) (Nagai et al., 2002, 2000a, b). Some cubozoan haemolysins appear to exhibit variable affinities for blood cells from different host species. For example, sheep erythrocytes were significantly more susceptible to lysis by the *C. marsupialis* haemolysin, CARTOX, than rabbit and human erythrocytes (Rottini et al., 1995), and guinea pig erythrocytes were more resistant to lysis by a 70 kDa haemolysin isolated from *C. fleckeri* tentacle extracts than human, mouse and rat erythrocytes (Keen and Crone, 1969a). In contrast, no significant difference in haemolytic activity was observed for the unfractionated nematocyst venoms of *C. xaymacana*, *C. bronzie* and *C. fleckeri* against sheep and human erythrocytes (Bailey et al., 2005).

Biochemical studies have indicated that cubozoan haemolysins are stabilised in high saline conditions and are dependent on divalent cations (Rottini et al., 1995; Nagai et al., 2000a; Chung et al., 2001). For example, the haemolytic activity of *C. alata* venom is dependent on the presence of divalent cations such as Mg^{2+}, Ca^{2+} or Zn^{2+} and optimal at a concentration of 10mM Ca^{2+} or Mg^{2+} (Chung et al., 2001). Likewise, purified CARTOX from *C. marsupialis* is dependent on the presence of Ca^{2+}, but optimal at lower concentrations (0.1-1mM Ca^{2+}) (Rottini et al., 1995).

The haemolytic activity of *C. alata* venom is also inhibited in the presence of some carbohydrates, particularly D-lactulose (Chung et al., 2001). Similar haemolytic inactivation of purified cnidarian cytolysins by various carbohydrates and lipids has been observed in other jellyfish including the scyphozoan, *R. nomadica* (Gusmani et al., 1997) and the box jellyfish, *C. marsupialis* (Rottini et al., 1995). In the case of *C. marsupialis*, the activity of purified

[§] Reported as *Chiropsalmus sp.* in Bailey et al. (2005); reclassified as *Chiropsella bronzie* (Gershwin, 2006).

CARTOX was inhibited by specific lipids (sphingomyelin and phosphatidylinositol) and carbohydrates (sialic acid and β-methylgalactopyranoside). Furthermore, erythrocytes pre-treated with particular glycosidases (neuraminidase and β-galactosidase) were significantly less susceptible to lysis by CARTOX, suggesting that carbohydrate moieties of cell surface proteins are involved in the cytolytic mechanism of action. A co-operative mechanism of cytolytic activity has therefore been proposed, which involves binding of the jellyfish toxins to specific cell surface carbohydrates and lipids (Rottini et al., 1995; Gusmani et al., 1997; Chung et al., 2001).

1.12 Clinical Effects of *C. fleckeri* Envenomation

1.12.1 Symptoms

The clinical manifestations of major or life-threatening human envenomations include excruciating pain, impaired consciousness, dyspnoea (difficulty breathing), cardiac dysfunction, pulmonary oedema, shock, hypertension followed by hypotension, and a rapid, acute inflammatory response of the skin (Lumley et al., 1988; Beadnell et al., 1992; Williamson et al., 1984). Respiratory and cardiac failure can occur within several minutes (Lumley et al., 1988).

When *C. fleckeri* tentacles contact the skin of a victim, millions of nematocysts are discharged into the skin. The tubules of the discharged nematocysts penetrate through the epidermis into the underlying vascular and nerve-rich dermis, progressively releasing venom into the surrounding tissue (Burnett, 1991). Venom is easily absorbed into lymphatic and capillary vessels, however the rate of uptake is dependent on the total dose of venom received and peripheral circulation, which is dictated primarily by muscular action of the surrounding tissues (Burnett, 1991). Children are more susceptible to serious envenomations than adults due to their greater surface area to mass ratio and their relative lack of body hair, which allows increased contact between the tentacles and the skin (Fenner and Harrison, 2000).

At the site of envenomation, symptoms include severe localised pain, oedema and the formation of raised, linear wheals, with white ischaemic centres surrounded by vivid erythematous flares (Burnett, 1991; Fenner, 1991; Burnett et al., 1996b). In humans, it has been suggested that the acute inflammatory response evoked by *C. fleckeri* venom may be caused by the release of compounds such as histamine, serotonin, bradykinin and perhaps prostaglandins (Williamson et al., 1984). However, no evidence of histamine, serotonin or vaso-active kinin liberation in the skin has previously been found in experimental animals (Keen and Crone,

1969b), so the cause of the extreme pain experienced by victims remains unclear. Alternatively, it has been suggested that the pain may be due to direct tissue damage or the release of potassium from damaged cells (Freeman and Turner, 1969). Within 24 hours the skin blisters and often blackens, and can lead to permanent scarring (Fenner, 1991; Baxter and Marr, 1975), particularly if secondary bacterial infection occurs (Burnett et al., 1996b). Delayed cutaneous hypersensitivity reactions are common although usually minor in *C. fleckeri* envenomations (Bailey et al., 2003).

1.12.2 Treatment

An envenomed victim should be removed from the water, reassured and ideally, placed at rest to limit enhanced circulation of the venom by muscular action (Fenner et al., 1993; Burnett, 1991). However, due to the excruciating pain usually felt by the victim, limitation of movement is usually difficult, especially in children. Urgent medical assistance should be obtained as soon as possible (Fenner et al., 1993). The victim's airway, breathing and heartbeat should be monitored and expired air resuscitation or cardiopulmonary resuscitation must be provided immediately if required (Fenner, 1991). If available, oxygen should be supplied (Currie, 1994).

Weak solutions of acetic acid and commercial vinegar are known to rapidly and irreversibly inhibit nematocyst firing (Hartwick et al., 1980) and liberal application of vinegar to the site of envenomation is still recommended during the provision of basic first aid. Once inactivated, adherent tentacles can be carefully and promptly removed (Fenner and Williamson, 1996). The use of compression bandages, doused in vinegar if possible, has been recommended for serious envenomations (Lumley et al., 1988). However, the efficacy of pressure immobilisation bandaging has yet to be clinically proven (Beadnell et al., 1992) and at worst, such treatment could be potentially dangerous (Bailey et al., 2003). The technique is also likely to be time consuming and impractical (Currie, 1994). Minor stings can be treated by application of ice or cold packs (Fenner et al., 1993). Microscopic examination of the nematocysts recovered from the victim's skin by scalpel-scraping or sticky tape sampling can aid in the identification of the envenoming jellyfish and thus may assist with patient management (O'Reilly et al., 2001; Currie and Wood, 1995).

In severe cases of human envenomation, administration of the commercially available CSL box jellyfish antivenom is recommended (Tibballs, 2006). The impressive analgesic effect of the antivenom following serious human envenomations has been noted by clinicians (Williamson et al., 1984; Boyd, 1984), however, additional pain relief using narcotics such as

morphine or pethidine is often required (Beadnell et al., 1992; Currie, 1994). CSL box jellyfish antivenom, which is prepared from the sera of sheep that have been hyperimmunised with *C. fleckeri* milked venom (Barnes, 1967), also significantly reduces the acute inflammatory response of the skin at the site of envenoming, however it is not known whether its effect lessens long-term scarring (Williamson et al., 1984). Minor side effects of the antivenom include cutaneous rashes, but the most severe effect of antivenom infusion has resembled an anaphylaxis reaction (Burnett and Calton, 1987). Therefore, administration of the antivenom is indicated only in significant envenomations, such as those associated with cardiorespiratory instability or severe pain, and its use restricted to trained personnel (Bailey et al., 2003; Williamson, 1985).

Although the administration of CSL box jellyfish antivenom is recommended in extreme cases, opinions vary as to the efficacy of the antivenom. Some studies suggest that the antivenom is only partially effective against the lethal effects of *C. fleckeri* venom. In prophylactic and rescue experiments undertaken on mice injected with partially purified *C. fleckeri* nematocyst toxins, CSL box jellyfish antivenom was more effective against a 150 kDa myotoxin than a 600 kDa myotoxin, suggesting that the 600 kDa component may not be present in milked venom, against which the antivenom is prepared (Endean and Sizemore, 1988).

More recently, researchers found that CSL box jellyfish antivenom was unable to prevent myotoxicity in chick cervicis nerve-muscle preparations when administered either prior to, or following exposure to *C. fleckeri* nematocyst venom (Ramasamy et al., 2003). Although prophylactic addition of antivenom was able to neutralise the neurotoxic effects of the *C. fleckeri* venom, it was unable to counter the effects if administered one hour after exposure to the venom. These results suggest that CSL box jellyfish antivenom is unlikely to be useful in the treatment of patients suffering myotoxic or neurotoxic effects unless administered soon after envenomations (Ramasamy et al., 2003). Subsequent *in vivo* experiments on rats showed that following prophylactic administration of antivenom, only 40% of rats survived challenge with *C. fleckeri* nematocyst venom, casting further doubt on the efficacy of the antivenom (Ramasamay et al., 2004).

In addition to administration of the CSL box jellyfish antivenom, some researchers have advocated the use of the Ca^{2+} antagonist, verapamil, in the treatment of seriously envenomed patients (Burnett and Calton, 1983; Burnett at al., 1990). However, its potential use remains controversial (Tibballs, 2006). Initial studies indicated that verapamil was able to correct

arrhythmias elicited by *C. fleckeri* tentacle extracts in mice (Burnett and Calton, 1983). Furthermore, during rescue experiments in mice, co-administration of antivenom and verapamil was found to be more effective in countering the actions of *C. fleckeri* venom than administration of either therapy on its own, suggesting the possible benefits of combination therapy in serious human envenomations (Burnett et al., 1990; Bloom et al., 1999). However in other studies, verapamil was ineffective against the lethal effects of *C. fleckeri* nematocyst venom in mice (Endean and Sizemore, 1987) and was only effective against one of two myotoxins partially purified from nematocyst venom (Endean, 1988b). Verapamil was also ineffective as a prophylactic agent against the fatal cardiovascular effects of *C. fleckeri* tentacle extracts in artificially ventilated piglets and caused increased morbidity and mortality (Tibballs et al., 1998). Similarly, verapamil was ineffective in preventing the fatal cardiovascular effects of *C. fleckeri* nematocyst venom in rats and potentially worsened the outcome (Ramasamy et al., 2004). From a clinical perspective, concerns have also been expressed regarding treatment of unstable patients with verapamil due to its ability to increase hypotension, induce cardiac dysrhythmias and thus potentially compromise resuscitation (Currie, 1994; Bailey et al., 2003; Tibballs, 2006). However, despite contradictory research findings and clinical concerns, researchers continue to argue the case for verapamil use in the treatment of serious jellyfish stings (Burnett, 1998; Burnett and Calton, 2004).

More recently, the use of magnesium sulphate ($MgSO_4$) has been suggested as an alternative adjunct therapy to the administration of CSL box jellyfish antivenom in the treatment of life-threatening envenomations (Ramasamy et al., 2004). Prophylactic administration of $MgSO_4$ alone was unable to prevent cardiovascular collapse in rats injected intravenously with *C. fleckeri* nematocyst venom. In comparison, prophylactic administration of CSL box jellyfish antivenom alone prevented venom-induced cardiovascular collapse in 40% of the animals. However, following pre-treatment of the rats with CSL box jellyfish antivenom and $MgSO_4$, cardiovascular collapse was prevented in all animals, thus supporting the use of $MgSO_4$ rather than verapamil as an adjunct therapy in serious *C. fleckeri* envenomations (Ramasamy et al., 2004).

Dermonecrosis is a common acute complication of *C. fleckeri* envenomations and current treatment is the same as that for a burn, with emphasis on the reduction of secondary bacterial infection (Bailey et al., 2003). In the occurrence of delayed cutaneous hypersensitivity, corticosteroid cream and oral antihistamines are the preferred method of treatment (Bailey et al., 2003). An old bush remedy for the treatment of box jellyfish stings, which has been used at

Bingil Bay (northern Queensland), involves rubbing crushed *Crinium pedunculatum* (Crinium Lily/Spider Lily) on the affected area (Cribb and Cribb, 1981). The relief of symptoms, although not complete, is thought to be due to the presence of the plant alkaloid, lycorine.

A topical cream is also currently available in several countries for the prevention of jellyfish stings. Safe Sea SPF15 (Nidaria Technology) was formulated based on the chemical properties of the mucous coating of the clownfish, *Amphiprion* sp., for the purpose of inhibiting nematocyst discharge during human envenomations (Boulware, 2006). Results from small-scale laboratory and field trials on humans indicate that although the cream does not prevent all jellyfish stings, the severity and frequency of stings from cubozoan and scyphozoan jellyfish such as *Chiropsalmus quadramanus* and *Chrysaora quinquecirrha* are significantly reduced by use of the prophylactic cream (Kimball et al., 2004; Boulware, 2006). However, the efficacy of Sea Safe SPF15 has not been clinically tested in the prevention of *C. fleckeri* stings.

Finally, due to the apparent pain-reducing effects of the TRPV1 receptor antagonist BCTC against jellyfish venom (Cuypers et al., 2006), the use of TRPV1 receptor antagonists has recently been patented with the aim to commercialise production of analgesics in the treatment and/or prophylaxis of cnidarian stings (Publication No. WO/2007/140551, http://www.wipo.int/pctdb/en/).

1.13 Project Aims and Research Outcomes

The venom of *C. fleckeri* is a complex mixture of proteins which is known to be toxic to heart and nerve tissue, destroy red blood cells, permanently damage skin cells and can cause death. Although significant research effort spanning four decades has been directed towards the biochemical and pharmacological characterisation of *C. fleckeri* venom, there has been limited success in achieving consistency in results between research groups, and the bioactive venom components and their mechanisms of action remain poorly elucidated.

To date, no amino acid sequence of the bioactive proteins contained in *C. fleckeri* venom or DNA sequence of the genes that encode them has been reported. However, knowledge of their amino acid sequences can assist in predicting the structure of the venom proteins, their inferred functions and potential mechanisms of action. Comparison of sequences with homologues found in other jellyfish or more primitive cnidarians can also provide information about the evolution of specific families of toxins and how the homologues relate to one another in sequence, structure and function. In addition, cloning the DNA sequence that encodes a particular venom protein provides an opportunity to synthesise the protein *in vitro* using

recombinant DNA technology. If synthesis of active recombinant venom proteins is achieved, the proteins can be used for many purposes to further elucidate their biochemical properties and mechanisms of action, including functional assays, bioassays and structural elucidation techniques such as NMR, circular dichroism and/or X-ray crystallography. The ability to produce pure, active recombinant proteins as required may also overcome sample shortages due to the seasonal variation and unpredictable occurrence of adult jellyfish populations for sampling – a major problem in current box jellyfish research.

Although an antivenom is currently available for the treatment of severe *C. fleckeri* stings in humans, its efficacy has been questioned with research indicating that the antivenom is only partially effective against components of *C. fleckeri* venom. This highlights the need for more comprehensive characterisation of the properties and actions of key venom proteins, which may assist in the improvement of contemporary treatments for *C. fleckeri* stings or the development of prophylaxes. Furthermore, individual venom components may be found to elicit some pharmacological or biochemical benefit to humans.

The purpose of this research project was to characterise the venom proteins contained in the nematocysts of *C. fleckeri* using biochemical and molecular biology techniques to achieve the following research outcomes:

- Modification and adaptation of current extraction and purification methods for the stable isolation of major bioactive proteins in *C. fleckeri* venom
- Characterisation of major venom proteins including, amino acid sequencing, cDNA sequencing and molecular mass analysis
- Investigation of potential sequence homology between *C. fleckeri* venom proteins and proteins from other organisms and prediction of the evolutionary relationships between homologues
- Identification of secondary structures and/or motifs to assist in 3-D structure prediction of key proteins and proposal of their possible mechanisms of action
- *in vitro* expression of one or more of the major venom proteins using recombinant DNA technology

The methodologies undertaken during this research project are summarised in Figure 1.12. Two research streams involving biochemistry and molecular biology were followed. Biochemical analyses were utilised to separate nematocyst-derived venom proteins, characterise their physical and chemical properties and to obtain amino acid sequence of selected proteins.

Molecular biology techniques including cDNA library screening, DNA cloning and sequencing were used to isolate venom clones from a *C. fleckeri* tentacle cDNA expression library or total tentacle RNA, deduce the complete amino acid sequence of major venom proteins and attempt recombinant expression of the venom proteins in a bacterial system. Functional assays and lethality tests were undertaken to monitor venom protein activity during extraction, fractionation or purification procedures.

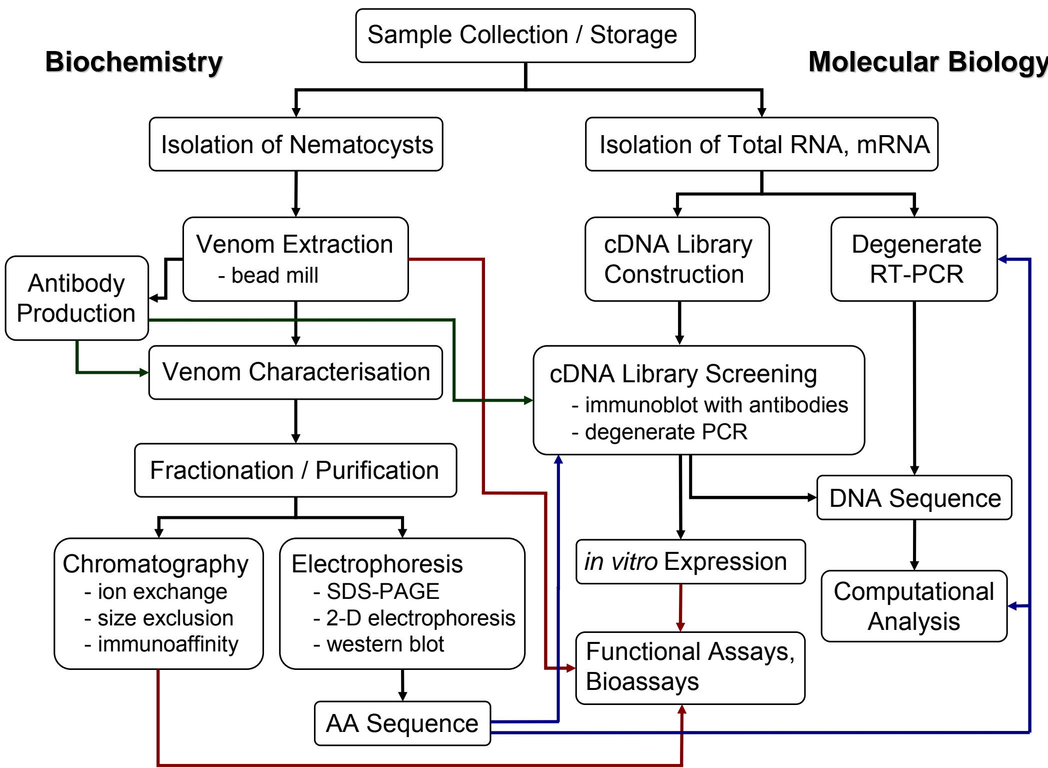


Figure 1.12 Flowchart of biochemical and molecular biology methods used to characterise the proteins contained in *C. fleckeri* venom.

CHAPTER 2

General Methods and Materials

2.1 Sample Collection and Storage

Jellyfish samples were collected from coastal waters near Cairns, Townsville and Weipa in northern Queensland, Australia, and kindly provided by Dr Jamie Seymour and Ms Griselda Ávila-Soria (James Cook University), Reef HQ (Great Barrier Reef Marine Park Authority) and Queensland Surf Life Saving Association.

For protein-related studies, nematocysts were isolated from excised *C. fleckeri* tentacles using an autolytic method previously described (Bloom et al., 1998). Some batches of nematocysts were subsequently lyophilised then stored at -20°C by Seymour and co-workers, however, most batches were received as non-lyophilised nematocysts and stored in filtered seawater or 35 g/L NaCl at 4°C. Non-lyophilised nematocysts were purified from tentacle debris by centrifugation (5.1K×*g*, 20°C, 30 min) in a discontinuous gradient of Percoll (Sigma) diluted with 35 g/L NaCl (bottom layer: 88% Percoll; top layer: 27% Percoll); Percoll was removed by washing with 35 g/L NaCl. Bell and tentacle tissue required for nucleic acid isolation was either stored at –80°C or immersed in RNAlater (Ambion) and stored at –20°C.

2.2 Preparation of *C. fleckeri* Venom

2.2.1 Bead Mill Homogenisation

C. fleckeri venom was prepared from lyophilised (0.01–0.04g dry weight) or non-lyophilised nematocysts (0.2–0.5g wet weight) by bead mill homogenisation, similar to a previously described method (Carrette and Seymour, 2004). Unless otherwise specified, the nematocyst extraction buffer (NEB) contained 25mM MOPS (pH 7.4), 100mM NaCl and protease inhibitors (100µM EDTA, 1µM leupeptin, 1µM pepstatin and 10µM PMSF). NEB-washed nematocysts were resuspended in ice-cold buffer (0.5–1mL) and ruptured using a Biospec Products Mini-BeadBeater bead mill and 0.5 mm glass beads. Nematocyst disruption was monitored microscopically; > 90% rupture of nematocysts was achieved by 4–6 cycles of homogenisation (10s burst at 5000 rpm and chilling on ice for 1 min). Crude nematocyst-derived venom (nematocyst venom) was clarified by centrifugation (20K×*g*, 4°C, 20 min).

2.2.2 Lyophilised Milked Venom

Lyophilised milked venom, donated by Jamie Seymour (JCU), was prepared using a method similar to Barnes (1967). Aliquots of milked venom were lyophilised and stored at -20°C.

2.2.3 Chemical Induction of Nematocyst Discharge

In one experiment (Chapter 4), SDS sample buffer was used to induce the discharge of venom from *C. fleckeri* nematocysts. NEB-washed nematocysts (~25µg) were incubated (1h, RT) in reducing 2X SDS sample buffer (100µL). Induction of nematocyst discharge was monitored microscopically. Following centrifugation (18K×*g*, 2 min, RT), the supernatant was transferred to another tube and the pellet, containing discharged nematocyst capsules, was resuspended in 2X SDS sample buffer (100µL).

2.3 General Protein Methods

2.3.1 Protein Concentration

Total protein concentration of *C. fleckeri* venom and chromatography fractions was estimated by measuring absorbance at 280nm (Beckman DU650, USA) or using the Bradford Assay (Bradford, 1976). Bovine serum albumin (BSA) was used as a protein standard. For the Bradford Assay, Bio-Rad Protein Assay Dye Reagent Concentrate (20µL) was added to triplicate protein samples diluted to 80µL in NEB. The solutions were mixed, incubated (15–30 min, RT) and absorbances measured at 595nm. A standard protein calibration curve was established using triplicate concentrations of BSA diluted in NEB (0.0, 2.5, 5.0, 7.5 and 10.0µg/mL).

2.3.2 Polyclonal Antibodies against *C. fleckeri* Venom

C. fleckeri antibodies were purchased (CSL Ltd, Australia), raised in mice (James Cook University) and raised in rabbits (IMVS Veterinary Services, Australia). Polyclonal antibodies were raised in 2 mice with four fortnightly injections of filtered (0.22µm Millex-GP filters, Millipore) and denatured (95ºC, 5 min) nematocyst extracts (~30µg protein for initial injection, ~50µg protein for subsequent boosters). Polyclonal antibodies were raised in a rabbit with three fortnightly injections of filtered and denatured nematocyst extracts (~50, 75 and 100µg protein,

successively). Polyclonal antibodies specific for CfTX-1 and CfTX-2 were raised in a rabbit with three fortnightly injections of combined and emulsified 43 and 45 kDa protein gel fragments (~50, 80 and 100µg total protein) obtained by reducing SDS-PAGE separation of nematocyst extracts. The CfTX-specific antiserum (2mg/mL) was purified using Sepharose protein A chromatography.

2.3.3 SDS-PAGE Analysis

SDS-PAGE analysis of proteins (Laemmli, 1970) was conducted on nematocyst extracts and chromatography fractions using a Bio-Rad Mini-PROTEAN II or PROTEAN IIxi electrophoresis system. Protein samples were diluted in 2X SDS sample buffer (62.5mM Tris-HCl pH 6.8, 10% glycerol, 2% SDS, 0.02% bromophenol blue, with or without 200mM DTT) and boiled (5 min). In some cases, 5X SDS sample buffer was used (component concentrations adjusted accordingly). Proteins were separated using 7.5–15% polyacrylamide gels in SDS-PAGE running buffer (25mM Tris, 192mM glycine, 0.1% SDS) and protein bands were visualised by Coomassie R-250 or silver staining. The molecular masses of the proteins were estimated using standard protein molecular mass markers (Sigma/Fermentas).

2.3.3.1 Coomassie Stain

Proteins in SDS-PAGE gels were stained with Coomassie R-250 according to Sambrook and Russell (2001). Gels were incubated (30 min–24h) in Coomassie stain (0.125% (w/v) Coomassie brilliant blue R-250, 50% methanol, 10% glacial acetic acid) then incubated several hours in destain solution (10% methanol, 10% glacial acetic acid).

2.3.3.2 Silver Stain

Proteins in SDS-PAGE gels were silver stained using a modified protocol (PlusOne Silver Staining Kit, Amersham Biosciences). Gels were soaked (30 min–24h) in fixative (40% ethanol, 10% glacial acetic acid), then incubated (30–60 min) in freshly prepared sensitising solution (30% ethanol, 0.5M sodium acetate, 8mM sodium thiosulphate, 0.5% glutaraldehyde). Gels were washed (3 × 5 min) in distilled water, incubated (40 min) in freshly prepared silver reagent (0.1% silver nitrate, 0.007% formaldehyde) and washed again in distilled water. Gels were immersed in freshly prepared developing solution (0.25M sodium carbonate, 0.004%

formaldehyde) until protein bands were visualised, incubated (2 × 10 min) in stop solution (40mM EDTA) and washed (3 × 5 min) in distilled water.

2.3.3.3 Concentration of Protein Samples by TCA Precipitation

Samples with low protein content were concentrated prior to SDS-PAGE analysis using trichloroacetic acid (TCA). Diluted TCA (30% in distilled water) was added to samples (10% TCA final concentration). The samples were incubated on ice (30 min) and centrifuged (16K×*g*, RT, 15 min). The resulting pellets were washed twice with cold acetone (1mL), recovered by centrifugation (16K×*g*, RT, 5 min) and air dried. Pellets were resuspended in 2X or 5X SDS reducing sample buffer and boiled (7–10 min) prior to SDS-PAGE analysis.

2.3.4 Western Blot Analysis

For western blot analysis, nematocyst proteins in unstained SDS-PAGE gels were transferred (30–40 min, 120V) to Immobilon-P or Hybond-P polyvinylidene fluoride (PVDF) membranes (Millipore or Amersham, respectively) in transfer buffer (25mM Tris, 192mM glycine, 0.1% SDS, 20% methanol) using a Bio-Rad Mini Trans-Blot electrophoretic transfer system. Membranes were incubated in blocking solution (5% skim milk powder in TBST, 1h), then incubated overnight in various dilutions of primary antibody in TBST (1:1000–1:3000). Membranes were washed (3 × 5 min in TBST) and incubated (1h) in an appropriate secondary antibody, conjugated to either alkaline phosphatase (AP) or horse radish peroxidase (HRP), diluted in TBST (1:5000–1:10,000).

Following membrane washing, antigenic proteins were visualised using NBT/BCIP (Promega) or *FAST* 3,3'-Diaminobenzidine (DAB) Tablets (Sigma). For AP-conjugated antibodies, membranes were immersed in alkaline phosphatase buffer (100mM Tris-HCl pH 9, 150mM NaCl, 1mM $MgCl_2$) containing NBT (6.6µL/mL) and BCIP (3.3µL/mL). For HRP-conjugated antibodies, membranes were immersed in a solution of DAB tablets dissolved in distilled water (1 set / 5mL). When protein visualisation was complete, colour development for both conjugate systems was stopped by rinsing the membranes in tap water and the membranes air-dried.

2.3.5 Two Dimensional Gel Electrophoresis (2DGE)

Proteins from non-lyophilised nematocyst extracts were analysed by 2DGE. Extract proteins (15–75µg) were solubilised and reduced (1h, RT) in rehydration buffer (8M urea, 2% CHAPS, 0.002% bromophenol blue, 2% IPG buffer pH 3–10 (Amersham BioSciences), 60mM DTT final concentrations). Following protein alkylation (1h, RT) by addition of acrylamide (120mM final concentration), samples (120µL) were applied to Immobiline DryStrip gels (7cm, pH 3–10, Amersham Biosciences), overlaid with mineral oil, and incubated (overnight, RT). Isoelectric focusing of rehydrated IPG gel strips was performed using a Bio-Rad PROTEAN IEF Cell and the parameters detailed in Table 2.1. Gel strips were incubated (15 min) in equilibration buffer (6M urea, 2% SDS, 0.375M Tris-HCl pH 8.3, 20% glycerol, 0.002% bromophenol blue), rinsed briefly in SDS-PAGE running buffer and embedded atop 15% SDS-PAGE gels with melted agarose sealing solution (0.5% agarose and 0.002% bromophenol blue in SDS-PAGE running buffer). Molecular markers (3–5µL) and nematocyst extracts (5–10µL), for one-dimensional separation, were applied to small pieces of IEF electrode wicks (Bio-Rad) and embedded adjacent to the immobilised IPG strips with agarose sealing solution. Following SDS-PAGE, the nematocyst proteins were silver-stained or subjected to western blot analysis using CSL antivenom.

Table 2.1

Isoelectric focusing programme for the separation of *C. fleckeri* nematocyst proteins using Immobiline DryStrips (7cm, pH 3-10) and the PROTEAN IEF Cell. Current was limited to 50µA/gel and the focus temperature was 20°C.

Step	Description	Voltage (V)	Voltage Slope	Duration (h)	Volt Hours (kVh)
1	conditioning	150	rapid	0.25-1	-
2	voltage ramping	3500	linear	variable	2.8
3	final focussing	3500	constant	variable	3-3.3
4	hold	500	rapid	16-24	-

2.3.6 N-Terminal Amino Acid Sequencing

N-Terminal amino acid sequences were determined for two major nematocyst proteins. Nematocyst extracts were separated by SDS-PAGE and transferred to PVDF, as described previously (Sections 2.3.3 and 2.3.4). BSA (0.5, 1.25, 2.5 and 5µg) was applied in lanes

adjacent to the extracts to estimate the amount of protein in the separated bands. Membranes were Coomassie stained (0.1% ($^{w}/_{v}$) Coomassie brilliant blue R-250, 50% methanol, 7% glacial acetic acid, 2 min), then destained (50% methanol, 7% glacial acetic acid). Protein bands (43 and 45 kDa) were excised from the membranes, air dried, and 10–50pmol of each protein was analysed by Edman degradation N-terminal amino acid sequencing (Biomolecular Research Facility, University of Newcastle, Australia). Resulting sequences were subjected to a search for short, nearly exact protein sequences using the basic local alignment search tool (BLAST; Altschul et al., 1997).

2.3.7 Peptide Mass Fingerprinting

The internal amino acid sequence was determined for a major 43 kDa nematocyst protein. Nematocyst extracts were filtered (0.22μm Millex-GP filters, Millipore), reduced with DTT (0.1M final concentration, 0.5 h, RT), then alkylated with a four-fold molar excess of acrylamide (0.5 h, RT) prior to boiling with 2X SDS-PAGE sample buffer. Care was taken to avoid contamination of the samples with keratin. Nematocyst proteins were separated by SDS-PAGE and protein bands corresponding to the 43 kDa protein were excised from the Coomassie-stained gel. The 43kDa protein bands (~70pmol) were combined, destained and digested with trypsin prior to analysis by mass spectrometry (Australian Proteome Analysis Facility). The digested protein (not HPLC purified) was analysed by MALDI-TOF/TOF mass spectrometry (MS) and the resulting peak list was subjected to Mascot Ion/Ion Search analysis (Matrix Sciences; http://www.matrixscience.com/). Further details of the preparation and MS analysis of protein samples by APAF are provided in Appendix 1.

2.3.8 Haemolytic Activity

2.3.8.1 Blood Agar Diffusion Test

The haemolytic activity of venom extracts and chromatography fractions were initially tested using a gel diffusion assay in sheep blood agar. Under sterile conditions, samples (15μL) were applied to wells (diameter ~3mm) cored into 5% sheep blood agar plates. Following incubation of the plates (2h, 37ºC), haemolytic activity was detected visually. Diluted Triton X-100 (1% in NEB) was included as a positive control, and NEB and chromatography buffers were included as negative controls. Haemolysis was indicated by the formation of cleared zones around the wells.

2.3.8.2 Spectrophotometric Haemolytic Assay

Quantitative haemolytic activity of extracts and fractions was determined spectrophotometrically. Heparinised sheep blood was centrifuged (3K×*g*, 4°C, 10 min) and the sedimented erythrocytes were washed (4 × 15 vol. sterile PBS) and recovered by centrifugation (3K×*g*, 4°C, 10 min). Washed, diluted erythrocytes (0.4% in PBS) were combined on ice in triplicate (1mL each) with various dilutions of samples in NEB (100µL) and incubated (30 min, 37ºC) with gentle shaking. Following incubation, samples were chilled on ice (10 min), centrifuged (3K×*g*, RT, 5 min) and the absorbance of released haemoglobin in the supernatants was measured (540nm). Diluted Triton X-100 (1% in NEB) and NEB were used as references for 100% lysis and 0% lysis, respectively. Complete lysis was normalised to A_{540} = 1.0 by adjusting the initial erythrocyte concentration (in this case, to 0.4%). Haemolysis results were calculated as a percentage relative to complete lysis.

2.4 General Nucleic Acid Methods

2.4.1 Quantification and Purity of Nucleic Acids

The concentration of DNA and RNA solutions was estimated using a Beckman DU650 Spectrophotometer. The absorbance of solutions was measured at 260 and 280nm, and DNA and RNA concentrations were calculated based on the assumption that 1 absorbance unit (AU) per cm cell path at 260nm corresponds to approximately 50µg/mL of DNA and 40µg/mL of RNA. The ratio of 260nm and 280nm absorbance measurements (A_{260}/A_{280}) provided an estimate of nucleic acid purity with respect to protein contamination. DNA and RNA solutions with A_{260}/A_{280} values ≥ 1.8 and ≥ 1.9, respectively, were considered sufficiently pure for downstream applications.

2.4.2 Isolation of Total RNA

All glassware, plasticware and reagents used for RNA isolation were RNase-free. With the exception of Tris-based buffers, all aqueous reagents were treated with diethylpyrocarbonate (DEPC). DEPC was added to distilled water or buffer (0.1% final concentration), the solutions shaken vigorously and stood for 24h, then autoclaved to inactivate the DEPC. Tris-based buffers were prepared with DEPC-treated distilled water and filter-sterilised. Glassware was

soaked in acid/methanol solution (1:1 conc. HCl : methanol, 2h), rinsed thoroughly in DEPC-treated distilled water and baked (300ºC, 4h). Plasticware, unless purchased as nuclease-free, was soaked overnight in 0.1% DEPC, rinsed thoroughly in DEPC-treated distilled water, autoclaved, oven-dried and stored in RNase-free containers.

Total RNA was isolated from *C. fleckeri* tentacle or bell tissue using TRIzol reagent (Invitrogen). Fragments of excised tissue were removed from RNAlater, drained briefly on blotting paper and weighed. The tissues were homogenised in TRIzol (~100mg tissue/mL reagent) using a bead mill or a DEPC-treated mortar-and-pestle. The homogenate was transferred to a nuclease-free tube and incubated (5 min, RT). Following addition of chloroform (0.2mL $CHCl_3$/mL TRIzol), the mixture was shaken vigorously (15s), allowed to settle (2–3 min, RT) and centrifuged (12K×*g*, 4°C, 10 min). The colourless upper supernatant was transferred to another tube, mixed with isopropanol (0.5mL isopropanol/mL TRIzol) and incubated (10 min, RT). Precipitated RNA was recovered by centrifugation (12K×*g*, 4°C, 10 min), washed with 75% ethanol (≥ 1mL 75% ethanol/mL TRIzol) and re-centrifuged (7.5K×*g*, 4°C, 5 min). The supernatant was discarded and the pellet air-dried (5–10 min) and resuspended in DEPC-treated distilled water. The concentration and purity of isolated total RNA was estimated spectrophotometrically (Section 2.4.1). Total RNA integrity was assessed by native agarose gel electrophoresis (Section 2.4.9).

2.4.3 Isolation of Poly(A)$^+$ RNA

Poly(A)$^+$ RNA was isolated from total tentacle RNA using a MicroPoly(A)Pure Kit (Ambion), according to the manufacturer's instructions. All tubes, columns and reagents, except ethanol, were provided with the kit. Nuclease-free 0.5M NaCl was added with mixing to a solution of total RNA (0.45M final concentration). The solution was mixed with Binding Buffer (1mL final volume) and incubated (65ºC, 5 min). A vial of oligo(dT) cellulose was added and the sample was incubated with gentle agitation on a rotary mixer (1h, RT). The poly(A)$^+$ RNA-bound cellulose was recovered by centrifugation (3K×*g*, RT, 10 min) and washed sequentially with high salt buffer (Binding Buffer, 3 × 1mL) and low salt buffer (Wash Buffer, 3 × 1mL). The cellulose was resuspended in wash buffer (400µL), transferred to a 2mL spin column and briefly centrifuged (3K×*g*, RT). The cellulose was washed (Wash Buffer, 2 × 0.5mL) and poly(A)$^+$ RNA was eluted from the cellulose by adding pre-warmed (65ºC) Elution Buffer (100µL) and centrifuging the column immediately (5K×*g*, RT, 30s). The elution step was repeated once. Poly(A)$^+$ RNA was precipitated from each eluent by adding 5M ammonium

acetate (20µL), glycogen (1µL) and 100% ethanol (550µL); the solutions were mixed and incubated overnight (-20ºC). Poly(A)$^+$ RNA was recovered by centrifugation (12K×*g*, 4ºC, 20 min) and resuspended in DEPC water/EDTA solution (10µL). The purity and yield of poly(A)$^+$ RNA was determined spectrophotometrically (Section 2.4.1).

2.4.4 Isolation of Plasmid DNA

Plasmid DNA was isolated from bacterial cultures using a QIAprep Spin Miniprep Kit (QIAGEN) according to the manufacturer's instructions (using a microcentrifuge) and eluted with 30-50µL pre-warmed TE buffer (65ºC).

2.4.5 Restriction Enzyme Digestion

Restriction enzymes were purchased from New England Biolabs or Fermentas. Typical digestion reactions (20µL final volume) contained 10 units of restriction enzyme per 1µg purified DNA, 100X BSA (10mg/mL; 0.2µL), a recommended 10X buffer (2µL) and distilled water. Unless otherwise specified, digestion reactions were incubated (37ºC, 2h) and inactivated by addition of Stop Mix solution (5µL; see Section 2.4.9).

2.4.6 DNA Ligation

DNA ligations (10µL final reaction volume) were performed by incubating DNA fragments with appropriately linearised cloning or expression vector (3:1 molar ratio of insert to vector DNA), T4 DNA ligase (3 Weiss Units; Promega), 10X Ligase buffer (1µL) and distilled water. Blunt-end and cohesive-end ligations were incubated at 4ºC and 16ºC, respectively, for times specified in subsequent chapters.

2.4.7 Polymerase Chain Reaction (PCR)

Amplification of DNA was achieved using a ThermoHybaid PCR Express thermocycler and *Taq* DNA polymerase (Fermentas). Reaction mixtures were prepared on ice; the reaction components are summarised in Table 2.2. Template DNA included plasmids, purified PCR products, phage DNA or single bacterial colonies. In general, plasmids, PCR products and phage DNA were pre-diluted (X0.1–0.01 in TE or SM Buffer). For colony PCR, a small amount of cells were transferred to the PCR reaction mix using sterile pipette tips or toothpicks. Positive and negative control reactions, containing known template or no template, respectively,

were included with each set of samples. Thermocycler programs are detailed in subsequent chapters.

Table 2.2

Components of PCR reaction mixtures for amplification of DNA with Taq DNA polymerase, based on a final reaction volume of 20µL.

Component	Volume (µL)	[Final]
10X Taq Buffer[a]	2.0	1X
$MgCl_2$ (25mM)	1.2	1.5mM
dNTPs (10mM ea.)	0.4	0.2
Taq DNA Pol. (5U/mL)	0.2	1U/20µL
Primer 1	1.0	0.25µM
Primer 2	1.0	0.25µM
DNA template	1.0	variable
ddH_20	13.2	-

[a]100 mM Tris-HCl (pH 8.8 at 25°C), 500 mM KCl, 0.8% (v/v) Nonidet P40

2.4.8 Reverse Transcription Polymerase Chain Reaction (RT-PCR)

Amplification of complementary DNA (cDNA) from mature mRNA template was performed using 2-Step RT-PCR and Ready-to-Go RT-PCR Beads (Amersham Biosciences). For each reaction (50µL final volume), a single RT-PCR bead was dissolved in DEPC-treated distilled water (46µL) on ice with occasional mixing. A poly$(dT)_{12-18}$ primer (0.5µg/µL in DEPC-treated distilled water) and total *C. fleckeri* tentacle RNA (1µg/µL) were added to the tube (1µL each) and the reaction mixture was mixed, centrifuged briefly, incubated (42°C, 40 min) and chilled on ice. Forward and reverse primers (5µM; 1µL each) were added to the mixture and hot start PCR amplification was performed using thermocycler programs detailed in subsequent chapters. Positive control reactions containing 1µg total RNA and specific primers known to reliably generate DNA products (e.g. *C. fleckeri* elongation factor-1α cDNA; see Chapter 3) were included with each set of samples. Negative controls, prepared without template, were also included with each set of samples.

2.4.9 Agarose Gel Electrophoresis

Plasmid DNA, total RNA and PCR products were analysed using a Bio-Rad Mini Sub-Cell GT Agarose Gel Electrophoresis System. Molecular biology-grade agarose was dissolved with heating in TBE (90mM Tris-HCl, pH8.3, 90mM boric acid, 2.5mM EDTA) to a final concentration of 0.8–2%, depending on expected DNA fragment lengths. Ethidium bromide (10mg/mL) was added (0.01% final concentration) and the agarose mixture was cast in 7–10cm casting trays with appropriately sized lane combs. Following addition of Stop Mix solution (60mM EDTA, pH8.0, 10mM Tris-HCl, pH 7.9, 30%$^{v}/_{v}$ glycerol, 0.01% bromophenol blue; 0.2 vol.), DNA samples were briefly vortexed and centrifuged. DNA samples were applied to separate lanes in agarose gels and DNA fragments were separated by electrophoresis (30–45 min, 100V). The length of the DNA fragments was estimated based on the migration distances of a standard 1kb DNA ladder (Fermentas).

For the analysis of RNA samples, gels were prepared using nuclease-free reagents and glassware (see Section 2.4.2 for further details). Dedicated electrophoresis tanks and casting trays for RNA analysis were pre-cleaned with ethanol. RNA samples were mixed 4:1 with 50% glycerol in DEPC-treated distilled water, applied alongside lanes containing aliquots of stop mix solution (used as a tracking dye) and a 1kb DNA ladder (used as a positive control for nucleic acid detection, not for length estimation), and subjected to electrophoresis, as previously described.

2.4.10 Purification of DNA Fragments

DNA fragments generated by PCR amplification or restriction enzyme digestion were purified using QIAquick PCR Purification or Gel Extraction Kits (QIAGEN) according to the manufacturer’s instructions (using a microcentrifuge). Purified DNA fragments were eluted from the spin columns with 30–50µL pre-warmed TE buffer (65ºC).

2.4.11 Preparation of Radiolabelled DNA Probes

Radiolabelled DNA probes were prepared to screen a *C. fleckeri* cDNA library or bacterial transformants. Purified DNA fragments, obtained by PCR amplification of targeted genes, were labelled with [α-^{32}P]-dATP in a random primed reaction using a DECAprime II Kit (Ambion). DNA (~25ng; ≤ 11.5µL) was denatured (95–100°C, 5 min) with 10X Decamer Solution (2.5µL) and chilled immediately on ice. The reaction mixture was supplemented with 5X Reaction

Buffer (-dATP; 5µL), exonuclease-free Klenow (2 Units/µL; 1µL), [α-^{32}P]-dATP (10mCi/mL; 5µL) and nuclease-free distilled water (to 25µL final volume), mixed and incubated (37°C, 10 min). The radiolabelled DNA probe was denatured (95–100°C; 5 min) and chilled immediately on ice in preparation for conventional DNA screening protocols (see Sections 2.4.14 and 2.4.16).

2.4.12 *C. fleckeri* cDNA Expression Library Construction and Titration

A cDNA expression library was constructed to facilitate the isolation of cDNA encoding putative *C. fleckeri* venom proteins. Total RNA from intact *C. fleckeri* tentacles was isolated by homogenisation in a bead mill or a mortar-and-pestle with TRIzol (Invitrogen) (see Section 2.4.2). Poly(A)$^+$ RNA was isolated from the total RNA using a MicroPoly(A)Pure Kit (Ambion), as previously described (Section 2.4.3). The cDNA library was constructed in lambda Uni-ZAP XR using 2µg poly(A)$^+$ and a ZAP-cDNA Gigapack II Gold Cloning Kit (Stratagene) according to the manufacturer's instructions.

To titrate the library, serial dilutions (10^{-1},…, 10^{-5}) of library phage were prepared in SM buffer (50mM Tris-HCl (pH 7.5), 100mM NaCl, 8mM $MgSO_4$, 0.01% gelatine). Aliquots of XL-Blue XRF' cells (200µL, OD_{600} ~0.5) were incubated (15 min, 37°C) with diluted library phage (10µL), mixed with 0.8% melted LB-agar (2–3mL, 50˚C), poured onto LB-tet (10µg tetracycline/mL) agar plates (90mm) and grown overnight at 37°C. The estimated packing titre of the library was 4.8×10^5 pfu/mL. After one round of amplification in XL-Blue MRF'*E. coli* cells, the estimated library titre was 1.4×10^9 pfu/mL.

2.4.13 Immunoscreening of Expressed cDNA Library Clones

The *C. fleckeri* cDNA library was screened for expressed clones that were antigenic to CSL box jellyfish antivenom and polyclonal antibodies raised in mouse and rabbit against *C. fleckeri* nematocyst extracts. Aliquots of XL-Blue XRF' cells (600µL; OD_{600} ~0.5) were incubated (15 min, 37°C) with diluted library phage (~2×10^4 plaques), mixed with 0.8% melted LB-agar (6–7mL, 50˚C), poured onto LB-tet (10µg/mL) agar plates (150mm) and grown overnight (37°C). Protein expression was induced by overlaying the plates with Hybond-C Extra membranes (137mm; Amersham Biosciences) that were pre-soaked in IPTG solution (10mM) and air-dried in darkness. The plates were incubated (4–5h, 37–42°C). Membranes were briefly washed (TBST), incubated in blocking solution (5% non-fat milk powder in TBST, 1h), then incubated

overnight in various dilutions of primary antibody in TBST (1:3000 for CSL antivenom; 1:1000 for mouse and rabbit antibodies). Membranes were removed, washed (3 × 5 min in TBST) and incubated (1h) in an appropriate AP-conjugated secondary antibody diluted in TBST (1:5000). Following membrane washing, antigenic proteins were visualised using NBT/BCIP (Promega) as described in the western blot protocol (Section 2.3.4). Positive plaques were cored, vortexed and diluted in SM buffer (200µL), incubated (2h at RT or overnight at 4ºC), and re-screened using smaller aliquots of cells (200µL) and melted agar (2–3mL), and smaller agar plates (90mm) and membranes (82mm). Independent positive plaques were isolated, vortexed and diluted in SM buffer (500µL), treated with chloroform (20µL) and stored as phage stock at 4°C.

2.4.14 DNA Screening of cDNA Library Phage

The *C. fleckeri* cDNA library was screened for target clones using sequence-specific radiolabelled oligonucleotide probes. Aliquots of XL-Blue XRF' cells (600µL; OD_{600} ~0.5) were incubated (15 min, 37°C) with diluted library phage (1–2 × 10^4 plaques), mixed with 0.8% melted LB-agar (6–7mL, 50˚C), poured onto LB-tet (10µg/mL) agar plates (150mm) and grown overnight at 37°C. The plates were cooled (2h, 4°C) and plaque DNA was transferred to Hybond N^+ membranes (137mm; Amersham Biosciences). Membranes were treated with denaturation buffer (0.5M NaOH, 1.5M NaCl; 5 min), neutralisation buffer (0.5M Tris-HCl pH 8.0, 1.5M NaCl; 3 min), washed briefly in 2X SSC (30mM trisodium citrate, 300mM NaCl, pH7) and baked (2h, 80°C). Membranes were incubated (2h, 42°C) in prehybridisation buffer (5X SSC, 5X Denhardt's solution (0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone), 50% formamide, 50mM K_2HPO_4, 0.1% SDS, 250µg/mL sheared, denatured herring sperm DNA), and following addition of a denatured sequence-specific radiolabelled probe (Section 2.4.11), hybridised overnight at 42°C. The membranes were washed (2 × 10 min each) in low stringency solution (2X SSC, 0.1% SDS) and high stringency solution (0.1X SSC, 0.1% SDS) at 42°C, then exposed to a phosphor screen (3–4h). Radioactivity was detected using a Storm 860 Phosphorimager (Molecular Dynamics). Positive plaque phage were cored, vortexed and diluted in SM buffer (200µL), incubated (2h at RT or overnight at 4ºC), and re-screened using smaller aliquots of cells (200µL) and melted agar (2-3mL) and smaller plates (90mm) and membranes (82mm). Independent positive plaques were isolated, vortexed and diluted in SM buffer (500µL), treated with chloroform (20µL) and stored as phage stock at 4°C.

2.4.15 *In vivo* Excision of pBluescript SK- Clones

Excision of pBluescript SK- (pBSK-) clones from the Uni-ZAP XR lambda vector was performed using ExAssist helper phage and SOLR *E. coli* cells (Stratagene). Aliquots of XL-Blue XRF' cells (200µL; OD_{600} ~1) were incubated (15 min, 37°C) with target phage stock (>1 × 10^5 phage particles in SM buffer; 250µL) and ExAssist helper phage (>1 × 10^6 pfu/µL; 1µL). Following addition of supplemented LB-tet (LB-tet (10µg/mL) + 0.01 vol. filter-sterilised 1M $MgSO_4$ / 20% maltose; 3mL), the mixture was incubated with shaking (37°C, 2.5-3h, 220 rpm). The XL-Blue cells and lambda phage particles were lysed by heating (70°C, 20 min) and the lysate was centrifuged (5K×*g*, 4ºC, 10 min). Aliquots of supernatant (10µL and 100µL) were mixed separately with aliquots of freshly grown SOLR cells (200µL each; OD_{600} ~1) and the cells were incubated (15 min, 37°C), plated onto LB-C (50µg carbenicillin/mL) agar and grown overnight at 37°C. Individual colonies were cultured overnight with shaking (37°C, 220rpm) in LB-C (50µg/mL) and their plasmid DNA isolated as previously described (Section 2.4.4).

2.4.16 DNA Screening Bacterial Colonies

To identify recombinant bacterial colonies carrying target expression constructs, colonies were screened with a radiolabelled sequence-specific oligonucleotide probe in a manner similar to screening cDNA library lambda phage. Transformed NM522 *E. coli* cells were grown overnight (37°C) on LB-C (50µg/mL) agar plates (90mm), transferred to Hybond N^+ membranes (82mm) and treated with lysis buffer (10% SDS, 2 min). Following colony lysis, DNA screening proceeded as previously described (see Section 2.4.14).

2.4.17 DNA Sequencing

Automated DNA sequencing was outsourced to the Genetic Analysis Facility (GAF), James Cook University, or Macrogen Inc., Korea. Plasmid DNA samples (150-300ng) were prepared using ET Terminator (Amersham Biosciences) or BigDye Terminator Cycle (Applied Biosystems) sequencing kits according to the manufacturer's instructions and analysed using MegaBACE 1000 (Amersham Biosciences) or ABI 3730xl (Applied Biosystems) DNA analysis systems.

2.5 Bacterial Methods

2.5.1 Preparation of Competent Cells

Chemically competent bacterial cells were prepared using calcium chloride. A glycerol stock of NM522 *E. coli* cells (Stratagene) was streaked onto LB agar and grown overnight (37°C). Aliquots of sterile LB (5mL) were inoculated with single NM522 colonies and the cultures grown overnight with shaking (37°C, 260 rpm). Sterile LB was inoculated with the saturated culture (0.01 vol.) and incubated with shaking (37°C, 260 rpm) until the cells reached logarithmic growth phase (OD_{600} ~0.5-0.6). The cells were harvested by centrifugation (5.1K×*g*, 4°C, 10 min), placed on ice and resuspended gently in sterile ice-cold 0.1M $CaCl_2$ (0.1 vol.). The washed cells were centrifuged (5.1K×*g*, 4°C, 10 min), placed on ice, and resuspended in sterile ice-cold 0.1M $CaCl_2$ (0.1 vol.). The cells were incubated on ice (1h), recovered by centrifugation (5.1K×*g*, 4°C, 10 min), replaced on ice and resuspended in sterile ice-cold 0.1M $CaCl_2$ / 15% glycerol (0.1 vol.). Aliquots of resuspended competent cells (100–200µL) were snap-frozen in liquid nitrogen and stored at –80°C.

2.5.2 Transformation of Competent Cells

Plasmids were introduced into competent bacterial cells by heat shock transformation. Aliquots of competent NM522 cells (100–200µL) were thawed on ice, gently mixed with purified plasmid DNA (20-50ng) and incubated on ice (20 min). The cells were heat-shocked (42°C, 90s), replaced immediately on ice (5–10 min), plated onto LB-C (50µg/mL) agar and grown overnight (37°C). For blue/white selection of recombinant colonies, cells were co-plated with 20mM IPTG / 2% X-Gal in *N, N*-dimethylformamide (50µL). To characterise individual transformants, LB (5mL) supplemented with an appropriate antibiotic was inoculated with a single bacterial colony and cultured overnight with shaking (37°C, 260rpm). Plasmid DNA was isolated from the cultures (Section 2.4.4) and characterised by restriction enzyme digestion (Section 2.4.5), PCR amplification (Section 2.4.7) and/or sequencing (Section 2.4.17).

2.5.3 Bacterial Glycerol Stocks

Bacterial glycerol stocks were prepared for long-term storage of recombinant *E. coli* cells by adding aliquots of saturated bacterial cultures (500µL) to sterile 80% glycerol (500µL). Stocks were mixed and stored at –80°C.

CHAPTER 3

Analysis of a *C. fleckeri* Tentacle cDNA Expression Library – the Search for Putative Venom Proteins

3.1 Introduction

The venom of *C. fleckeri* contains a variety of bioactive proteins that are stored and discharged by nematocysts. In the last 40 years, several researchers have examined the biochemical properties and biological activities of a variety of *C. fleckeri* venom proteins (see Table 1.1), however, no studies have yet reported the amino acid sequences of nematocyst-derived *C. fleckeri* venom proteins or the DNA sequences of genes that encode them.

Cnidarian nematocysts are non-living nematocyte secretions that lack nucleic acids (Kass-Simon and Scappaticci, 2002), thereby making it impossible to clone venom protein genes directly from isolated nematocysts. However, previous researchers have successfully isolated transcripts that encode venom proteins from the surrounding tentacle tissues of other box jellyfish species, thus enabling the cDNA cloning and deduced amino acid sequencing of lethal and haemolytic toxins from *C. rastoni*, *C. alata* and *C. quadrigatus* (Nagai et al., 2000a, b; 2002). In this study, a *C. fleckeri* cDNA expression library was constructed using tentacle mRNA to facilitate the cDNA cloning and sequencing of major *C. fleckeri* venom proteins. As described in this chapter, the first approach to isolate and characterise major *C. fleckeri* venom proteins involved screening the expressed cDNA library with a variety of antibodies raised against *C. fleckeri* venom. In previous studies, similar immunoscreening strategies were successfully used for the detection and cDNA cloning of protein toxins expressed by other venomous organisms such as snakes (e.g. Welton and Burnell, 2005; Assakura et al., 2003) and insects (e.g. Fang et al., 1988; Jones et al., 1992). This experiment also presented the opportunity to discover other novel and/or existing families of proteins that are expressed in cubozoan jellyfish, while providing preliminary insight into the transcriptome profile of *C. fleckeri* tentacles.

3.2 Methods

The expressed *C. fleckeri* tentacle cDNA library was screened in three separate experiments using CSL box jellyfish antivenom (1:3000 in TBST) and polyclonal antibodies raised against *C. fleckeri* nematocyst venom in mice or rabbits (1:1000 in TBST each) (see Sections 2.3.2 and 2.4.13 for further details). Antigenic phage plaques were re-screened, isolated, excised *in vivo* (see Section 2.4.15) and the 5'- and/or 3'-ends of the clones were sequenced. Nucleotide sequences were processed and aligned (Sequencher, Gene Codes Corp.) and subjected to BLAST analysis (Altschul et al., 1997) using nucleotide, translated nucleotide and amino acid queries (BLASTN, BLASTX and BLASTP, respectively). Selected full-length clones were sequenced in both directions.

In an attempt to obtain the full-length sequences of three incomplete cDNA clones isolated during the CSL antivenom immunoscreening experiment, the *C. fleckeri* cDNA library was screened with gene-specific radiolabelled oligonucleotide probes. Forward and reverse primers were designed according to the nucleotide sequences of pBluescript SK- (pBSK) clones partially encoding thypedin, an actin-binding protein (ABP) and a heat shock protein (HSP70) (Table 3.1). DNA was amplified using hot start PCR (95°C for 5 min, 35 cycles of 95°C for 45s, 49-52°C for 1 min and 72°C for 2 min, then 72° for 5 min). For thypedin and ABP, undiluted cDNA library (1µL) was added to the PCR reactions; whereas for HSP70, diluted clone 3-6-1 pBSK DNA (1:9 in TE buffer; 1µL) was used (see Section 2.4.7 for further details).

Table 3.1

Forward (F) and reverse (R) primers designed according to the nucleotide sequences of pBSK clones partially encoding thypedin, ABP or HSP70.

Protein	Clone ID	Primer ID	Primer Sequence (5'→3')	T_m (°C)
Thypedin	1-2-1	Th-F1	TTCAAGAAGAGGCTACGG	54
		Th-R1	GACATAGAGGCAGCATTAGC	60
ABP	4-3-3	ABP-F1	AAGGGGAGACAGCACGATAAGG	56
		ABP-R1	TCCTACGGTTTTGTCTTCG	56
HSP70	3-6-1	HSP-F1	GTCGTCAAGCAAACAAGC	54
		HSP-R1	CTTTGCTTGGATTCGGTC	54

Three PCR products (1487, 801 and 796bp, respectively) were obtained. Each purified product (25ng) was labelled with [α-^{32}P]-dATP in a random primed reaction using a

DECAprime II Kit (Ambion) and used to screen ~1 × 10^5 plaques from the *C. fleckeri* cDNA library (see Sections 2.4.11 and 2.4.14). The longest clones of positive plaques were determined by PCR amplification of recovered and diluted phage DNA (1:200 in SM buffer) using the primer SK together with Th-R1, ABP-R1 or HSP-R1, and PCR conditions as described previously. Selected phage were rescreened with the radiolabelled probes and independent positive plaques were isolated, excised *in vivo* (Section 2.4.15) and sequenced in both directions.

3.3 Results and Discussion

3.3.1 Immunodetection of cDNA Clones using CSL box jellyfish antivenom

Immunoscreening the *C. fleckeri* cDNA library (~1 × 10^5 plaques) with CSL box jellyfish antivenom resulted in the detection of 164 antigenic clones, constituting ~0.2% of total library clones. The majority of clones were strongly antigenic and clearly visible in secondary library screens (Figure 3.1).

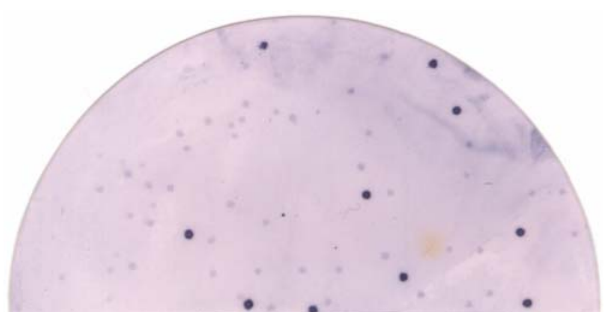


Figure 3.1 Secondary screen of clone 1-2 using CSL box jellyfish antivenom immunodetection. The strongly antigenic proteins corresponding to expressed clone 1-2 (elongation factor-1α) are indicated by dark purple spots on the nitrocellulose membrane. Non-antigenic proteins are indicated by pale purple spots.

Of the 164 isolated clones, 2 clones (1-14-1 and 4-8-1) were not successfully excised as Bluescript plasmids from the Uni-Zap XR lambda vector. Of the remainder, 100 independent clones were fully or partially sequenced. BLAST analysis revealed that 6 clone sequences share no significant homology (i.e. *E*-value ≥ 1) with any sequence available in non-redundant databases. No clearly defined open reading frames were detected in 3 of these clones (1-1-1, 1-4-1 and 1-7-1), suggesting that the clones may represent either long UTRs (3' or 5') or

regulatory RNAs. The remaining translated clone sequences were significantly homologous to a variety of proteins, including ubiquitous transcription and translation co-factors, structural proteins and enzymes, however, none of the sequences were similar to any previously reported protein toxins (Table 3.2). The functional distribution of the non-venom proteins is summarised in Figure 3.2 (see also Appendix 2 for further details). Six clusters, representing 62% of the total sequenced clones, contained multiple cDNA clones and the remaining clusters (38) represented singletons. The identification numbers of clones within Clusters 1-5 are listed in Appendix 3. Twenty-eight clones obtained by immunodetection were full-length, as indicated by BLAST alignments of translated sequences with homologous full-length proteins and the presence of putative initiating methionine residues. The nucleotide and deduced amino acid sequences of the most commonly detected clones, elongation factor-1α (EF1α) and 3-phosphoglycerate dehydrogenase (3-PGDH), are shown in Figure 3.3 and Figure 3.4, respectively. Following screening the cDNA library with gene-specific probes, the full-length sequences of HSP70 (designated HSP70-1) and ABP were also obtained (Figure 3.5 and Figure 3.6). However, in the case of thypedin, the longest clone obtained from the library (clone 3-9; 2378bp) still lacked the initiating start codon of the protein coding region (Figure 3.7). The nucleotide and deduced amino acid sequences of the remaining cDNA clones listed in Table 3.2 are provided in Appendix 4. In each sequence (where applicable), the putative translation initiation codon (ATG) is highlighted in bold and the stop codon is indicated with an asterisk.

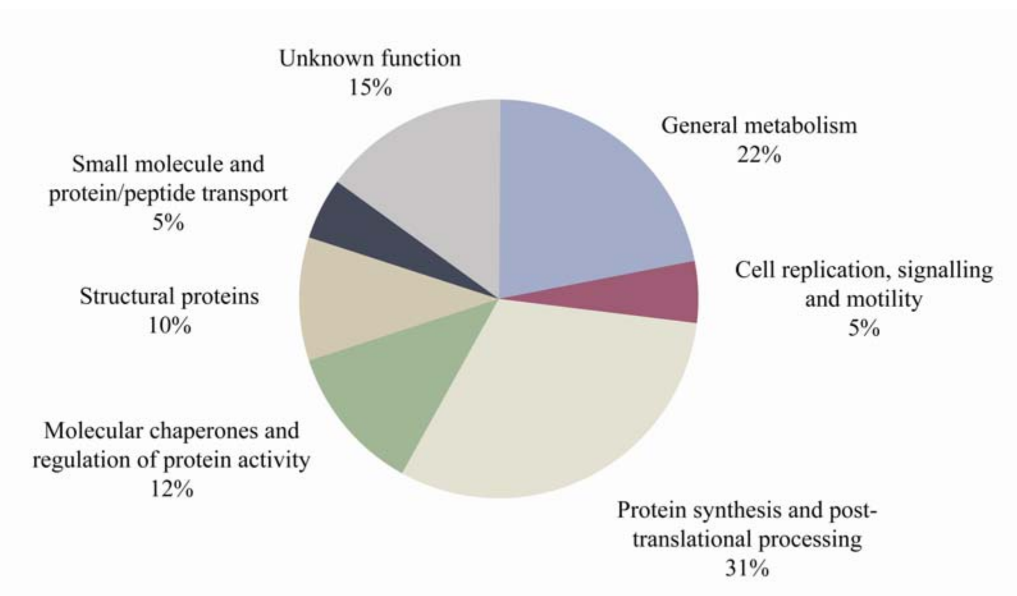


Figure 3.2 Expression profile of *C. fleckeri* tentacles categorised according to general protein function. The abundance of antigenic non-venom proteins is indicated as a percentage of 100 clone sequences.

Table 3.2

Summary of antigenic *C. fleckeri* cDNA clones obtained by immunodetection with CSL box jellyfish antivenom. Putative full-length protein entries are highlighted in green.

Cluster No.	No. Clones	Clone ID	Accession No.	Insert Length (kb)	No. Deduced Residues	Homologous Protein (Best BLAST Match)	Species of Origin	Accession No.	E-Value	No. Residues	% Match	Comments
1	29	4-5-1[a]	FJ460241	1.8	464	EF1α	*Axinella verrucosa* (sponge)	CAE45763	0.E+00	462	87	**F**
2	16	3-9-1[a]	FJ460242	1.7	520	3-PGDH	*Nematostella vectensis* (sea anemone)	XP_001629480	2.E-114	487	50	**F**
3	7	7-1[b]	FJ460243	2.0	652	HSP70	*Alligator mississippiensis* (alligator)	BAF94143	0.E+00	646	86	**F** (DNA screen)
4	5	consensus[c]		2.5	no contig	filamin	*Homo sapiens* (human)	BAC03408	7.E-111	2651	48	1524bp (T3), 593bp (T7)
5	3	13-5[b]	FJ460244	2.1	511	ABP, coronin 1B	*Danio rerio* (zebra fish)	NP_001103177	4.E-111	499	41	**F** (DNA screen)
6	2	8-3, 32-1		1.3	330	trans-homoaconitate synthase	*Pyrococcus abyssi* GE5 (archaea)	NP_126127	4.E-01	361	27	
7	1	1-5-1		1.5	305	replication factor C	*Bos taurus* (cattle)	AAI09607	4.E-92	356	60	
8	1	1-6-1		1.1	167	alkylhydroperoxidase (AhpD) core	*Burkholderia phytofirmans* (bacteria)	EAV06712	5.E-56	201	65	**F**
9	1	1-9-1		1.1	168	cyclophilin A	*Argopecten irradians* (scallop)	ABM92916	7.E-66	164	79	**F**
10	1	1-10-1		1.5	367	ATPase-like protein (contains a cation transport ATPase domain)	*Nematostella vectensis* (sea anemone)	EDO44727	3.E-107	1177	59	
11	1	2-1-1		1.8	no contig	non-muscle myosin heavy chain-like	*Danio rerio* (zebra fish)	XP_683046	1.E-115	2093	56	782bp (T3), 836bp (T7)
12	1	3'-2		1.3	no contig	DnaJ (HSP40) b11 homologue	*Strongylocentrotus purpuratus* (sea urchin)	XP_793217	9.E-66	358	68	577bp (T3), 645bp (T7)
13	1	3-9	FJ460245	2.4	666	thypedin	*Hydra vulgaris* (hydra)	AAW82079	0.E+00	1089	62	
14	1	3-18-1		1.5	468	proprotein convertase-like protein	*Nematostella vectensis* (sea anemone)	EDO38851	3.E-115	627	75	
15	1	3-21-1		2.0	no contig	peroxisomal biogenesis factor 14	*Mus musculus* (mouse)	NP_062755	4.E-22	376	33	813bp (T3), 757bp (T7)
16	1	3-23-1		2.2	no contig	metaxin-like	*Nematostella vectensis* (sea anemone)	EDO43101	4.E-37	230	41	718bp (T3), 837bp (T7)
17	1	4-1-1		1.5	463	chromosome segregation ATPase-like	*Apis mellifera* (bee)	XP_001120388	5.E-24	2064	22	
18	1	4-2-1		0.8	89	hypoxia-induced gene 1 (HIG1)-like	*Nematostella vectensis* (sea anemone)	EDO46809	1.E-08	85	62	
19	1	6/7-6-4		0.8	no contig	B-cell translocation gene 1 (BTG family)	*Danio rerio* (zebra fish)	NP_956314	4.E-26	182	34	**F**; 760bp (T3 only)
20	1	10-1-1		1.5	no contig	polyA binding protein	*Monodelphis domestica* (opossum)	XP_001369327	3.E-52	306	70	**F**; 706bp (T3), 441bp (T7)
21	1	12-2		2.1	no contig	KIF27A homologue (kinesan family)	*Monodelphis domestica* (opossum)	XP_001379231	1.E-27	1408	36	663bp (T3), 732bp (T7)
22	1	12'-2		1.2	287	TDC-1 (tyrosine decarboxylase family)	*Caenorhabditis elegans* (nematode)	NP_495743	2.E-68	705	47	1200bp
23	1	14-1-1		1.2	no contig	memo-like protein	*Danio rerio* (zebra fish)	AAI55285	1.E-18	297	67	651bp (T3), 796bp (T7)
24	1	16-1-1		2.5	no contig	viral A-type inclusion protein or myosin heavy chain-like	*Caenorhabditis elegans* (nematode)	AAA28124	3.E-07	1966	23	621bp (T3), 705bp (T7)
25	1	17-2		0.6	87	mitochondrial inorganic phosphate carrier protein	*Nematostella vectensis* (sea anemone)	EDO31806	2.E-30	355	86	
26	1	18-3		1.6	no contig	short chain oxidoreductase	*Nematostella vectensis* (sea anemone)	EDO42526	3.E-79	415	75	**F**; 618bp (T3), 524bp (T7)
27	1	26-2		1.8	no contig	calreticulin	*Strongylocentrotus purpuratus* (sea urchin)	XP_001177910	1.E-52	591	62	**F**; 659bp (T3), 545bp (T7)
28	1	28-2		1.0	261	ornithine aminotransferase	*Homo sapiens* (human)	BAD92054	2.E-113	439	73	
29	1	29-1		1.3	no contig	cell wall anchor family protein	*Dictyostelium discoideum* (amoeba)	AAO52512	5.E-03	1806	23	573bp (T3), 549bp (T7)
30	1	31-1		1.5	no contig	transportin-3 homologue	*Nematostella vectensis* (sea anemone)	EDO32037	4.E-47	934	44	621bp (T3), 624bp (T7)
31	1	37-1		0.9	159	translationally-controlled tumour protein	*Branchiostoma belcheri* (Japanese lancelet)	AAK84394	8.E-27	169	39	
32	1	38-1		1.2	no contig	calponin	*Branchiostoma belcheri* (Japanese lancelet)	BAC16745	5.E-25	187	35	324bp (T3), 697bp (T7)
33	1	39-1		1.2	335	embryonic-1	*Hydra vulgaris* (hydra)	AAW71478	2.E-23	342	29	**F**
34	1	41-1		1.5	no contig	folliculin	*Gallus gallus* (fowl)	XP_414807	7.E-23	579	37	502bp (T3), 554bp (T7)
35	1	46-1		1.7	no contig	serine/threonine protein kinase	*Aedes aegypti* (mosquito)	EAT36306	6.E-39	667	44	670bp (T3), 645bp (T7)
36	1	3-3		1.1	214	predicted protein	*Nematostella vectensis* (sea anemone)	EDO41333.	2.E-07	173	29	
37	1	3-7-1		1.2	294	predicted protein	*Nematostella vectensis* (sea anemone)	EDO33301	1.E-61	176	72	
38	1	20-1		1.3	110	predicted protein (contains neurotransmitter-gated ion-channel domain)	*Nematostella vectensis* (sea anemone)	EDO33747	8.E-04	131	38	
39	1	1-8-1		1.5	~336	nil						
40	1	3-3-3		1.7	343	nil						
41	1	30-1		1.1	169	nil						
42	1	1-1-1		0.8	no ORF	nil						691bp
43	1	1-4-1		0.9	no ORF	nil						827bp
44	1	1-7-1		0.6	no ORF	nil						527bp
-	-	1-14-1		-								failed SCE
-	-	4-8-1		-								failed SCE
	100	Total Clones										

[a] selected full-length cDNA clones representing each cluster that were sequenced in both directions
[b] full-length cDNA clones obtained by DNA screening that were sequenced in both directions
[c] consensus of aligned partial nucleotide sequences of cDNA clones belonging to the same cluster (see Appendix 3 for clone IDs)

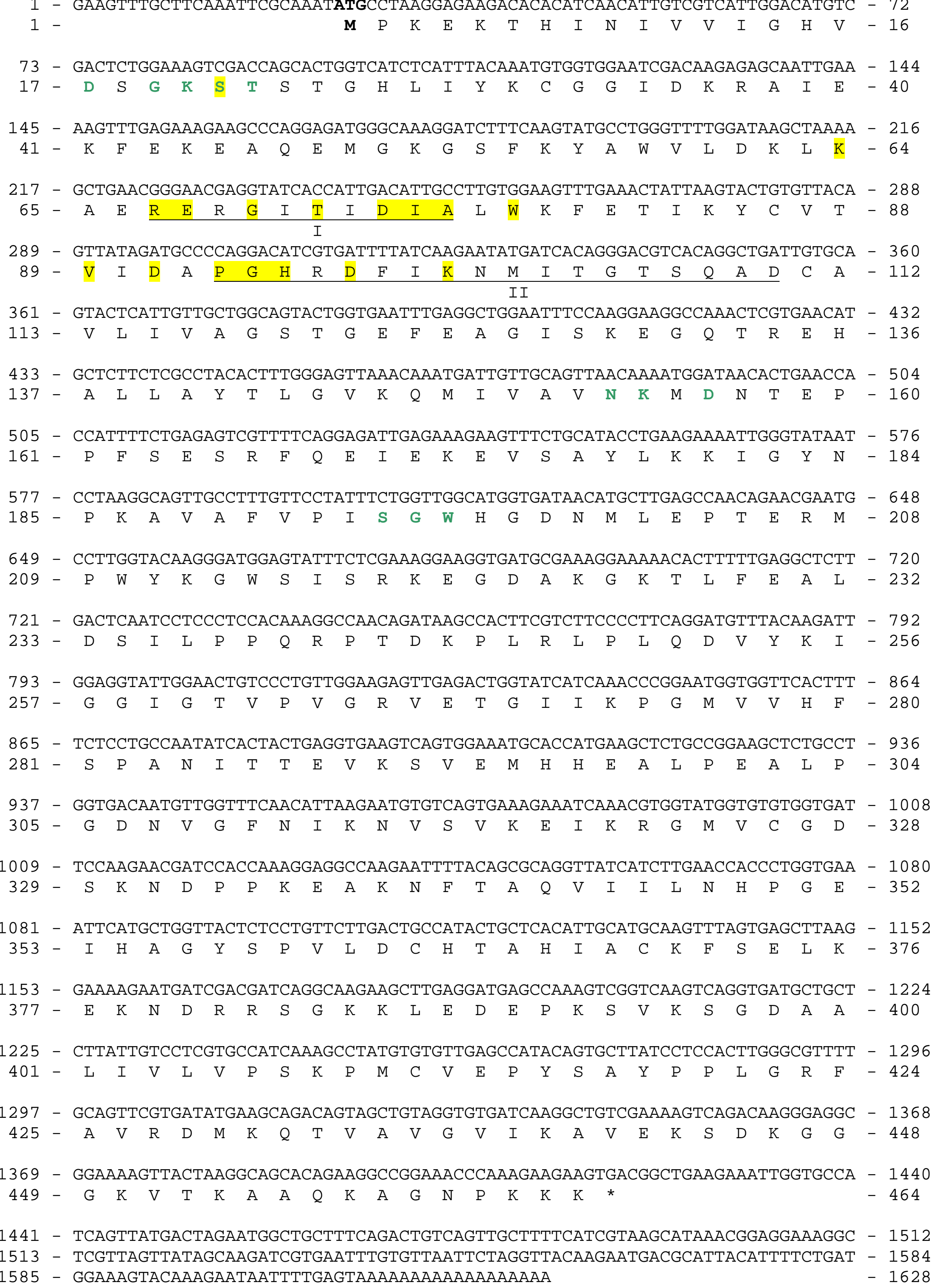


Figure 3.3 Full-length cDNA (clone 4-5-1) encoding EF1α detected by immunoscreening the *C. fleckeri* cDNA library with CSL box jellyfish antivenom. Accession No. FJ460241. Green residues correspond to a highly conserved GTP/Mg^{2+} binding site, residues highlighted in yellow correspond to an EF1Bα binding site, underlined residues indicate switch regions I and II.

```
   1 - AGATCTGGAACTTCAAAATTTTCCTGTCACACTTGAAACATGGCTGACCTAACAATTAAACGTGTCCTCATA - 72
   1 -                                         M  A  D  L  T  I  K  R  V  L  I  - 11

  73 - TCCGATAAAGTAGATGAAGTTTGCAAAACGATATTTGAAAGCAATAAAATTAACGTAGACTACCGCCCAGGG - 144
  12 -  S  D  K  V  D  E  V  C  K  T  I  F  E  S  N  K  I  N  V  D  Y  R  P  G  - 35

 145 - ATTTCAAAGGAAGAACTTCTTGCAATCATCAAGGATTATGACTGTCTTGTTGTTCGTTCTGCTACAAAAGTT - 216
  36 -  I  S  K  E  E  L  L  A  I  I  K  D  Y  D  C  L  V  V  R  S  A  T  K  V  - 59

 217 - ACAGCTGAAGTCTTTTCTGCTGCACCAAACCTTAAATTGGTTGGTCGTGCAGGCACAGGTGTTGACAACATA - 288
  60 -  T  A  E  V  F  S  A  A  P  N  L  K  L  V  G  R  A  G  T  G  V  D  N  I  - 83

 289 - GACTTGAAAGCAGCAACAAATAATGGTGTCTTTGTGATGAATACTCCTGGAGGAAACACAATGAGTGCTGCA - 360
  84 -  D  L  K  A  A  T  N  N  G  V  F  V  M  N  T  P  G  G  N  T  M  S  A  A  - 107

 361 - GAACATACATGTGCAATGATGTTTAGTTTGGCAAGGCACATTCCGCAGGGTTATATGTCTATGCAAGAAGGA - 432
 108 -  E  H  T  C  A  M  M  F  S  L  A  R  H  I  P  Q  G  Y  M  S  M  Q  E  G  - 131

 433 - AAATGGGAGAGATCCAAGTTTATGGGCTGTGAGCTGAAGGGAAAAACTCTGGGGATCATTGGACTAGGGCGT - 504
 132 -  K  W  E  R  S  K  F  M  G  C  E  L  K  G  K  T  L  G  I  I  G  L  G  R  - 155

 505 - ATTGGAAAAGAAGTTGCATTAAGAATGCAAGCATGTGAAATGAGGACCGTTGGATATGACCCAATAGTTCCG - 576
 156 -  I  G  K  E  V  A  L  R  M  Q  A  C  E  M  R  T  V  G  Y  D  P  I  V  P  - 179

 577 - AAGGAAGCAGCTGCTGAATTTGGTGTCGAGTTTATGGAGCTTGAAGAGATGTGGCCCATTGTGGATTTCATC - 648
 180 -  K  E  A  A  A  E  F  G  V  E  F  M  E  L  E  E  M  W  P  I  V  D  F  I  - 203

 649 - ACAGTTCACACACCATTAATTCCACAAACGAAAGGATTGGTCAATGCAAAAACGTTTGACAAATGTAAGAAG - 720
 204 -  T  V  H  T  P  L  I  P  Q  T  K  G  L  V  N  A  K  T  F  D  K  C  K  K  - 227

 721 - ACGATCCGTGTTATCAATGTGGCCCGAGGAGGCATTATTGATGAATCTGATCTACTCGACGCGCTAAATGCT - 792
 228 -  T  I  R  V  I  N  V  A  R  G  G  I  I  D  E  S  D  L  L  D  A  L  N  A  - 251

 793 - GGTAAATGCGCTGGTGCAGGCCTGGACGTATTTTCCAGTGAACCACCATCAGGTGTTGTCAGTGAATTAGTT - 864
 252 -  G  K  C  A  G  A  G  L  D  V  F  S  S  E  P  P  S  G  V  V  S  E  L  V  - 275

 865 - AAGCACCCGAGGGTGGTATGCACACCACATCTTGGTGCGAGCACAGCGGAAGCTCAAGTGCGTGTTGCTAAG - 936
 276 -  K  H  P  R  V  V  C  T  P  H  L  G  A  S  T  A  E  A  Q  V  R  V  A  K  - 299

 937 - GAGATTGCTGAACAAATCGTTGACGCATGCAATGGCAAGACAGCCGTTGGATTAGTTAATGCTCCGGCCATT - 1008
 300 -  E  I  A  E  Q  I  V  D  A  C  N  G  K  T  A  V  G  L  V  N  A  P  A  I  - 323

1009 - AGTGAGGCGGGAAAGGAAGATGTTAAGCCTTGGATGGCTCTTGGACAAGCCCTTGGAGCAGTCTTATGCAAG - 1080
 324 -  S  E  A  G  K  E  D  V  K  P  W  M  A  L  G  Q  A  L  G  A  V  L  C  K  - 347

1081 - ATTTCTCCAACACTTCCAAAATTAGTCAACATTAAGACACGAGGAGAGAAACCGAAAGGTTTAACAAGGGCG - 1152
 348 -  I  S  P  T  L  P  K  L  V  N  I  K  T  R  G  E  K  P  K  G  L  T  R  A  - 371

1153 - TTGACTTCAGCCGTCAGTCTTGGGCTCTTGAAATTCCAAGGAGGAAATATTAATCTCCTAAATGCGCCATCT - 1224
 372 -  L  T  S  A  V  S  L  G  L  L  K  F  Q  G  G  N  I  N  L  L  N  A  P  S  - 395

1225 - GTGGCAAAGGAAAAGGGGATCCAGATTCAAGTGTCGGATGATGTCGTCCAAAAATGTTGTTTTTCAACATGT - 1296
 396 -  V  A  K  E  K  G  I  Q  I  Q  V  S  D  D  V  V  Q  K  C  C  F  S  T  C  - 419

1297 - CTGTCGTTGTCAAGTGACAATGGTACCGTGCAAGCCAGTGGTACAGTTTTAGATAACAAAGCCGTTTTAACC - 1368
 420 -  L  S  L  S  S  D  N  G  T  V  Q  A  S  G  T  V  L  D  N  K  A  V  L  T  - 443

1369 - AACTGTTTAGGTGTAGCTCTTGACAATCCGTTAATCCTTGGTAATTGCCTTGTGATTGGATCGGGAAAATTG - 1440
 444 -  N  C  L  G  V  A  L  D  N  P  L  I  L  G  N  C  L  V  I  G  S  G  K  L  - 467

1441 - TCAGAAAAATCTAGCCTGATCTCAGCTTTAGTTGGGAAAATTACTTCTGTGAAGGCATTTGTTACTGGGGCA - 1512
 468 -  S  E  K  S  S  L  I  S  A  L  V  G  K  I  T  S  V  K  A  F  V  T  G  A  - 491

1513 - GACAATGATAGATGCGTTATTGCTGCCTCTGTTGCTGAAGTCTGTGAAGAAGCTATAGGGAATGGGTTTTGT - 1584
 492 -  D  N  D  R  C  V  I  A  A  S  V  A  E  V  C  E  E  A  I  G  N  G  F  C  - 515
```

Figure 3.4 Full-length cDNA (clone 3-9-1) encoding 3-PGDH obtained by immunoscreening the *C. fleckeri* cDNA library with CSL box jellyfish antivenom. Accession No. FJ460242.

```
1585 - TGTATCTCTCTTGTATGAGATGTTTAAGTGTGAAGTTGGACTTCGTCTTGTTAAGATGAAACCCTGTGAGTG - 1656
 516 -  C  I  S  L  V  *                                                       - 520

1657 - AGAACCAGTATAAAAAAAAAAAAAAAAAA                                            - 1685
```

Figure 3.4 (*continued*)

```
   1 - TTAAGTGCAGAATTCTCTAAACGTCCGTAATACGGTCAATTTTGCGTCAAACCGAAAATGTCGAAAGGACCA - 72
   1 -                                                           M  S  K  G  P  - 5

  73 - GCTATAGGAATCGACCTTGGAACAACCTATTCATGCGTTGGCGTATTTCAACATGGCAAAGTCGAAATTATT - 144
   6 -  A  I  G  I  D  L  G  T  T  Y  S  C  V  G  V  F  Q  H  G  K  V  E  I  I  - 29

 145 - GCAAATGACCAAGGTAATAGGACAACGCCAAGCTACGTCGCCTTCACTGATACCGAAAGGTTGATTGGTGAT - 216
  30 -  A  N  D  Q  G  N  R  T  T  P  S  Y  V  A  F  T  D  T  E  R  L  I  G  D  - 53

 217 - TCAGCCAAAAACCAGGTTGCATTGAATCCTACGAACACAGTATTTGATGCAAAACGTCTGATTGGAAGAAGA - 288
  54 -  S  A  K  N  Q  V  A  L  N  P  T  N  T  V  F  D  A  K  R  L  I  G  R  R  - 77

 289 - TTTGATGATGCTGGTGCCCAGGCTGATATGAAACACTGGCCATTCAAGGTTATAAATGATGGTGGTCGGCCA - 360
  78 -  F  D  D  A  G  A  Q  A  D  M  K  H  W  P  F  K  V  I  N  D  G  G  R  P  - 101

 361 - AAAATTGAAGTTTCTTACAAGGGTGAGATAAAATCGTTTTATGCTGAAGAAATCTCCTCCATGGTTTTAACA - 432
 102 -  K  I  E  V  S  Y  K  G  E  I  K  S  F  Y  A  E  E  I  S  S  M  V  L  T  - 125

 433 - AAGATGAAAGAAGTTGCCGAGGCTTACCTTGGAAAGAACTGTCAAAATGTTGTTGTCACTGTTCCAGCATAT - 504
 126 -  K  M  K  E  V  A  E  A  Y  L  G  K  N  C  Q  N  V  V  V  T  V  P  A  Y  - 149

 505 - TTCAATGATTCTCAGAGGCAAGCTACAAAGGATGCCGGTACTATAGCTGGGTTAAATGTGATGCGGATTATC - 576
 150 -  F  N  D  S  Q  R  Q  A  T  K  D  A  G  T  I  A  G  L  N  V  M  R  I  I  - 173

 577 - AATGAACCTACTGCAGCTGCTATTGCCTATGGCTTGGACAAAAAGGTTGGAGCTGAAAGAAATGTTCTTATC - 648
 174 -  N  E  P  T  A  A  A  I  A  Y  G  L  D  K  K  V  G  A  E  R  N  V  L  I  - 197

 649 - TTCGACCTCGGAGGAGGAACATTTGATGTGTCAATCCTTACAATTGAGGATGGTATCTTTGAGGTCAAAGCC - 720
 198 -  F  D  L  G  G  G  T  F  D  V  S  I  L  T  I  E  D  G  I  F  E  V  K  A  - 221

 721 - ACAGCTGGTGATACACATTTAGGTGGTGAGGACTTTGATAACAGGATGGTCAATCACTTTACCAAGGAATTC - 792
 222 -  T  A  G  D  T  H  L  G  G  E  D  F  D  N  R  M  V  N  H  F  T  K  E  F  - 245

 793 - AGGACCAAGTACAAGAAAGATATTACAGGTAACAAGAGAGCAGTAAGAAGACTTCGCACTGCATGTGAGCGA - 864
 246 -  R  T  K  Y  K  K  D  I  T  G  N  K  R  A  V  R  R  L  R  T  A  C  E  R  - 269

 865 - GCCAAAAGAACCTTGTCATCAAGCAAACAAGCTAGCATTGAGATTGATTCCTTGTTTGAAGGCATTGACTTC - 936
 270 -  A  K  R  T  L  S  S  S  K  Q  A  S  I  E  I  D  S  L  F  E  G  I  D  F  - 293

 937 - TACACATCAATTACAAGGGCCAAATTTGAGGAACTGAATGGAGACCTATTTAGAGGCACTATTGAGCCTGTG - 1008
 294 -  Y  T  S  I  T  R  A  K  F  E  E  L  N  G  D  L  F  R  G  T  I  E  P  V  - 317

1009 - GAAAAGGCCATGCGTGATACTAAGCTTGAAAAGAAGGAAATTCATGACATTGTACTTGTTGGTGGATCAACA - 1080
 318 -  E  K  A  M  R  D  T  K  L  E  K  K  E  I  H  D  I  V  L  V  G  G  S  T  - 341

1081 - CGTATACCAAAGATACAGCAACTTTTATCAGATACTTTTAATGGGAAGGAGTTAAACAAATCCATTAACCCT - 1152
 342 -  R  I  P  K  I  Q  Q  L  L  S  D  T  F  N  G  K  E  L  N  K  S  I  N  P  - 365

1153 - GATGAGGCTGTTGCTTATGGTGCTGCTGTGCAAGCTGCCATCCTCCAAGGTGACAAAAGTGAAGAAGTTTCA - 1224
 366 -  D  E  A  V  A  Y  G  A  A  V  Q  A  A  I  L  Q  G  D  K  S  E  E  V  S  - 389

1225 - GATCTGCTTCTTTTGGATGTTGCACCGCTGTCTCTGGGTATCGAGACAGCAGGAGGAGTCATGACTGCACTG - 1296
 390 -  D  L  L  L  L  D  V  A  P  L  S  L  G  I  E  T  A  G  G  V  M  T  A  L  - 413

1297 - ATTAAAAGAAATTCCACAATCCCAACCAAACAGCAACAGGTTTTCACAACTTATGCAGACAATCAGCCTGGT - 1368
 414 -  I  K  R  N  S  T  I  P  T  K  Q  Q  Q  V  F  T  T  Y  A  D  N  Q  P  G  - 437

1369 - GTCTTGATTCAAGTTTATGAAGGTGAAAGGGGGATGACTAAGGATAACAACTTGTTGGGTAAATTTGAATTG - 1440
 438 -  V  L  I  Q  V  Y  E  G  E  R  G  M  T  K  D  N  N  L  L  G  K  F  E  L  - 461

1441 - AGTGGAATTCCTCCAGCACCCCGTGGTGTACCACAGATTGAAGTAACTTTTGATATTGATGCCAATGGTATT - 1512
 462 -  S  G  I  P  P  A  P  R  G  V  P  Q  I  E  V  T  F  D  I  D  A  N  G  I  - 485

1513 - TTGAACGTGTCTGCTGTAGACAAAAGTACTGGCAAGGAAAACAAGATCACAATCACAAATGACAAAGGTCGT - 1584
 486 -  L  N  V  S  A  V  D  K  S  T  G  K  E  N  K  I  T  I  T  N  D  K  G  R  - 509
```

Figure 3.5 Full-length cDNA (clone 7-1) encoding HSP70-1, obtained by DNA screening the *C. fleckeri* cDNA library with a gene-specific radiolabelled oligonucleotide probe. Accession No. FJ460243.

```
1585 - TTGTCTAAAGAAGATATAGAACGCATGGTCAATGAAGCTGAAAAATTCAAACAAGAAGATGAGCAGCAACGA - 1656
 510 -  L  S  K  E  D  I  E  R  M  V  N  E  A  E  K  F  K  Q  E  D  E  Q  Q  R  - 533

1657 - GACCGAATCCAAGCAAAGAACAGTTTGGAAAGTTATGCTTACCAGATGAAAAGTACAGTCGAAGATGACAAA - 1728
 534 -  D  R  I  Q  A  K  N  S  L  E  S  Y  A  Y  Q  M  K  S  T  V  E  D  D  K  - 557

1729 - GTGAAAGACAAAATAAGTGAAGAGGACAGAAAGACAATTGTGGACAAATGTAAGGAGACTTTGGACTGGTTA - 1800
 558 -  V  K  D  K  I  S  E  E  D  R  K  T  I  V  D  K  C  K  E  T  L  D  W  L  - 581

1801 - GACAGAAACCAGATGGCTGAGAAGGACGAATTTGATAGTCAGCAGAAAGAACTGGAAAAAGTTTGTGCACCA - 1872
 582 -  D  R  N  Q  M  A  E  K  D  E  F  D  S  Q  Q  K  E  L  E  K  V  C  A  P  - 605

1873 - ATCATAACCAAAATGTACCAAAGTGCGGGTGGAATGCCTGGTGGTGGTATGCCTGGCGGTATGCCTGGTGGA - 1944
 606 -  I  I  T  K  M  Y  Q  S  A  G  G  M  P  G  G  G  M  P  G  G  M  P  G  G  - 629

1945 - TTTCCTGGTGGACCTGGAGGCCAGCCTGAGGGGGGTAGTGCTGGTGGACCAACAATTGAAGAAGTTGATTGA - 2016
 630 -  F  P  G  G  P  G  G  Q  P  E  G  G  S  A  G  G  P  T  I  E  E  V  D  *  - 652

2017 - ACATTGTTTAGCATCGGGGCTGTTTATTGTACACGGATGTTGTCGAAAATGTTAGAGCCCAGTCTTTAAAAA - 2088
2089 - GAAAAAAAAAATGACGCGGTTCCCGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA            - 2149
```

Figure 3.5 (*continued*).

```
   1 - CTCGATCCGGCATCTGAATTAGCAGGCCGTGTTACTAGCAGACTTCTTCGTTTTGACTTTCAAGTGGCCAGA - 72

  73 - ATGAGCGGAAGGTTTGTCAGAGCGTCGAAATTTCGCCATGTCTTCGGGACGGCGGCAAGGAAGGAACAGTGT - 144
   1 -  M  S  G  R  F  V  R  A  S  K  F  R  H  V  F  G  T  A  A  R  K  E  Q  C  - 24

 145 - TATGAGAATGTGACCATCACAAGAAGTTCCTTGGAGGGCAAATTTTGTGCTGCCAATTCTAAGTTTGTCGCT - 216
  25 -  Y  E  N  V  T  I  T  R  S  S  L  E  G  K  F  C  A  A  N  S  K  F  V  A  - 48

 217 - GTTTGCATTTATGCTGGTGGTGGGGGAAGTTTTCTGGTTATTCCTCATGATAGGTATGGCAGACAAGACATA - 288
  49 -  V  C  I  Y  A  G  G  G  G  S  F  L  V  I  P  H  D  R  Y  G  R  Q  D  I  - 72
                1→
 289 - AATGCACCCAAGGTTGCTGGACACTCAAATGATGTCTTGGACCTAGAGTGGAACCCTTTTAATGATAACATG - 360
  73 -  N  A  P  K  V  A  G  H  S  N  D  V  L  D  L  E  W  N  P  F  N  D  N  M  - 96

 361 - ATTGCCTCAAGTGATGAAGATGGTGCCATTAAAATTTGGGAGATCCCTGATGGTGGACTTACAACAAACCTA - 432
  97 -  I  A  S  S  D  E  D  G  A  I  K  I  W  E  I  P  D  G  G  L  T  T  N  L  - 120
                                                 2→
 433 - AATTCACCACTCCTTACAGTCAACCACCAGAAAAAGGTGTCTCAGCTCAGTTGGCATCCAGTTGCTGCCAAT - 504
 121 -  N  S  P  L  L  T  V  N  H  Q  K  K  V  S  Q  L  S  W  H  P  V  A  A  N  - 144

 505 - ATCCTGCTAAGTGTTTCATATGAACCATGTGTAGCAGTGTGGAATCTTGAAACAGGGGAACAGGTTACGGAA - 576
 145 -  I  L  L  S  V  S  Y  E  P  C  V  A  V  W  N  L  E  T  G  E  Q  V  T  E  - 168
                                                 3→
 577 - ATTGAACATCCTGACCAAGTGTATAATGCCGAATGGAATGAAATTGGAAGCAAAATTGTATCTGCATGCAAA - 648
 169 -  I  E  H  P  D  Q  V  Y  N  A  E  W  N  E  I  G  S  K  I  V  S  A  C  K  - 192

 649 - GATAAGTTCTTCAGAATCATAAACCCACGAACAGGAGAGGTTATTCAGAAATTTAAGGGACATGTTGGAGGA - 720
 193 -  D  K  F  F  R  I  I  N  P  R  T  G  E  V  I  Q  K  F  K  G  H  V  G  G  - 216
                               4→
 721 - AAACCACAGCGTGTCATTTTCCTTACTGATGAAAGACTATTTTCAACTGGATTTACAAAGATGAGTTCAAGA - 792
 217 -  K  P  Q  R  V  I  F  L  T  D  E  R  L  F  S  T  G  F  T  K  M  S  S  R  - 240

 793 - GAATATGCTGTTTGGGACTTGAAAAACACAAGCGAGGCCTTGTTTGAAGACCAATTGGATTCAGCAAATGGT - 864
 241 -  E  Y  A  V  W  D  L  K  N  T  S  E  A  L  F  E  D  Q  L  D  S  A  N  G  - 264
                         5→
 865 - ACTTTGGTACCATACTATGATCGAGATACCAAAATTCTCTACGTCGCTGGAAGGGGAGACAGCACGATAAGG - 936
 265 -  T  L  V  P  Y  Y  D  R  D  T  K  I  L  Y  V  A  G  R  G  D  S  T  I  R  - 288

 937 - TATTATGAAATAACTGACGAGGAACCATATGCTCATTGGATTACAAATTTCCAATCCAAGTTACCGCACCGT - 1008
 289 -  Y  Y  E  I  T  D  E  E  P  Y  A  H  W  I  T  N  F  Q  S  K  L  P  H  R  - 312

1009 - GGGGTTTGTATGTCATCTAAGCGTGCAGTTGATGTGAATCGCAATGAAGTAGCACTTTTTTATCGCCTCCTG - 1080
 313 -  G  V  C  M  S  S  K  R  A  V  D  V  N  R  N  E  V  A  L  F  Y  R  L  L  - 336

1081 - AGTGACAAACACATGGTTGAACCAGTATCATTCACTGTGCCGAGAAAGGGGGATACTTTCCAACAAGATATT - 1152
 337 -  S  D  K  H  M  V  E  P  V  S  F  T  V  P  R  K  G  D  T  F  Q  Q  D  I  - 360

1153 - TACCCGGATACTCCTGGAGATGAACCGGCTTTGACTGCTGATGAATGGTGTCAGGGGAAAGACGCCGAACCA - 1224
 361 -  Y  P  D  T  P  G  D  E  P  A  L  T  A  D  E  W  C  Q  G  K  D  A  E  P  - 384

1225 - AAGAGAGTTTCTATGCAAGATTTTTTCATGTCCAAGGACAAAAAGAAAGCAACAGGATTAGGTCACAAGGTC - 1296
 385 -  K  R  V  S  M  Q  D  F  F  M  S  K  D  K  K  K  A  T  G  L  G  H  K  V  - 408

1297 - GGATCTAAAACAGGTTTGAAACCTGCTGCTGTTTCAGCTGCTCCAGCCGCTTCAGCCACCTCAGCCTCAACA - 1368
 409 -  G  S  K  T  G  L  K  P  A  A  V  S  A  A  P  A  A  S  A  T  S  A  S  T  - 432

1369 - CCGGCTGCGACAGCTGCAGATTCCGCTGAGGTAGTTCACCTTCAAAGCGAGATTAAGGGCCTGAAAGATAAC - 1440
 433 -  P  A  A  T  A  A  D  S  A  E  V  V  H  L  Q  S  E  I  K  G  L  K  D  N  - 456
```

Figure 3.6 Full-length cDNA (clone 13-5) encoding ABP (coronin 1B), obtained by DNA screening the *C. fleckeri* cDNA library with a gene-specific radiolabelled oligonucleotide probe. Accession No. FJ460244. A highly basic N-terminal amino acid signature sequence (pI 12.5) characteristic of coronin proteins is underlined in blue. WD40 peptide repeats (5) are underlined in red. Conserved residues in each WD40-repeat domain are also highlighted in red. A putative C-terminal coiled coil region is underlined in green.

```
1441 - GAGAAGAAACTTAAAGACGAAATTACATCGCTGCGTAGTGAGGTCAGTGAGTTGAAGAAACTAAAGGATGAA - 1512
 457 -  E  K  K  L  K  D  E  I  T  S  L  R  S  E  V  S  E  L  K  K  L  K  D  E  - 480

1513 - TTTGCAGCAATCAAAGGAGCAGTAAAAAGGAATGCGGAGCGAGTTGACGCTCTTGAATCCATGGTACAAGAA - 1584
 481 -  F  A  A  I  K  G  A  V  K  R  N  A  E  R  V  D  A  L  E  S  M  V  Q  E  - 504

1585 - GTTAGCGATGATGAGGCGAAATAAAGTTGTGTGCAGTCAATCTTGTACAAGATTCCGTATAGAAGTGATTTG - 1656
 505 -  V  S  D  D  E  A  K  *                                                  - 511

1657 - TTAGCCGAATTTGATACATGACACTATTTTTTTCAGAAAATCGAAGACAAAACCGTAGGACAAATTTAGGAA - 1728
1729 - TATACTGTGGACTTGATCTCTTGTTATCGCAACACTGCAGTGCTTGTTGTAACACGCACTCTGCTTGATTAT - 1800
1801 - ACGAGTGAATATTCTGCTCTAAGTACGCCATTGGAATTGTACAACTTCATAAAAGAACAAGACAGATGGAAC - 1872
1873 - AAGGGGGTAGATCAAAAAAAAAAAAAAAAAAA                                         - 1903
```

Figure 3.6 (*continued*)

```
   1 - CATGTTGAGACACAAGAAAAAGTAGTCCTGCCAACCAAAGAAGTGATACAAGAGGAGGCAACAGATTCAAGG - 72
   1 -  H  V  E  T  Q  E  K  V  V  L  P  T  K  E  V  I  Q  E  E  A  T  D  S  R  - 24

  73 - GCTGAAGTCAAATCATTTGATCATTCCAAGCTCAAGCATGTAGAGACACAGGAGAAGAATGAATTACCAACA - 144
  25 -  A  E  V  K  S  F  D  H  S  K  L  K  H  V  E  T  Q  E  K  N  E  L  P  T  - 48

 145 - CCACAAATGTTAAAATTAGAAATGAAACCAGATGAGCTTCCTGATGTTTCTGAAGTGAGCAAGTTTGACACT - 216
  49 -  P  Q  M  L  K  L  E  M  K  P  D  E  L  P  D  V  S  E  V  S  K  F  D  T  - 72

 217 - GGAAAGTTAAAGCATGTAGAGACACAGGAAAAAGTAGTCCTGCCAACAAAAGAAGTAATACAAGAGGAGGCA - 288
  73 -  G  K  L  K  H  V  E  T  Q  E  K  V  V  L  P  T  K  E  V  I  Q  E  E  A  - 96

 289 - ACAGATTCAAGAGCTGAGGTCAAATCATTTGATCATTCCAAACTTAAGCATGTAGAGACACAGGAGAAGAAT - 360
  97 -  T  D  S  R  A  E  V  K  S  F  D  H  S  K  L  K  H  V  E  T  Q  E  K  N  - 120

 361 - GAATTACCAACACCACAAATGTTAAAATTAGAAATGAAACCAGATGAGCTTCCTGATGTTTCGGGAGTGAGC - 432
 121 -  E  L  P  T  P  Q  M  L  K  L  E  M  K  P  D  E  L  P  D  V  S  G  V  S  - 144

 433 - CATTTTGACACAGGAAAGCTAAAGCATGTTGAGACACAAGAAAAAGTAGTCCTTCCAACAAAAGAGGTTATT - 504
 145 -  H  F  D  T  G  K  L  K  H  V  E  T  Q  E  K  V  V  L  P  T  K  E  V  I  - 168

 505 - CAAGAAGAGGCTACGGATTCAAGGGCTGAAGTCAAATCCTTTGATCATTCCAAACTTAAGCATGTAGAGACA - 576
 169 -  Q  E  E  A  T  D  S  R  A  E  V  K  S  F  D  H  S  K  L  K  H  V  E  T  - 192

 577 - CAAGAGAAGAATGAATTACCAACACCACAAATGTTAAAATTAGAAATGAAACCAGATGAGCTTCCTGATGTT - 648
 193 -  Q  E  K  N  E  L  P  T  P  Q  M  L  K  L  E  M  K  P  D  E  L  P  D  V  - 216

 649 - TCTGAAGTTAGCAAGTTTGACACTGGAAAGTTGAAACATGTAGAGACGCAAGAAAAGGTAGTCCTCCCTACA - 720
 217 -  S  E  V  S  K  F  D  T  G  K  L  K  H  V  E  T  Q  E  K  V  V  L  P  T  - 240

 721 - AAAGAGGTAATACAAGAGGAGGCAACAGACTCAAGAGCTGAGGTCAAATCCTTTGATCATTCCAAACTTAAG - 792
 241 -  K  E  V  I  Q  E  E  A  T  D  S  R  A  E  V  K  S  F  D  H  S  K  L  K  - 264

 793 - CATGTAGAGACACAGGAGAAGAATGAATTACCAACACCACAAATGTTGAAATTAGAAATGAAACCAGATGAG - 864
 265 -  H  V  E  T  Q  E  K  N  E  L  P  T  P  Q  M  L  K  L  E  M  K  P  D  E  - 288

 865 - CTTCCTGATGTTTCAGAAGTGAACAAGTTTGACACAGGAAAATTAAAGCATGTAGAGACACAAGAGAAAAAC - 936
 289 -  L  P  D  V  S  E  V  N  K  F  D  T  G  K  L  K  H  V  E  T  Q  E  K  N  - 312

 937 - ATGTTACCAACGGCTGATGTCATTAAGGAGGAAGTGGCAGACTCCAGAGCGGAAGTCAAGTCTTTTGATCAT - 1008
 313 -  M  L  P  T  A  D  V  I  K  E  E  V  A  D  S  R  A  E  V  K  S  F  D  H  - 336

1009 - ACGAAGCTCAAGCATGTGGAAACGCAAGAGAAGAATCCATTGCCTACTCCACAAACATTGAAAGAAGAATTG - 1080
 337 -  T  K  L  K  H  V  E  T  Q  E  K  N  P  L  P  T  P  Q  T  L  K  E  E  L  - 360

1081 - CGACCAGATTCCCTGCCCGATGTATCAGAAGTGAGCGAATTTGATTCTGGAAAATTGAAACATGTTGAAACA - 1152
 361 -  R  P  D  S  L  P  D  V  S  E  V  S  E  F  D  S  G  K  L  K  H  V  E  T  - 384

1153 - CAAGAAAAAAGTGTGCTTCCAACAAAGGAAGTAATCCAGGAAGAGGCAGCAGATTCCAGGGCAGAGGTCAAG - 1224
 385 -  Q  E  K  S  V  L  P  T  K  E  V  I  Q  E  E  A  A  D  S  R  A  E  V  K  - 408

1225 - TCATTCGACCATTCGAAACTAAAACACGTTCAAACTGAAGAGAAGAACCCGTTGCCAACAGCCGGAACAATT - 1296
 409 -  S  F  D  H  S  K  L  K  H  V  Q  T  E  E  K  N  P  L  P  T  A  G  T  I  - 432

1297 - CGAGAAGAGCTTAGACCCGACGAATTTCCTGATCGGTCAGAAGTCGCCGATTTCGATACCAGAAAACTAAAA - 1368
 433 -  R  E  E  L  R  P  D  E  F  P  D  R  S  E  V  A  D  F  D  T  R  K  L  K  - 456

1369 - CATGTAGAAACGACAGAGAAAAATGTATTACCTACGAAAGAAGTGATCAAAGAGGAAGCAGTTGATAGCAGA - 1440
 457 -  H  V  E  T  T  E  K  N  V  L  P  T  K  E  V  I  K  E  E  A  V  D  S  R  - 480

1441 - GCTGAGGTTAAGTCGTTTGATCATGGCAAGTTAAAGCATGTGGAAACAGAAGAGAAGAACACGTTACCGTCT - 1512
 481 -  A  E  V  K  S  F  D  H  G  K  L  K  H  V  E  T  E  E  K  N  T  L  P  S  - 504
```

Figure 3.7 Incomplete cDNA (clone 3-9) encoding thypedin obtained by DNA screening the *C. fleckeri* cDNA library with a gene-specific radiolabelled oligonucleotide probe. Accession No. FJ460245. Approximately 423 residues are missing from N-terminus of the complete protein.

```
1513 - GCTGGTACTTTGAAGGAAGAAATGCGACCAGAATTTCTGCCAGATGTGTCAGAAGTAAAGGACTTCAATACT - 1584
 505 -  A  G  T  L  K  E  E  M  R  P  E  F  L  P  D  V  S  E  V  K  D  F  N  T  - 528

1585 - GGAAAGTTGAAACACGTTCAAACCGAAGAGAAAACTGTCCTCCCCACAAAGGAAGTCATCAATCAAGAGGCG - 1656
 529 -  G  K  L  K  H  V  Q  T  E  E  K  T  V  L  P  T  K  E  V  I  N  Q  E  A  - 552

1657 - CTTGATAGTCGAGCAGAGGTGAAAACATTCGATCACACCAAATTGAAACACGTGGAAACACAAGAGAAGAAC - 1728
 553 -  L  D  S  R  A  E  V  K  T  F  D  H  T  K  L  K  H  V  E  T  Q  E  K  N  - 576

1729 - GTCCTACCCACCCCTCAAGTATTAAAGGAAGAACTTAGGCCGGATGCCCTAGCTGACGTCTCGGAGGTCAAG - 1800
 577 -  V  L  P  T  P  Q  V  L  K  E  E  L  R  P  D  A  L  A  D  V  S  E  V  K  - 600

1801 - GATTTCGATCAAGGAAAATTAAAACATGTGGAAACCAAAGAAAAAGTCACGGTACCGTCAAAGGAGGTCATA - 1872
 601 -  D  F  D  Q  G  K  L  K  H  V  E  T  K  E  K  V  T  V  P  S  K  E  V  I  - 624

1873 - GACGAGGAAGCCAAAAACACGATGGCAGAAGTAACAAAGTTCGACAAAGGCAAACTGAAGCACGCCGATACA - 1944
 625 -  D  E  E  A  K  N  T  M  A  E  V  T  K  F  D  K  G  K  L  K  H  A  D  T  - 648

1945 - AAAGAAAAAGACAAACTACCAACAGCGGAGGATATCAAGAACGAAAAGAGCGCTTGACCTCTCGTTCCTCAA - 2016
 649 -  K  E  K  D  K  L  P  T  A  E  D  I  K  N  E  K  S  A  *                 - 666

2017 - TACTCTTGGATATTGGAAGTCGGTTACCCTGCACGGAGCTGGATTTGATTTCTGCATTTTCTCCGTTGTACT - 2088
2089 - ACAGACTACTGGAATCAATGCATCTTAAACTCTCCTATTCAGGAGGATATTTTTCTATTTTCTACGCGATTC - 2160
2161 - AATAATTGAACTTTTTCGGCAATTCACATCTGGAATTTTCATACTATTTAATAATGCTTTAAAATACGATTC - 2232
2233 - ATGCACGTGCTTACGCACCTGGCGATACCCTGCATGTAGTGATTAAGTCATGCAGTTCAGTTTGTCGTTTGA - 2304
2305 - TTGTGTACTGTTTTTACTGAATCGGATATGTGGTCGTCTCACTTTTACTTCATAAAAAAAAAAAAAAAAAAA - 2376
2377 - AA                                                                       - 2378
```

Figure 3.7 (*continued*)

3.3.2 Immunodetection of cDNA Clones using Mouse and Rabbit Antibodies Raised Against *C. fleckeri* Nematocyst Venom

In comparison to the CSL box jellyfish antivenom experiment, immunoscreening ~1.4 × 10^5 cDNA library plaques with mouse antibodies raised against *C. fleckeri* nematocyst extracts resulted in the detection of 20 antigenic clones, constituting 0.014% of total library clones. However, when the 20 original clones were rescreened, only 4 expressed clones (14, 17, 19 and 20) were strongly reactive to the mouse antibodies. One of the antigenic clones (19-1) was not successfully excised from the Uni-Zap XR lambda vector. Of the remainder, 3 independent clones (14-1, 17-2 and 20-2) were partially sequenced. BLAST analysis revealed that all the translated clone sequences shared significant homology with HSP70 and the best match was isolated from the scallop, *Clamys farreri* (655 residues; Accession No. AAO38780; *E*-value = 4E-100). The translated consensus sequence of the 3 independent clones, designated HSP70-2 (Appendix 4), was also similar (~80%) to HSP70-1 (Figure 3.5), which was initially detected by immunoscreening the cDNA library with CSL box jellyfish antivenom.

Immunoscreening ~1 × 10^5 cDNA library plaques with rabbit antibodies raised against *C. fleckeri* nematocyst venom also resulted in the detection of 20 antigenic clones, constituting 0.02% of total library clones. However, when the clones were rescreened, the plaques of only 7 original clones (7, 8, 9, 10, 11, 13, and 15) were strongly antigenic. Ten clones, representing the 7 original antigenic clones, were partially sequenced (clones 9, 11 and 13 were isolated and sequenced in duplicate). BLAST analysis revealed that 7 translated clone sequences (9-1, 9-2, 10-3, 11-5, 11-7, 13-8 and 13-10) shared significant homology with calreticulin and the translated sequences of the remaining clones (7-3, 8-10, and 15-2) were homologous to a DNA excision repair protein. The best match for the calreticulin homologue (69% amino acid similarity) was isolated from the sea urchin, *Strongylocentrotus purpuratus* (421 residues; Accession No. AAD55725; *E*-value = 4E-149), whereas the best match for the DNA excision repair homologue originated from the monkey, *Macaca mulatto* (1496 residues; Accession No. XP_001107536; *E*-value = 2E-116). Although a putative translation initiating start site and an ORF were detected in the nucleotide sequence encoding calreticulin (Appendix 4), no in-frame stop codon was detected. As the translated *C. fleckeri* clone contains only 377 residues, the 3'-end of the cDNA clone may have been prematurely truncated during reverse transcription for library construction. The consensus nucleotide and translated nucleotide sequences of clones partially encoding the DNA excision repair protein are provided in Appendix 4.

3.3.3 Failure to Detect Putative Venom Clones

During this study, a cDNA expression library constructed from *C. fleckeri* tentacle mRNA was independently screened three times using polyclonal antibodies raised in sheep, mouse and rabbit against either milked venom or nematocyst venom. In each case, no putative venom clones were detected. In particular, immunoscreening the library using CSL box jellyfish antivenom resulted in the detection of at least 38 different genes out of 41 (>90%) that were unrelated to previously reported venom proteins or toxins. These results clearly indicate that the CSL box jellyfish antivenom, which is currently used by clinicians for the treatment of life-threatening *C. fleckeri* stings, is not highly specific – a finding that may have important implications for the efficacy and potential side-effects of the antivenom. The majority of identified *C. fleckeri* proteins are homologous to eukaryotic proteins and the best matches are often proteins derived from marine organisms including cnidarians such as sea anemones and hydra (see Table 3.2). Therefore, the strong antigenicity of these non-venom proteins suggests that the milked venom used for CSL antivenom production may be contaminated with cellular

proteins from *C. fleckeri* tentacle tissue other than the contents of nematocysts. In addition, the majority of full-length proteins obtained in the study are predicted to lack signal peptides or retention signals (SignalP 3.0, Bendtsen et al., 2004; PSORT II, Horton and Nakai, 1997), thus suggesting that these proteins are cytoplasmic and unlikely to be localised within nematocysts. Interestingly, many of the identified proteins such as EF1α, HSP70-1, filamin, ABP, thypedin, non-muscle myosin and calponin, exhibit actin-binding properties. Therefore, a substantial number of the sequenced clones ($\geq$ 42%) may have been detected due to strong binding of the CSL antibodies to epitopes in a conserved actin-binding domain.

In an effort to overcome the apparent lack of specificity conferred by the commercially-available antivenom, small-scale production of mouse and rabbit antibodies against *C. fleckeri* nematocyst venom was undertaken. The nematocyst venom was prepared from tentacle-free nematocysts to avoid contamination with non-venom jellyfish proteins. Moreover, the nematocyst venom and antigen emulsions were prepared aseptically to minimise bacterial protein contamination. However, despite a dramatic increase in the apparent specificity of the mouse and rabbit antibodies (~10-fold), no venom clones were identified.

Several reasons could explain the failure of the immunoscreening technique to detect cDNA clones that express *C. fleckeri* venom proteins. Firstly, the venom clones may have been under-represented in the cDNA library because whole tentacle mRNA, containing cells other than nematocytes, was used for its construction. In contrast, previous researchers that successfully used immunodetection strategies to isolate venom protein cDNA clones (e.g. Welton and Burnell, 2005; Assakura et al., 2003; Fang et al., 1988; Jones et al., 1992), used venom glands of the target organisms for construction of the cDNA expression libraries, thus excluding non-venom-producing cell types. Secondly, the venom clones may have been under-represented in the cDNA library due to the age of the box jellyfish used for mRNA extraction. At the time of library construction, only juvenile *C. fleckeri* specimens were available, hence the expression levels of venom proteins may have been lower than those of adult jellyfish. Thirdly, transcription and/or translation of the venom clones may have been inhibited in the *E. coli* host cells, resulting in very low levels of protein expression that could not be detected with the sheep, mouse or rabbit antibodies. Alternatively, small amounts of the transcription and/or translation products may have been toxic to the bacterial host, causing inhibition of cell growth and/or cell death.

If the failure to detect immunogenic venom clones is indeed the result of low protein expression levels in the prokaryote host, random nucleotide sequencing of the library may offer

a better alternative for the detection of cDNA encoding *C. fleckeri* venom proteins. Also, as demonstrated in Chapter 5, another effective approach to isolate *C. fleckeri* venom proteins involves screening the library with gene-specific oligonucleotide probes encoding partial peptide sequences of major nematocyst proteins.

3.3.4 Elongation Factor-1α

The most commonly isolated cDNA in the CSL immunodetection experiment, representing 29% of the sequenced clones, encoded EF1α. Three conserved EF1α structural domains (I, II and III) were predicted in *C. fleckeri* EF1α spanning deduced residues 9-239, 242-232 and 335-438, respectively. Furthermore, several specific features/motifs were detected in Domain I, including a GTP/Mg^{2+} binding site, ạ binding site for the nucleotide exchange factor, EF1β (also known as EF1Balpha), and two highly conserved surface loops (Switch Regions I and II) that undergo conformational changes upon GTP binding (Figure 3.3).

EF1α (also known as eEF1A) is the alpha subunit of EF1 and the eukaryotic paralogue of bacterial EF-Tu and archaeal aEF1α. It is a multifunctional G-protein that is not only involved in numerous translation-related functions but also interacts with components of signal transduction pathways and the cytoskeleton (Moore and Cyr, 2000). EF1α is highly abundant in the cytosol and its role during protein synthesis is to bind and transport aminoacyl-tRNA to the A site of the ribosome in a GTP-dependent mechanism (Liu et al., 2002). EF1α also interacts strongly with the actin and plays an essential role in cytoskeleton organisation and cell morphology (Gross and Kinzy, 2005).

EF1α is widely used as a phylogenetic marker for studying eukaryote evolution (e.g. Kobayashi et al., 1996; Moreira et al., 1999; Gaucher et al., 2001). Therefore, to investigate the evolutionary position of cubozoan jellyfish in EF1α-based phylogeny, the phylogenetic relationships of 29 EF1α amino acid sequences from a diverse range of organisms were assessed by phylogenetic tree construction. The tree was constructed using the distance matrix method, Neighbour Joining, and Poisson correction for estimation of evolutionary distances (MacVector, Accelrys) (Figure 3.8). Eukaryote sequences were selected based on a previous phylogenetic study using elongation factors (Gaucher et al., 2001), but several full-length cnidarian EF1α sequences were also included. Sequences were retrieved from GenBank (www.ncbi.nlm.nih.gov) and aligned prior to tree construction using ClustalW (Thompson et al., 1994).

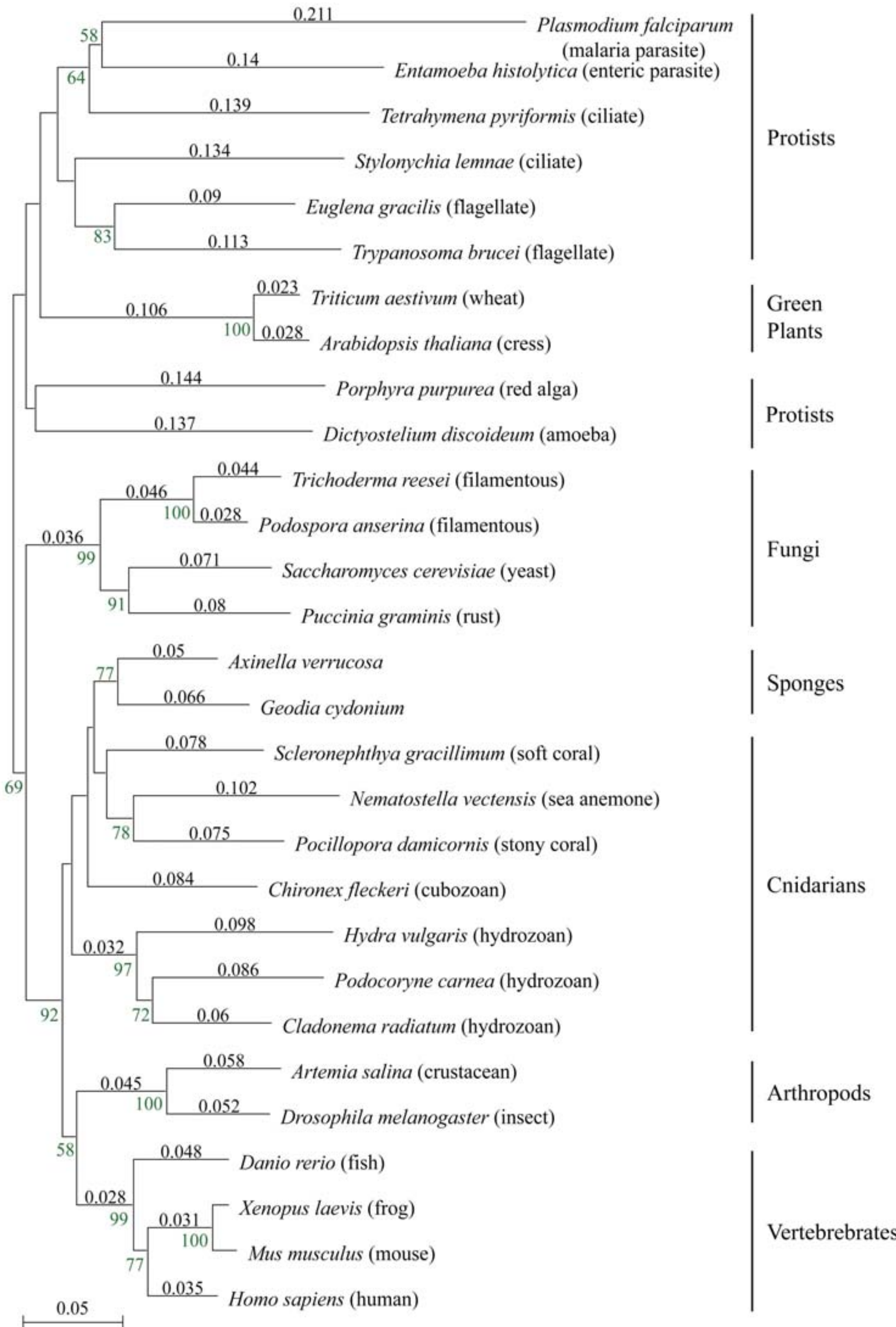


Figure 3.8 Molecular phylogenetic tree of eukaryotes based on Neighbour Joining analysis of EF1α amino acid sequences. Details of the biological origin, accession numbers and alignment of the protein sequences used for phylogenetic analysis are provided in Appendix 5. Estimates of evolutionary distance are indicated above internal branches. The scale bar indicates an evolutionary distance of 0.05 amino acid substitutions per position in the sequence. Bootstrap values (based on 1000 replicates) are indicated in green below internal branches and denoted as a percentage. Weak bootstrap values (< 50%) are not shown.

This alignment and the full list of species and accession numbers of the EF1α protein sequences used for phylogenetic analysis are provided in Appendix 5. The reliability of the phylogeny was tested by bootstrap analysis (1000 replicates) and bootstrap values calculated for each internal branch were reported as the percentage of trees that contained that branch.

Phylogenetic analysis revealed that *C. fleckeri* clusters with other cnidarians and poriferans (sponges), forming a diploblastic sister group to the triploblastic animals, which includes the vertebrates and arthropods (Figure 3.8). Similar localised topology was reported in a previous EF1α-based phylogeny study that included a small selection of cnidarian and poriferan sequences (Berney et al., 2000). Ostensibly, sub-branching within the diploblastid clade suggests that *C. fleckeri* is more closely related to anthozoans (corals and sea anemones) and sponges than to hydrozoans, however, this inference was not well supported by bootstrap analysis (i.e. bootstrap value <50%).

In an attempt to improve the reliability of the phylogenetic predictions, the tree was reconstructed using an alignment in which 4 regions of hypervariable sequence were removed (i.e. consensus residues 1-17, 204-238, 403-406 and 450-480). Reanalysis produced a tree with similar topology to the original tree, except *C. fleckeri* and the sponges were placed more closely to the hydrozoans than to the anthozoans (data not shown). However again, bootstrap analysis indicated that the branches supporting this pattern of evolutionary divergence were only weakly supported (bootstrap values <50%). These results indicate that the evolutionary position of *C. fleckeri* relative to anthozoans and hydrozoans cannot be satisfactorily resolved using the phylogenetic analysis parameters described in this study. The addition of new cnidarian sequences, could improve the reliability of the inferred phylogeny, however, no additional full-length cnidarian EF1α sequence was publicly available at the time of analysis, except for that of *Hydra magnipapillata*, which is significantly homologous to *Hydra vulgaris*. Other methods of analysis, such as Maximum Parsimony, Maximum Likelihood, etc., could also be assessed as alternatives to the Neighbour-Joining approach. Similarly, Bayesian inferences of phylogeny based upon the posterior probability distribution of the trees (e.g. MrBayes; mrbayes.csit.fsu.edu) could also be tested.

3.3.5 General Features of Other Commonly Detected Full-length cDNA Clones

The second most commonly detected clone in the CSL immunodetection experiment, representing 16% of the sequenced clones, encoded 3-PGDH. A conserved domain search (CD-search; Marchler-Bauer and Bryant, 2004) of the Conserved Domain Database (CDD; Marchler-Bauer et al., 2007) predicted that deduced residues 9-319 of *C. fleckeri* 3-PGDH contain a conserved catalytic domain characteristic of NAD^+-dependent D-2-hydroxycarboxylate dehydrogenases (Accession No. pfam00389). Within this region (deduced residues 111-287), the conserved NAD^+ binding domain of the D-2-hydroxycarboxylate dehydrogenases was also detected (Accession No. pfam02826). 3-PGDH plays an important role in amino acid metabolism by catalysing the NAD^+-dependent conversion of phosphoglycerate to 3-phosphohydroxypyruvate during the first and rate-limiting step in L-serine biosynthesis (Cho et al., 2000). Serine is a non-essential amino acid that is necessary for protein biosynthesis and as a precursor of many important metabolites, including glycine, cysteine, serine phospholipids, sphingomyelins, and cerebrosides (Pineda et al., 2000). It is also a major source of methylenetetrahydrofolate (a methionine precursor) and other one-carbon donors.

The third most commonly detected clone in the CSL immunodetection experiment encoded HSP70-1 (6% of sequenced clones). A conserved domain search indicated that deduced residues 6-612 of the translated clone contain a conserved multifunctional HSP70 domain (Accession No. pfam00012). The calculated molecular mass of *C. fleckeri* HSP70-1 (71392 Da) is also similar to that of HSP70 protein family members (Pelham, 1988). The partial deduced amino acid sequence of a closely related homologue (HSP70-2) was also isolated during the rabbit immunodetection experiment (~80% similarity between 266 aligned residues).

HSP70 proteins (HSP70s) belong to a ubiquitous family of highly conserved housekeeping proteins that are present in archaea, bacteria and eukaryotes (Gupta and Golding, 1993). The HSP70s are molecular chaperones that play important roles in the folding and assembly of newly translated proteins, repair of misfolded or aggregated proteins, translocation of protein across membranes and regulation of protein activity (Mayer and Bukau, 2005). All HSP70s contain 3 distinct domains: an N-terminal ATPase domain (~400 residues), a central substrate binding domain (~200 residues) and a hypervariable C-terminal domain (Renner and Waters, 2007). The activity of HSP70s is ATP-dependent and involves the interaction of hydrophobic peptide segments of target proteins with the substrate binding domain (Mayer and Bukau, 2005). In eukaryotes, at least 4 types of HSP70s are localised to different cellular

compartments, including the cytoplasm, mitochondrion, endoplasmic reticulum and, where applicable, the chloroplasts (Renner and Waters, 2007). In the case of *C. fleckeri* HSP70-1, computational analysis of the amino acid sequence (SignalP 3.0, Bendtsen et al., 2004; PSORT II, Horton and Nakai, 1997), predicted that the protein is cytoplasmic.

Another commonly detected clone in the CSL box jellyfish antivenom screening experiment encoded an ABP similar to coronin 1B (Figure 3.6). A conserved domain search predicted that the translated *C. fleckeri* clone contains 5 clustered WD40 domain repeats (Accession No. cd00200) between deduced residues 18-337 and 2 conserved domains of unknown function, DUF1899 (Accession No. pfam08953) and DUF1900 (Accession No. pfam08954), spanning deduced residues 4-68 and 253-398, respectively.

Coronins are a conserved family of ABPs that are involved in the regulation of several actin-mediated processes including cytoskeleton organisation, cell motility, neural development, vesicular transport and Golgi trafficking (Rybakin and Clemen, 2005; de Hostos, 1999). All coronins contain a common central WD40-repeat domain core, a short conserved N-terminal motif and a conserved 70-residue region located downstream of the WD40 repeats (Rybakin and Clemen, 2005). Coronin homologues form two distinct subfamilies, short proteins (450-650 residues) containing a coiled-coil region near the C-terminus and longer proteins (920-1080 residues) that contain 2 two core domains, rather than one, and lack a C-terminal coiled-coil region. The short coronins also contain a highly basic N-terminal motif (~20 residues) located at the N-terminus (Rybakin and Clemen, 2005). Amino acid sequence analysis of *C. fleckeri* ABP (501 residues) suggests that the box jellyfish protein is a short coronin homologue, due to the prediction of a C-terminal coiled-coil region, spanning residues ~446-483 (COILS Server; Lupas et al., 1991), and the presence a conserved, highly basic N-terminal motif (pI 12.5; ProtParam, Gasteiger et al., 2005) similar to other short coronin homologues (Figure 3.9).

```
1   MSGRFVRASKFRHVFGTAARK 21 Chironex fleckeri

1    MSKVVRSSKYRHVFAAQPKK 20 Dictyostelium discoideum
1   MSGKFVRASKYRHVFGQAAK  20 Saccharomyces cerevisiae
1   MSFRVVRSSKFRHVYGQALK  20 Drosophila melanogaster
1 MSFRRGVVRQSKFRHVFAQAWK  22 Danio rerio 1B
1   MSRQVVRSSKFRHVFGQPAK  20 Homo sapiens  1A
1    MRRVVRQSKFRHVFGQAVKN 20 Homo sapiens  1C
        .** **.***.    .
```

Figure 3.9 Multiple sequence alignment of highly basic N-terminal signature motifs in short coronins. Identical residues are marked by an asterisk and similar residues are marked with a stop. Highly conserved residues are highlighted in bold. Accession numbers for *D. discoideum* (amoeba), *S. cerevisiae* (yeast), *D. melonogaster* (fruit fly), *D. rerio* (zebra fish, isoform 1B) and *H. sapiens* (isoforms 1A and 1C) are CAA43707, EDN59327, NP_610242, NP_001103177, NP_009005, and CAO94663, respectively.

3.3.6 Thypedin – a Thymosin-like Protein or a Pedin Precursor?

One of the more interesting clones isolated during the CSL box jellyfish antivenom immunoscreening experiment encoded thypedin. Analysis of the deduced amino acid sequence of the partial thypedin clone revealed the presence of two similar, yet distinct, alternating repeat sequences classified as type 1 and type 2 (Figure 3.10a). BLAST analysis revealed that the type 1 repeat consensus sequence (29 residues) was significantly homologous to type 1 repeats of thypedin from *Nematostella vectensis* (76%) and *Hydra vulgaris* (64%), thymosin β4 from *Holothuria glaberrima* (73%) and β-thymosin from *Paracentrotus lividus* (66%). In comparison, BLAST analysis revealed that the *C. fleckeri* type 2 repeat consensus sequence (37 residues) shared significant homology with type 2 repeats of thypedin from *N. vectensis* (72%) and *H. vulgaris* (62%), and β-thymosin from *Triatoma infestans* (58%).

Thypedin was first reported for *Hydra vulgaris* and its amino acid sequence contains 13 copies of the peptide pedin (EELRPEVLPDVSE) interspersed between 27 β-thymosin-like repeats (Hermann et al., 2005). Pedin is implicated in the stimulation of foot formation and regulation of bud outgrowth in hydra, whereas β-thymosin and β-thymosin repeat-containing proteins, including actobindin, ciboulot and thymosin β4, are actin-binding proteins that play important roles in cell migration, differentiation, proliferation, regeneration and survival (Hoffmeister, 1996; Hermann et al., 2005; Hannappel et al., 2007). Therefore the function(s) of thypedin in cnidarians remains unclear. While thypedin may act solely as a precursor for pedin, it may also, or instead, interact with actin in its unprocessed form and exert a physiological function via the multiple β-thymosin-like repeats (Hermann et al., 2005).

However, unlike thypedin from *H. vulgaris*, in the thypedin homologues from *C. fleckeri* and *N. vectensis*, the corresponding pedin-like sequences are not as highly conserved (Figure 3.11). Indeed, conservation of residues in *C. fleckeri* thypedin appears to be greater in other regions of sequence that contain the β-thymosin-like repeats (Figure 3.10). Thus, the function of thypedin in anthozoans and cubozoans may differ from that of hydrozoans, wherein the role of pedin may be secondary to that of the β-thymosin-like domain. Further investigations are therefore necessary to better understand the role(s) of thypedin in cnidarians.

(a)

```
  1 -               HVETQEKVVLPTKEVIQEEATDSR   - 24   type 1 repeat (partial)
 25 -    AEVKSFDHSKLKHVETQEKNELPTPQMLKLEMKPDEL - 61   type 2 repeat
 62 - PDVSEVSKFDTGKLKHVETQEKVVLPTKEVIQEEATDSR  - 100  type 1 repeat
101 -    AEVKSFDHSKLKHVETQEKNELPTPQMLKLEMKPDEL - 137
138 - PDVSGVSHFDTGKLKHVETQEKVVLPTKEVIQEEATDSR  - 176
177 -    AEVKSFDHSKLKHVETQEKNELPTPQMLKLEMKPDEL - 213
214 - PDVSEVSKFDTGKLKHVETQEKVVLPTKEVIQEEATDSR  - 252
253 -    AEVKSFDHSKLKHVETQEKNELPTPQMLKLEMKPDEL - 289
290 - PDVSEVNKFDTGKLKHVETQEKNMLPTADVIKEEVADSR  - 328
329 -    AEVKSFDHTKLKHVETQEKNPLPTPQTLKEELRPDSL - 365
366 - PDVSEVSEFDSGKLKHVETQEKSVLPTKEVIQEEAADSR  - 404
405 -    AEVKSFDHSKLKHVQTEEKNPLPTAGTIREELRPDEF - 441
442 - PDRSEVADFDTRKLKHVETTEKNVLPTKEVIKEEAVDSR  - 480
481 -    AEVKSFDHGKLKHVETEEKNTLPSAGTLKEEMRPEFL - 517
518 - PDVSEVKDFNTGKLKHVQTEEKTVLPTKEVINQEALDSR  - 556
557 -    AEVKTFDHTKLKHVETQEKNVLPTPQVLKEELRPDAL - 593
594 - ADVSEVKDFDQGKLKHVETKEKVTVPSKEVIDEEAKNTM  - 632
633 -    AEVTKFDKGKLKHADTKEKDKLPTAEDIKNEKSA    - 666
         *  *   **** .* **  .*.   .  *
```

(b)

```
124 - PDRSEVAKFDTSKLKHVETKEKVVMPTKDVIEAEAIDSR  - 162 N. vectensis thypedin-like (76%)
  5 - PDVSEVTKFDKGKLKTVETQEKNTLPTKDTIDQE       -  38 H. glaberrima thymosin β4 (73%)
  5 - PDVSEVASFDKSKLKKAETQEKNTLPTKETIEQEKT     -  40 P. lividus β-thymosin (66%)
122 - PDVSEVEKFDSSKLKNVTPEVKCHLPTKDVIQEEAVATR  - 160 H. vulgaris thypedin type 1 β-thymosin repeat (64%)

  1 -               HVETQEKVVLPTKEVIQEEATDSR   -  24 C. fleckeri thypedin type 1 β-thymosin repeat
 62 - PDVSEVSKFDTGKLKHVETQEKVVLPTKEVIQEEATDSR  - 100
138 - PDVSGVSHFDTGKLKHVETQEKVVLPTKEVIQEEATDSR  - 176
214 - PDVSEVSKFDTGKLKHVETQEKVVLPTKEVIQEEATDSR  - 252
290 - PDVSEVNKFDTGKLKHVETQEKNMLPTADVIKEEVADSR  - 328
366 - PDVSEVSEFDSGKLKHVETQEKSVLPTKEVIQEEAADSR  - 404
442 - PDRSEVADFDTRKLKHVETTEKNVLPTKEVIKEEAVDSR  - 480
518 - PDVSEVKDFNTGKLKHVQTEEKTVLPTKEVINQEALDSR  - 556
594 - ADVSEVKDFDQGKLKHVETKEKVTVPSKEVIDEEAKNTM  - 632
       * * *  *.  ***       *  .*. . *  *   .
```

(c)

```
239 -    AEVKSFDHSKLKHVETVEKNPLPSAAVLKEEMRPEVL - 275 N. vectensis thypedin-like (72%)
  9 -    AEVKTFDQSNLKHVETEEKNPLPTAATLREELRPEVL -  45 H. vulgaris thypedin type 2 β-thymosin repeat(62%)
136 -    AGIENFDHKKLKHTETQEKNPLPTKEAIDQE       - 166 T. infestans β-thymosin (58%)

 25 -    AEVKSFDHSKLKHVETQEKNELPTPQMLKLEMKPDEL - 61  C. fleckeri thypedin type 2 β-thymosin repeat
101 -    AEVKSFDHSKLKHVETQEKNELPTPQMLKLEMKPDEL - 137
177 -    AEVKSFDHSKLKHVETQEKNELPTPQMLKLEMKPDEL - 213
253 -    AEVKSFDHSKLKHVETQEKNELPTPQMLKLEMKPDEL - 289
329 -    AEVKSFDHTKLKHVETQEKNPLPTPQTLKEELRPDSL - 365
405 -    AEVKSFDHSKLKHVQTEEKNPLPTAGTIREELRPDEF - 441
481 -    AEVKSFDHGKLKHVETEEKNTLPSAGTLKEEMRPEFL - 517
557 -    AEVKTFDHTKLKHVETQEKNVLPTPQVLKEELRPDAL - 593
633 -    AEVTKFDKGKLKHADTKEKDKLPTAEDIKNEKSA    - 666
         * .  **. .*** .* **. **.   .  *   . *
```

Figure 3.10 (a) Alignment of two distinct β-thymosin-like repeat sequences (types 1 and 2) in the partial sequence of thypedin (*C. fleckeri*). Multiple sequence alignment of *C. fleckeri* β-thymosin-like repeats (b) type 1 and (c) type 2 with their respective best matches obtained by BLAST analysis, including thypedin homologues and different members of the β-thymosin family. The similarity of the sequences is indicated in brackets as a percentage of identical residues. Identical residues are marked by an asterisk and similar residues are marked with a stop. Highly conserved residues are highlighted in bold. Putative pedin peptides are underlined. Accession numbers for *N. vectensis* (sea anemone), *H. glaberrima* (sea cucumber), *H. vulgaris* (hydra), *P. lividus* (sea urchin) and *T. infestans* (blood-sucking bug) are EDO48924, ABS29642, AAW82079, CAD29144 and ABR27867, respectively.

```
            EELRPEVLPDVSE         H. vulgaris (13 identical peptides)

116 - AEMMPEVLPDRSE - 128  N. vectensis
192 - EELIPKNKPDRSE - 204
268 - EEMRPEVLPDVSA - 280

 54 - LEMKPDELPDVSE - 66   C. fleckeri
130 - LEMKPDELPDVSG - 142
206 - LEMKPDELPDVSE - 218
282 - LEMKPDELPDVSE - 294
358 - EELRPDSLPDVSE - 370
434 - EELRPDEFPDRSE - 446
510 - EEMRPEFLPDVSE - 522
586 - EELRPDALADVSE - 598
       *. *    * *
```

Figure 3.11 Multiple alignment of the repeat sequence for pedin in thypedin from *H. vulgaris* and corresponding sequences in thypedin homologues from *N. vectensis* and *C. fleckeri*. Identical residues are marked by an asterisk and similar residues are marked with a stop. Accession numbers for *H. vulgaris* (hydra) and *N. vectensis* (sea anemone) are AAW82079 and EDO48924, respectively.

3.4 Commentary

Characterisation of proteins using biochemical methods usually requires a significant amount of proteins and this issue is particularly relevant to the characterisation of box jellyfish nematocyst proteins. Live specimens are often difficult to locate and capture in their marine habitat and due to the venomous nature of the animals, skilled sampling is required. In addition, sampling is hampered by the relatively small geographical range of box jellyfish populations, prevailing weather conditions and the seasonal variation in abundance of free-swimming jellyfish, which is generally sporadic and unpredictable (Bloom et al., 1998; Fenner and Harrison, 2000; Hartwick, 1991). Thus, the use of molecular biology techniques can provide opportunities to overcome some of the problems associated with a lack of samples for protein extraction, and comparatively, the procedures usually require much smaller amounts of nucleic acids to achieve similar research outcomes. Certainly, the beginning of this research project was marked by a severe shortage of *C. fleckeri* samples with which to isolate nematocysts for biochemical analyses. However, sufficient amounts of jellyfish tissue were available to construct a novel *C. fleckeri* tentacle cDNA expression library.

Although initial screening of the expressed cDNA library to isolate putative *C. fleckeri* venom proteins using antibodies was unsuccessful, the experiment enabled the isolation and cDNA sequencing of 41 different *C. fleckeri* proteins (3 of which were novel). Indeed, prior to this study, only 7 microsatellite loci DNA sequences (Coughlan et al., 2006) and 1 ribosomal DNA sequence (Collins et al., 2006) obtained from *C. fleckeri* were available in non-redundant sequence databases. Therefore, the results of this study contribute significantly to our current

knowledge and understanding of cubozoan proteins and their functions. Furthermore, due to the significant proportion of cDNA clones isolated in this study that were full-length (28%), this unique cDNA expression library constitutes a useful long-term resource for the cDNA cloning and sequencing of many other *C. fleckeri* proteins.

CHAPTER 4

Biochemical Characterisation of *Chironex fleckeri* Venom Proteins

4.1 Introduction

Historically, the isolation and characterisation of the bioactive components in *C. fleckeri* venom has been difficult and results have varied significantly between research groups (for reviews see Chapter 1; Bloom et al., 1998; Tibballs, 2006). Discrepancies in results are attributed primarily to variability in the extract source and/or the extraction and fractionation techniques adopted by researchers. In particular, the use of whole tentacles has been criticised because envenomed animals are never injected with tentacle material other than the contents of the nematocysts (Endean et al., 1993; Bloom et al., 1998). Problems encountered during venom isolation and characterisation, such as venom thermolability, adhesion to preparative surfaces and aggregation/disaggregation of venom components have further complicated investigations (Bloom et al., 1998; Endean et al., 1993; Othman and Burnett, 1990).

In recent years, bead mill homogenisation of isolated *C. fleckeri* nematocysts has been adopted by several researchers as a rapid and reproducible method of venom extraction. Consequently, several toxicological and pharmacological studies using bead mill-derived nematocyst venom have been reported (e.g. Carrette and Seymour, 2004, 2005; Kitner et al., 2005; Ramasamy et al., 2003, 2004; Winter et al., 2007a; Bailey et al., 2005), however, the physical and biochemical characteristics of venom proteins extracted using this method have yet to be adequately described.

The aim of this study was to comprehensively characterise the proteins contained in whole *C. fleckeri* bead mill nematocyst venom by electrophoretic techniques such as one-dimensional sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE), two-dimensional gel electrophoresis (2DGE) and western blot analysis using a variety of different polyclonal antibodies. The efficiency of protein extraction using bead mill homogenisation was examined and differences in protein composition between bead mill nematocyst venom and milked venom were compared. The effects of extraction buffer composition and nematocyst lyophilisation on protein extraction were also investigated. The lethality and haemolytic activity of bead mill

extracts obtained in this study were determined to ensure that biologically active venom proteins were present in the nematocyst extracts.

4.2 Methods and Materials

4.2.1 Preparation of Nematocyst Venom

Nematocyst-derived venom (nematocyst venom) was prepared from isolated nematocysts using bead mill homogenisation (Section 2.2.1). Unless otherwise specified, non-lyophilised nematocysts were used. Lyophilised milked venom was donated by Dr Jamie Seymour (Section 2.2.2). Chemically discharged nematocyst venom was prepared in SDS sample buffer (Section 2.2.3). All extracts were freshly prepared before analysis.

4.2.2 SDS-PAGE and Western Blot Analysis

The proteins of non-lyophilised nematocyst *C. fleckeri* extracts were initially examined by SDS-PAGE and western blot analysis using CSL box jellyfish antivenom. Nematocyst proteins (2–20µg) were applied to 15% SDS-PAGE gels and following electrophoresis for variable times at constant voltage (0.75 or 2h, 170V), the proteins were Coomassie- or silver-stained, or subjected to western blot analysis using CSL box jellyfish antivenom, as previously described (Sections 2.3.3 and 2.3.4).

4.2.3 Two Dimensional Gel Electrophoresis (2DGE)

The native proteins of nematocyst venom (15–75µg) were solubilised in urea/CHAPS, reduced and alkylated with DTT and acrylamide, respectively, then separated by 2DGE; a process involving the one-dimensional separation of proteins according to their isoelectric points (pI) by isoelectric focusing, then separation of the proteins according to their molecular mass by reducing SDS-PAGE in a second dimension (for further details refer to Section 2.3.5). Following SDS-PAGE using 15% gels, nematocyst proteins were silver-stained or subjected to western blot analysis using CSL box jellyfish antivenom.

4.2.4 Examination of Bead Mill Extraction Efficiency

The effectiveness of bead mill homogenisation to extract proteins contained in *C. fleckeri* nematocysts was examined by comparing the protein profile of bead mill nematocyst venom

with the protein profile of *C. fleckeri* venom obtained by chemical induction of nematocyst discharge. Samples were analysed by SDS-PAGE (Coomassie-stained) and western blot using CSL or rabbit antibodies (whole nematocyst venom or CfTX-specific). Details of the preparation of rabbit antibodies are provided in Section 2.3.2.

4.2.5 Effects of NaCl Concentration on Nematocyst Protein Extraction

The effect of NaCl on nematocyst protein extraction was examined by comparing the protein profiles of nematocyst venom in buffers containing various concentrations of NaCl. Three nematocyst extraction buffers were prepared containing 25mM MOPS and protease inhibitors (see Section 2.2.1) and different concentrations of NaCl (0, 100 and 600mM). Three aliquots of non-lyophilised nematocysts (labelled 1, 2 and 3) were washed (3 × 2mL) with ice-cold buffers (0, 100 and 600mM, respectively). The NEB-washed nematocysts (wet weights 0.0447, 0.0730 and 0.1017g, respectively) were extracted with their respective buffers (0.5mL) by bead mill homogenisation. Samples were transferred to clean pre-chilled tubes, centrifuged (12.1K *g*, 4°C, 20 min) and the A_{280} of supernatants were measured to estimate total extracted protein. Supernatants were dialysed in centrifugal filters (10K MWCO, Millipore; 12.1 K *g*, 4°C) against cold NEB and protein concentrations were estimated at A_{280}. Supernatants, diluted in reducing 2X SDS sample buffer, and pellets, washed (2 × 1mL) with their respective buffers and resuspended in reducing 2X SDS sample buffer (50µL each), were analysed by SDS-PAGE. Separated nematocyst proteins were Coomassie-stained.

4.2.6 Lethal Activity of Whole Nematocysts Extracts

In collaboration with Dr Jamie Seymour and Ms Teresa Carrette (James Cook University), the lethality of venom from non-lyophilised *C. fleckeri* nematocysts was tested on the crayfish, *Cherax quadricarinatus* (Animal Ethics No. A901). Nematocyst extracts (1 and 2) were prepared by bead mill homogenisation using two buffers (NEB 1 and 2), containing 25mM MOPS and protease inhibitors (see Section 2.2.1), but differing in NaCl concentration (0 and 800mM, respectively). Venom proteins were analysed by SDS-PAGE and western blot analysis using CSL box jellyfish antivenom. Undiluted extracts were stored on ice and extract dilutions in cold PBS were prepared immediately prior to crayfish testing. Extract 2 was re-tested 31h later, following storage at 4°C. Lethality tests were undertaken by Seymour and Carrette using established methods (Carrette and Seymour, 2004, 2005). Briefly, nematocyst extracts were

injected intramuscularly into the ventral second abdominal segment of the crayfish and time to cardiac stand-still was measured using an Elite Vascular Doppler, positioned dorsally above the heart. Time measurements were restricted to three minutes to avoid prolonged pain and stress caused by a sub-lethal dose; any surviving animal was euthanised. Injection volumes were proportional to animal weight (1µL/g crayfish), but the doses of venom proteins varied (ng total protein/g crayfish). Protein concentration was determined using the Bradford Assay (see Section 2.3.1). Unless otherwise specified, each concentration of nematocyst extract was tested in quadruplicate. Sets of 4 control animals included animals injected with buffer only (PBS or NEB ± 0.8M NaCl) and animals that were not injected. Control animals were maintained for 24 hours after injection under identical conditions.

4.2.7 Comparison of *C. fleckeri* Venom Proteins from Different Extract Sources

The protein profiles of lyophilised milked venom and nematocyst venoms, obtained from lyophilised and non-lyophilised nematocysts, were compared. Lyophilised milked venom was resuspended in ice-cold NEB (250µL) and following centrifugation (20K*g*, 5 min, 4°C), an aliquot of the dark orange supernatant was analysed by SDS-PAGE and western blot using CSL (sheep) antibodies. In a subsequent experiment (21 months later), another lyophilised milked venom sample was resuspended in ice-cold NEB (1mL) and an aliquot (15µL) was set aside on ice. Following centrifugation of the remaining sample (18K*g*, 2 min, RT), the supernatant (colourless on this occasion) was transferred to another tube on ice and the orange-brown pellet was resuspended in 2X SDS sample buffer (100µL). An aliquot of supernatant (15µL) was set aside on ice prior to concentration of the remaining sample by TCA precipitation (refer to Section 2.3.3.3) and resuspension in 2X SDS sample buffer (80µL). The milked venom supernatant and pellet samples were subjected to SDS-PAGE and western blot analysis using CSL and rabbit antibodies (whole nematocyst and CfTX-specific). Aliquots of nematocyst venom, resuspended milked venom and milked venom supernatant were tested for haemolytic activity on 5% sheep blood agar (see Section 2.3.8.1).

4.2.8 Antibody Specificity towards *C. fleckeri* Venom Proteins

The antigenicity of *C. fleckeri* venom proteins was examined using antibodies raised against milked venom or nematocyst venom in sheep, mouse or rabbit (see Section 2.3.2 for details on antibody preparation). *C. fleckeri* nematocyst venom proteins (~20µg/lane) were applied to six lanes in 15 and 10% SDS-PAGE gels and following electrophoresis, two lanes containing the molecular marker and a sample of nematocyst proteins were Coomassie-stained. The nematocyst proteins in the remaining lanes were transferred to PVDF and the membrane was cut into sections corresponding to each lane. Individual sections were subjected to western blot analysis using rabbit (1:1000), mouse (1:1000) or sheep (CSL antivenom; 1:5000) antibodies. Normal mouse and rabbit sera were used as control antibodies; normal sheep serum was unavailable at the time of the experiment.

4.2.9 Interspecies Comparison of Nematocyst Extracts

Venom extracted from non-lyophilised *C. fleckeri* nematocysts and lyophilised *C. bronzie* nematocysts using bead mill homogenisation were compared by SDS-PAGE and western blot analysis. *C. fleckeri* venom proteins (~18µg), *C. bronzie* venom proteins (~8µg) and TCA-concentrated *C. bronzie* venom proteins (CF = 10) were applied to 12% SDS-PAGE gels. Following electrophoresis, proteins were Coomassie-stained or subjected to western blot analysis using CSL and rabbit antibodies (whole nematocyst venom and CfTX-specific). Native nematocyst venom from each species was tested for haemolytic activity on 5% sheep blood agar (see Section 2.3.8.1).

4.3 Results and Discussion

4.3.1 SDS-PAGE and Western Blot Analysis

SDS-PAGE and western blot analyses provided important insight into the protein composition of *C. fleckeri* venom. SDS-PAGE analysis of *C. fleckeri* nematocyst venom proteins obtained by bead mill homogenisation revealed the presence of several protein bands ranging in mass from 10 to 180 kDa (Figure 4.1a-b and Figure 4.2a-b). The protein profiles included two major proteins (43 and 45 kDa, respectively), referred to as CfTX-1 and CfTX-2, and several relatively abundant protein bands (~97, 65, 30, 24 and 20 kDa). A few minor protein bands (~160 kDa, 120/125 kDa doublet) and clusters of protein bands in the molecular mass ranges 35–40 kDa and 25–30 kDa were also observed. Similar protein patterns were observed between different batches of nematocysts, except in some cases, the abundance of a 12 kDa protein was significantly greater. Also, in one batch of nematocysts, an additional doublet of major proteins (39 and 41 kDa) was consistently observed in the nematocyst venom profiles. Venom extracts from this batch of nematocysts were subsequently used in size exclusion and cation exchange chromatography experiments (see Chapter 6).

Western blot analysis using CSL box jellyfish antivenom (Figure 4.1c and Figure 4.2c) indicated the presence of several strongly antigenic proteins including CfTX-1 and -2, two minor proteins (~97 and 20 kDa) and a cluster of minor proteins in the molecular mass range 30–35 kDa. A number of proteins were not antigenic to the antivenom (e.g. ~65 and 24 kDa etc.). Notably, upon dilution of the extract, the antigenicity of the 20 kDa protein diminished.

Under reducing and non-reducing SDS-PAGE conditions, the apparent molecular masses of most nematocyst proteins, including the CfTX proteins, remained unchanged (Figure 4.3). A comparison of SDS-PAGE profiles of reduced nematocyst proteins that were unmodified or alkylated following reduction, revealed that the molecular masses at least three proteins (~20, 24 and 36 kDa) increased significantly upon alkylation, due to modification of cysteine residues to form cysteine-S-β-proprionamide (71 Da increase per residue) (Figure 4.4).

The SDS-PAGE and western blot results obtained in this study are comparable to earlier research in which the SDS-PAGE protein profile of *C. fleckeri* venom, obtained from thawed and ground nematocysts, contained numerous protein bands, including a dominant 40 kDa protein (likely to correspond to unresolved CfTX-1 and -2) and at least two major proteins with

molecular masses greater than 106 kDa (Wiltshire et al., 2000). Intense binding of CSL box jellyfish antivenom to the 40 kDa protein was also noted.

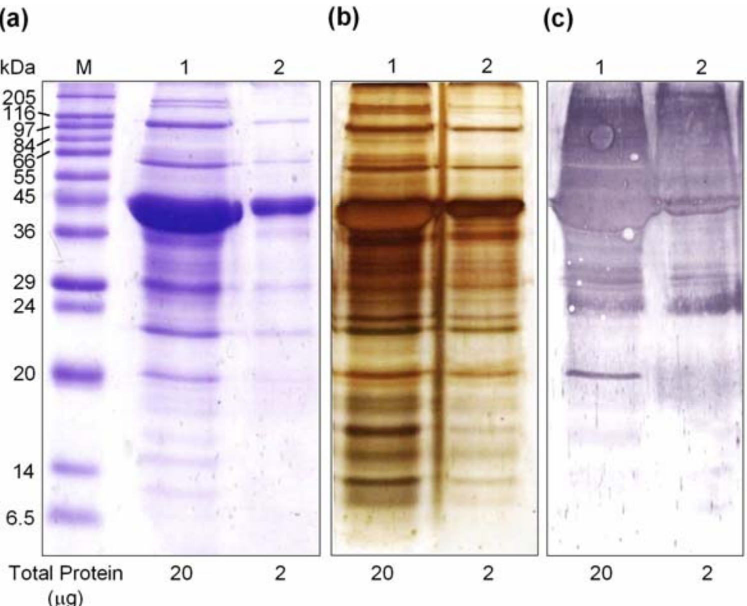


Figure 4.1 Typical profiles of non-lyophilised nematocyst venom proteins following SDS-PAGE and western blot analysis using 15% polyacrylamide gels (electrophoresis time = 45min). (a) Coomassie-stained; (b) silver-stained; (c) CSL box jellyfish antivenom. Total venom protein applied is indicated below each lane.

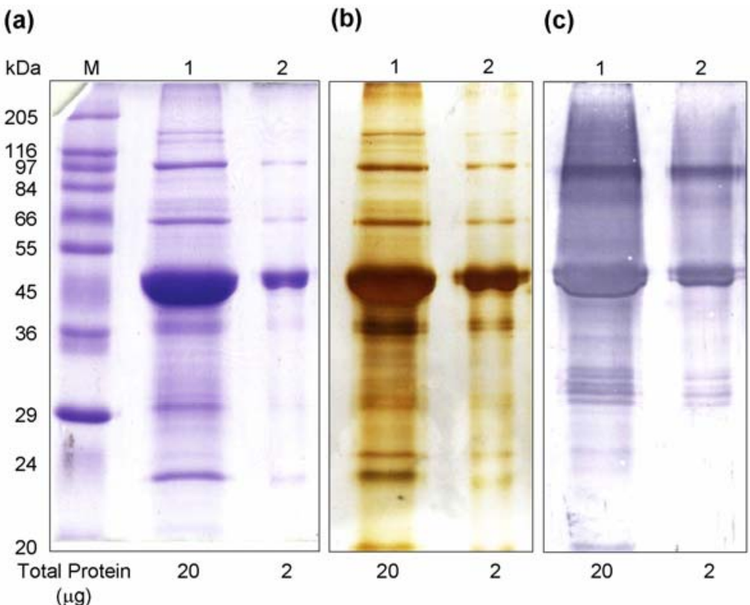


Figure 4.2 Typical profiles of non-lyophilised nematocyst venom proteins following SDS-PAGE and western blot analysis using 15% polyacrylamide gels (electrophoresis time = 2h). (a) Coomassie-stained; (b) silver-stained; (c) CSL box jellyfish antivenom. Total venom protein applied is indicated below each lane.

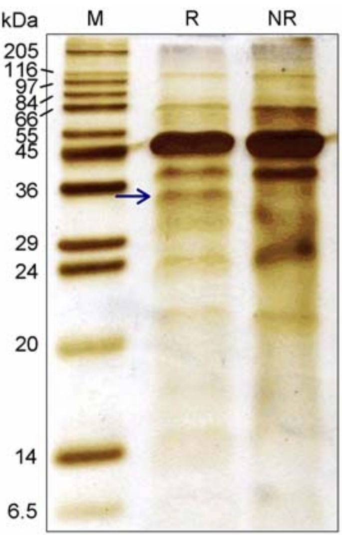


Figure 4.3 Reducing (R) and non-reducing (NR) SDS-PAGE analysis of non-lyophilised nematocyst venom proteins. Venom proteins were boiled in SDS sample buffer in the presence and absence of DTT (100mM final concentration), respectively, separated on a 15% polyacrylamide gel and silver stained. A blue arrow indicates a protein band, present in the reduced extract (~35 kDa), but absent in the non-reduced extract.

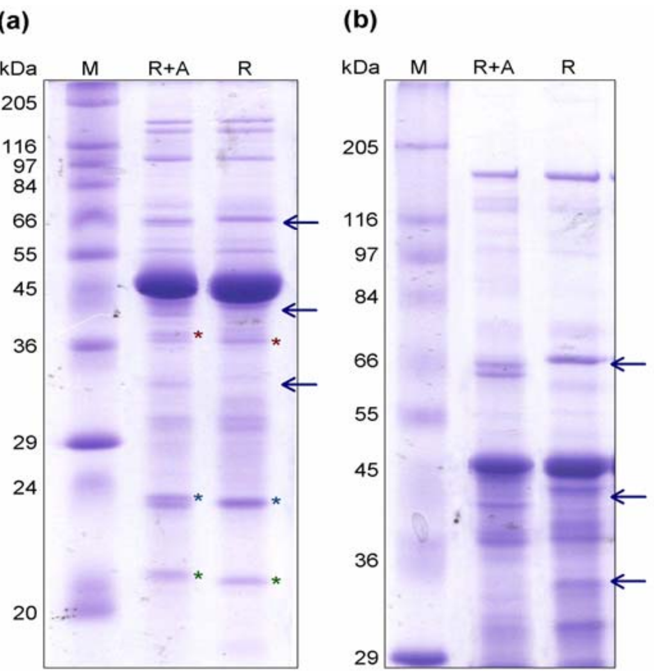


Figure 4.4 Effects of cysteine alkylation on the molecular masses of non-lyophilised nematocyst venom proteins. Extract proteins were reduced and alkylated (R+A) or reduced only (R), separated on large (a) 15% and (b) 10% SDS-PAGE gels and Coomassie-stained. Paired asterisks (red, blue and green) indicate proteins likely to contain multiple cysteine residues, and their corresponding alkylated forms. The expected increase in molecular mass due to alkylation is 71 Da/cysteine residue. Blue arrows indicate regions where alkylation of reduced proteins has resulted in a protein profile change; but where pairing of related proteins is less certain.

4.3.2 Two Dimensional Gel Electrophoresis (2DGE)

Separation of proteins from *C. fleckeri* nematocyst venom using 2DGE provided an overview of total protein distribution across the pH gradient 3–10. Resolution of several proteins as discrete spots in the gels was achieved (Figure 4.5). Protein streaks were visible for the CfTX proteins (~43–45 kDa) and a 65 kDa protein, particularly at the acidic end of the IEF gel, suggesting either poor solubility of the proteins or the presence of contaminants in the extracts that were affecting the 2DGE results. Although separation of CfTX-1 and -2 was relatively poor, a “train” of at least five closely positioned protein spots corresponding to the molecular masses of the CfTX proteins was observed (Figure 4.5 inset a), suggesting the presence of several CfTX isoforms with different pI values (pH 8–9.2). Some evidence of a second train of CfTX isoforms was also observed between pH 5.5–6.5, however, the protein spots were poorly resolved. Most of the other relatively abundant proteins found in nematocyst venom exhibited pI values of 5.7–7.5. However, several minor proteins in the 25–35 kDa molecular mass range were more acidic (pI ~4.3–5.5) and appeared to form multiple protein trains (Figure 4.5 insets b and c). Western blot analysis results were less informative, due to poor transfer of low molecular mass proteins (Figure 4.6), however, the antigenicity of the CfTX proteins and the 65 kDa protein were confirmed.

The calculated pI values for mature, unmodified CfTX-1 and -2 are calculated as 8.3 and 7.7, respectively (see Chapter 5), while the 2DGE data suggests the presence of putative CfTX-1 and -2 isoforms with pI values ranging between pH 8–9.2 (Figure 4.5). Since the protein spots in the train increase only slightly in molecular mass with decreased pI, the presence of multiple CfTX isoforms may be due to various post-translational modifications that mainly influence the charge of the proteins (e.g. phosphorylation, deamidation, glycosylation and N-terminal acetylation, etc.) rather than modifications that cause significant molecular mass changes (e.g. C- or N-terminal truncations, endogenous protein degradation or oligomerisation etc.) (Harry et al., 2000). The former post-translational modifications characteristically cause an acidic shift in unmodified protein pI (Zhu et al., 2005). Hence, the theoretical pI values for modified CfTX-1 and CfTX-2 isoforms are expected to be less than 8.3 and 7.7, respectively. This expected pI shift varies from the observed results, in which the CfTX isoforms are distributed across a higher pH range. However, the estimated pI values observed in this study may be relatively inaccurate, due to the broad pH range of the IEF gel strip used for the analysis (pH 3–10). Therefore, repetition of the 2DGE experiments using IEF gels with a narrower pH range may allow more accurate pI determination of potential CfTX isoforms.

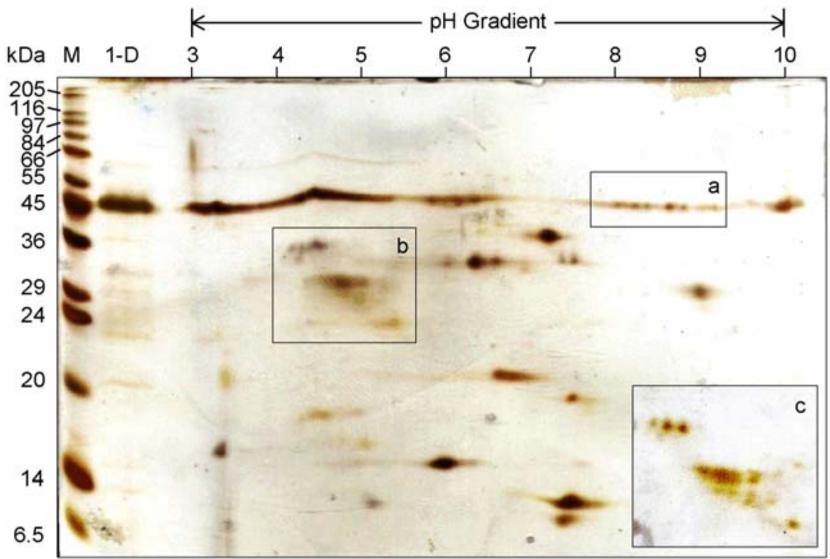


Figure 4.5 Separation of nematocyst venom proteins by 2DGE. Solubilised, reduced and alkylated nematocyst proteins (~17µg) were isoelectrically focussed across a pH gradient (pH 3-10), separated by reducing SDS-PAGE and visualised by silver staining. Nematocyst proteins (~6µg), separated by reducing SDS-PAGE in one dimension only, are indicated by 1-D. M indicates the molecular mass marker. Boxed insets contain a) putative CfTX-1 and -2 protein isoforms; b) putative trains of acidic protein isoforms in the molecular mass range 24-36 kDa; c) magnification of the region corresponding to inset b obtained in another 2DGE experiment revealing increased resolution of protein trains.

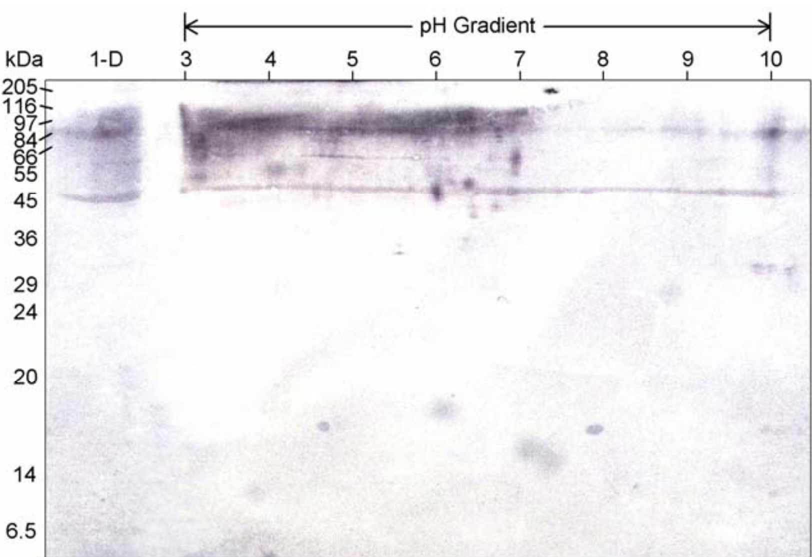


Figure 4.6 Western blot of nematocyst venom proteins separated by 2DGE. Solubilised, reduced and alkylated nematocyst proteins (~75µg) were isoelectrically focussed across a pH gradient (pH 3-10), separated by reducing SDS-PAGE (15% polyacrylamide gel) and transferred proteins were blotted with CSL box jellyfish antivenom. Nematocyst proteins (~30µg), separated by reducing SDS-PAGE in one dimension only, are indicated by 1-D.

In addition to separating *C. fleckeri* nematocyst proteins by 2DGE, analysis of individual protein spots by mass spectrometry would assist the identification of nematocyst-derived venom proteins and localisation of potential post-translational modification sites. Other techniques, such as the use of phospho-specific protein stains (Schaefer et al., 2006) or lectins that bind specifically to glycosylated proteins (e.g. Concanavalin A; Hawkes, 1982), would also complement conventional 2DGE analysis and mass spectrometry data for the characterisation of post-translational protein modifications to CfTX-1 and -2 and other nematocyst proteins.

4.3.3 Examination of Bead Mill Extraction Efficiency

In previous research on hydrozoans, proteins contained in the nematocysts of *Hydra* were examined by electrophoretic methods, following the resuspension of nematocysts in SDS sample buffer to induce their chemical discharge (Szczepanek et al., 2002). This method is particularly useful because the discharged nematocyst proteins are immediately protected from proteolytic degradation in the SDS sample buffer and less likely to adsorb to preparative surfaces, thus providing a more accurate profile of protein composition. Therefore, this method was adopted to assess the effectiveness with which *C. fleckeri* nematocyst proteins are extracted using bead mill homogenisation.

Bead mill extraction efficiency was examined by qualitative comparison of SDS-PAGE protein profiles of bead mill extracts with nematocyst proteins discharged directly into SDS sample buffer. Resuspension of *C. fleckeri* nematocysts in SDS sample buffer triggered vigorous nematocyst discharge, as monitored microscopically. SDS-PAGE protein profiles of bead mill nematocyst extracts and the contents of chemically discharged nematocysts were similar (Figure 4.7a), indicating that the majority of soluble proteins contained within nematocysts are effectively extracted by the bead mill homogenisation method. The pellet, which was not washed prior to analysis, also contained the same suite of proteins observed in the bead mill extracts and the discharge. At least four additional protein bands (~55, 50, 36 and 18 kDa), likely to be structural nematocyst proteins, were also observed. Western blot analysis, using CfTX-specific antibodies, confirmed that CfTX-1 and -2 are more abundant in the nematocyst discharge than in the nematocyst capsules (Figure 4.7b), thus indicating that CfTX-1 and -2 are venom-derived proteins. Western blots, using CSL or rabbit anti-whole nematocyst extract antibodies, were less informative due to high background colouration (particularly in the direct discharge samples), which made comparisons difficult (Figure 4.7c and d).

The results of this experiment indicate that bead mill homogenisation of box jellyfish nematocysts is an efficient method of native protein extraction. Moreover, direct discharge of nematocysts into SDS sample buffer provides an effective method of protein extraction, provided that subsequent bioactive/functional assays are not required (e.g. for amino acid sequencing). The technique is also particularly useful if supply of nematocysts is limited, because significantly less sample is required than for bead mill homogenisation and other extraction methods.

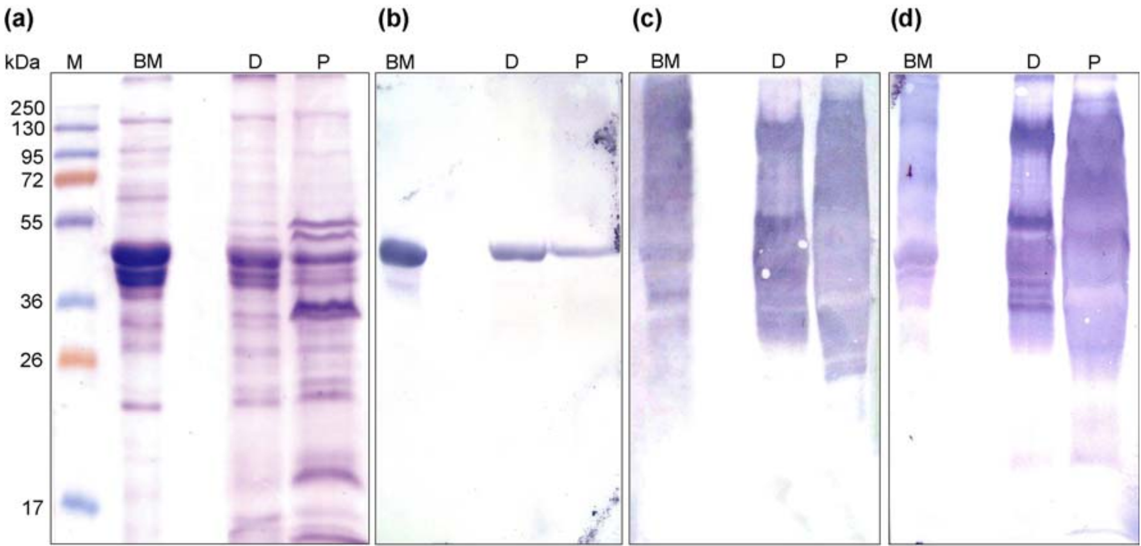


Figure 4.7 Comparison of nematocyst proteins from non-lyophilised nematocysts obtained using bead mill (BM) and direct extraction methods. Direct extraction involved the induction of the nematocyst discharge in SDS sample buffer. The nematocyst proteins in the discharge and the remaining pellet are indicated by D and P, respectively. (a) Coomassie-stained nematocyst proteins; Western blot analyses of the corresponding proteins using (b) CfTX-specific antibodies (c) Rabbit anti-whole nematocyst extract antibodies and (d) CSL box jellyfish antivenom.

4.3.4 Effects of NaCl Concentration on Nematocyst Protein Extraction

The effects of different NaCl concentrations in extraction buffers on total protein extraction from nematocysts were assessed by estimation of protein concentration and SDS-PAGE analysis. Amounts of total protein extracted from non-lyophilised nematocysts using buffers containing 0, 100 and 600mM NaCl were ~11.3, 12.6 and 8.7 mg protein/g nematocysts (wet weight), respectively. Therefore, no correlation between NaCl concentration and estimated amounts of total extracted nematocyst protein was established in this study. However, qualitative comparisons of SDS-PAGE protein profiles of each nematocyst extract and their

corresponding pellets revealed slight differences between extracts (Figure 4.8). The abundance of the CfTX proteins in the three supernatants (relative to the total protein applied to the gel) appeared to increase with increasing salt concentration. A reverse pattern was observed in the protein profiles of the pellets, wherein the abundance of the CfTX proteins diminished with increasing salt concentration. The relative amounts of a 97 kDa protein also increased significantly in extracts containing 100 and 600mM NaCl, but the reverse pattern was not observed in the pellet samples. This finding may indicate that the 97 kDa protein is extracted efficiently using each of the buffers, but is less stable in relatively low salt concentration once extracted.

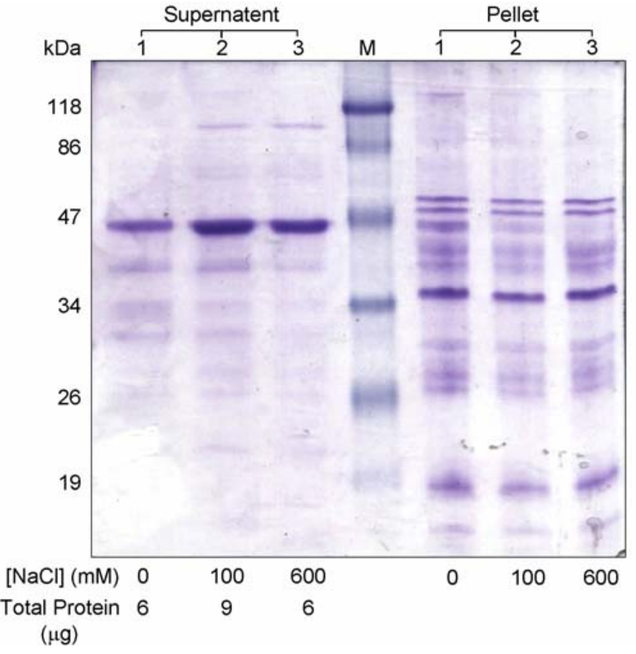


Figure 4.8 Effect of NaCl concentration on the extraction of non-lyophilised nematocyst proteins using bead mill homogenisation. Extracts 1, 2, and 3 were obtained using NEB containing 0, 100 and 600mM NaCl, respectively. Total extract supernatant protein applied to the gel is indicated below each lane. Equal volumes of pellet samples were applied. Following SDS-PAGE, nematocyst proteins were Coomassie-stained. M indicates the pre-stained molecular marker (New England Bioloabs).

The lack of correlation between salt concentration and yield of total nematocyst proteins could be due to various factors, including inaccurate wet weight determination of nematocysts, limited reproducibility of the homogenisation process, and inaccuracies in estimation of protein concentrations. Alternatively, addition of salt may not increase total yield during extraction, but instead, act to reduce adsorption of venom proteins to preparative surfaces and minimise

protein-protein interactions which could lead to aggregation of venom proteins during subsequent sample processing (e.g. filtration, centrifugation, sample transfer etc.).

Examination of the proteins in the pellets of disrupted nematocysts separated by SDS-PAGE (see Figure 4.8), confirm that at least four major proteins (~51, 49, 36 and 19 kDa), and several minor proteins (~17, 26–30 and 38–42 kDa) are likely to be structural proteins from *C. fleckeri* nematocyst capsules or their internal structures, such as tubules. Moreover, the protein profiles of the pellet samples in this experiment correlated well with the profile of the pellet obtained from the direct discharge of nematocysts into SDS sample buffer (see Figure 4.7).

Little is known about the structural protein composition of box jellyfish nematocysts to compare with the findings of this study. However, in the hydrozoan, *Hydra attenuata*, the outer wall of a stenotele nematocyst predominantly contains proteins >40 kDa and the inner wall consists primarily of proteins ranging from 12–16 and 25–40 kDa (Holstein et al., 1994). The major components of stenotele nematocyst capsules are minicollagens and NOWA, which have molecular masses of ~30–40 kDa and 88 kDa, respectively (Engel et al., 2002). Spinalin, which is a major constituent of spines and opercula in *Hydra* nematocysts, is a 24 kDa protein (Hellstern et al., 2006). Although the presence of homologues similar in molecular mass to NOWA and spinalin is not evident in *C. fleckeri* nematocyst protein profiles, a number of observed *C. fleckeri* proteins could be minicollagens. Peptide mass fingerprinting or N-terminal amino acid sequencing of the four dominant protein bands could provide further insight into the major constituents of cubozoan nematocysts.

4.3.5 Lethal Activity of Whole Nematocysts Extracts

The lethality of *C. fleckeri* venom has been demonstrated in a wide variety of animals including fish, prawns, crayfish, mice, rats, guinea pigs, sheep, rabbits and monkeys (Endean et al., 1969; Carrette and Seymour, 2004; Baxter and Marr, 1969; Baxter and Marr, 1975; Comis et al. 1989; Endean et al. 1993; Wiltshire et al., 2000). Lethality tests undertaken in this study by Seymour and Carrette confirmed that nematocyst venom was lethal to the crayfish, *C. quadricarinatus*, but its potency was dependent on ionic strength. Following injection of undiluted extract 1 (429 ng total protein/g crayfish; 0mM NaCl), one animal survived and the times until death of the remaining animals were relatively long (187 ± 20s; n = 5). In contrast, injection of diluted extract 2 (235 and 23 ng/g; 800mM NaCl) caused more rapid death of all test animals (73 ± 15s and 109 ± 6s, respectively; n = 4). Following injection of more dilute

extract 2 (10 ng/g), one animal survived and the times until death of the remaining animals were extended (172 ± 25s; n = 4). All control animals used in the experiment survived (> 24h).

As reflected by the lethality results, nematocyst extracts in high-salt buffer (800mM NaCl) were significantly more lethal to crayfish than nematocyst extracts in buffer without salt. Using high-salt buffer extracts, doses as low as 10 ng total venom protein/g crayfish were sufficient to kill the majority of crayfish within three minutes (smaller doses were not tested). These results are comparable to a previous study in which similar lethality assays were conducted on *C. quadricarinatus* using *C. fleckeri* venom derived from lyophilised nematocysts (Carrette, 2002). Through interpolation of a dose-response curve obtained in the previous study, a dose of ~5 ng total venom protein/g crayfish was estimated as sufficient to kill crayfish within three minutes. This is an interesting finding, given that the extraction buffer used in the earlier study was only deionised water. However, the similarity of the results could be explained by the method in which the nematocysts were prepared in the previous study. During preparation, the isolated nematocysts were lyophilised from suspensions in seawater, so reconstitution of the nematocysts in deionised water would also have redissolved the lyophilised seawater salts, thus increasing the ionic strength of the extraction buffer.

Despite reports that the bioactivity of partially purified carybdeid toxins (CrTX-A/B) are stabilised at 4°C in high salt concentration (Nagai et al., 2000a), in this study, a decrease in protein concentration and lethality was observed in high-salt nematocyst extracts (800mM NaCl) following storage at 4°C. After storage of undiluted extract 2 (31h, 4°C), total protein concentration decreased by 14% (2.34 to 2.02 mg/mL). No deaths occurred following injection of the extract when subsequently diluted (10 ng/g), however, following injection of more concentrated extracts (52, 105 and 222 ng/g), death occurred for all test animals (121 ± 18s, 93 ± 10 and 81 ± 13s, respectively; all n = 4).

SDS-PAGE protein profiles for extracts 1 and 2 (containing 0 and 800mM NaCl, respectively) were similar, however, there appeared to be a disproportionate increase in the abundance of some proteins in extract 2 (relative to total protein applied), compared to extract 1 (Figure 4.9a). For example, the abundance of one or more proteins in the molecular mass range ~39–41 kDa greatly exceeded the abundance of proteins in the corresponding range of extract 1, despite consideration that 2.4 times more extract 2 proteins were applied to the gels than extract 1 proteins. Some high molecular mass proteins also appeared to be relatively more abundant in extract 2 than in extract 1. Western blot analysis of the nematocyst proteins revealed little difference in the profiles of antigenic proteins in extracts 1 and 2; except for the presence of a

minor but strongly antigenic protein (~20 kDa) that was present in extract 2 only (Figure 4.9b). Notably, two more abundant extract 2 proteins (39–41 kDa) observed in the silver-stained gel, were not reactive to CSL box jellyfish antivenom.

The findings of this study suggest that increased ionic strength results in more efficient extraction of one or more lethal components. However, the same variations were not observed in the SDS-PAGE gel from another experiment, using extraction buffers with different NaCl concentrations (0, 100 and 600mM) (see Figure 4.8). Unfortunately, no western blot analysis was undertaken in the previous experiment for comparison. Therefore, repetition of the experiment using a concerted approach is necessary to definitively determine the effects of increasing salt concentration on total protein yield, nematocyst extract composition, lethality and cytolysis.

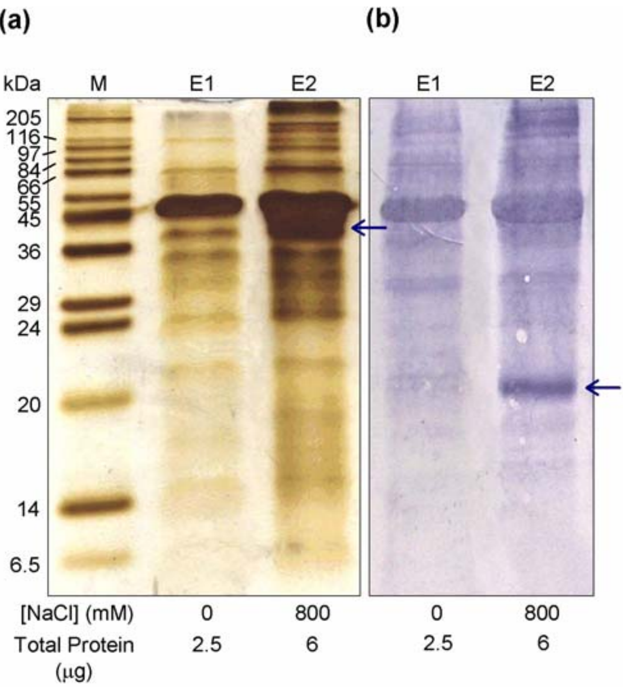


Figure 4.9 Comparison of protein profiles of nematocyst extracts 1 and 2 (E1 and E2), used in lethality studies on *C. quadricarinatus*. Extracts were obtained by bead mill homogenisation of non-lyophilised nematocysts using different buffers, NEB 1 and 2, containing 0 and 800mM NaCl, respectively. (a) silver-stained nematocyst proteins; a blue arrow indicates at least one major protein (~39-41 kDa), which was present in E2 but absent in E1. (b) Western blot of corresponding proteins using CSL box jellyfish antivenom; a blue arrow indicates an antigenic protein (~20 kDa), which was present in E2 but absent in E1. M indicates the molecular marker. Total extract protein applied to SDS-PAGE gels is indicated below each lane.

4.3.6 Comparison of *C. fleckeri* Venom Proteins from Different Extract Sources

Lyophilised nematocysts isolated from *C. fleckeri* tentacles have been used extensively in recent box jellyfish venom research (e.g. Carrette and Seymour, 2004, 2005; Kitner et al., 2005; Ramasamy et al., 2003, 2004; Winter et al., 2007a; Bailey et al., 2005). As most nematocysts remain intact during the freeze-drying process (Bloom et al., 1998), lyophilisation facilitates more convenient (and safer) transport of nematocysts to research laboratories. However, the potential impacts of lyophilisation on venom protein composition have not been examined. Therefore in this study, the protein profiles of extracts derived from lyophilised and non-lyophilised nematocysts were compared to determine any change in venom composition.

The SDS-PAGE protein profiles of bead mill nematocyst venoms obtained from lyophilised and non-lyophilised nematocysts were relatively similar (see Figure 4.10a). CfTX-1 and -2 (~43 and 45 kDa, respectively) were the most abundant proteins in both venoms. However, variations in the patterns of minor protein bands were observed, particularly between the molecular mass range 30–45 kDa. Western blot analyses of the two venoms using CSL box jellyfish antivenom revealed similar patterns of antigenic proteins. However, some variation in the pattern of antigenic proteins was observed in the 20–30 kDa range and the lyophilised nematocyst extracts produced a higher background of non-specific antibody binding.

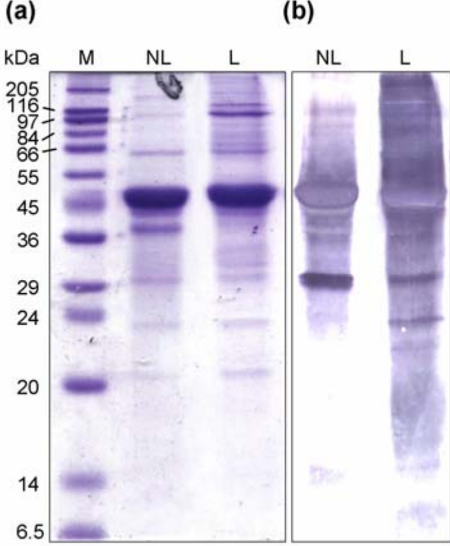


Figure 4.10 Comparison of *C. fleckeri* bead mill venom proteins from non-lyophilised (NL) and lyophilised (L) nematocysts. (a) Coomassie-stained nematocyst proteins; (b) western blot of corresponding proteins using CSL box jellyfish antivenom. Western blots using CfTX-specific antibodies and rabbit antibodies against whole nematocyst extracts were not performed.

Although the SDS-PAGE protein profiles of venoms from lyophilised and non-lyophilised nematocysts were generally similar, variations in the abundances of minor proteins and the pattern of antigenic proteins, suggest the potential for the bioactivity of some minor components to be altered by the lyophilisation process. Hence, whenever practicable, the use of non-lyophilised nematocysts is recommended.

Milked venom, obtained by the electrical stimulation of excised *C. fleckeri* tentacles to induce *in situ* nematocyst discharge (Barnes, 1967), is currently used for the commercial production of CSL box jellyfish antivenom. However, in recent years, the efficacy of the CSL antivenom has been questioned, with some researchers proposing that the antivenom lacks the ability to neutralise the biological effects of one or more components from *C. fleckeri* nematocyst venom (e.g. Endean and Sizemore, 1988; Ramasamy et al., 2003; Winter et al., 2007a). In this study, the SDS-PAGE and western blot profiles of proteins from nematocyst venom and milked venom were compared to examine whether significant differences in venom composition could explain the suggested limitations of CSL box jellyfish antivenom.

An initial comparison of the SDS-PAGE protein profiles of bead mill-derived nematocyst venom and reconstituted lyophilised milked venom indicated a significant difference in protein composition between the two venom sources (Figure 4.11a). Although CfTX-1 and -2 were present in the milked venom, the profile was dominated by a broad and diffuse 16–20 kDa band, which was absent from the nematocyst venom. Western blot profiles of nematocyst venom and reconstituted milked venom were also significantly different (Figure 4.11b). In the milked venom, the dominant 16-20 kDa protein band and the CfTX proteins were strongly antigenic to CSL box jellyfish antivenom, suggesting the antivenom has the capability to attenuate any potential biological activity associated with these proteins. Similar results were obtained for milked venom in a previous study. Following non-reducing SDS-PAGE analysis of milked venom, a major doublet of protein bands at 17 kDa and a minor band at 40 kDa (likely to be unresolved CfTX-1 and -2) were observed (Wiltshire et al., 2000). Western blot analysis revealed that the proteins were also strongly antigenic towards CSL box jellyfish antivenom. In this study, several minor proteins in the milked venom, only faintly detected in the corresponding Coomassie-stained gel, were also strongly antigenic. The pattern of these antigenic proteins did not correspond well to the pattern of minor antigenic proteins detected in nematocyst venom.

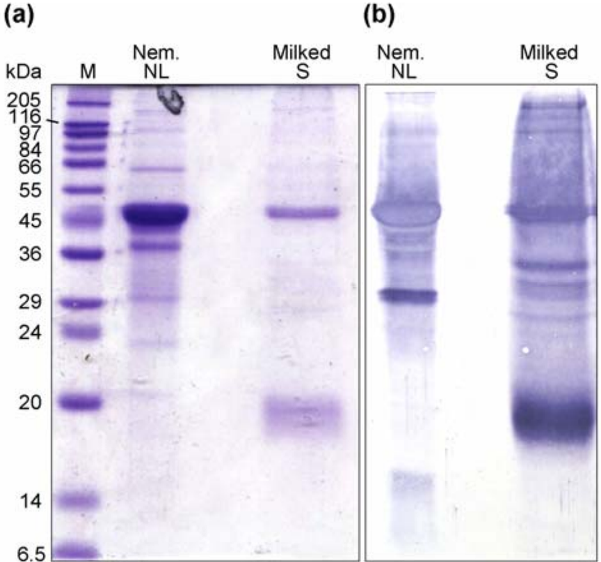


Figure 4.11 Comparison of nematocyst extracts (Nem.) and milked venom (Milked). Nematocyst extract was obtained by bead mill homogenisation of non-lyophilised nematocysts (NL). Milked venom was obtained by a method similar to Barnes (1967) and involved electrical stimulation of excised *C. fleckeri* tentacles, then collection and lyophilisation of the nematocyst discharge. S represents the soluble fraction (supernatant) obtained from the lyophilised venom (a) Coomassie-stained bead mill and milked venom proteins; (b) western blot of corresponding proteins using CSL box jellyfish antivenom.

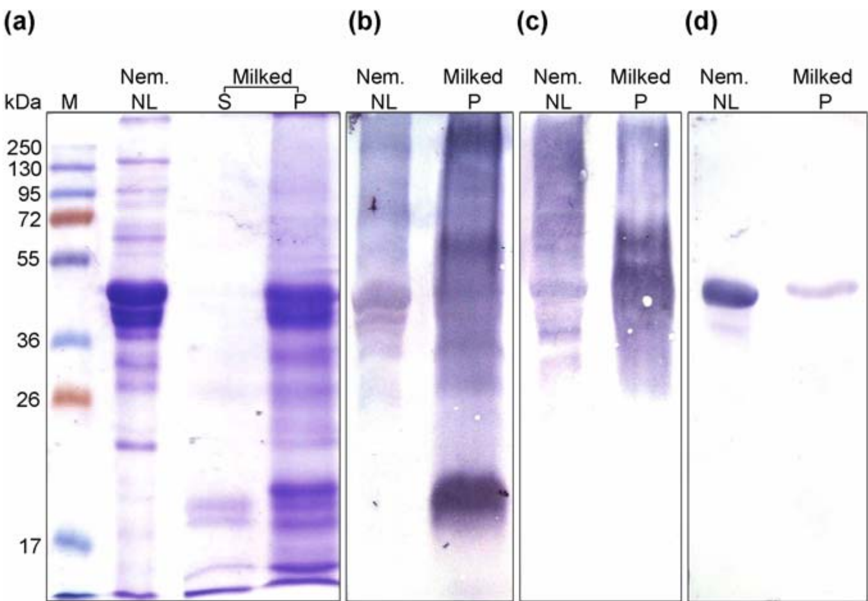


Figure 4.12 Further comparison of bead mill nematocyst extracts (Nem.) from non-lyophilised nematocysts (NL) and milked venom (Milked). The soluble (S) fraction and the insoluble pellet (P) of lyophilised milked venom were analysed by SDS-PAGE, but only the pellet was used for western blot analyses. (a) Coomassie-stained bead mill and milked venom proteins; western blot analyses of corresponding proteins using (b) CSL box jellyfish antivenom, (c) rabbit antibodies against nematocyst extracts and (d) rabbit CfTX-specific antibodies.

Following long-term storage (21 months, –20°C), reconstitution of lyophilised milked venom was much more difficult, suggesting that the solubility of the venom had decreased over time, possibly due to oxidation of the venom proteins. Although reconstituted milked venom and the soluble fraction (supernatant) of the milked venom were both haemolytic to sheep erythrocytes (see also Section 4.3.7), the supernatant appeared colourless rather than dark orange, as observed in the first experiment. Despite concentration of the soluble fraction (~3 times more than the previous experiment), only a doublet of proteins (~17 and 18 kDa) was clearly visible in the SDS-PAGE protein profile (see Figure 4.12a).

In contrast, SDS-PAGE analysis of the insoluble fraction of milked venom proteins revealed that its protein composition was more similar to that of bead mill nematocyst venom than originally considered. Moreover, CfTX-1 and -2 and several other proteins were present in similar proportions. However, the insoluble milked venom contained at least three additional proteins (~17–20 kDa) that were not present in the nematocyst venom, but lacked several higher molecular mass proteins (> 55 kDa) that were present in the nematocyst venom (Figure 4.12a). Visually, the SDS-PAGE profile of the insoluble milked venom components obtained in this study was remarkably similar to that obtained in a previous study in which reconstituted milked venom was analysed (see Olson et al., 1984). The researchers subsequently isolated a 20 kDa lethal, pore-forming cardiotoxin from the milked venom, which possibly corresponds to one (or more) of the 17–20 kDa proteins observed in this study.

The origin of the 17–20 kDa proteins in milked venom remains unclear. A comparison of western blots using CSL box jellyfish antivenom and rabbit antibodies raised against nematocyst venom revealed that the 17–20 kDa insoluble milked venom proteins are strongly antigenic to the CSL antivenom, but not to the rabbit antibodies (Figure 4.12b and c). These results suggest that the 17–20 kDa milked venom proteins are not present or abundant in typical bead mill nematocyst venom, nor subunits/degradation products of other antigenic nematocyst venom proteins. In western blots using CfTX-specific antibodies, CfTX-1 and -2 were detected in both nematocyst and milk venoms. In contrast, no binding of the CfTX-specific antibodies to the 17–20 kDa insoluble milked venom proteins was observed (Figure 4.12d), indicating that the proteins are unlikely to be subunits of the CfTX proteins, as otherwise proposed for proteins of similar molecular mass in a previous study (Wiltshire et al., 2000). There is some evidence to suggest that the 17–20 kDa proteins may be due to contamination of the milk venom with nematocyst debris, due to the presence of proteins similar in molecular mass in pelleted empty nematocyst capsules (see Figure 4.7 and Figure 4.8). However, without amino acid sequence

information or the availability of specific antibodies, the identity, function and origin of the 17–20 kDa proteins is difficult to establish.

The results of this study also highlight the dramatic impact that venom protein solubility following lyophilisation can have on milked venom composition. Inadequate reconstitution of milked venom has the potential to affect venom bioactivity and/or the quality of antivenom production, and may therefore explain some of the discrepancies observed in pharmacological results between previous researchers (see Chapter 1). Indeed, milked venom may be lacking in one or more bioactive components due to inadequate solubilisation of the full suite of venom proteins, rather than inefficiencies in the milking method itself. If reduced venom protein solubility is caused by protein oxidation, this problem could be overcome by preparation and subsequent storage of venom in an inert atmosphere (such as N_2).

4.3.7 Interspecies Comparison of Nematocyst Extracts

C. fleckeri and *C. bronzie* are closely related box jellyfish belonging to family Chirodropidae (see Chapter 1, Figure 1.2). The cnidome of the two species are very similar, except in mature jellyfish, the proportion of venom-injecting mastigophores is greater in *C. fleckeri* than in *C. bronzie* (Carrette et al., 2002a). Pharmacological studies *in vitro* and *in vivo* suggest that nematocyst venoms from both species also share similar neurotoxic and cardiotoxic activities, but the activities of *C. bronzie* venom are relatively less potent than those of *C. fleckeri* venom (Ramasamy et al., 2003, 2005a; Winter et al., 2007a). Consequently, the composition of *C. fleckeri* and *C. bronzie* venoms may be quite similar and the observed differences in biological activities could be related simply to variations in total injectable venom concentration in each jellyfish. However, cardiovascular studies undertaken in rats have revealed that while prophylactic administration of CSL box jellyfish antivenom prevented *C. fleckeri* venom-induced cardiovascular collapse in 40% of animals, the antivenom was ineffective against the cardiovascular effects of *C. bronzie* nematocyst venom (Ramasamy et al., 2004, 2005a). And, despite reports of similar cardiac and lethal effects of *C. fleckeri* and *C. bronzie* nematocyst venoms in crayfish (Carrette, 2002), *C. fleckeri* venom was significantly more lethal in fish than *C. bronzie* venom and produced different effects on cardiac function (Kintner et al., 2005). These findings suggest that *C. fleckeri* and *C. bronzie* venoms may also differ significantly in composition as well as overall venom concentration. Therefore, the aim of this study was to compare the protein composition of nematocyst venoms from *C. fleckeri* and *C. bronzie* and determine whether differences in the pattern and abundance of venom

proteins could potentially explain the differences in venom-induced biological activities reported for the two species.

Unfortunately, the study was hindered by the limited amount of *C. bronzie* nematocysts that were available for venom protein extraction. Consequently, the use of relatively small amounts of *C. bronzie* proteins in SDS-PAGE and western blot analyses made comparisons between the two venoms difficult. In addition, unlike the *C. fleckeri* nematocysts, the nematocysts used to extract *C. bronzie* venom had been lyophilised prior to analysis, which may have slightly altered the *C. bronzie* venom protein composition (see Section 4.3.6).

Despite these limitations to the study, a comparison of the SDS-PAGE protein profiles of *C. fleckeri* and *C. bronzie* nematocyst venoms revealed significant differences in overall venom composition between the two chirodropids (Figure 4.13a). Only four proteins of similar molecular mass were common to both species (~28, 32 43 and 45 kDa), suggesting that the differences in venom composition may contribute to the differences in biological activities observed between species. Two of the proteins (~43–45 kDa) were similar in molecular mass to CfTX-1 and -2 in *C. fleckeri* venom, initially suggesting the presence of putative *C. bronzie* homologues. However, western blot analysis of *C. bronzie* venom revealed that the venom proteins were not antigenic towards CfTX-specific antibodies (Figure 4.13b). The inability of the CfTX-specific antibodies to detect homologues of CfTX-1 and -2 in *C. bronzie* venom is an interesting finding because not only are the CfTX proteins the major proteins in *C. fleckeri* venom, but as detailed in Chapter 5, lethal, haemolytic homologues of CfTX-1 and -2 have been isolated from other venomous box jellyfish, including *C. quadrigatus*, *C. rastoni* and *C. alata* (Nagai et al., 2002, 2000a, b). Therefore, since *C. bronzie* is more closely related to *C. fleckeri* than *C. rastoni* and *C. alata* (see Chapter 1, Figure 1.2), it seems unlikely that homologues of CfTX-1 and -2 are non-existent in *C. bronzie* venom. However, a reduced abundance of these major proteins in *C. bronzie* venom could explain why *C. fleckeri* venom is significantly more lethal than *C. bronzie* venom (see Ramasamy et al., 2005a) and thus emphasises the importance of further research to investigate the biological role(s) of CfTX-1 and -2 in *C. fleckeri* nematocyst venom.

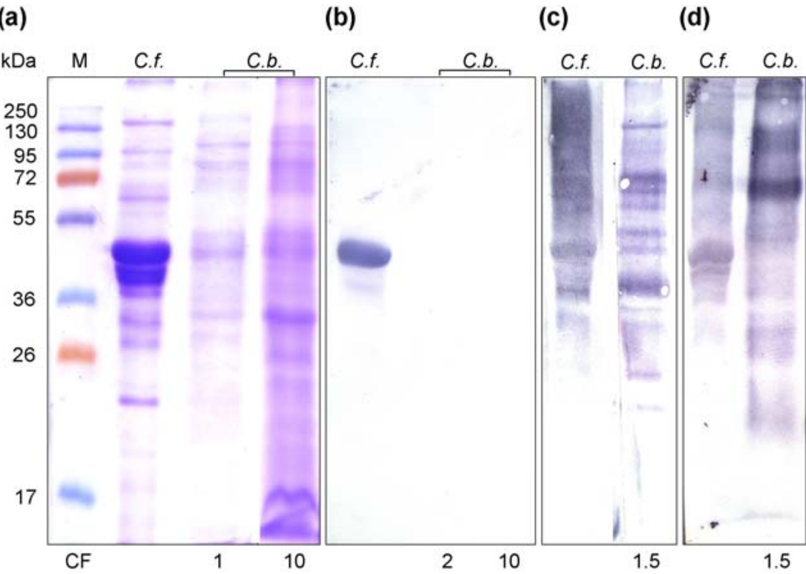


Figure 4.13 Comparison of nematocyst venom proteins from *C. fleckeri* (*C.f.*) and *C. bronzie* (*C.b.*). Variable amounts of *C. bronzie* venom proteins were applied to the SDS-PAGE gels, due to concentration of the nematocyst venom by TCA precipitation or by application of different sample volumes. The concentration factor (CF), indicated below each *C. bronzie* lane, represents the amount of applied extract proteins relative to 10µL of the original extract (CF = 1). The amount of applied *C. fleckeri* venom proteins was constant. (a) Coomassie-stained nematocyst proteins; western blot analyses of the corresponding proteins using (b) rabbit CfTX-specific antibodies, (c) rabbit antibodies against nematocyst extracts and (d) CSL box jellyfish antivenom.

A comparison of the western blot results also showed that although several discrete *C. bronzie* nematocyst venom proteins were antigenic to rabbit antibodies raised against *C. fleckeri* nematocyst venom (Figure 4.13c), relatively few proteins were distinctly antigenic to CSL box jellyfish antivenom (Figure 4.13d). These results suggest that antivenom prepared from *C. fleckeri* nematocyst may be more effective in neutralising or inhibiting the venom-induced effects of *C. bronzie* envenoming than antivenom prepared from milked venom. However, this hypothesis requires further investigation.

Similar to a previous study (Bailey et al., 2005), haemolytic activity was detected in the nematocyst venoms of both *C. fleckeri* and *C. bronzie*. However, as shown in Figure 4.14, the cleared zone of haemolysis caused by *C. bronzie* nematocyst venom (~12µg total protein) was very small compared to that of *C. fleckeri* nematocyst venom (~5µg total protein). These results suggest that *C. fleckeri* venom contains one or more cytolytic proteins that are significantly more abundant or potent than those of *C. bronzie* venom. The dominance of CfTX-1 and -2 in *C. fleckeri* nematocyst venom could explain the significant difference in cytolytic activity between species, however, further research is necessary to fractionate the nematocyst venoms

from each species to compare the biochemical and molecular characteristics of individual cytolysins contained in each venom.

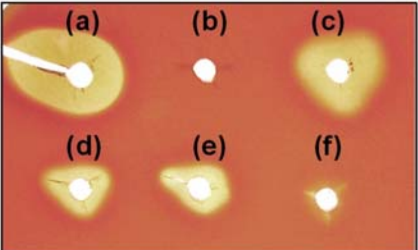


Figure 4.14 Haemolytic activity of *C. fleckeri* and *C. bronzie* venoms on 5% sheep blood agar. (a) Positive control (1% Triton X-100 in NEB); (b) negative control (NEB); (c) *C. fleckeri* nematocyst venom (~5µg); (d) resuspended *C. fleckeri* milked venom; (e) *C. fleckeri* milked venom supernatant. For further details on (d) and (e) see Section 4.3.6; (f) *C. bronzie* nematocyst venom (~12µg). Crude venoms and controls were applied to wells in blood agar and the plates were incubated ~2h at 37ºC. Haemolysis was indicated by the formation of cleared zones in the agar surrounding the wells, similar to the positive control. No evidence of clearing was observed in the negative control.

4.3.8 Antibody Specificity towards *C. fleckeri* Venom Proteins

Western blot analyses of nematocyst venom proteins were performed using rabbit and mouse antibodies raised against nematocyst venom, and CSL box jellyfish antivenom raised in sheep against milked venom. A comparison of the western blot results with the SDS-PAGE protein profile of nematocyst venom revealed that several but not all of the nematocyst-derived proteins were antigenic (Figure 4.15 and Figure 4.16). Some variations in the pattern of antigenic proteins were also observed between the three different antibody types. However, all three antibodies were effective in the detection of antigenic proteins such as CfTX-1 and -2 and several proteins in the 20–35 kDa molecular mass range. Antigenic protein patterns for the rabbit antibodies and CSL box jellyfish antivenom appeared to be most similar, despite the different source of antigen used for their preparation. In comparison, the mouse antibodies strongly bound to least two additional proteins in the 20–35 kDa molecular mass range. Some non-specific binding of normal rabbit and mouse sera to the nematocyst proteins observed in this study, however, the amount of non-specific binding was relatively low compared to the binding observed for the antibody preparations. Normal sheep serum was unavailable at the time of the study for comparison.

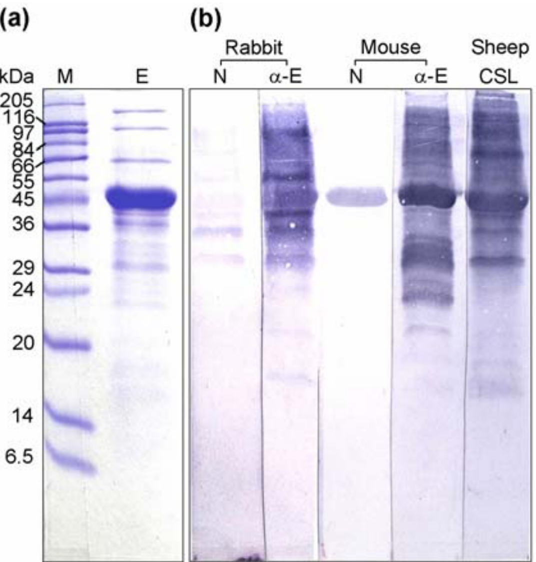


Figure 4.15 Binding of rabbit, mouse and sheep antibodies to *C. fleckeri* bead mill nematocyst proteins (E). Nematocyst proteins (~20μg/lane) were separated in a 15% SDS-PAGE gel. (a) Coomassie-stained nematocyst proteins; western blot analyses of the corresponding proteins using (b) normal (N) sera from rabbit and mouse, antibodies raised against bead mill nematocyst proteins (α-E) in rabbit and mouse, and CSL box jellyfish antivenom raised in sheep. M indicates the molecular marker.

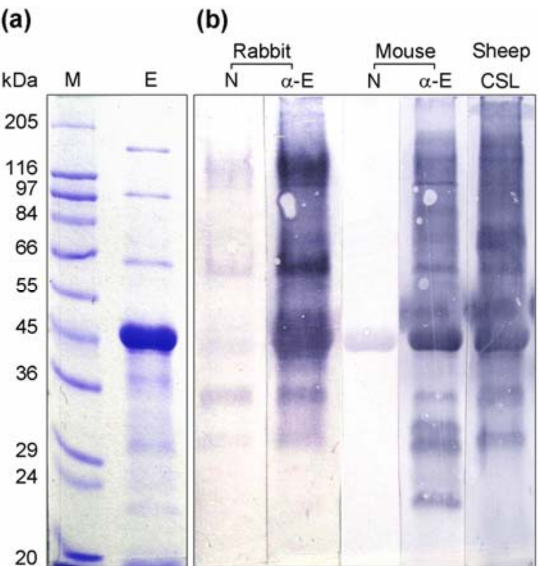


Figure 4.16 Binding of rabbit, mouse and sheep antibodies to *C. fleckeri* bead mill nematocyst proteins (E). Nematocyst proteins (~20μg/lane) were separated in a 10% SDS-PAGE gels. (a) Coomassie-stained nematocyst proteins; western blot analyses of the corresponding proteins using (b) normal (N) and antibodies raised against bead mill nematocyst proteins (α-E) in rabbit and mouse, and CSL ox jellyfish antivenom raised in sheep. M indicates the molecular marker.

The variation in results between the mouse and rabbit antibodies is an interesting finding, particularly because both types of antibodies were prepared using the same venom. Moreover, the rabbit and mouse antibodies were raised against nematocyst venom, but did not bind to all the proteins present in nematocyst venom. These results suggest that the immunological response of mammalian hosts to some *C. fleckeri* venom proteins is species-dependent and in some cases, no significant immunological response is elicited by the proteins. Such findings could have important implications for the production of box jellyfish antivenom for the treatment of envenomed humans, particularly because the efficacy of the antivenom may depend on the host species in which the antibodies are raised. Furthermore, the antivenom would be ineffective in neutralising or attenuating the biological activities (if any) of proteins that are not significantly immunogenic.

In the past, some researchers have suggested that CSL box jellyfish antivenom is unable to neutralise the biological effects of one or more components from *C. fleckeri* nematocyst venom (e.g. Endean and Sizemore, 1988; Ramasamy et al., 2003; Winter et al., 2007a). However, the results of this study indicate that the CSL antivenom binds to several nematocyst venom proteins as effectively as the rabbit and mouse antibodies, in a relatively similar pattern of binding. Consequently, if the CSL antivenom lacks the ability to neutralise all the bioactive components in *C. fleckeri* venom, then antibodies raised against bead mill nematocyst venom may also be partially ineffective, and thus not represent a better treatment for *C. fleckeri* envenoming in humans. Further pharmacological studies are therefore necessary to investigate the efficacy of antibodies raised against bead mill nematocyst venom to inhibit or inactivate the biological activities of *C. fleckeri* venom.

CHAPTER 5

Identification and Molecular Characterisation of Two Major Venom Proteins from *Chironex fleckeri*

5.1 Introduction

The Australasian box jellyfish *Chironex fleckeri* is dangerous to humans due to its ability to inflict extremely painful and potentially life-threatening stings. The rapid onset of symptoms following serious envenomations is well documented (Williamson et al., 1984; Beadnell et al., 1992) and in the most severe cases, death from cardiac and/or pulmonary failure has occurred within minutes (Lumley et al., 1988).

The venom of *C. fleckeri* is a complex mixture of proteins, which is known to be myotoxic, haemolytic, dermonecrotic and lethal (Tibballs, 2006). Previous researchers have attempted to fractionate and characterise the bioactive components of *C. fleckeri* venom (Bloom et al., 1998; Tibballs, 2006 for summaries) and more recently, research has focussed on the biological activities of unfractionated nematocyst extracts *in vitro* and *in vivo* (e.g. Winter et al., 2007a; Bailey et al., 2005; Kintner et al., 2005; Ramasamy et al., 2004; Cuypers et al., 2006). However, despite the accumulation of biochemical, toxicological and pharmacological data over several decades, the underlying mechanism(s) of lethal action of *Chironex* venom at the molecular level remains largely unknown. To date, no amino acid sequence of the toxic proteins contained within *C. fleckeri* venom, or DNA sequence of the genes that encode them, has been reported. In contrast, researchers have isolated, cloned and sequenced lethal, haemolytic proteins from the nematocysts of other box jellyfish including, CaTX-A (43 kDa) from *Carybdea alata* (Nagai et al., 2000a), CrTXs (43-45 kDa) from *Carybdea rastoni* (Nagai et al., 2000b) and CqTX-A (44 kDa) from *Chiropsalmus quadrigatus* (Nagai et al., 2002). Amphiphilic α-helices predicted in the N-terminal regions of the three jellyfish proteins have been suggested as the potential mechanism of cytolytic activity ascribed to the toxins (Nagai et al., 2000b; Nagai et al., 2002).

In this study a proteomic approach was used to identify and sequence the major proteins from the nematocysts of *C. fleckeri* and, with extensive computational analysis of the amino

acids sequences, gain some insight into the possible structures, functions and potential mechanisms of action of the nematocyst proteins.

5.2 Methods and Materials

5.2.1 SDS-PAGE and Western Blot Analysis

The protein profile of *C. fleckeri* nematocyst venom, obtained by bead-mill homogenisation (see Section 2.2.1), was examined using SDS-PAGE, as described in Section 2.3.3. The antigenicity of the separated nematocyst proteins was assessed by western blot analysis (see Section 2.3.4) using CSL box jellyfish antivenom and polyclonal antibodies raised in mice and a rabbit against nematocyst extracts (see Section 2.3.2).

5.2.2 N-Terminal Amino Acid Sequencing and Peptide Mass Fingerprinting

N-Terminal amino acid sequences were determined for two major proteins (43 and 45 kDa) by Edman degradation (see Section 2.3.6) and internal peptide sequences were also determined for the 43 kDa protein by mass spectrometry (see Section 2.3.7).

5.2.3 Isolation of Full-Length cDNA Clones Encoding Two Major Venom Proteins

The full-length cDNA clones of the 43 and 45 kDa proteins were isolated by screening a *C. fleckeri* tentacle cDNA library with a radiolabelled oligonucleotide probe that was specific for both clones. Construction of the cDNA library is described in Section 2.4.12. A series of forward and reverse degenerate oligonucleotide primers were designed according to the 43 kDa protein internal peptides, FIAMVVQR and NDDLEKFNFVPVK, respectively (Table 5.1). RT-PCR was performed for each combination of primers using 2-step RT-PCR, 1μg total tentacle RNA, a poly(dT)$_{12\text{-}18}$ primer for first-strand cDNA synthesis and Ready-to-Go RT-PCR Beads (Amersham Biosciences). Further details on RT-PCR are provided in Section 2.4.8 Amplification was performed using hot start PCR (95°C for 5 min, 35 cycles of 95°C for 1 min, 49°C for 1 min and 72°C for 2 min, then 72°C for 5 min). A 614 bp product obtained using the primer combination 43-F6 / 43-R1 was purified, ligated overnight into pGEM-T (Promega), and the nucleotide sequence was determined using a MegaBACE DNA analysis system and a

DYEnamic ET Dye Terminator Cycle Kit (Amersham Biosciences). The 5'-end of cDNA encoding the 43 kDa protein was amplified from the *C. fleckeri* cDNA library by nested PCR using *Taq* DNA polymerase (Fermentas), under the PCR conditions described earlier, with primers T3 and 5'-GCAAATGAGAGCAAGGTAG-3' (43-R5) for the first round of PCR, and SK and 5'-TCTTTGAACCACCATCGC-3' (43-R6) for second round of amplification. Further details on PCR are provided in Section 2.4.7. A 700bp product was purified, ligated overnight into pGEM-T and several recombinant plasmids were sequenced.

Table 5.1

Degenerate forward (F) and reverse (R) primers designed according to the internal peptide sequences, FIAMVVQR and NDDLEKFNFVPVK, respectively, present in a 43 kDa major *C. fleckeri* nematocyst protein.

ID	Primer Sequence (5'→3')
43-F5	TTYATHGCNATGGTNGTACA
43-F6	TTYATHGCNATGGTNGTTCA
43-F7	TTYATHGCNATGGTNGTGCA
43-F8	TTYATHGCNATGGTNGTCCA
43-R1	GGNACRAARTTRAACTTCTC
43-R2	GGNACRAARTTRAATTTTTC
43-R3	GGNACRAARTTRAACTTTTC
43-R4	GGNACRAARTTRAATTTCTC

Sequence data revealed that two distinct cDNA populations were being amplified. To isolate full-length clones of both cDNAs simultaneously, DNA was amplified from 1μg total RNA by 2-step RT-PCR with poly(dT)$_{12-18}$, a gene-specific forward primer common to both clones, 5'-AAGTTCCAGTCAGGCGAC-3' (43-F10), and 43-R5 under hot start PCR conditions (95°C for 5 min, 35 cycles of 95°C for 45s, 51°C for 1 min and 72°C for 2 min, then 72° for 5 min). A purified 806bp RT-PCR product (25ng) was labelled with [α-^{32}P]-dATP, in a random primed reaction using a DECAprime II Kit (Ambion), and used to screen ~5 × 10^4 plaques from the *C. fleckeri* cDNA library (see Sections 2.4.11 and 2.4.14). The longest clones of 24 randomly selected positive plaques were determined by PCR using the primers SK and 43-R5 (95°C for 5 min, 35 cycles of 95°C for 45s, 51°C for 1 min and 72°C for 2 min, then 72° for 5 min). Six plaques were re-screened with the radiolabelled probe and four independent positive plaques were isolated, excised *in vivo* (see Section 2.4.15) and sequenced in both directions.

5.3 Results

5.3.1 SDS-PAGE and Western Blot Analysis

SDS-PAGE analysis of *C. fleckeri* nematocyst venom revealed the presence of several protein bands ranging in mass from 10 to 180 kDa, but two proteins (43 and 45 kDa) dominated the profile (Figure 5.1). In extracts of high protein concentration, the 43 and 45 kDa proteins often appeared as a single large band. Under reducing and non-reducing analysis conditions, the apparent molecular masses of the two major proteins remained unchanged (Figure 4.3). Although the 43 and 45 kDa proteins could not be completely separated by SDS-PAGE, the bands were sufficiently resolved to enable N-terminal and internal amino acid sequencing. Western blot analysis using the CSL box jellyfish antivenom and the rabbit antiserum showed that the 43 and 45 kDa proteins were strongly antigenic (Figure 5.1; see also Figure 4.15 and Figure 4.16).

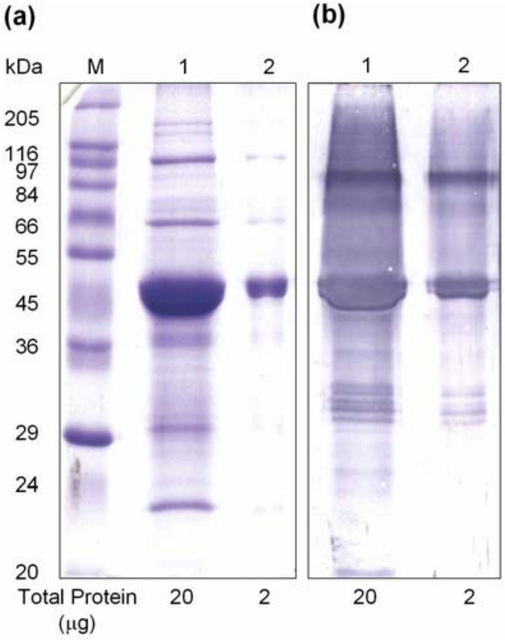


Figure 5.1 SDS-PAGE and western blot analysis of *C. fleckeri* nematocyst venom. (a) Nematocyst venom proteins (20 and 2µg total protein, respectively) were separated in a 15% polyacrylamide gel and Coomassie-stained. M indicates the protein molecular marker (Sigma) and the molecular masses of the protein standards are shown in kilodaltons (kDa); (b) western blot of the corresponding nematocyst venom proteins using CSL antivenom (raised in sheep). Similar results were observed using polyclonal antibodies raised in rabbit against nematocyst venom obtained by bead-mill homogenisation. Total venom protein applied is indicated below each lane.

5.3.2 N-Terminal and Internal Amino Acid Sequencing

The N-terminal amino acid sequences of the 43 and 45 kDa proteins were very similar (ESTISSGLNSLKAKIDAKMP(T)G and ESTISSGIASLKAKI, respectively). During analysis of the 43 kDa protein, a minor signal for alanine (A) was detected at residue 9 (asparagine, N), and residue 21 (threonine, T) was only tentatively called. For the 45 kDa protein, minor signals for leucine (L) and asparagine were detected at residues 8 and 9, respectively. BLAST analysis of the N-terminal sequences showed significant similarity (>80%) to the N-terminal sequence of the jellyfish toxin, CqTX-A, isolated from *C. quadrigatus* (Nagai et al., 2002).

Peptide mass fingerprinting data for trypsin-digested CfTX-1 are provided in Appendix 6. A Mascot Ion/Ion Search identified five peptide ions (out of 30 unique peptide ions) that potentially matched tryptic peptides of CqTX-A: (1) AFLDGVR (777.41), (2) FIAMVVQR (963.52; 979.51, M_o), (3) VLSMLELFTDLCSLR (1811.00, C_a; 1826.98, M_o+C_a), (4) FFDLMKVK (1043.53, M_o) and (5) NDDLEKFNFVPVK (1564.76); observed monoisotopic peptide ion masses (MH^+, Da) and the variable modifications that were considered, namely oxidation of methionine (M_o) and acrylamide-modified cysteine (C_a), are indicated in brackets after each peptide.

5.3.3 Isolation of Full-Length cDNA Clones Encoding Two Major Venom Proteins

A radiolabelled 806 bp RT-PCR product was used to screen the *C. fleckeri* cDNA library and 55 positive plaques, representing ~0.1% of the cDNA clones in the library, were detected. Following isolation and sequencing of 4 independent positive plaques, two distinct cDNA populations were obtained. One cDNA clone was 1789 bp in length and was composed of a 60 bp 5'-untranslated region (UTR), a 1368 bp open reading frame (ORF) and a 342 bp 3'-UTR attached to a poly(A) tail (Figure 5.2). The ORF translated into a 456-residue protein referred to as CfTX-1. The second cDNA clone was 1624 bp long with a 42 bp 5'-UTR, a 1386 bp ORF and a 180 bp 3'-UTR attached to a poly(A) tail (Figure 5.3). The ORF translated to a 462-residue protein referred to as CfTX-2.

The N-terminal sequences of the 43 and 45 kDa proteins were present in the deduced amino acid sequences of CfTX-1 and -2, respectively, and located 20 and 17 residues downstream from their respective initiating methionine. Minor variations in residues between the N-terminal sequences and the deduced amino acid sequences were observed, however, these

variations were attributed to cross contamination resulting from the incomplete separation of the two proteins using SDS-PAGE. Computational analysis of the deduced amino acid sequences (SignalP 3.0, Bendtsen et al., 2004; PSORT II, Horton and Nakai, 1997), predicted the presence of a 20- and 17-residue signal peptide in the N-terminal regions of CfTX-1 and CfTX-2, respectively, consistent with the N-terminal amino acid sequence data. Following signal peptide cleavage, the masses of mature CfTX-1 and CfTX-2 were calculated to be 49,146 Da and 49,883 Da, respectively, and their isoelectric points (pI) were estimated to be 8.3 and 7.8, respectively (ProtParam; Gasteiger et al., 2005).

The five tryptic peptides of the 43 kDa protein, identified by mass spectrometry, were present in the deduced amino acid sequence of CfTX-1, however, the CfTX-1 sequence corresponding to peptide 3 varied by two residues, as indicated in Figure 5.2. Despite repeated nucleotide sequencing of the isolated clones obtained by DNA screening, and sequencing of clones obtained by library PCR amplification, these variations could not be attributed to sequence inaccuracies. However, the variations may be explained by the presence of another previously unmatched peptide mass (MH^+=1888.86 Da) in the mass spectrometry data (Appendix 6), which is similar to the calculated mass (MS-Digest, Clauser et al., 1999) of the deduced CfTX-1 sequence, MLTMLELFTDLCSLR (MH^+=1888.92 Da, $2M_o$+C_a). Also, peptide 4 (FFDLMKVK) was present in CfTX-1, but not proceeded by arginine (R) or lysine (L), as expected for a tryptic peptide. Again, no nucleotide sequencing errors were detected, however, the calculated mass of a peptide from CfTX-2 (LSDGFYTIK, MH^+=1043.54 Da) matched the mass expected for peptide 4. Hence, detection of the peptide ion may have been due to contamination of CfTX-1 with CfTX-2, as observed for the N-terminal sequence data. Moreover, eight other unassigned peptide masses in the mass fingerprinting data for CfTX-1 matched calculated masses of CfTX-2 tryptic peptides (see Appendices 6 and 7). Alternatively, the presence of asparagine (N) rather than R or L could be due to mutation(s) in the codon sequence of the residue preceding peptide 4, which may have occurred during reverse-transcription of tentacle mRNA while constructing the cDNA library. In this case, a single nucleotide mutation at the third position of the codon may have resulted in a change in amino acid-coding from K (AAA or AAG) to N (AAT).

A pairwise sequence alignment of the mature CfTX-1 and -2 amino acid sequences (ClustalW; Thompson et al., 1994) revealed that the two proteins are significantly homologous (73%). BLAST analysis of the mature CfTX-1 and -2 amino acid sequences showed that the two CfTX isoforms also share significant homology with only three known proteins: CqTX-A

from *C. quadrigatus* (89% and 72%, respectively), CrTXs (CrTX-A/-B) from *C. rastoni* (27% and 26%) and *C. alata* (24% and 25%, respectively). Multiple sequence alignment of the five similar box jellyfish amino acid sequences (ClustalW) highlighted the presence of several conserved amino acids, and in particular, short sections of highly conserved amino acids located towards the N-terminal end of the proteins (see Figure 5.4).

Five cysteine residues are present in the CfTX-1 and -2 sequences, and all are conserved in the three chirodropid sequences (CfTX-1, CfTX-2 and CqTX-A). Under reducing and non-reducing conditions, the apparent molecular masses of CfTX-1 and -2 remained unchanged, which suggests that any disulphide bond formation in the proteins is intramolecular. Three of the five cysteine residues are also conserved in the carybdeid sequences (CrTXs and CaTX-A), which each contain four cysteine residues. Computational analysis of the chirodropid sequences (DISULFIND, Ceroni et al., 2006) predicted that none of the conserved cysteine residues are likely to form intramolecular disulphide bridges. In comparison, one disulphide bridge was predicted in CrTXs, connecting C_{319} to C_{332}, and two disulphide bridges were predicted in CaTX-A, connecting C_{82} to C_{277} and C_{328} to C_{341}. However in both cases, confidence in the disulphide bonding state prediction was low.

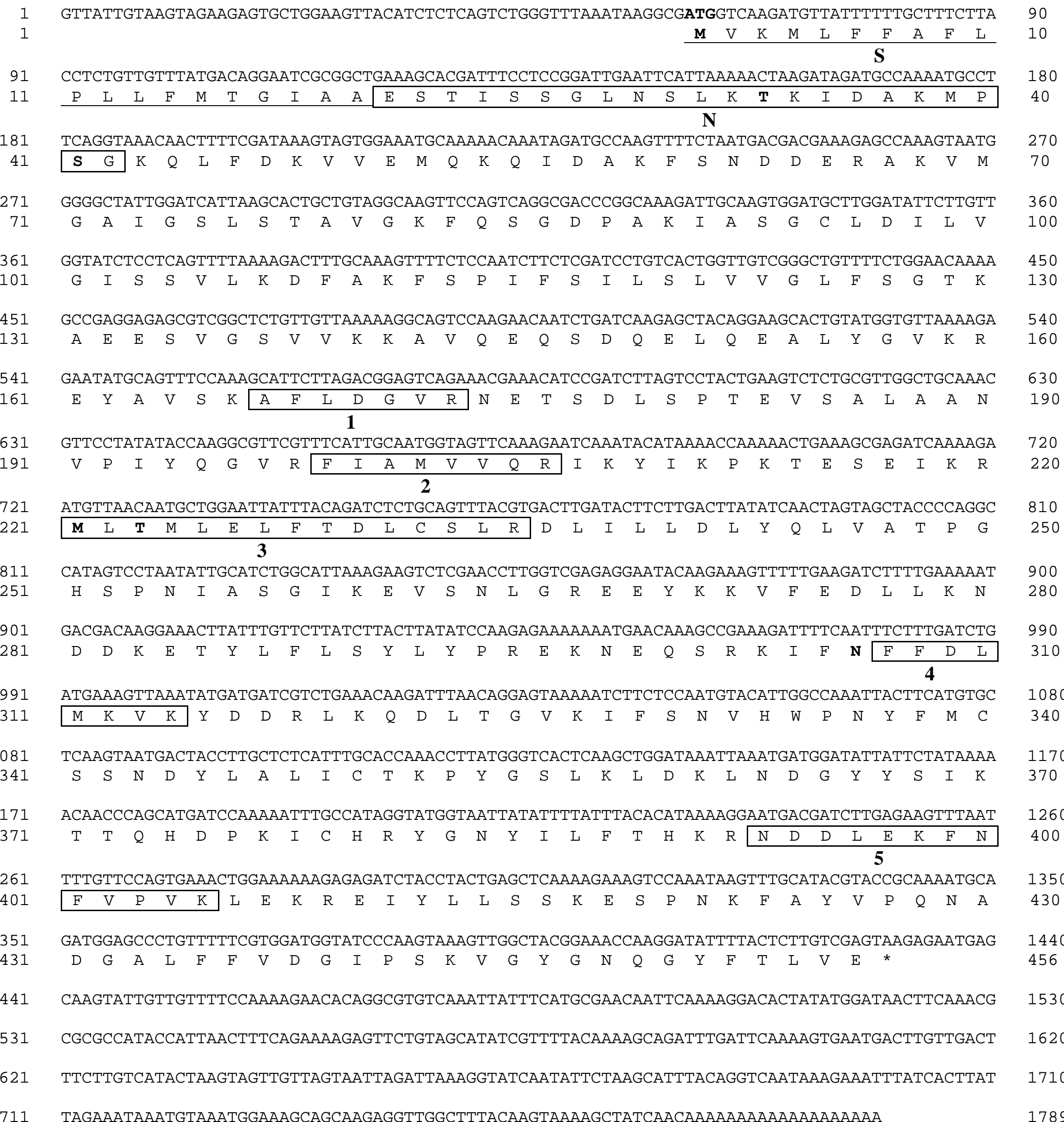


Figure 5.2 Nucleotide and deduced amino acid sequences of CfTX-1 (GenBank Accession No. EF636902). The initiating methionine (M_i) in bold is included in the underlined 20-residue signal peptide (S), which precedes the boxed N-terminal peptide sequence (N). Boxed amino acids, numbered 1 to 5, correspond to internal amino acid peptides determined by mass spectrometry. Other amino acids in bold correspond to variations in amino acid sequence between deduced and N-terminal/MS data. An asterisk indicates the first stop codon in-frame with M_i.

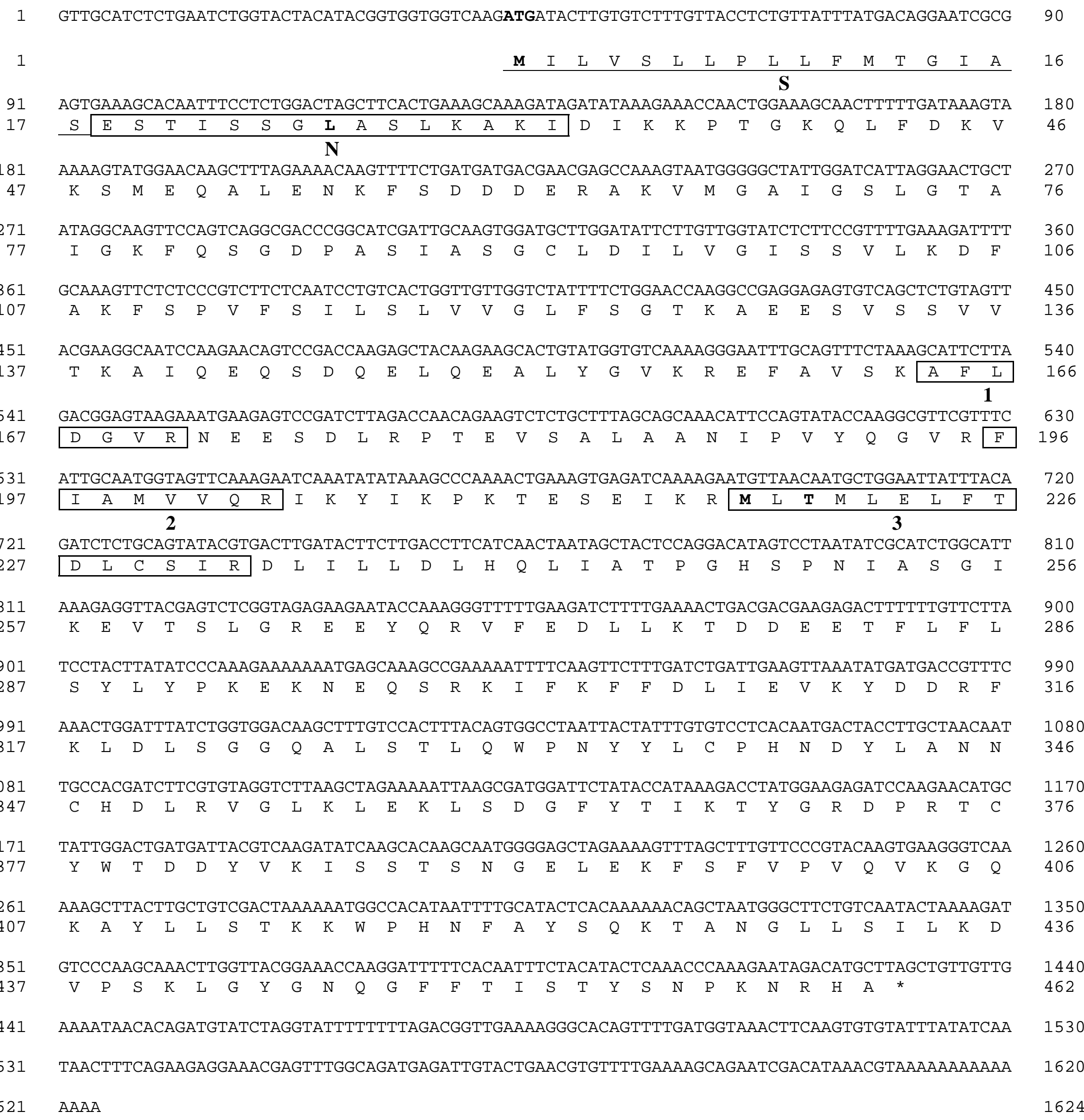


Figure 5.3 Nucleotide and deduced amino acid sequences of CfTX-2 (GenBank Accession No. EF636903). The initiating methionine (M_i) in bold is included in the underlined 17-residue signal peptide (S), which precedes the boxed N-terminal peptide sequence (N). Boxed amino acids, numbered 1 to 3, correspond to internal amino acid peptides determined by mass spectrometry for CfTX-1. Variations in amino acid sequence between deduced and N-terminal/MS data are highlighted in bold. An asterisk indicates the first stop codon in-frame with M_i.

```
                                  amphiphilic α-helices
CfTX-1     1 -----ESTISSGLNSLKTKIDAKMPSGKQLFDKVVEMQKQIDAKFSNDDER--AKVMGAIGSLSTAVGKF  63
CfTX-2     1 -----ESTISSGLASLKAKIDIKKPTGKQLFDKVKSMEQALENKFSDDDER--AKVMGAIGSLGTAIGKF  63
CqTX-A     1 -----EGTISSGLASLKAKIDAKRPSGKQLFDKVANMQKQIEEKFSNDDER--AKVMGAIGSLSTAVGKF  63
CrTXs      1 -----KHGKRSDVNSLLTKVETALKEASGSNEAALEALEGLKGEIQTKPDR-VGQATKILGSVGSALGKL  64
CaTX-A     1 RLSRYRRSAADAVSTDIDGIIGQLNDLGTDTKRLKEALQGVQEAVKKEPATTIAKVSTIVGSVGGSLSKF  70
                          . .    .                 .          .     .**.  .. *

                                         TSR1
CfTX-1    64 QSGDPAKIASGCLDIILVGISSVLKD------FAKFSPIFSILSLVVGLFSGTKAEESVGSVVKKAVQEQS 127
CfTX-2    64 QSGDPASIASGCLDIILVGISSVLKD------FAKFSPVFSILSLVVGLFSGTKAEESVSSVVTKAIQEQS 127
CqTX-A    64 QSGDPAKIASGCLDIILVGISSVLKD------FAKFSPIFSILSMVVGLFSGTKAEESVGSVVKKVVQEQS 127
CrTXs     65 NSGDATKIISGCLDIVAGIATTFGGPVGMGIGAVASFVSSILSL----FTGSSAKNSVAAVIDRALSKHR 130
CaTX-A    71 KSGDPFDVASGCLDIIASVATTFGGPYGIAIGAVASLISSILSL----FSGNSMGSAIKQVIDDAFKKYR 136
               ***   . ******.  ...                *   *  . ****.     *.*      ..   *.

CfTX-1   128 DQELQEALYGVKREYAVSKAFLDGVRNETSDLSPTEVSALAANVPIYQGVRFIAMVVQRIKYIKPKTES- 196
CfTX-2   128 DQELQEALYGVKREFAVSKAFLDGVRNEESDLRPTEVSALAANIPVYQGVRFIAMVVQRIKYIKPKTES- 196
CqTX-A   128 DQELQEALYGVKREYAVSKAFLDGVRNETSDLSPTEVSALGANVPVYQGVRFIAMVVQRIKNRKPRTES- 196
CrTXs    131 DEAIQRHAAGAKRDFAESSAFIQVMK-QQSNLTDSDLSIIAANVPVYKFSNFIGQLESRISQGAATTSLS 199
CaTX-A   137 DQELEDNVKGAKRTFNAVITFVNSVS-KTENLTEVHLDSVRDAVRVDAFTNMLGVLESRINRGSVSTDNN 205
             *. ..     * ** .     .*.  .        *     .   .    . .       .  .  **       *

CfTX-1   197 EIKRMLTMLELFTDLCSLRDLILLDLYQLVATPGHSPNIASGIKEVSNLGREEYKKVFEDLLKNDDKETY 266
CfTX-2   197 EIKRMLTMLELFTDLCSIRDLILLDLHQLIATPGHSPNIASGIKEVTSLGREEYQRVFEDLLKTDDEETF 266
CqTX-A   197 EIKRVLSMLELFTDLCSLRDLILLDLYQLVATPGHSPNIASGIKEVSNLGREEYKKVFEDLLKTNDKETY 266
CrTXs    200 DAKRAVDFILLYCQLVVMRETLLVDLAILYRKGN--AEHVASAVENANRVNKELAADTLDFLHKLIPEQA 267
CaTX-A   206 EAMRTINFIFLYLQLSVMRETLLTQVILLYKRAGGAYDELALSLSLTSDQNKEATRETVTFLHQMETKYS 275
             .   *  .   . *.   *   .*.  .*   .   *        .         .        *           *.

CfTX-1   267 LFLSYLYPREKNEQSRKIFNFFDLMKVKYDDRLKQDLTGVKIFSNVHWPNYFMCSS---NDYLALICTKP 333
CfTX-2   267 LFLSYLYPKEKNEQSRKIFKFFDLIEVKYDDRFKLDLSGGQALSTLQWPNYYLCPH---NDYLANNCHDL 333
CqTX-A   267 LFLSYLYPRERNEQSQKIFKFFDLMKVKYDDRLKQDLTGIQVFSSLHWPNYFLCSS---KDYLALICTKP 333
CrTXs    268 LIGAVYHPISASETSKAILNYTKYFGVPDVPRPIGN--RRYKFTNSYWNTYSICSEAYMGNYMFRGCSNV 335
CaTX-A   276 LCGSYYYPIDHSKAAIGILKLTKFFGVPDPARYTFDG-LYYRMQNRAWNRYSICKESYAGNHMFRGCKDS 344
             *   .    *        .   *          *     *                 *   *  .*          .    *

CfTX-1   334 -YGSLKLDKLNDGYYSIKTTQHDPKICHRYGNYILFTHKRND--DLEKFNFVPVKLEKREIYLLSSKESP 400
CfTX-2   334 -RVGLKLEKLSDGFYTIKTYGRDPRTCYWTDDYVKISSTSNG--ELEKFSFVPVQVKGQKAYLLSTKKWP 400
CqTX-A   334 -YGSLRLDKLNDGFYSIKTTQSNPKVCHRYGEYILFTHDRND--DLEKFNFVPVKLGERKIYLLSSKASP 400
CrTXs    336 RNPNIRVSKMSDGFYTMENSDRRKLYITKHDQGWGWGTLDEDPGDQGHMRFIPLRHG---KYMVSSKRWP 402
CaTX-A   345 SYHGIRIKKLENGYHTITLRS-KAMYVTKHAQGWGWGTADEDPGEQGYFTFIPLTNG---FYMVSTKKWP 410
                 ... *.  *. ..                                .       *.*.         *..*.*   *

CfTX-1   401 NKFAYVPQNADGALFFVDGIPSKVGYGNQGYFTLVE--------- 436
CfTX-2   401 HNFAYSQKTANGLLSILKDVPSKLGYGNQGFFTISTYSNPKNRHA 445
CqTX-A   401 NKFAYVPKTAKGDLFFVDGIPSQLGYGNQGYFTLATDENEQT--- 442
CrTXs    403 NWFMYMESSASGYIRSWENNP-----GPQGHWSIT---------- 432
CaTX-A   411 DYFVYMESSAHGYIRSWHYNP-----DPQGQWKIL---------- 440
               * *     * * .       *        **     .
```

Figure 5.4 Comparison of five homologous box jellyfish proteins using ClustalW multiple amino acid sequence alignment: CfTX-1 and -2 (*C. fleckeri*), CqTX-A (*C. quadrigatus*), CrTXs (*C. rastoni*) and CaTX-A (*C. alata*). Identical amino acids are marked by an asterisk and similar amino acids are marked with a stop. Cysteine residues (C) and sections of highly conserved and consecutive residues are highlighted in bold. Dashes indicate gaps introduced for better alignment. Double lines above and below the five sequences indicate a transmembrane spanning region (TSR1). Underlined sequences indicate transmembrane helices predicted outside TSR1 and broken lines indicate putative amphiphilic α-helices. GenBank Accession numbers for CqTX-A, CrTXs and CaTX-A are AB045319, AB015878 and AB036714, respectively.

5.3.4 Phylogenetic Analysis

The phylogenetic relationships of the five CfTX-like jellyfish proteins were assessed by constructing an unrooted phylogenetic tree using the distance matrix method, neighbour joining, and Poisson correction for estimation of evolutionary distances (MacVector, Accelrys). The reliability of the phylogeny was tested by *bootstrap* analysis using 1×10^5 replications and values of 100% were obtained for each branch. Phylogenetic analysis revealed divergence in the amino acids sequences, separating the chirodropid proteins (CfTX-1, CfTX-2 and CqTX-A) from the carybdeid proteins (CrTXs and CaTX-A), and that among the chirodropid proteins, CqTX-A appears to be more closely related to CfTX-1 than to CfTX-2 (Figure 5.5).

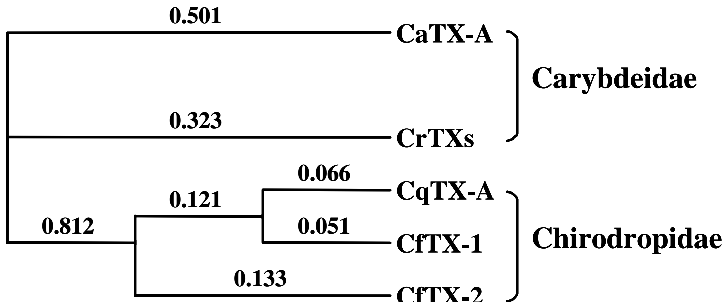


Figure 5.5 Phylogenetic relationships of five homologous box jellyfish proteins: CfTX-1 and -2 (*C. fleckeri*), CqTX-A (*C. quadrigatus*), CrTXs (*C. rastoni*) and CaTX-A (*C. alata*). Sequence divergence is observed between members of Family Chirodropidae (*C. fleckeri* and *C. quadrigatus*) and members of Family Carybdeidae (*C. rastoni* and *C. alata*). The unrooted tree was constructed by a neighbour joining Poisson-corrected distance matrix method, with gaps distributed proportionally (MacVector, Accelrys). *Bootstrap* analysis (1×10^5 replications) returned values of 100% for each branch.

5.3.5 Secondary Structure Analysis

Secondary structure analysis by four different methods predicted that CfTX-1 and -2 contain a mixture of α-helix, β-strand and loop structures (JPred, Cuff and Barton, 1999; SAM-T99, Karplus et al., 1998; PredictProtein, Rost et al., 2004; PsiPred, Jones, 1999). The consensus of the prediction methods suggests that the first 300 amino acids of the mature CfTX isoforms predominantly comprise α-helices and loop regions with at least 12 α-helices and one tentatively predicted β-strand. In contrast, the remainder of the CfTX sequences were predicted to comprise exclusively β-strands (~12) and loop structures. Similar secondary structure prediction patterns were observed for the amino acid sequences of CqTX-A, CrTXs and CaTX-A using the same methods.

One of the longest α-helices (> 30 residues) consistently predicted for the five cubozoan sequences spanned amino acids 195–227 in the three chirodropid sequences, 198–229 in CrTXs and 204–236 in CaTX-A. Hydropathy plots showed the helix to be relatively hydrophobic. An amphiphilic α-helix was predicted in the N-terminal region of CfTX-1 and -2 spanning amino acids 24–33 (MacVector). In comparison, putative N-terminal amphiphilic α-helices were predicted in CqTX-A, CaTX-A, and less confidently in CrTXs, spanning amino acids 24–37, 33–44 and 11–19, respectively.

Five different methods were used to predict the presence of transmembrane helices in the mature CfTX isoforms. Four methods predicted the presence of one transmembrane helix in CfTX-1 and -2, at similar locations within the amino acid range 78–112 (PHDhtm, Rost et al., 1996; TopPred, Claros and von Heijne, 1994; ALOM2, Nakai and Kanehisa, 1992; HMMTop, Tusnády and Simon, 1998). One method predicted the presence of two transmembrane helices in CfTX-2, spanning amino acids 68–86 and 89–107 (TMpred, Hofmann and Stoffel, 1993). Therefore, the broader amino acid range 68–112 and the corresponding aligned sequences of CqTX-A, CrTXs and CaTX-A were defined as transmembrane spanning region 1 (TSR1), as shown in Figure 5.4. In comparison, four methods predicted at least one transmembrane helix within TSR1 for CqTX-A, CrTXs and CaTX-A (PHDhtm, TMpred, TopPred and HMMTop) and one method predicted the presence of one transmembrane helix in TSR1 for the carybdeid proteins but not for CqTX-A (ALOM2). Construction of hydrophobicity profiles for the five homologous amino acid sequences (MacVector) supported the prediction of a single transmembrane helix in TSR1, however in most cases, two peak maxima were observed within the region, suggesting that two closely positioned transmembrane helices may be present in the jellyfish proteins. Additional transmembrane helices were predicted in the cubozoan sequences at various locations outside TSR1 by three of the methods (PHDhtm, HMMTop and TMPred), however, the locations of the helices varied and predictions were less consistent than those within TSR1.

5.3.6 Tertiary Structure Analysis

No putative conserved domains for CfTX-1 or -2 were detected using CDART, the Conserved Domain Architecture Retrieval Tool (Geer et al., 2002), however, remote protein homology searches of the Structural Classification of Proteins (SCOP) database (SUPERFAMILY; Madera et al., 2004) predicted that the chirodropid proteins (CfTX-1, CfTX-2 and CqTX-A) each contain a central domain within amino acids 212–333 with structural

similarity to the cytokine superfamily (*E*-values[**] < 6.6e–06). Sub-family classification indicated that the domains might be related to interleukin-1 (IL-1), however classifications were weak (*E*-values ranging from 0.021 to 0.032), indicating that the domains are possibly members of a sub-family for which no structure has yet been solved (Madera et al., 2004). In comparison, no significant superfamily matches were found for the carybdeid proteins (CrTXs and CaTX-A), but CaTX-A produced weak matches to the δ-endotoxin, N-terminal domain superfamily (*E*-value = 2.9e–03) and the porin superfamily (*E*-value = 0.47) in amino acid ranges 74–146 and 37–439, respectively.

Remote protein homology searches against the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) by another method (SAM-T06; Karplus et al., 2005) found no significant homology between CfTX-1, CqTX-A and CrTXs and other known protein structures. However, the best remote homology matches to CfTX-2 were three insecticidal δ-endotoxins from *Bacillus thuringiensis*, Cry1Aa (PDB ID 1ciy; *E*-value = 0.838), Cry3Bb1 (PDB ID 1ji6A; *E*-value = 0.841) and Cry3A (PDB ID 1dlc; *E*-value = 2.0), spanning the CfTX-2 amino acid range 68–445, although predictions were very weak. Similarly for CaTX-A, the top three weak remote homology matches were Cry3Bb1, Cry1Aa and Cry3A (*E*-values of 0.298, 0.401 and 1.81, respectively) in the CaTX-A amino acid range 78–438.

Since publication of these predictions (Brinkman and Burnell, 2007), a modified version of the SAM Server (SAM-T08; compbio.soe.ucsc.edu/SAM_T08/T08-query.html) has been developed. Re-analysis of the protein sequences using the newly released program produced some variation in the results obtained using SAM-T06. On this occasion, all the box jellyfish toxins were predicted to share weak structural homology with the N-terminal domains of *B. thuringiensis* δ-endotoxins. The top three matches to CfTX-1 and CfTX-2 were predicted to be Cry1Aa (*E*-values = 3.65 and 4.56, respectively), Cry3Bb1 (*E*-values = 4.75 and 5.89, respectively) and Cry3A (*E*-values = 5.60 and 6.70, respectively). Similarly, Cry1Aa, Cry3Bb1 and Cry3A were the top three predicted matches to CqTX-A, CrTXs and CaTX-A (*E*-values = 1.1175–7.7908). Tentative three dimensional (3-D) models of CfTX-1 and -2 generated by SAM-T08 are depicted in Figure 5.6.

[**] The Expect (*E*) value is an estimate of how many sequences would score this well by chance in the database searched. The lower the *E*-value the more significant the score.

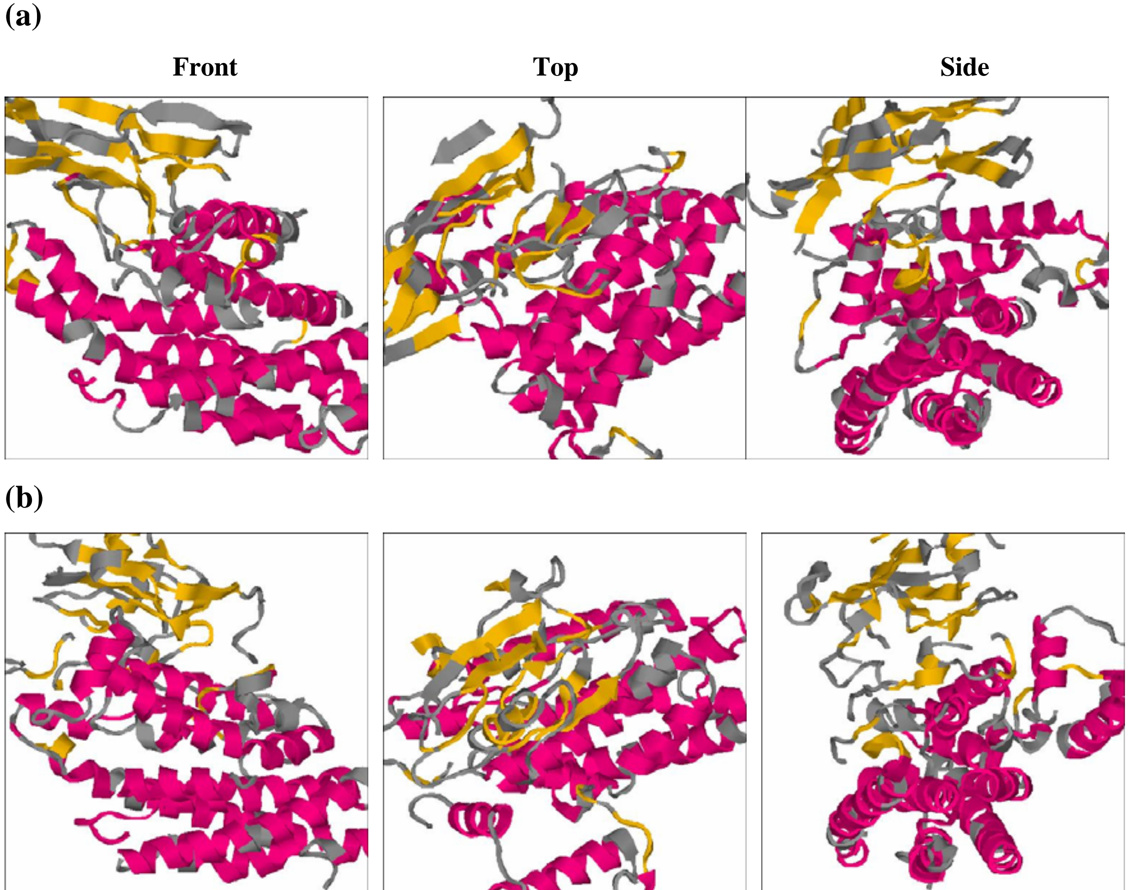


Figure 5.6 Crude 3-D model predictions for (a) CfTX-1 and (b) CfTX-2. Front, top and side views of the predicted structures are ordered from left to right. The models were created using the remote protein homology program, SAM-T08 (Karplus et al., 2005), and are considered to be crude backbone models, not full 3-D models. Colouring is used to show predicted secondary structure (red for helix, yellow for strand, blue for turn, grey for other).

The five cubozoan sequences were also re-analysed using the remote homology prediction server HHPred (Söding et al., 2005); a high-performer in the recently completed 7th community wide experiment for the critical assessment of techniques for protein structure prediction (CASP7) (Trapane and Lattman, 2007). Similarly, following a search against the RCSB PDB, weak remote structural homology was predicted in all cases between the box jellyfish proteins (amino acid ranges ~70–299) and the N-terminal domains of various insecticidal δ-endotoxins. Cry3A (PDB 1dlc), Cry3Bb (PDB 1ji6), Cry2Aa (PBD li5p), Cry1Aa (1ciy) and Cry4Aa (PDB 2c9k) were among the top five hits for each protein query. More specifically, the top hit for CfTX-1, CfTX-2, CqTX-A and CaTX-A was Cry3A (*E*-values = 0.036, 0.019, 0.042, 0.005, respectively) and the top hit for CrTXs was Cry3Bb (*E*-value = 0.044).

5.4 Discussion

This study has enabled for the first time, the isolation and sequencing of venom proteins from the box jellyfish *C. fleckeri*. The two major proteins, CfTX-1 and CfTX-2, are similar in molecular mass (~43 kDa and 45 kDa, respectively) and both proteins are strongly antigenic to CSL box jellyfish antivenom and rabbit antibodies raised against *C. fleckeri* nematocyst venom. Historically, little information about these two major venom proteins has been reported (see Chapter 1 for a review). However, the discovery of CfTX-1 and -2 as the most abundant components in *C. fleckeri* venom is similar to the findings of a previous study in which the most abundant protein in *C. fleckeri* nematocyst venom was reported to be approximately 40 kDa and antigenic against CSL box jellyfish antivenom (Wiltshire et al., 2000). In this case, the doublet of proteins may have appeared as a single large band, as often observed in this study. Also, since reporting the findings of this study (Brinkman and Burnell, 2007), other researchers have published a similar SDS-PAGE protein profile of *C. fleckeri* nematocyst venom in which the dominant doublet of CfTX-1 and -2 is clearly visible (Winter et al., 2008). Importantly, CfTX-1 and CfTX-2 are also similar in mass to lethal, haemolytic box jellyfish proteins, CqTX-A, CrTXs and CaTX-A, previously isolated from *C. quadrigatus*, *C. rastoni* and *C. alata*, respectively (Nagai et al., 2002, 2000a, b).

Protein database searches indicate that CfTX-1 and CfTX-2 belong to a newly emerging family of unique jellyfish toxins, which includes CqTX-A, CrTXs and CaTX-A. Discrepancies are noted between the estimated molecular masses of CfTX-1 and -2 (~43 and 45 kDa), determined by SDS-PAGE analysis, and the calculated molecular masses of their deduced sequences (~49 and 50 kDa, respectively). These discrepancies may be due to either inaccurate molecular mass estimation using the SDS-PAGE protein standards or post-translational cleavage of the proteins (in addition to signal peptide cleavage), which may occur during preproprotein processing. Similar molecular mass discrepancies are observed for CqTX-A, CrTXs and CaTX-A.

In an effort to identify regions in the mature CfTX proteins where post-translational cleavage(s) could have occurred, peptide mapping of CfTX-1 and -2 was attempted by comparing the ion masses of peptides obtained by MS analysis of trypsin-digested CfTX-1 (Appendix 6) with the calculated monoisotopic ion masses of peptides obtained from theoretically digested CfTX-1 and -2 amino acid sequences (MS-Digest, Clauser et al., 1999; Appendix 7). N-terminal amino acid sequence data obtained by Edman degradation analysis was also included (Section

5.3.2). However, as shown in Figure 5.7, insufficient peptide mass fingerprinting data was available to adequately map the proteins and confidently identify unmapped regions of amino acids that could account for a 5-6 kDa loss in the molecular mass of the proteins.

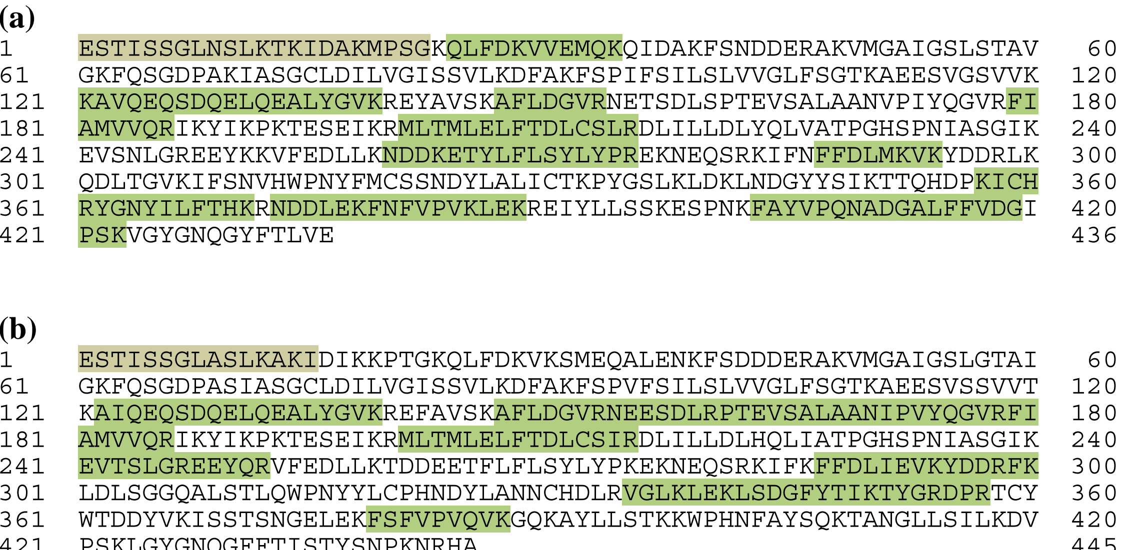


Figure 5.7 Peptide maps of (a) CfTX-1 and (b) CfTX-2. Theoretical peptides similar in monoisotopic mass to the monoisotopic peptide ion masses obtained by MS analysis of trypsin-digested CfTX-1 (Appendix 6) are shaded in green. N-terminal amino acid sequences are shaded in brown. Calculated tryptic peptide ion masses for CfTX-1 and -2 are provided in Appendix 7. Peptide mapping of CfTX-2 was performed based on the assumption that the sample of CfTX-1 analysed by MS was contaminated with CfTX-2 (see Section 5.3.3).

Two remote protein structural homology programs were used in this study to predict the tertiary structure (and hence inferred function) of the CfTX proteins and their three cubozoan homologues (CqTX-A, CrTXs and CaTX-A). Using one program (Superfamily, Madera et al., 2004), the chirodropid proteins (CfTX-1, CfTX-2 and CqTX-A) were each predicted to contain a central domain with structural similarity to the cytokine superfamily. Sub-family classification of the domains was assigned to IL-1, however, classifications were weak.

The IL-1 family of cytokines plays an important role in the regulation of inflammatory response and immunity in mammals and many other biological functions such as cell maturation, proliferation and activation (Beck and Habicht, 1986). IL-1 proteins have also been found in other vertebrates and tunicates (Pestarino et al., 1997) and IL-1-like proteins have been isolated from invertebrates including molluscs (Stefano et al., 1991) and sea stars (Beck and Habicht, 1986), suggesting that IL-1 is an ancient and functionally conserved family of proteins that emerged early during the evolution of metazoans. Structurally, IL-1 family members have a 12-

strand β-trefoil topology (Evans et al., 1995). In contrast, secondary structure analysis of CfTX-1 and -2 amino acid sequences, predicted the presence of only three consecutive β-strands in their putative cytokine domains, despite the prediction of approximately twelve β-strands at the C-terminal end of each protein. These apparent structural variations may explain the weak IL-1 sub-family assignments obtained in the Superfamily database searches. The prediction of remote homology, albeit weak, between the IL-1 family and the chirodropid proteins is interesting given that acute inflammatory response is an important feature in human envenomations by chirodropids, but not as significant in envenomations by carybdeids, the homologues of which, appear to lack structural similarity to the cytokine superfamily. However without structural elucidation or immunological characterisation of CfTX-1 and CfTX-2, the role of these putative cytokine-like domains in the chirodropid proteins remains speculative.

Using another remote protein homology prediction program (SAM-T06) and its successor (SAM-T08), weak structural homology was detected between the five box jellyfish toxins and the N-terminal domains of three insecticidal δ-endotoxins (Cry1Aa, Cry3Bb and Cry3A). Similarly, using the Superfamily prediction server, CaTX-A produced a weak match to the δ–endotoxin, N-terminal domain superfamily.

Insecticidal δ-endotoxins, produced by *B. thuringiensis*, belong to a family of proteins with specific toxicities against insects from the Orders Lepidoptera, Diptera and Coleoptera (Li et al., 1991). The structures of lepidopteran-specific Cry1Aa and coleopteran-specific Cry3Bb and Cry3A are similar and each protein comprises three domains (Grochulski et al., 1995; Galitsky et al., 2001; Li et al., 1991). The N-terminal domain (Domain I) forms a seven- or eight-helix bundle, in which a central helix is surrounded by the remaining outer helices, and the middle and C-terminal domains (Domains II and III) contain mostly β-sheets. The N-terminal domains of δ-endotoxins have been strongly implicated in membrane insertion and pore formation whereas the other two domains are important in receptor recognition (Grochulski et al., 1995; Pardo-López et al., 2006). Therefore, the box jellyfish toxins may also be involved in a pore-forming mechanism similar to that of δ-endotoxins.

The mechanism underlying Cry3A activity has been proposed to involve a complex process of protoxin-receptor binding, proteolytic processing and the formation of an oligomeric Cry3A pre-pore intermediate, possibly a tetramer, capable of penetrating insect membranes (Rausell et al., 2004). More recent studies of another δ-endotoxin, Cry1Ac, have also indicated that interaction of its pre-pore tetramer with the carbohydrate moiety of a second receptor facilitates membrane insertion and pore formation by the N-terminal domains of each monomer (Pardo-

López et al., 2006), thus emphasising the complexity of the underlying mechanisms of δ-endotoxin cytolytic activity.

Secondary structure analysis of the CfTX-1 and -2 amino acid sequences has identified a putative transmembrane spanning region, referred to as TSR1, that is common to the other cubozoan sequences, CqTX-A, CrTXs and CaTX-A. The importance of this region appears to be significant as it coincides with several highly conserved, consecutive amino acids (Figure 5.4). The presence of this common transmembrane spanning region also suggests the potential for the CfTX proteins and their cubozoan homologues to act as pore-forming toxins, disrupting normal transmembrane ion concentration gradients in susceptible cells. Following recent pharmacological studies involving *C. fleckeri* nematocyst venom, researchers have also suggested an ionophoric effect as the most likely mechanism of cardiotoxic action of *C. fleckeri* venom (Winter et al., 2007a). Furthermore, purified CqTX-A, CrTXs and CaTX-A are haemolytic to sheep erythrocytes (Nagai et al., 2002, 2000a, b), demonstrating the propensity of this family of jellyfish toxins to act as cytolysins. Also, whole and partially purified *C. fleckeri* nematocyst extracts are haemolytic to sheep and human erythrocytes, and pore formation has been observed in cultured rat cardiocytes exposed to whole nematocyst venom (Bailey et al., 2005; Endean et al., 1993).

Previous sequence analysis of CqTX-A, CrTXs and CaTX-A predicted the presence of an amphiphilic α-helix in the N-terminal region of each toxin that may relate to their potent haemolytic activity (Nagai et al., 2002, 2000b). In this study, the prediction of amphiphilic α-helices in the N-terminal regions of CqTX-A, CrTXs and CaTX-A correlates well with previous predictions and amphiphilic α-helices were also predicted for the CfTX proteins in the same region (Figure 5.4). However, lack of conservation in amino acid sequence in the region where the helices are predicted and variation in the location of the helices between the chirodropid and the carybdeid proteins, suggest that the amphiphilic helices are perhaps less important to the mechanism of action of the cubozoan toxins than TSR1.

Further research is therefore necessary to determine the importance of TSR1 and whether it confers the cytolytic activity attributed to this family of jellyfish toxins. As more sequence data becomes available for homologues in other cubozoan species or more primitive cnidarians, it will also be possible to gain more information about the evolution of this unique class of jellyfish toxins and how they relate to one another in sequence, structure and function.

CHAPTER 6

Fractionation of *Chironex fleckeri* Venom Proteins

6.1 Introduction

As described in Chapter 5, the two major proteins contained in the nematocyst venom of *C. fleckeri* (CfTX-1 and -2) have been isolated, cloned and sequenced, but little is known about their biochemical characteristics and biological activities. Secondary structure predictions suggest that the CfTX proteins and their cubozoan homologues (CqTX-A, CrTXs and CaTX-A) are potential cytolysins, due to the presence of a conserved transmembrane spanning region (TSR1), which may be involved in a pore-forming mechanism of action (Brinkman and Burnell, 2007). Therefore, the aim of this study was to fractionate *C. fleckeri* nematocyst venom, investigate the biochemical characteristics of CfTX-1 and -2 (and other nematocyst proteins) and determine whether the CfTX proteins exhibit cytolytic activity, as previously hypothesised.

Initially, immunoaffinity chromatography, using CSL antivenom and CfTX-specific antibodies coupled to Sepharose-4B was attempted to fractionate nematocyst venom proteins and purify the CfTX proteins for further biochemical characterisation. Secondly, size exclusion chromatography was used to fractionate nematocyst venom and investigate the possible quaternary structure of native CfTX-1 and -2. Thirdly, due to the basic nature of mature CfTX-1 and -2 (pI 8.3 and 7.8, respectively), cation exchange chromatography was used in an effort to separate the two homologues based on differences in their intrinsic charge and to determine the biochemical characteristics and biological activities of the individually purified venom proteins. Furthermore, CqTX-A, CrTXs and CaTX-A were successfully isolated and purified using similar methodology (Nagai et al., 2002, 2000a, b).

Haemolysis tests and assays were used to assess the cytolytic activity of crude and fractionated nematocyst extracts. A simple blood agar haemolysis test provided a rapid and sensitive method for initial detection of haemolytic activity in chromatography fractions, whereas, haemolytic assays allowed quantitative comparisons in cytolytic activity between distinct bioactive components.

6.2 Methods and Materials

6.2.1 Preparation of Nematocyst Extracts

Nematocysts were isolated from excised jellyfish tentacles and purified using methods previously described (see Section 2.1). Nematocyst extracts were freshly prepared by bead mill homogenisation (see Section 2.2.1), except the nematocyst extraction buffer (NEB) differed for immunoaffinity chromatography (20mM PO_4^{3-}, pH 7.4; 25mM MOPS, 0.5M NaCl, pH 7.4; or 25mM MOPS, 0.2M NaCl, pH 7.4), ion exchange chromatography (20mM PO_4^{3-}, pH 6.0) and size exclusion chromatography experiments (20mM PO_4^{3-}, 0.15M NaCl, pH 7.2), but included the same protease inhibitors.

6.2.2 SDS-PAGE and Western Blot Analysis

Reducing SDS-PAGE and western blot analysis of proteins was conducted on crude nematocyst extracts and chromatography fractions, as previously described (Sections 2.3.3 and 2.3.4). Samples containing low amounts of protein were concentrated by TCA precipitation prior to analysis (Section 2.3.3.3). For western blot analyses, proteins were blotted with CfTX-specific rabbit antibodies, rabbit antibodies raised against whole bead mill nematocyst extracts and CSL box jellyfish antivenom (1:1000, 1:1000 and 1:3000, respectively).

6.2.3 Immunoaffinity Chromatography

Immunoaffinity chromatography experiments were undertaken using two different ligands: CSL antivenom and purified rabbit antibodies specific for CfTX-1 and -2. Cyanogen bromide-activated Sepharose-4B (1.5–3g, Amersham Biosciences), rehydrated and washed in dilute HCl (1mM), was combined with coupling buffer (0.1M $NaHCO_3$, 0.5M NaCl, pH 8.3; 15mL) containing CSL antivenom or CfTX-specific rabbit antibodies (5–7mg IgG/mL media). Following gentle mixing (1h, RT), the IgG-coupled Sepharose (> 96% ligand binding achieved based on supernatant A_{280} measurements) was washed in coupling buffer (10 gel vol.), resuspended in blocking buffer (0.1M Tris-HCl, pH 8.0; 4 gel vol.) and incubated without mixing (2h, RT). The media was washed alternately with high and low pH buffers (0.1M Tris-HCl, 0.5M NaCl, pH 8.0 and 0.1M NaOAc, 0.5M NaCl, pH 4.0, respectively) three times each (5 gel vol. / wash), resuspended in binding buffer (20mM PO_4^{3-}, pH 7.4) and packed in Bio-Rad

Econo-Columns (final column volumes (CV) ~10mL for CSL antivenom-coupled and 2mL for CfTX-specific IgG-coupled Sephadex-4B).

Crude non-lyophilised nematocyst extract was filtered (0.22µm Millex-GP filters, Millipore) and applied (~0.5mg total protein) to the CSL antivenom immunoaffinity column, pre-equilibrated with binding buffer, and connected to a Bio-Rad peristaltic pump (1 mL/min, 4°C). Following sample application, the column was washed with binding buffer (~3 CV) and retained proteins were eluted by switching to elution buffer (75mM glycine-HCl, 0.5M NaCl, pH 2.7; ~3 CV). Protein elution was monitored by UV detection (280nm) and fractions (0.5 mL), numbered according to the elution volume (V_e), were collected, neutralised (0.1 vol. Tris-HCl, pH 8.0) and stored on ice. Fractions corresponding to protein peaks 1–4 (V_e = 5.5–9.5 mL, 10–14 mL, 47–51.5 mL and 52–56 mL, respectively) were pooled and concentrated (~100µL) using centrifugal filters (10K MWCO, Millipore; 12.1 K×*g*, 4°C), and stored on ice prior to analysis by SDS-PAGE and western blot using CSL antivenom. No bioactivity assays were undertaken in conjunction with the CSL antivenom immunoaffinity experiment.

Crude lyophilised nematocyst venom (~1.2mg total protein) was applied to the gravity-fed CfTX-specific immunoaffinity column, pre-equilibrated with binding buffer (25mM MOPS, 0.5M NaCl, pH 7.4). Following sample application, the column was washed with binding buffer (2 CV) and low-salt buffer (25mM MOPS, 0.1M NaCl, pH 7.4; 1 CV). Retained proteins were eluted with low-pH buffer (75mM glycine-HCl, pH 2.7; 2 CV). Protein elution was monitored by UV detection (280nm) and numbered fractions (1-26; 0.5 mL each) were collected, neutralised (0.1 vol. Tris-HCl, pH 8.0) and stored on ice. Selected fractions (2–5, 10–14, 18–19) were analysed by SDS-PAGE. This experiment was repeated using a lower-salt binding buffer (25mM MOPS, 0.2M NaCl, pH 7.4) and application of crude lyophilised nematocyst extract (~0.3mg). Selected fractions (2–5 and 8) were analysed by SDS-PAGE. No western blot analysis or bioactivity assays were undertaken in conjunction with the CfTX-specific immunoaffinity experiments.

6.2.4 Size Exclusion Chromatography

Crude nematocyst venom (2.5–4mg total protein) was applied to a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) that was pre-equilibrated and eluted with NEB (20mM PO_4^{3-}, 0.15M NaCl, pH 7.2), using a Bio-Rad Econo FPLC system (0.5 mL/min, 4°C). The column was calibrated using a size exclusion standard mixture (Bio-Rad), which included thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and

vitamin B_{12} (1.4 kDa), plus a combination of individual standards (GE Healthcare/Boehringer Mannheim/Sigma), which included blue dextran (2000 kDa), ferritin (440 kDa), aldolase (158 kDa), albumin (45 kDa) and cytochrome c (12 kDa). Protein elution was monitored by UV detection (280nm) and fractions (1mL), numbered according to V_e, were collected and stored on ice. The approximate molecular masses of native proteins were calculated using the linear regression of the partition coefficient (K_{av}) of the size exclusion standards plotted as a function of their log molecular mass.

6.2.5 Cation Exchange Chromatography

Crude nematocyst venom (0.6–1.5mg total protein) was applied to a 1mL Uno-S1 column (GE Healthcare), pre-equilibrated with NEB (20mM PO_4^{3-}, pH 6.0), and connected to a Bio-Rad BioLogic Duo-Flow HPLC system (0.5 mL/min, 4°C). Following sample application, the column was washed with application buffer (5 CV) and retained proteins were eluted by stepwise increases in elution buffer concentration (5–10 CV each of 10, 20, 30, 40, 50, 60 and 100% 20mM PO_4^{3-}, 1M NaCl, pH 6.0). Protein elution was monitored by UV detection (280nm) and fractions (0.5–1 mL), numbered according to V_e, were collected and stored on ice.

6.2.6 Haemolytic Activity

The haemolytic activity of crude nematocyst extracts and chromatography fractions was initially tested on 5% sheep blood agar (see Section 2.3.8.1 for further details). Quantitative haemolytic activity of extracts and fractions was determined spectrophotometrically (see Section 2.3.8.2 for further details). Haemolysis results were calculated as a percentage relative to complete lysis. Size exclusion chromatography fractions corresponding to partially purified 370 and 145 kDa haemolysins were subsequently pooled (V_e = 58–61 and 68–70mL, respectively). HU_{50} values, defined as the concentration of protein that causes 50% lysis, were determined for the crude venom and pooled chromatography fractions from plots of percentage haemolysis as a function of the concentration of assayed protein.

The dependence of haemolytic activity on temperature was examined for the 370 kDa haemolysin, which contained partially purified CfTX-1 and -2. Diluted protein (equiv. $60HU_{50}$, 3mL) was added to 0.4% sheep erythrocytes in NEB (30mL) on ice and mixed. Sample aliquots (30 × 1mL), each containing ~$2HU_{50}$ protein, were prepared on ice. A set of triplicate samples was retained on ice (time = 0 min) and the remaining samples were incubated (37ºC) with gentle

agitation. Triplicate samples were removed from incubation at set time intervals (2, 4, 6, 8, 10, 15, 20, 25 and 30 min) and immediately chilled on ice (≥ 10 min). Samples were centrifuged (3000*g*, RT, 5 min) and the A_{540} of the supernatants were determined. The experiment was repeated at two different temperatures (4ºC and 18ºC).

Haemolysis kinetic and dose-response plots were each fitted with a four parameter logistic curve (SigmaPlot, Systat Software Inc.).

$$y = \min + \frac{(\max - \min)}{1 + \left(\frac{x}{EC_{50}} \right)^{Hillslope}}$$

6.3 Results

6.3.1 Immunoaffinity Chromatography

Limited separation of *C. fleckeri* nematocyst venom proteins was achieved using CSL immunoaffinity chromatography. Nematocyst proteins eluted from the column in four peaks (Figure 6.1a). SDS-PAGE and western blot analyses (Figure 6.1b and Figure 6.1c, respectively) revealed that CfTX-1 and -2 were mostly unretained by the column and eluted predominantly in the column flow-through (peak 1). In the SDS-PAGE protein profile, partial retention of some proteins (~65, 24, 20 and 18 kDa) was observed in peak 2. Although western blot results were poor due to diffusion of proteins (> 45 kDa) from the membrane, four low molecular mass antigenic proteins (~32, 30, 20 and 18 kDa) were detected in peak 2, in addition to the CfTX proteins. While no proteins were detected in peaks 3 and 4 by Coomassie-staining, western blot analysis detected the presence of the CfTX proteins in peak 3 and one antigenic protein (~55 kDa) in both peaks.

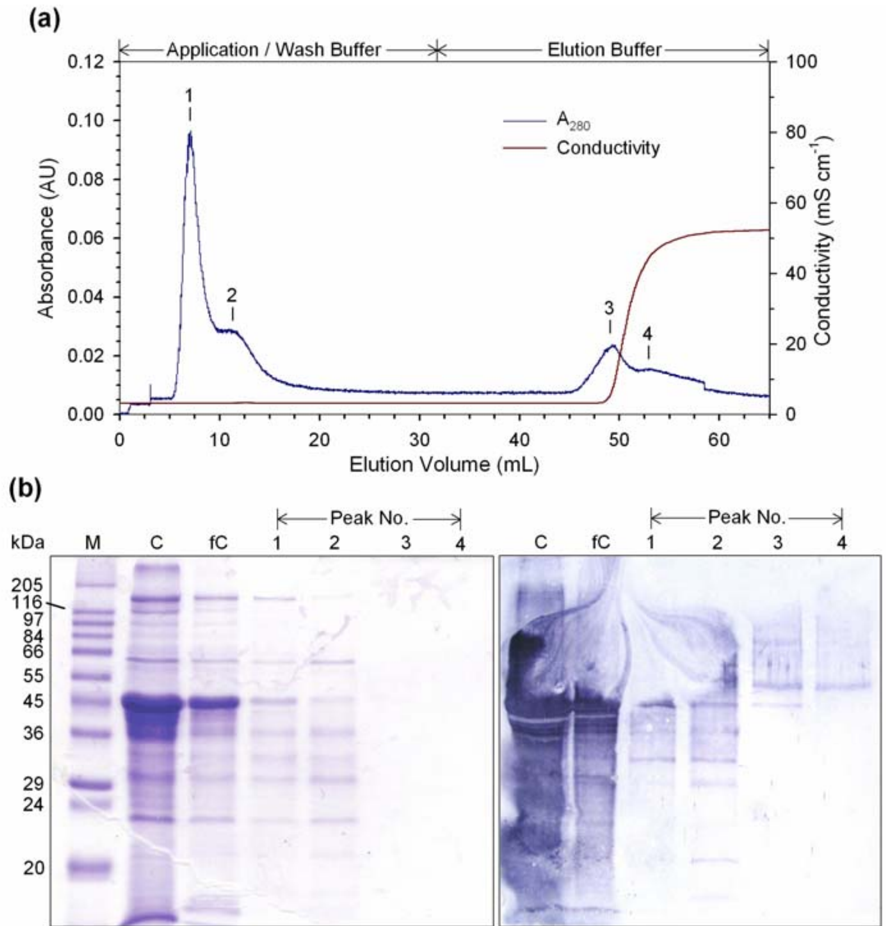


Figure 6.1 Separation of nematocyst extract proteins by CSL antivenom immunoaffinity chromatography. Nematocyst proteins (~0.5mg) were applied to a 10mL ligand-coupled Sepharose 4B column in application buffer (20mM PO_4^{3-}, pH 7.4), washed with application buffer and eluted with low-pH buffer (0.2M glycine-HCl, pH 2.7) (a) Chromatogram of proteins eluted from the CSL antivenom immunoaffinity column; (b) Coomassie-stained proteins from crude nematocyst venom (C), filtered crude venom (fC) and pooled, concentrated chromatography fractions (peaks 1-4); (c) western blot of corresponding proteins using CSL antivenom. M indicates the molecular mass marker (Sigma) and the molecular masses of the standard proteins are shown in kilodaltons (kDa).

In second attempt to separate *C. fleckeri* venom proteins using immunoaffinity chromatography, a CfTX-specific immunoaffinity column was prepared and tested. Unfortunately, no separation or purification of CfTX-1 and -2 (or other nematocyst proteins) was achieved. SDS-PAGE analysis of peak fractions revealed that all nematocyst proteins detected by Coomassie-staining were unretained and eluted in the column flow-through (Figure 6.2). No apparent improvement in protein retention was achieved by decreasing the salt concentration in the binding buffer from 0.5M to 0.2M NaCl (Figure 6.3).

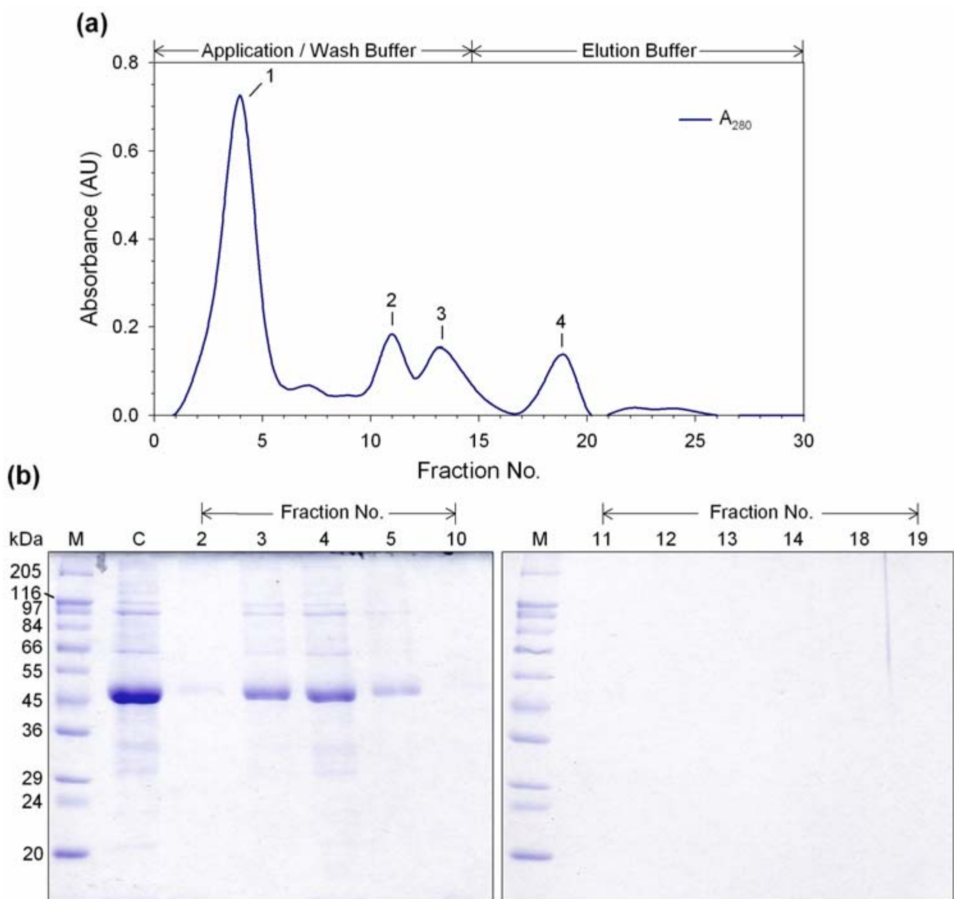


Figure 6.2 Separation of nematocyst venom proteins using CfTX-specific immunoaffinity chromatography. Nematocyst proteins (~1.2mg) were applied to a 2mL ligand-coupled Sepharose 4B column in high-salt application buffer (25mM MOPS, pH 7.2, 0.5M NaCl), the column was washed with low-salt buffer (25mM MOPS, pH 7.2, 0.1M NaCl), and retained proteins were eluted with low-pH buffer (0.2M glycine-HCl, pH 2.7). (a) Chromatogram of proteins eluted from the CfTX-specific immunoaffinity column; (b) Coomassie-stained proteins from crude nematocyst venom (C) and chromatography fractions coinciding with peaks 1-4. M indicates the molecular mass marker (Sigma) and the molecular masses of the standard proteins are shown in kDa.

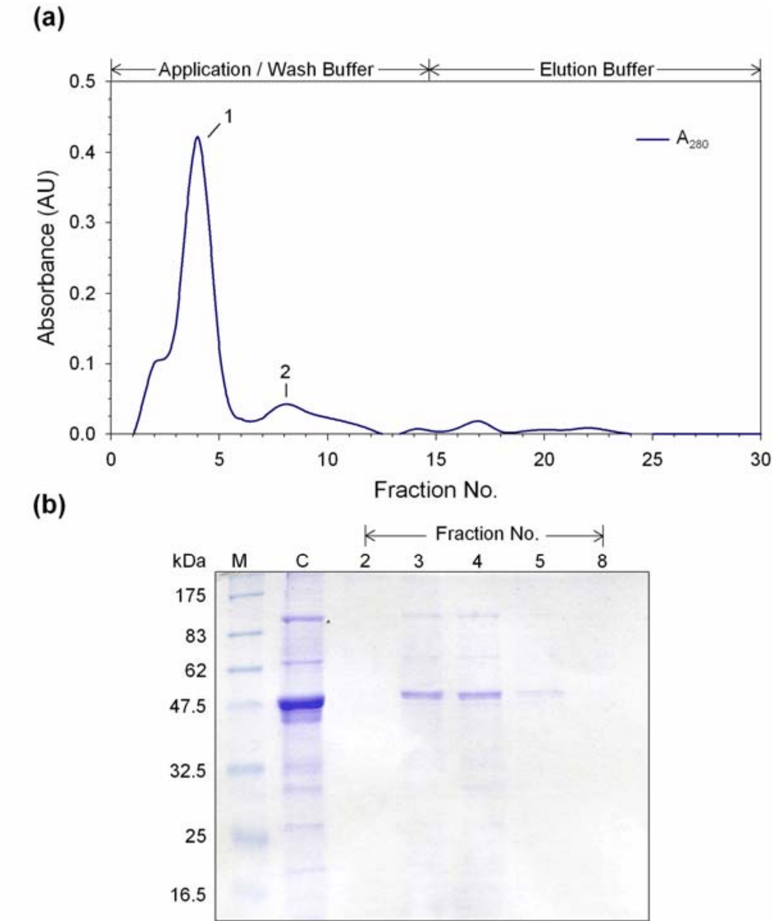


Figure 6.3 Separation of nematocyst venom proteins by CfTX-specific immunoaffinity chromatography . Nematocyst proteins (~1.2mg) were applied to a 2mL ligand-coupled Sepharose 4B column in application buffer (25mM MOPS, pH 7.2, 0.2M NaCl), the column was washed with low-salt buffer (25mM MOPS, pH 7.2, 0.1M NaCl), and retained proteins were eluted with low-pH buffer (0.2M glycine-HCl, pH 2.7). (a) Chromatogram of proteins eluted from the CfTX-specific immunoaffinity column; (b) Coomassie-stained proteins from crude nematocyst venom (C) and chromatography fractions coinciding with peaks 1 and 2. M indicates the prestained molecular mass marker (NEB) and the molecular masses of the standard proteins are shown in kDa.

6.3.2 Size Exclusion Chromatography

Separation of native *C. fleckeri* nematocyst venom proteins was achieved by size exclusion chromatography across the molecular mass separation range, 10–600 kDa (see Figure 6.4a). Peak 1, which eluted as a discrete peak at the void volume (V_0) of the column, corresponded to native proteins with molecular masses ≥ 600 kDa. The estimated native molecular masses of the other major peaks (2–6) were calculated as 370, 145, 90, 55 and 27 kDa, respectively;

although peaks 4 and 5 were poorly resolved. The native molecular masses of the two minor peaks (7 and 8) were ≤ 10 kDa.

Reducing SDS-PAGE analysis of the crude venom and chromatography fractions corresponding to peaks in protein concentration (Figure 6.4b), revealed that peak 2 (370 kDa) contained significantly purified CfTX-1 and -2 (~43 and 45 kDa, respectively). Western blot analysis of the peak fractions with CfTX-specific antibodies, confirmed that CfTX-1 and -2 were the major proteins in peak 2, and that lesser amounts of the two proteins were also present in the other peaks (Figure 6.4c). Peak 3 (145 kDa) contained significant amounts of two major venom proteins (~39 and 41 kDa). Although the 39 and 41 kDa proteins were abundant in crude nematocyst venom, western blot analysis revealed that they were not significantly antigenic towards rabbit anti-whole venom antibodies (or CSL antivenom), compared to other proteins such as CfTX-1 and -2 (Figure 6.4d). Peak 1 (≥ 600 kDa) contained exclusively a 160 kDa nematocyst protein plus small amounts of the CfTX proteins and the 39 and 41 kDa proteins. Peaks 4–6 (90, 55 and 27 kDa, respectively) contained several proteins, yet, partial purification of some proteins was evident (e.g. ~97, 65, 38, 30, 28, 27 and 24 kDa). The complexity of crude extract proteins in the 26–36 kDa molecular mass range was also significantly decreased, due to the differential elution of proteins in peaks 3–6. Negligible amounts of proteins were observed in SDS-PAGE profiles for peaks 7 and 8.

Blood agar haemolysis tests indicated that all the chromatography fractions collected between V_0 (~45mL) and the total bed volume (V_t; 120mL) exhibited haemolytic activity, particularly within the range V_e = 55–79mL. Haemolysis was evident by the formation of a colourless zone in the blood agar surrounding the sample wells, in a pattern similar to bacterial β-haemolysis (Balashova et al., 2006). Significant haemolysis was also observed for crude venom and positive controls, but no haemolysis was detected for the negative controls. Quantitative haemolytic assays revealed that two haemolysis maxima coincided with native protein peaks 2 and 3 (~370 and 145 kDa, respectively), and a minor peak in haemolysis between native protein peaks 4 and 5 (~70 kDa) was also observed (see Figure 6.4a).

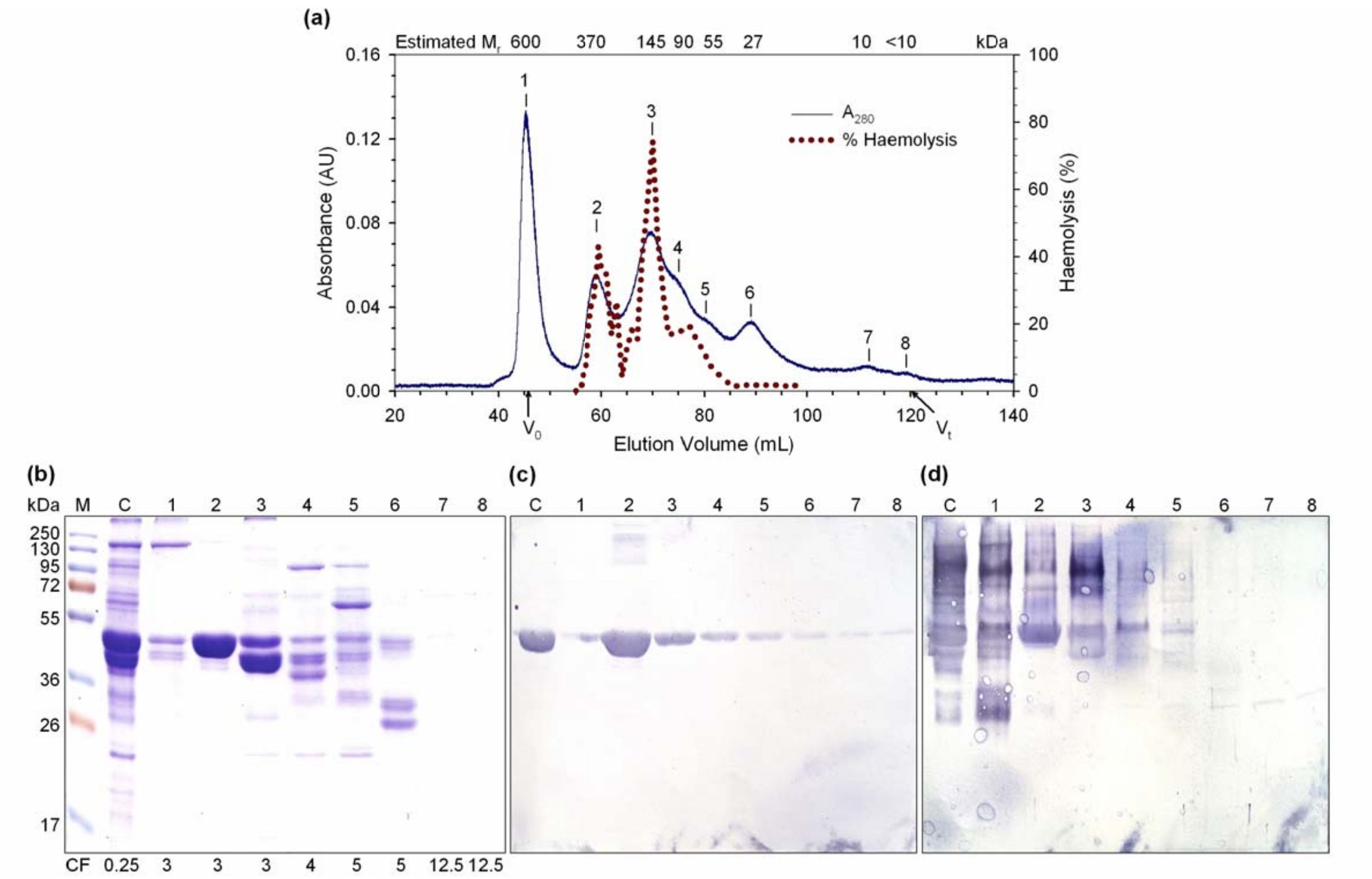


Figure 6.4 Fractionation of crude nematocyst extracts by size exclusion chromatography (a) Chromatogram of native proteins eluted from a Superdex 200 column. Protein concentration, monitored at A_{280}, is represented by a blue line. The estimated molecular mass (M_r) of native proteins corresponding to peaks numbered 1–8 are indicated in kilodaltons (kDa) above each peak. Haemolysis measurements of selected fractions are represented by a red dotted line; (b) Coomassie-stained SDS-PAGE protein profiles of crude venom (C) chromatography fraction corresponding to peaks 1–8. M indicates the protein molecular marker (Fermentas); the M_r of the standards are shown in kDa. Crude venom was diluted prior to analysis (DF = 4). Chromatography fractions were concentrated by TCA precipitation prior to analysis; the concentration factor (CF) is indicated below each lane; (c) western blot of corresponding samples using rabbit CfTX-specific antibodies; (d) western blot of corresponding samples using rabbit anti-whole nematocyst venom antibodies.

Haemolysis experiments revealed that the haemolytic activities of the crude venom and partially purified 370 and 145 kDa haemolysins were dependent on protein concentration (Figure 6.5). Calculated HU_{50} values of the crude venom, 370 kDa and 145 kDa haemolysins were 5, 14 and 7 ng/mL, respectively. The HU_{50} value of the minor haemolysis peak was not quantitatively determined, however, based on experimental data, the chromatography fraction with maximum haemolytic activity (V_e = 77mL), which contained ~70 ng/mL protein, caused complete lysis (i.e. HU_{50} < 70 ng/mL).

Losses in the relative activities of the 370 and 145 kDa haemolysins were observed during size exclusion purification (see Table 6.1). Total recovery of activities were also low, due in part to the small number of fractions that were combined and assayed for each haemolysin, and the broad distribution of CfTX-1 and -2 across the entire molecular mass range (see Figure 6.4c).

Table 6.1

Purification of two *C. fleckeri* haemolysins (370 and 145 kDa) using Superdex 200 size exclusion chromatography.

Sample	Total Protein (mg)	Total HU	Specific Activity (HU/mg)	Relative Activity	Recovery (%)
Crude	2.58	515400	200000	1	100
370 kDa	0.44	31429	71429	0.36	6
145 kDa	0.47	66429	142857	0.71	13

Kinetic studies revealed that the haemolytic activity of purified CfTX-1 and -2 was temperature-dependent. Complete and ~40% haemolysis was measured following incubation of 0.4% sheep erythrocytes in solution with $2HU_{50}$ of the 370 kDa haemolysin for 30 min at 37ºC and 18ºC, respectively, however, no increase in haemolytic activity was detected in samples incubated at 4ºC. At 37ºC and 18ºC, a lag phase (~6–7 min) preceded initiation of haemolysis (Figure 6.6). At 37°C, the maximum rate of haemolysis (R_{max} = 6.5%/min) occurred after 24 minutes.

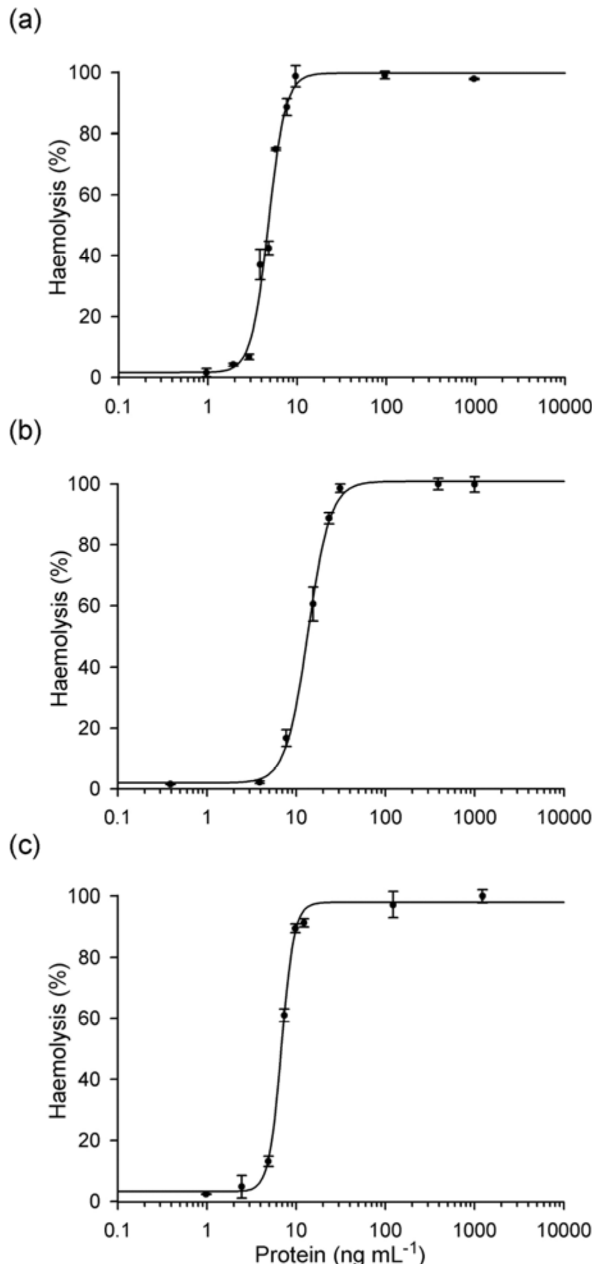


Figure 6.5 Dose-response curves of sheep erythrocyte lysis by crude venom and partially purified *C. fleckeri* haemolysins obtained using size-exclusion chromatography (a) unfractionated crude nematocyst extract; (b) 370 kDa haemolysin, containing predominantly CfTX-1 and -2 (peak 2, Figure 7.3a); (c) 145 kDa haemolysin (peak 3, Figure 7.3a). Vertical bars represent the standard error from three independent assays at each protein concentration.

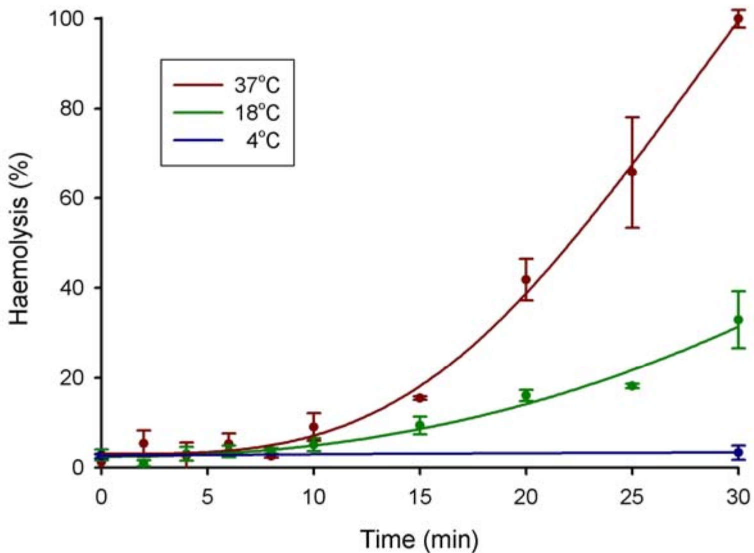


Figure 6.6 Kinetics of purified CfTX-1 and -2 haemolytic activities at different temperatures. Three sets of samples (10 × 1mL), with each sample containing ~$2HU_{50}$ of partially purified CfTX-1 and -2 in a solution of ~0.4% sheep erythrocytes in NEB, were incubated at three temperatures (4, 18 and 37°C). Triplicate samples were removed from incubation at various time intervals from 0 – 30 minutes and haemoglobin release was measured at 540nm. Vertical bars represent standard errors at each time point.

6.3.3 Cation Exchange Chromatography

Fractionation of *C. fleckeri* nematocyst venom using cation exchange chromatography also enabled the partial separation of several venom proteins, however, resolution of CfTX-1 and -2 was not achieved (Figure 6.7a). Comparison of the SDS-PAGE protein profile of the crude venom with the profiles of chromatography fractions corresponding to peaks in protein concentration (see Figure 6.7b) showed that the major proteins, CfTX-1 and -2, co-eluted mostly in peak 4 (0.3M NaCl), whereas the doublet of the other major protein bands (39 and 41 kDa) co-eluted predominantly in peak 5 (0.4M NaCl). Minor venom proteins (65, 38 and 20 kDa) were also partially purified in the 0.2–0.4M NaCl fractions and a high molecular mass protein (160 kDa) eluted only in 1M NaCl fractions. Unretained minor proteins included a partially purified 97 kDa protein and most of the extract proteins found within the 26–36 kDa molecular mass range.

Western blot analysis of the chromatography fractions, using CfTX-specific antibodies, confirmed the presence of the CfTX proteins mainly in peak 4 (see Figure 6.7c), but also in peaks 3, 5, 6 and 7/8 (0.2, 0.4, 0.5 and 1M NaCl, respectively). Western blot analysis, using rabbit antibodies raised against whole nematocyst venom, indicated that several proteins (97, 38 and 20

kDa proteins and the CfTX proteins) were significantly more antigenic than the 39 and 41 kDa proteins (see Figure 6.7d).

Blood agar haemolysis tests revealed that all the chromatography fractions exhibited haemolytic activity, except those collected at the tail-end of the column flow-through and at 0.1M NaCl concentration. Significant haemolysis was observed for the positive controls, but no haemolysis was evident in the negative controls. Haemolysis maxima corresponded to protein peaks 4 and 5 (0.3 and 0.4M NaCl, respectively) (Figure 6.7a). Quantitative HU_{50} values were not determined for this experiment; however, based on experimental data, 3.2 and 32 ng/mL protein from the 0.3M NaCl peak chromatography fraction (V_e = 24.5 mL) caused 0% and 100% lysis, respectively (i.e. 3.2 ng/mL < HU_{50} < 32 ng/mL), and 6.1 ng/mL protein from the 0.4M NaCl peak chromatography fraction (V_e = 30.5mL) caused 37% lysis (therefore, estimated HU_{50} ~ 8 ng/mL).

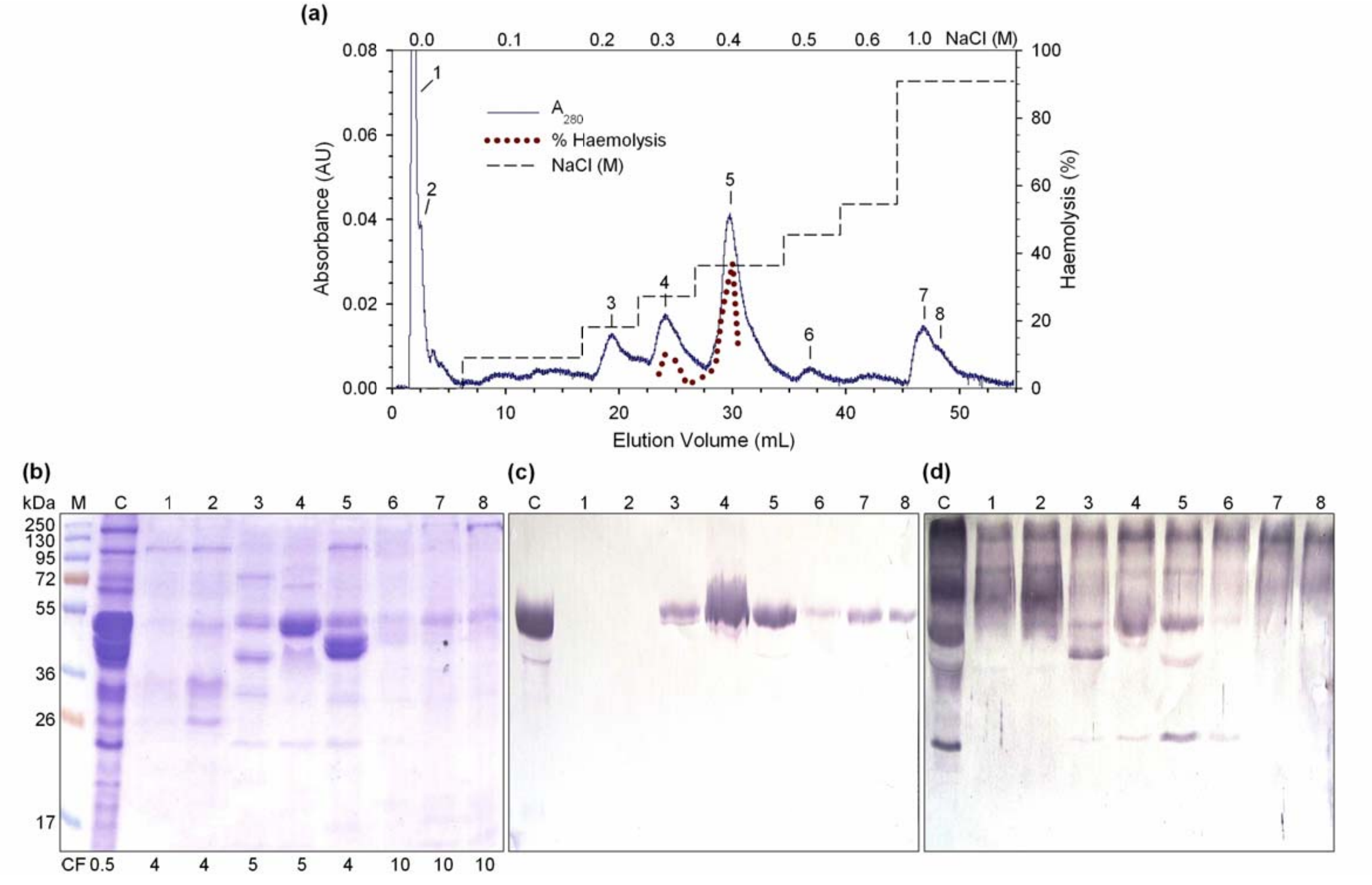


Figure 6.7 Fractionation of crude nematocyst extracts by cation exchange chromatography (a) Chromatogram of proteins eluted from a Uno S1 column by stepwise increases in NaCl concentration. Protein concentration, monitored at A_{280}, is represented by a blue line. NaCl concentration is represented by a black line; the concentration at each step is indicated at the top of the chromatogram in molar units (M). Haemolysis measurements of selected fractions are represented by a red dotted line; (b) Coomassie-stained SDS-PAGE protein profiles of crude nematocyst venom (C) and chromatography fractions corresponding to peaks numbered 1–8. M indicates the protein molecular marker (Fermentas); the molecular masses of the standards are shown in kDa. Crude venom was diluted prior to analysis (DF = 2). Chromatography fractions were concentrated by TCA precipitation prior to analysis; the concentration factor (CF) is indicated below each lane; (c) western blot of corresponding samples using rabbit CfTX-specific antibodies; (d) Western blot of corresponding samples using rabbit anti-whole nematocyst extract antibodies.

6.4 Discussion

The nematocyst venom of *C. fleckeri* is a mixture of several bioactive proteins and, due to its complexity, the identity and biochemical characteristics of individual components can only be established once the crude venom has been fractionated and the venom proteins significantly purified. Early researchers have previously used immunoaffinity chromatography to successfully isolate bioactive proteins from *C. fleckeri* venom. Initially, a 20 kDa cardiotoxin was isolated from milked venom using α-*Physalia physalis* (Portuguese man-o'-war) toxin immunoaffinity chromatography (Olson et al., 1984). In a subsequent study, a lethal, cardiotoxic factor, which contained two proteins (50 and 150 kDa) was purified from frozen tentacle extracts using CSL antivenom immunoaffinity chromatography (Calton and Burnett, 1986). In this study, fractionation of *C. fleckeri* nematocyst venom proteins using a CSL antivenom immunoaffinity column resulted in the partial purification of five minor antigenic proteins (~55, 32, 30, 20 and 18 kDa). However, no apparent purification of the two major proteins present in nematocyst venom, CfTX-1 and -2, was achieved and the proteins were mostly unretained by the immunoaffinity column (Figure 6.1). Similarly, no retention or partial purification of CfTX-1 and -2 was achieved using CfTX-specific immunoaffinity chromatography (see Figure 6.2 and Figure 6.3). Although some chromatography fractions could have been concentrated to identify minor proteins retained by the columns, from a preparative perspective, the usefulness of the immunoaffinity columns prepared for this study was very limited. Furthermore, the inability of the immunoaffinity columns to effectively retain nematocyst proteins contradicts earlier research findings which clearly demonstrate that the antibodies used for immunoaffinity column preparation bind strongly to nematocyst proteins, such as CfTX-1 and -2, in western blot analyses (see Chapter 4). Therefore, based on these observations, it seems likely that the columns prepared for this study were flawed.

Despite the lack of success using immunoaffinity chromatography, this study reports for the first time the partial purification of CfTX-1 and -2 from the nematocyst venom of *C. fleckeri*. Partial purification of CfTX-1 and -2 was successfully achieved using two different methods of separation (ion exchange and size exclusion chromatography), however, the two homologues could not be resolved to enable biochemical characterisation of the individual proteins. Size exclusion chromatography, in particular, yielded highly purified CfTX proteins (Figure 6.4b). The CfTX proteins co-eluted predominantly in a discrete protein peak corresponding to a native molecular mass of ~370 kDa, which suggests that under the described

experimental conditions, an aggregated form of the proteins is present in the native venom, possibly an octamer or a mixture of oligomers. Typical Superdex 200 fractionation profiles of *C. fleckeri* venom also included five other significant protein peaks corresponding to native molecular masses of ≥ 600, 145, 90, 55 and 27 kDa. Subsequent assays revealed that the 370 kDa protein, composed of CfTX-1 and -2 subunits, and the 145 kDa protein, composed primarily of two other major proteins (~39 and 41 kDa) and smaller amounts of the CfTX proteins, were haemolytic.

Few research groups have previously fractionated nematocyst-derived *C. fleckeri* venom and their results vary significantly (Baxter and Marr, 1969; Endean, 1987, 1993; Bloom et al., 1998). The infrequent use of SDS-PAGE and western blot analyses to examine the protein composition of the crude venom and chromatography fractions has also made comparison of results difficult. However, in one of the studies, the fractionation profile of extracts from thawed and ground non-lyophilised nematocysts, obtained using Sephadex G-200 chromatography (fractionation range 5–600 kDa), resulted in the elution of four protein peaks with molecular masses equivalent to 600, 150, 70 and < 14 kDa (Endean et al., 1993). The two larger peaks (600 and 150 kDa) exhibited myotoxic activity, whereas the 70 kDa protein was haemolytic. Reducing SDS-PAGE profiles of the 600 and 150 kDa proteins peaks contained one major band (~18 kDa) and three minor bands (~59, 43 kDa and 25 kDa), one of which is similar in molecular mass to the CfTX proteins (~43–45 kDa). Notably, a 370 kDa protein corresponding to purified CfTX proteins was absent from the elution profile. Conversely, no major 18 kDa venom protein was observed in this study.

The differences between past and present results could be due to several factors. Inaccurate molecular mass calculations explain the variations between some protein masses. Variation in the quality of the extract source and/or the preparation of nematocyst extracts (e.g. freeze/thawing of nematocysts, lack of protease inhibitors etc.) may have resulted in the degradation of the CfTX proteins prior to analysis in the previous study. Differences in chromatographic conditions (e.g. ionic strength) may also have affected the results. In this study, the ionic strength of the size exclusion buffers was elevated to minimise non-specific protein-protein and/or protein-stationary phase hydrophobic interactions. Moreover, stabilisation of purified *C. rastoni* homologues (CrTX-A and -B) was achieved at high salt concentration (Nagai et al., 2000a). However, in the previous study (Endean et al., 1993), the ionic strength of the extraction/elution buffers was low. Therefore, the CfTX proteins may have

adsorbed to the column media or formed insoluble aggregates, which were unable to pass through the column.

In the present study, western blot analysis using CfTX-specific antibodies provided a useful method to examine the stability of the CfTX proteins during nematocyst venom preparation and fractionation by size exclusion or ion exchange chromatography. In crude venom and chromatography fractions, the highly specific antibodies bound only to the 43/45 kDa protein bands, indicating that CfTX-1 and -2 were not significantly degraded during extraction or purification. Interestingly, the CfTX proteins were detected in all size exclusion peak fractions, suggesting that native CfTX-1 and -2 may have dissociated during purification and/or interacted with the stationary phase (despite the addition of salt to the buffers), and eluted in a manner unrelated to size. Similarly, CfTX-1 and -2 were detected in all specific cation exchange peak fractions, indicating strong interactions of residual CfTX proteins with the stationary phase, despite increasing ionic strength.

Western blot analyses using rabbit antibodies raised against nematocyst venom confirmed that several proteins in the crude venom and chromatography fractions, including the CfTX proteins, are strongly antigenic. In contrast, the 39 and 41 kDa protein components of the 145 kDa cytolysin did not significantly bind the rabbit antibodies (see Figure 6.4d and Figure 6.7d) or CSL antivenom (data not shown), suggesting that neither antibody type could neutralise the potential biological effects of the 39 and 41 kDa proteins. However, the lack of antigenicity of the 39/41 kDa proteins could relate to the particular batch of nematocysts that was used for the size-exclusion and ion-exchange chromatography experiments. In previous molecular characterisation studies, which utilised different batches of nematocysts, the 39/41 kDa proteins were not major components in typical SDS-PAGE protein profiles of the crude venom (see Chapter 5 for comparison). As these nematocyst extracts were subsequently used for rabbit antibody production, it is likely that insufficient amounts of the 39/41 kDa proteins were present in the antigen preparations to produce a significant immunological response. Similarly, ‘‘milked venom’’, used for the production of CSL antivenom in sheep (Barnes, 1967), may not typically contain the 39/41 kDa proteins. Conversely, the 39/41 kDa proteins may not be strongly immunogenic in rabbit or sheep. The partially purified 39 and 41 kDa proteins did not bind the CfTX-specific antibodies, indicating that the proteins are unlikely to be derived from CfTX-1 and -2. As the 39 and 41 kDa proteins were twice as haemolytic as the CfTX proteins, these previously uncharacterised proteins may represent a novel class of cytolytic jellyfish proteins.

The observed difference in venom composition between nematocyst batches is an interesting finding and may be due to variations in the gender of the sampled jellyfish or differences in the geographical location or environmental conditions during jellyfish collection. For example, the nematocysts used for initial SDS-PAGE analysis (Figure 4.1 and Figure 4.2), were collected from jellyfish near Townsville (Queensland, Australia), whereas nematocysts used for size exclusion and ion exchange chromatography experiments (see Figure 6.4 and Figure 6.7) were collected from jellyfish near Mission Beach (~200 km north of Townsville). Alternatively, the different protein profiles may reflect an ontogenetic shift in venom composition, as reported for the venoms of other organisms, including the box jellyfish *Carukia barnesi* (Underwood and Seymour, 2007), snakes (e.g. Mackessy et al., 2006) and spiders (e.g. Herzig et al., 2004). Unfortunately, the batches used in the current and previous studies contained nematocysts from several *C. fleckeri* jellyfish of undocumented size or age, so the observed difference in venom composition cannot be explained under these circumstances.

Haemolytic activity has been demonstrated in whole nematocyst venom from a number of box jellyfish, including *C. alata* (Chung et al., 2001), *C. marsupialis* (Rottini et al., 1995), *C. bronzie*[††], *C. xaymacana*, as well as *C. fleckeri* (Bailey et al., 2005). In this study, the potent haemolytic activity of crude *C. fleckeri* nematocyst extracts was also observed. Both ion exchange and size exclusion chromatography enabled the partial purification of two major haemolysins (370 and 145 kDa) from nematocyst-derived venom. In the size exclusion studies, a minor 70 kDa haemolysin was also detected. Isolation of three haemolysins varies from previous reports of a single 70 kDa haemolysin in *Chironex* tentacle or nematocyst extracts (Endean et al., 1993; Crone and Keen, 1969) or two haemolysins, with native molecular masses of 70 and 120 kDa (Naguib et al., 1988). However, as significant amounts of the CfTX proteins or the 39/41 kDa proteins were not previously reported in the SDS-PAGE protein profiles of crude extracts and chromatography fractions (e.g. Endean et al., 1993), these haemolytic proteins may not have been abundant in the previous crude extracts.

Based on the amino acid sequences of CfTX-1, -2 and homologues from three other species of box jellyfish, CqTX-A, CrTX-A and CaTX-A (Nagai et al., 2002, 2000a, b), secondary structure predictions suggest that the CfTX proteins and their homologues are potential cytolysins, due to the presence of a conserved TSR1 (Brinkman and Burnell, 2007). Haemolysis assays undertaken in this study on partially purified CfTX proteins support the

[††] Reported as *Chiropsalmus sp.* in Bailey et al. (2005); reclassified as *Chiropsella bronzie* in Gershwin (2006).

predictions of a cytolytic mechanism of action. Furthermore, the potent haemolytic activity of individually purified CqTX-A, CrTX-A and -B, and CaTX-A and -B has also been confirmed (Nagai et al., 2002, 2000a, b).

Kinetic results from the current study indicate that the *in vitro* haemolytic activity of co-purified CfTX-1 and -2 (~370 kDa) is temperature-dependent and that the mechanism of haemolysis involves a pre-lytic lag period of ~6-7 minutes. In contrast, death of seriously envenomed humans can occur more rapidly (Lumley et al., 1988). According to one study, there is no apparent difference in the haemolytic activity of whole *C. fleckeri* nematocyst extracts on human and sheep erythrocytes (Bailey et al., 2005). Therefore, unless the cytolytic activity of the CfTX proteins is activated or potentiated *in vivo*, CfTX-1 and -2 may not be the lethal components associated with *Chironex* venom. However, the observed lag phase may be inversely proportional to cytolysin concentration, as observed for other haemolysins (e.g. Koyama, 1965; Cooper et al., 1964), and it is possible that doses of CfTX proteins injected during lethal envenomation may exceed the doses used in this study. The mechanism of cytolytic activity may also depend on the presence of ions or other compounds, which were omitted from the extraction buffers in this study. For example, the haemolytic activity of the *C. alata* venom is dependent on the presence of divalent cations (Mg^{2+}, Ca^{2+} and Zn^{2+}) and is optimal at a concentration of 10mM Ca^{2+} or Mg^{2+} (Chung et al., 2001). The dependence of haemolytic activity on Ca^{2+} concentration has also been observed for the *C. marsupialis* toxin, CARTOX (Rottini et al., 1995).

Finally, as extensive haemolysis has not been documented in clinical situations (Bailey et al., 2005; Tibballs, 2006), the cytolytic effects of the CfTX proteins may be more specific, fast-acting and damaging to different cells types, such as muscle or nerve cells, due to variations in membrane composition or affinity to specific membrane-surface receptors. The potent cardiotoxicity, neurotoxicity and nociceptive reactivity of whole *Chironex* venom *in vitro* and *in vivo* are already well established (Winter et al., 2007a; Ramasamy et al., 2003, 2004; Cuypers et al., 2006). Therefore, as CfTX-1 and -2 are the major constituents of nematocyst venom, the proteins could also be responsible for these painful and potentially life-threatening effects.

Despite recent advances in elucidating the mechanism of action associated with the severe pain sensation of *C. fleckeri* stings (Cuypers et al., 2006), the mechanism underlying the lethal effects of *C. fleckeri* venom in seriously envenomed humans remains largely unknown. Although the lethal factors involved were not identified, recent independent pharmacological studies suggested a primary cardiotoxic role in fatal jellyfish stings and that an ionophoric effect

is the most likely mechanism of action (Winter et al., 2007a). The pore-forming potential of purified CfTX proteins on mammalian cardiocytes has yet to be examined, however, exposure of intact cultured rat cardiocytes to unfractionated *C. fleckeri* nematocyst venom caused the formation of numerous circular lesions with internal diameters of ~50-80nm (Bailey et al., 2005), indicating the presence of at least one stoichiometric cytolysin. In addition, assays on selected CfTX homologues demonstrated that CrTX-A, CaTX-A and CqTX-A are lethal to crayfish, CrTX-A and -B are lethal to mice and CrTX-A causes pain, inflammation and necrosis of the skin similar to that observed in envenomed humans (Nagai et al., 2002, 2000a, b). Due to the significant sequence homology between the CfTX proteins and their cubozoan homologues, it is likely that this novel family of box jellyfish toxins share similar structure and function. Hence, CfTX-1 and -2 could also be inflammatory, dermonecrotic and lethal.

Further research is therefore necessary to determine the clinical or ecological importance of CfTX-1 and -2. Molecular characterisation of the newly discovered 39/41 kDa cytolytic proteins is also required. Lethality studies and *in vitro* / *in vivo* cardiotoxicity and neurotoxicity studies on fractionated nematocyst venom would complement the haemolytic activity data acquired in this study and provide more information about the specificity of the CfTX proteins and the other venom components. Further method development is also necessary to enable the separation of CfTX-1 and -2, and the 39 and 41 kDa proteins, for biochemical characterisation of the individual proteins. Investigation of factors, which may explain apparent shifts in venom composition, such as ontogeny, gender, spatial distribution of jellyfish populations and environmental variations, etc., would also benefit future *C. fleckeri* venom research.

CHAPTER 7

Expression of Major *C. fleckeri* Venom Proteins in *E. coli*

7.1 Introduction

In previous studies investigating the toxicity and biochemical characteristics of box jellyfish venom, difficulties encountered during the purification of bioactive venom proteins are a major concern, as exemplified by the inconsistent results of studies on *C. fleckeri* (see Chapter 1 for a review). These problems also limit the ability of researchers to undertake critical functional and structural studies that are necessary to elucidate the molecular mechanisms of isolated venom proteins. Obtaining sufficient amounts of crude nematocyst venom is also challenging because the occurrence of *C. fleckeri* is seasonal, sporadic and unpredictable, and collection of samples can be time-consuming and dangerous.

For these reasons, recombinant expression technology may offer an alternative to the isolation of native *C. fleckeri* venom proteins. However, no expression studies of cubozoan venom proteins have yet been reported. In contrast, several researchers have used recombinant protein expression to elucidate the structures, functions and/or molecular mechanisms of toxins from venomous organisms such as sea anemones (see Anderluh and Maček, 2002 for a review), stonefish (Ghadessy et al., 1996), cone snails (e.g. Pi et al., 2007), snakes (e.g. Singhamatr and Rojncuckarin, 2007) and spiders (e.g. da Silveira et al., 2007). Therefore, similar outcomes could also be achieved for the characterisation of *C. fleckeri* toxins and those of other box jellyfish species. Furthermore, if protein yields are significant, recombinantly-expressed venom proteins could be used for commercial antivenom production or other therapeutic applications. From a conservation or human safety perspective, the demand for collecting jellyfish samples for the isolation of wild-type venom proteins could also be minimised.

The aim of this study was to undertake novel expression studies of the cytolytic *C. fleckeri* toxins CfTX-1 and -2 in *E. coli.* Expression constructs were designed for the synthesis of N-terminal (histidine)$_6$-tagged CfTX proteins to simplify purification of the venom proteins using nickel-affinity chromatography. The challenges in achieving high-level expression of soluble CfTX-1 and -2 are described and potential modifications to future expression studies are discussed.

7.2 Methods

7.2.1 Subcloning CfTX-1 into pProEX HTc

CfTX-1 was subcloned from pBSK- into pProEX HTc via 2 existing *Bst*UI restriction sites in the 5'- and 3'-UTR of CfTX-1 that were compatible with a unique *Stu*I site in the pProEX HTc vector. pBSK- DNA containing the 1789bp CfTX-1 fragment (see Figure 5.2) was digested with *Bst*UI (2h, 60ºC). The resulting CfTX-1 DNA fragment (1415bp) was purified (QIAquick Gel Purification Kit, Qiagen) and blunt-end ligated (4 days, 4ºC) with pre-digested (*Stu*I; 2h, 37ºC) and purified pProEX HTc (Life Technologies). The ligation products were transformed into competent NM522 *E. coli* cells (Stratagene) and the cells grown overnight (37ºC) on LB-C (50µg/mL carbenicillin) agar. Further details on restriction enzyme digestion, DNA ligation and bacterial transformation are provided in Sections 2.4.5, 2.4.6 and 2.5.2).

NM522 colonies containing the CfTX-1 expression construct were identified by screening the colonies with a radiolabelled gene-specific oligonucleotide probe, prepared as previously described (see Sections 2.4.11 and 2.4.14). Four positive NM522 colonies were subjected to colony PCR to select pProEX HTc clones containing CfTX-1 inserts in the correct orientation with respect to the initiating start codon of the vector. PCR was performed using primers M13R (5'-GCGGATAACAATTTCACACAGG-3') and 43-R5 under hot start PCR conditions (95°C for 5 min, 35 cycles of 95°C for 45s, 54°C for 1 min and 72°C for 2 min, then 72° for 5 min). All 4 colonies produced a PCR product (~1.8kb). The colonies were cultured overnight (37ºC, 260rpm) in LB-C (50µg/mL), and their plasmid DNA was extracted (QIAprep Spin Miniprep Kit; Qiagen) and sequenced in both directions using primers M13R, 43-F9 (5'-GACTACCTTGCTCTCATTTG-3'), 43-F10, 43-R5, 43-R10 (5'-ACACCATACAGTGC-TTCC-3') and a pProEX reverse primer, ProEX-R (5'-TTCACTTCTGAGTTCGGCATGG-GG-3').

7.2.2 Subcloning CfTX-2 into pProEX HTc

CfTX-2 was subcloned from pBSK- into pProEX HTc via an existing *Bst*UI restriction site in the 5'-UTR of CfTX-2 that was compatible with *Stu*I in pProEX HTc and a downstream *Kpn*I restriction site common to both pBSK- and pProEX HTc. pBSK- DNA containing the 1624bp CfTX-2 fragment (see Figure 5.3) was restriction digested sequentially with *Bst*UI (2h, 60ºC) and *Kpn*I (2h, 37ºC). The resulting CfTX-2 DNA fragment (1584bp) was purified and ligated

overnight (16ºC) with double-digested (*Stu*I / *Kpn*I; 2h, 37ºC) and purified pProEX HTc. The ligation products were transformed into competent NM522 cells and the cells grown overnight (37ºC) on LB-C (50µg/mL) agar. Six colonies were screened by PCR using primers 43-F10 and pProEX-R under hot start PCR conditions (95°C for 5 min, 35 cycles of 95°C for 45s, 51°C for 1 min and 72°C for 2 min, then 72° for 10 min). Three colonies that produced a PCR product (1543bp) were cultured overnight (37ºC, 260rpm) in LB-C (50µg/mL). Plasmid DNA was extracted from the cultures and sequenced in both directions using primers M13R, 43-F10, 43-F12 (5'-CTCACAATGACTACCTTGC-3'), 43-R8 (5'-GGTAGTCATTGTGAG-GACAC-3'), 43-R9 (5'-GTTCTTGGATTGCCTTCG-3') and pProEX-R.

7.2.3 Small-Scale Expression of CfTX-1 and -2

Prokaryotic expression of CfTX-1 and -2 was initially undertaken in the *E. coli* strain NM522. Aliquots of competent NM522 cells were retransformed with CfTX-1 and CfTX-2 in pProEXc and grown overnight (37ºC) on separate LB-C (50µg/mL) agar plates. Single colonies from each plate were cultured overnight (37ºC, 260rpm) in LB-C (50µg/mL). Bacterial glycerol stocks were prepared from each NM522 saturated culture (Section 2.5.3) and stored at -80ºC. Aliquots of LB-C (50µg/mL; 20mL) were inoculated with the cultures of CfTX-1 or -2 (0.01 vol.) and the cells were grown (37ºC, 260rpm) to mid-logarithmic phase (OD_{600} ~0.5). Plasmid DNA in the remainder of each culture was extracted and the presence of the CfTX-1 and -2 inserts was confirmed by restriction enzyme analysis (*Nco*I / *Kpn* I; 2h, 37ºC). When the NM522 cultures reached OD_{600} ~0.5, an aliquot from each sample (10mL) was removed and centrifuged (5K×*g*, 4ºC, 10 min). The pellet was washed briefly (10mM Tris-HCl buffer, pH 7.5), re-centrifuged and the resulting pellet (uninduced sample) was frozen (-20°C). Protein expression in the remainder of each culture was induced by addition of IPTG (0.6mM final concentration), and incubation of the cultures was continued (2.5h at 37ºC then overnight at 23ºC, 260rpm). In parallel, a culture of NM522 cells transformed with uncut non-recombinant pProEXc (in LB-C) was induced as a negative control. The induced samples and negative control (10mL) were centrifuged (5K×*g*, 4ºC, 10 min) and the pellets were briefly washed. All pellets (induced and uninduced) were resuspended in 2X SDS-sample buffer.

In an attempt to optimise expression of CfTX-1 and -2, the experiment was repeated several times by adjusting experimental parameters including the optical density of cultures prior to induction (OD_{600} ~0.5–2), final IPTG concentration (0.6–1mM), incubation temperature (20–37°C) and the time of incubation following induction (1, 2, 2.5, 3, 3.5, 4 and 20h). To

assess whether the expressed proteins were toxic to the host cells, the OD_{600} of the cultures was monitored before and after induction of expression in selected experiments. Cultures containing pProEX-HT clones without CfTX inserts were used as controls.

7.2.4 Large-Scale Expression of CfTX-1 and -2

Large-scale prokaryotic expression of CfTX-1 and -2 was undertaken in *E. coli* strain BL21(DE3)-RIPL (Stratagene). Aliquots of competent BL21 cells were transformed with CfTX-1 or CfTX-2 plasmid DNA (pProEXc) and grown overnight (37ºC) on LB-C-Cam (50µg/mL carbenicillin and chloramphenicol) agar. Cultures of each clone were incubated overnight (37ºC, 260rpm) in LB-C-Cam (50µg/mL each) and aliquots (5mL) were retained for plasmid purification and restriction enzyme analysis (*Eco*RI and *Kpn*I or *Sac*I). For each clone, 2L LB-C (50µg/mL) was inoculated with the remaining culture (0.05 vol.) and incubated (37ºC, 260rpm) until the cells reached early stationary growth phase (OD_{600} ~2). An aliquot of each culture (10mL) was retained as an uninduced sample. The uninduced cells were harvested (5K×*g*, 4ºC, 10 min) and the pellets were washed twice with pellet wash buffer (20mM PO_4^{3-}, 0.5M NaCl, pH 8.0), resuspended in 5X SDS sample buffer and stored at -20ºC until analysis by SDS-PAGE and western blot. Protein expression in the remaining cultures was induced by addition of IPTG (1mM final concentration) and incubation of the cultures continued (4h, 37ºC, 260rpm). The induced cells were harvested and the pellets were washed twice with pellet wash buffer. A small amount of each pellet was retained for SDS-PAGE and western blot analysis; the remainder was stored at -80ºC.

7.2.5 Purification of Recombinant CfTX-1 and -2

7.2.5.1 Soluble Protein Purification under Native Conditions

The soluble fraction of expressed proteins was extracted under non-denaturing conditions and purified using nickel-affinity column chromatography. Induced BL21 cells were thawed on ice and resuspended (~3-5mL/g pellet) in application buffer (Buffer A; 20mM PO_4^{3-}, 0.5M NaCl, 5mM imidazole, 1mM PMSF, 1% Tween 20, pH 8.0). Following addition of lysozyme (10mg/mL, 80µL/g pellet), the cells were incubated on ice (30 min). The cells were lysed on ice using 10-15 cycles of pulsed sonication (1 pulse/s, 10 s/cycle), interspersed with brief cooling in liquid nitrogen. The cell lysate was centrifuged (48K×*g*, 4ºC, 30 min) and the supernatant filtered (0.22µm Millex-GP filters). The filtered native cell lysate (C) and the

remaining pellet (P) were stored on ice. A small amount of each sample was set aside for SDS-PAGE and western blot analysis.

The filtered crude native lysate was applied to a 5mL HiTrap IMAC HP column (GE Healthcare), pre-equilibrated with application buffer (5–10 CV), and connected to a Bio-Rad BioLogic Duo-Flow HPLC system (0.5 mL/min, 4°C). Following sample application, the column was washed with application buffer (50–60 CV, 2.5mL/min) and retained proteins were eluted by stepwise or gradual increases in elution buffer concentration (Buffer B; 20mM PO_4^{3-}, 0.5M NaCl, 0.5M imidazole, 1mM PMSF, 1% Tween 20, pH 8.0) using isocratic or linear gradient HPLC. Elution programs for CfTX-1 and CfTX-2 are outlined in Table 7.1 and Table 7.2, respectively. Protein elution was monitored by UV detection (280nm) and fractions (1 mL), numbered according to elution volume (V_e), were collected and stored on ice.

Table 7.1

Program for the elution of recombinant CfTX-1 extracted under native conditions from a nickel-affinity column.

[Imidazole] (mM)	% A	% B	V_e (mL)
40	93	7	120
80	85	15	50
100	77	23	50
100→500	77→0	23→100	150
500	0	100	50

Table 7.2

Program for the elution of recombinant CfTX-2 extracted under native conditions from a nickel-affinity column.

[Imidazole] (mM)	% A	% B	V_e (mL)
40	93	7	100
80	85	15	100
120	77	23	50
500	0	100	100

7.2.5.2 Insoluble Protein Purification under Denaturing Conditions

Following extraction of soluble proteins from the induced BL21 cultures under native conditions (see previous section), the remaining insoluble proteins (inclusion bodies) were solubilised and extracted under denaturing conditions prior to nickel-affinity purification. The post-native extraction pellet was resuspended in application buffer, treated with lysozyme, re-extracted using sonication (5 cycles) and centrifuged, as described previously, to remove remnant soluble proteins. The pellet was resuspended in 2M urea wash buffer (20mM PO_4^{3-}, 0.5M NaCl, 2M urea, 2% Triton X-100, pH 8.0), sonicated (5 cycles) and recovered by centrifugation. An aliquot of the urea wash supernatant (UW) was retained on ice for SDS-PAGE and western blot analysis. The remaining pellet was washed sequentially with urea wash buffer and pellet wash buffer. Purified inclusion bodies (0.4–0.8g) were resuspended in solubilisation buffer (10mL; 20mM PO_4^{3-}, 0.5M NaCl, 5mM imidazole, 8M urea, pH 8.0), stirred gently (1h, RT) and centrifuged (76K×*g*, 4ºC, 30 min). The denatured crude lysate (DC) and remaining pellet (DP) were stored on ice. A small amount of each sample was set aside for SDS-PAGE and western blot analysis.

Denatured crude lysate was applied (without filtering) to the 5mL HiTrap IMAC HP column, pre-equilibrated with solubilisation buffer (5–10 CV, 0.5 mL/min, 4°C). Following sample application, the column was washed with solubilisation buffer (10-15 CV, 2.5 mL/min) and 6M urea wash buffer (20mM PO_4^{3-}, 0.5M NaCl, 5mM imidazole, 6M urea, pH 8.0; 2.5 mL/min). Retained proteins were refolded slowly during buffer exchange to application buffer (i.e. 6→0M urea, linear gradient 0→100% A in 30 CV, 1 mL/min) and eluted with stepwise increases in elution buffer concentration. The elution programs for solubilised CfTX-1 and CfTX-2 are outlined in Table 7.3 and Table 7.4, respectively. Protein elution was monitored by UV detection (280nm) and fractions (1 mL), numbered according to V_e, were collected and stored on ice.

Table 7.3 Program for the elution of recombinant CfTX-1 extracted under denaturing conditions from a nickel-affinity column.

[Imidazole] (mM)	% A	% B	V_e (mL)
40	93	7	85
80	85	15	50
120	77	23	50
200	60	40	50
500	0	100	100

Table 7.4

Program for the elution of recombinant CfTX-2 extracted under denaturing conditions from a nickel-affinity column.

[Imidazole] (mM)	% A	% B	V_e (mL)
40	93	7	120
80	85	15	50
120	77	23	50
500	0	100	100

7.2.6 SDS-PAGE and Western Blot Analysis

Reducing SDS-PAGE analysis of proteins was conducted on cell lysates, pellets and chromatography fractions, as previously described (Section 2.3.3). Crude cell lysates and high protein concentration chromatography fractions were diluted prior to analysis. Chromatography fractions containing low amounts of protein were concentrated by TCA precipitation prior to analysis (Section 2.3.3.3). Proteins were separated on 10 or 12.5% polyacrylamide gels and Coomassie-stained. For western blot analysis, transferred proteins were blotted with rabbit α-CfTX and mouse α-$(His)_4$ primary antibodies diluted in blocking solution (1:1000 or 1:2000, respectively) (see Section 2.3.4). In most cases, aliquots (1 μL) of crude nematocyst extract and $(His)_6$-tagged PPDK (pyruvate orthophosphate dikinase), respectively, were used as positive controls on the western blot membranes.

7.3 Results

7.3.1 Subcloning CfTX-1 and -2 from pBSK- into pProEX HTc

Successful subcloning of CfTX-1 and -2 into pProEX HTc was achieved through restriction enzyme digestion and ligation reactions. The correct orientation and frame of the CfTX inserts with respect to the initiating start codon of pProEX HTc were confirmed by sequencing, and the expected molecular masses of $(His)_6$-tagged CfTX-1 and -2 fusion proteins were calculated as 53,167 and 53,920 Da, respectively (ProtParam; Gasteiger et al., 2005). Nucleotide and deduced amino acid sequences of the $(His)_6$-tagged CfTX-1 and -2 coding regions in the expression constructs are provided in Appendix 8.

7.3.2 Small-Scale Expression of CfTX-1 and -2

Preliminary small-scale expression studies in NM522 revealed that CfTX-1 and -2 were poorly expressed in their bacterial host. No significant expression of $(His)_6$-CfTX proteins (~53-54 kDa) was detected by comparing SDS-PAGE protein profiles of uninduced and induced NM522 cultures (Figure 7.1a). Western blot analysis using CfTX-specific antibodies provided mixed results in which expression of CfTX-1 was detected, but little or no expression of CfTX-2 was observed (Figure 7.1b). In contrast, no $(His)_6$-tagged proteins were detected for either expression construct using α-$(His)_4$ antibodies (Figure 7.1c). Attempts to optimise CfTX-1 and -2 expression by varying experimental parameters did not significantly improve these results. Restriction analysis of the expression constructs prior to induction consistently confirmed the presence of the CfTX inserts, and nucleotide sequencing verified that the insert sequences were unchanged since initial cloning. Therefore, the low or negligible expression of the CfTX proteins was not attributable to the loss of the plasmid and/or insert, or sequence mutations that may have caused a frame shift in the insert coding region. In addition to the low expression yields, cell density data obtained following induction revealed that CfTX-1 and -2 expression caused significant inhibition of bacterial growth compared to the control cultures (Figure 7.2).

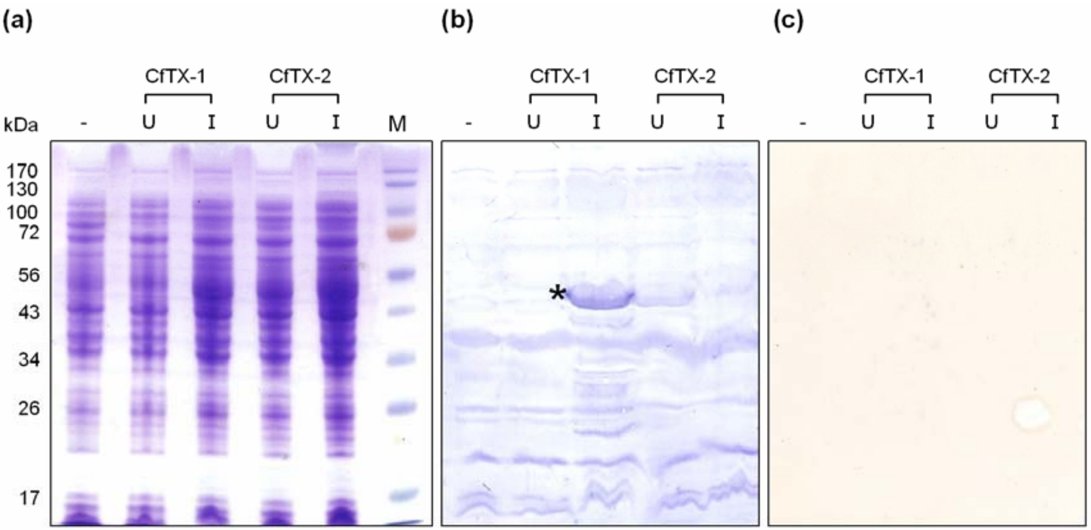


Figure 7.1 Small-scale expression of CfTX-1 and -2 in *E. coli* NM522. (a) Coomassie-stained SDS-PAGE protein profiles of uninduced (U) and induced (I) BL21 cell cultures transformed with CfTX-1 and CfTX-2 pProEX HTc clones. NM522 cells carrying only the pProEX HTc vector were used as a negative control (-). M indicates the protein molecular marker (Fermentas); the molecular masses of the protein standards are shown in kDa; (b) western blot of corresponding samples using α-CfTX antibodies; (c) western blot of corresponding samples using α-(His)$_4$ antibodies.

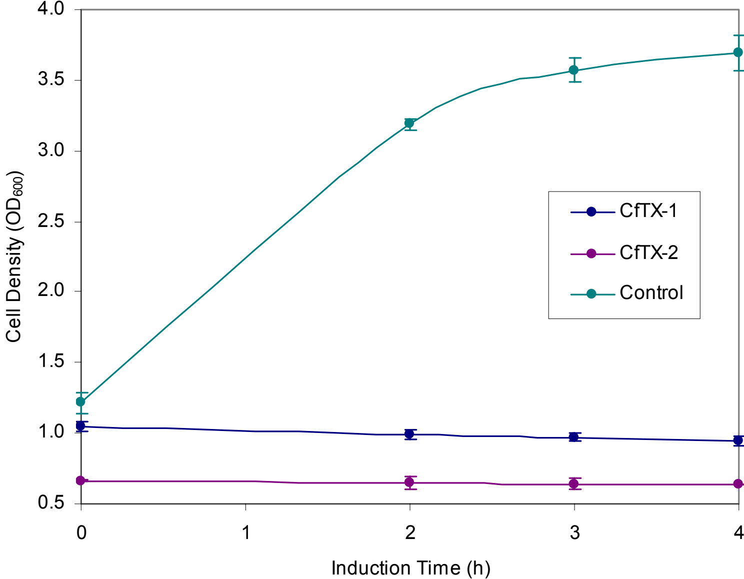


Figure 7.2 Bacterial growth following induction of protein expression. Bacterial cultures containing CfTX-1, CfTX-2 or control expression constructs were induced with IPTG (1mM) and incubated at 37°C with vigorous shaking (260 rpm). The cell density of cultures was monitored at OD_{600}. Error bars indicate the standard error of replicate samples (n=3).

7.3.3 Large-Scale Expression of CfTX-1 and -2

CfTX-1 and -2 were successfully expressed in large-scale BL21 cultures (2L each), however, estimated expression levels were very low (ng protein/g cell pellet). Extraction and nickel-affinity purification of CfTX-1 under native conditions demonstrated that the majority of expressed CfTX-1 was retained within the cell pellet as insoluble inclusion bodies (Figure 7.3). In contrast, significant concentration of proteins eluted from the nickel-affinity column with 500mM imidazole (peak 4; CF = 50) was required to visualise the minimal amount of soluble CfTX-1 present in the native cell lysate. Western blot analysis using α-CfTX antibodies confirmed the correct size and antigenicity of CfTX-1 (~50 kDa). However, the less sensitive α-$(His)_4$ antibodies were unable to detect an associated $(His)_6$-tag. Due to the lack of CfTX-1 purity in the 500mM imidazole fraction, no bioactivity assays such as sheep blood haemolysis were untaken.

In an effort to recover and renature insoluble CfTX-1 in the cell pellet, an additional chromatography step was required. The inclusion bodies in the cell pellet were isolated, solubilised under denaturing conditions and the proteins were purified by nickel-affinity chromatography (Figure 7.4). The proteins retained by the column were slowly refolded *in situ* by gradual removal of urea from the chromatography buffer, then eluted from the column. Although recombinant CfTX-1 was barely visible in the SDS-PAGE gel, western blot analyses using α-CfTX and α-$(His)_4$ antibodies both confirmed the presence of $(His)_6$-tagged CfTX-1 in most of the specific chromatography fractions, thus indicating the successful solubilisation of CfTX-1 from the inclusion bodies. This finding was also confirmed by comparing the protein profiles of the pellet before and after solubilisation (Figure 7.5). However, the amount of protein in peaks 5-7 (Figure 7.4) was very low and required significant concentration (CF = 85–210) prior to SDS-PAGE and western blot analysis. Western blot analyses also revealed significant CfTX-1 degradation in both the denatured crude lysate and eluted chromatography fractions, despite the addition of protease inhibitors to the extraction and chromatography buffers. The purity of CfTX-1 in the eluted peaks was also compromised by the presence of several, more abundant, non-recombinant bacterial proteins that were also retained by the column. Consequently, no bioactivity assays were attempted.

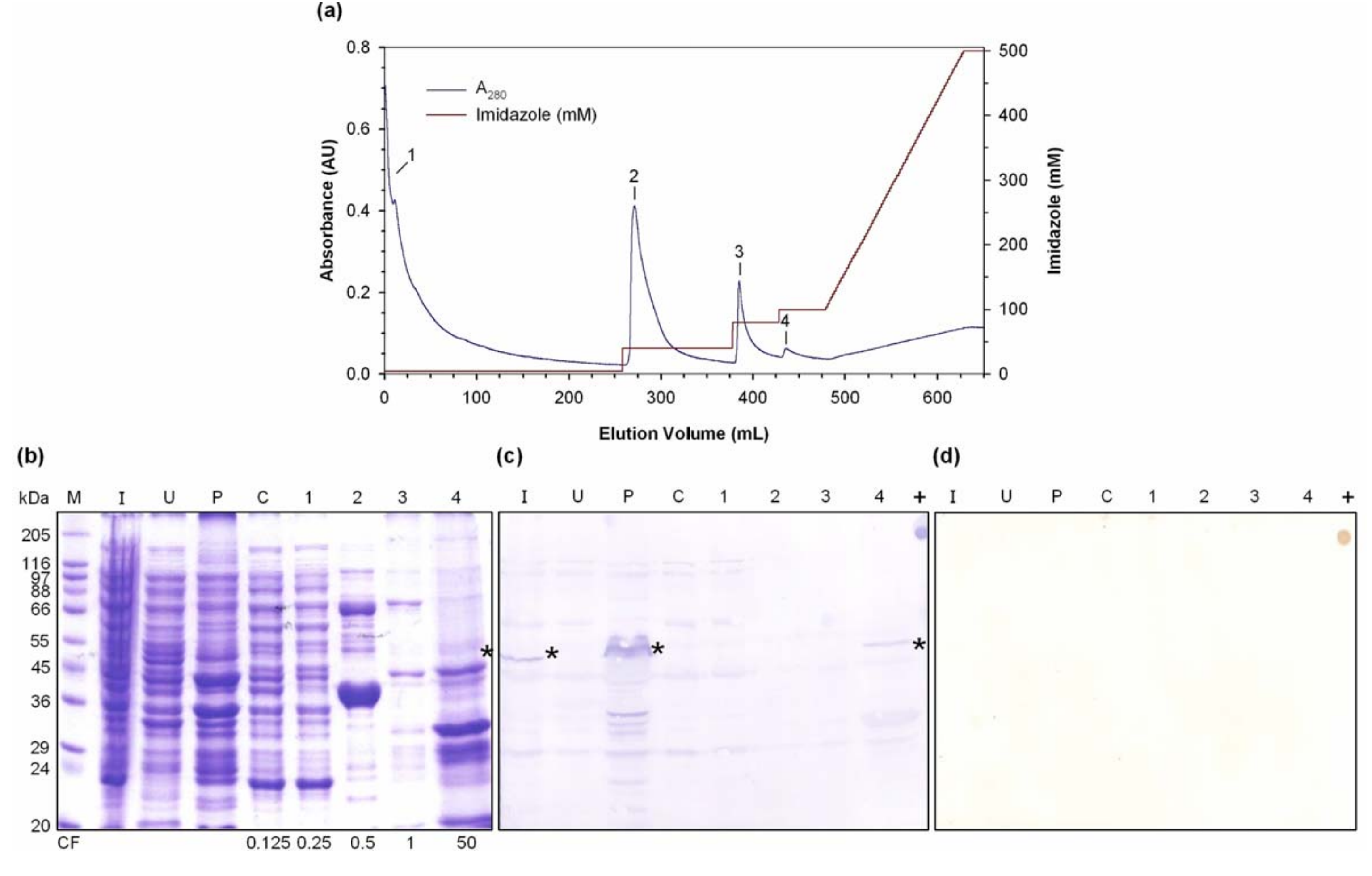


Figure 7.3 Purification of soluble CfTX-1 under native conditions (a) Chromatogram of native proteins eluted from a 5mL HiTrap IMAC HP column. Imidazole concentration and protein concentration (monitored at A_{280}) are indicated by red and blue lines, resp.; (b) Coomassie-stained SDS-PAGE protein profiles of induced (I) and uninduced (UI) BL21 cell cultures, the pellet remaining after protein extraction (P), native crude lysate (C) and chromatography fractions corresponding to eluted peaks 1–4. M indicates the protein molecular marker. The crude extract and peak 1 and 2 samples were diluted prior to analysis. Proteins in peak 4 sample were concentrated by TCA precipitation prior to analysis. The relative concentration factor (CF) is indicated below each lane; (c) western blot of corresponding samples using α-CfTX antibodies; (d) western blot of corresponding samples using α-$(His)_4$ antibodies. Recombinant CfTX-1 is indicated with an asterisk and positive controls are indicated by +.

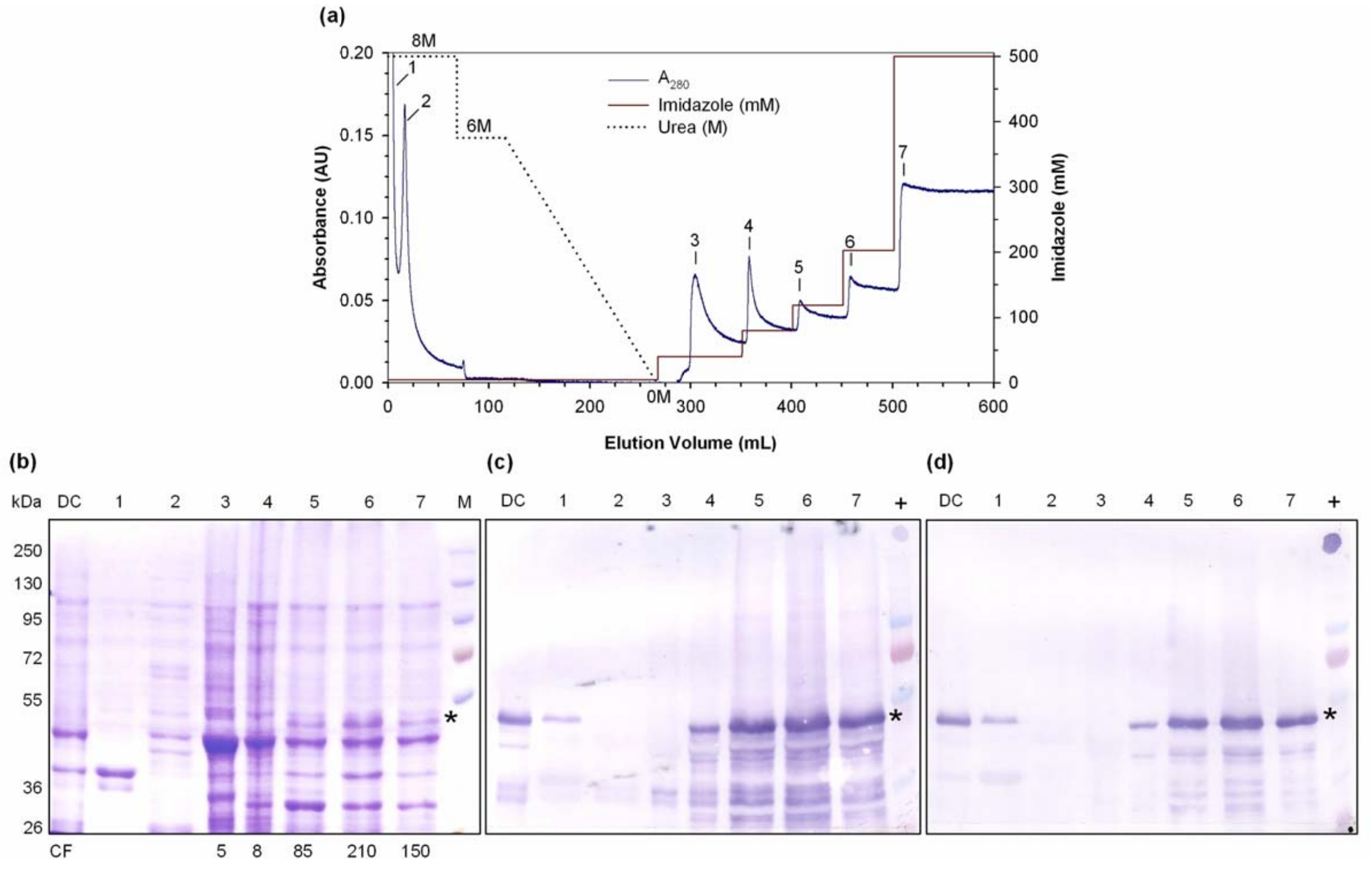


Figure 7.4 Purification of CfTX-1 from inclusion bodies under denaturing conditions. (a) Chromatogram of solubilised proteins eluted from a 5mL HisTrap column. Imidazole and protein (A_{280}) concentrations are indicated by red and blue lines, resp. Urea concentration is represented by a dotted line; (b) Coomassie-stained SDS-PAGE protein profiles of denatured crude lysate (DC) and chromatography fractions corresponding to eluted peaks 1–7. M indicates the protein molecular marker. The crude lysate was diluted prior to analysis. Proteins in peaks 2–7 were concentrated by TCA precipitation prior to analysis. The CF is indicated below each lane; (c) western blot of corresponding samples using α-CfTX antibodies; (d) western blot of corresponding samples using α-$(His)_4$ antibodies. Recombinant CfTX-1 is indicated with an asterisk and positive controls are indicated by +.

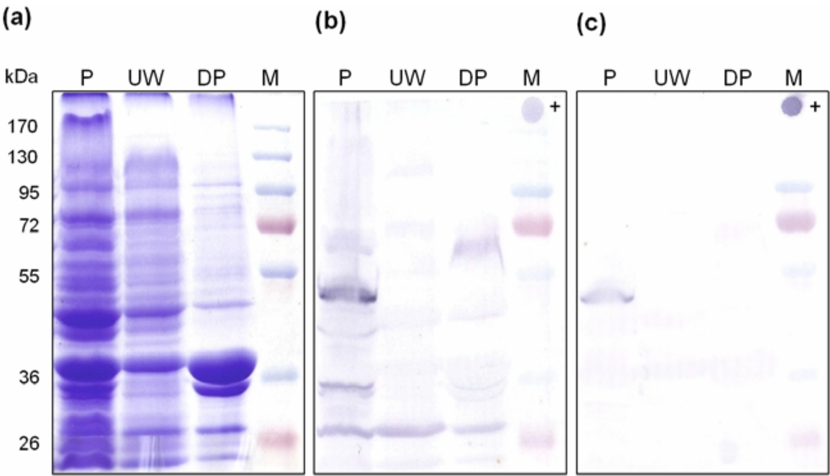


Figure 7.5 A comparison of proteins in the pellet before and after solubilisation of CfTX-1 inclusion bodies under denaturing conditions. (a) Coomassie-stained SDS-PAGE protein profiles of the pellet remaining after extraction of proteins under native conditions (P), a urea wash of P prior to solubilisation (UW), and the pellet remaining after solubilisation of inclusion bodies (DP). M indicates the protein molecular marker; the molecular masses of the protein standards are shown in kDa. (b) western blot of corresponding samples using α-CfTX antibodies; (c) western blot of corresponding samples using α-$(His)_4$ antibodies. Positive controls are indicated by +.

Success in the large-scale expression and purification of CfTX-2 was also limited. Although soluble $(His)_6$-tagged CfTX-2 was detected in the concentrated 120 and 500mM imidazole fractions of native cell lysates purified by nickel-affinity chromatography (Figure 7.6), the amount and purity of recovered CfTX-2 were still very low. No CfTX-2 was detected in the cell pellet or the native cell lysate using α-CfTX or α-$(His)_4$ antibodies, possibly due to the extremely low level of CfTX-2 expression. However, when the inclusion bodies in the cell pellet were purified and solubilised, both α-CfTX or α-$(His)_4$ antibodies weakly detected $(His)_6$-tagged CfTX-2 in the denatured crude lysate (Figure 7.7). Following purification of the solubilised proteins by nickel-affinity chromatography, small amounts of CfTX-2 were detected in concentrated chromatography fractions eluted with 80, 120 and 500mM imidazole. However, the amount and purity of recovered CfTX-2 were again very low and no bioactivity assays were undertaken.

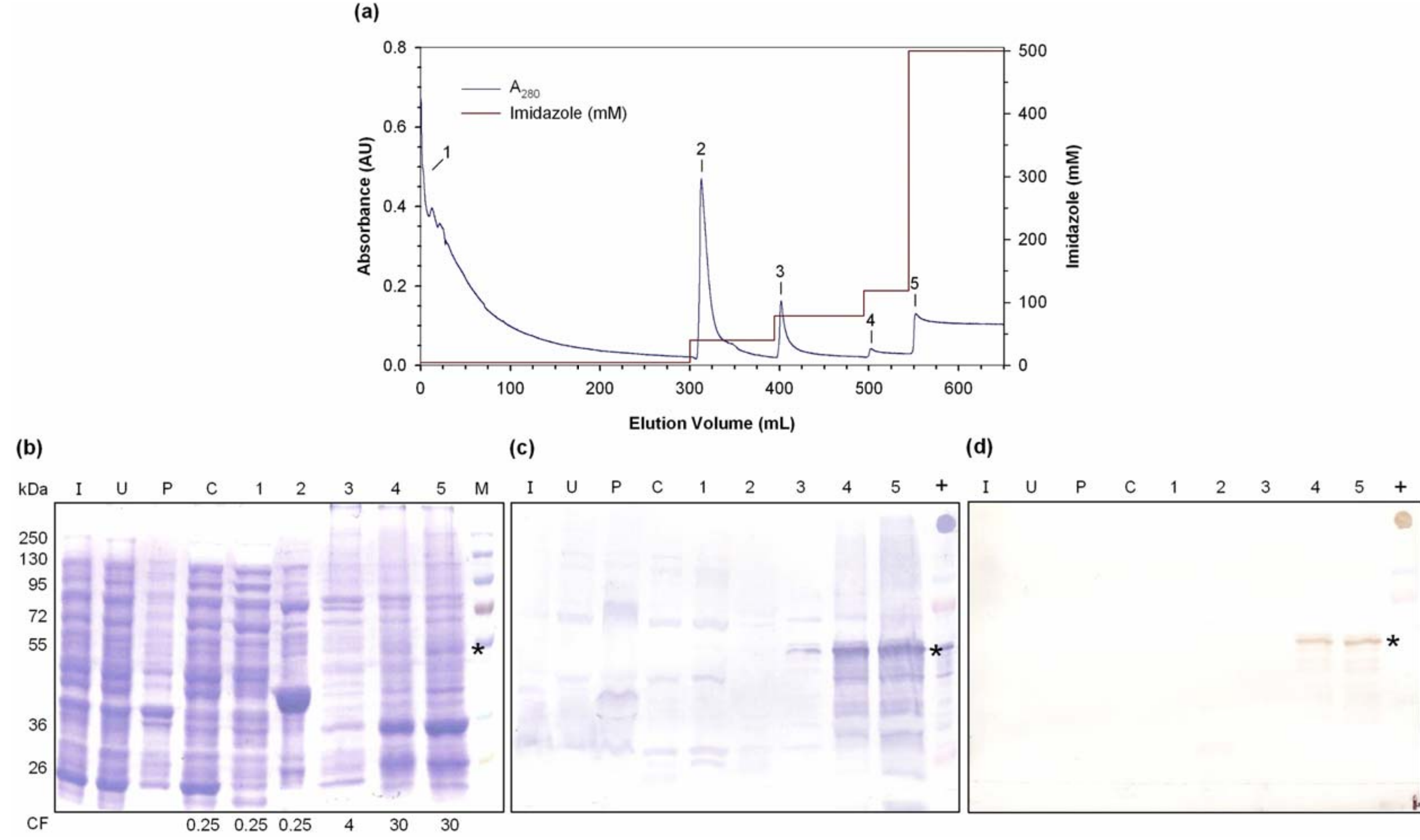


Figure 7.6 Purification of soluble CfTX-2 under native conditions (a) Chromatogram of native proteins eluted from a 5mL HiTrap IMAC HP column. Imidazole concentration and protein concentration (monitored at A_{280}) are indicated by red and blue lines, resp.; (b) Coomassie-stained SDS-PAGE protein profiles of induced (I) and uninduced (UI) BL21 cell cultures, the pellet remaining after protein extraction (P), native crude lysate (C) and chromatography fractions corresponding to eluted peaks 1–5. M indicates the protein molecular marker. The crude extract and peak 1 and 2 samples were diluted prior to analysis. Peak 3–5 samples were concentrated by TCA precipitation prior to analysis. The CF is indicated below each lane; (c) western blot of corresponding samples using α-CfTX antibodies; (d) western blot of corresponding samples using α-$(His)_4$ antibodies. Recombinant CfTX-2 is indicated with an asterisk and positive controls are indicated by +.

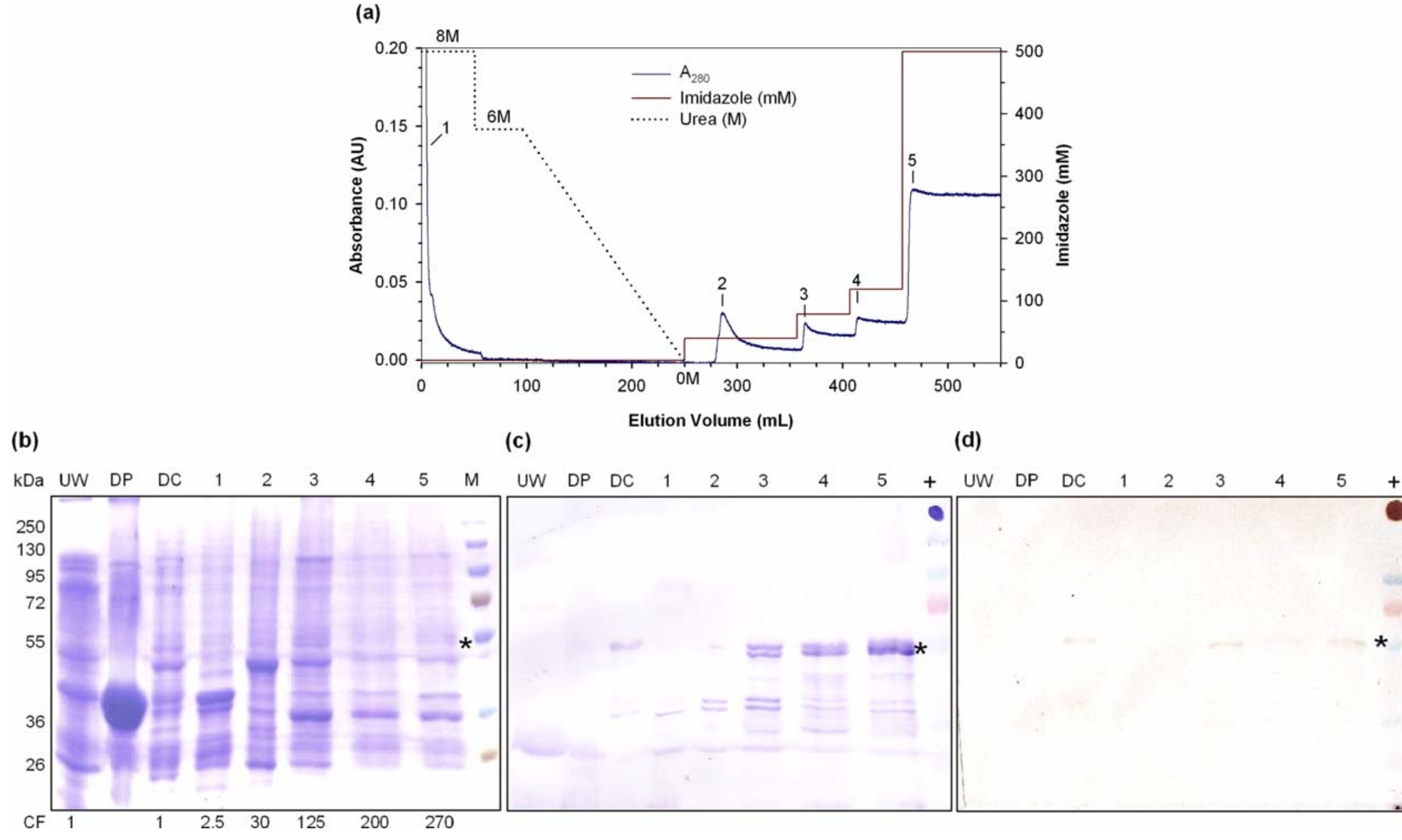


Figure 7.7 Purification of CfTX-2 from inclusion bodies under denaturing conditions. (a) Chromatogram of solubilised proteins eluted from a 5mL HisTrap column. Imidazole and protein (A_{280}) concentrations are indicated by red and blue lines, resp. Urea concentration is represented by a dotted line; (b) Coomassie-stained SDS-PAGE protein profiles of the 2M urea wash (UW), the pellet remaining after denaturing protein extraction (DP), denatured crude lysate (DC) and chromatography fractions corresponding to eluted peaks 1–5. M indicates the protein molecular marker. DC was diluted prior to analysis. Proteins in peaks 1–5 were concentrated by TCA precipitation prior to analysis. The CF is indicated below each lane; (c) western blot of corresponding samples using α-CfTX antibodies; (d) western blot of corresponding samples using α-$(His)_4$ antibodies. Recombinant CfTX-2 is indicated with an asterisk and positive controls are indicated by +.

7.4 Discussion

The results of this study demonstrate that the major nematocyst proteins, CfTX-1 and -2, can be expressed in a bacterial host, however, without further method development and optimisation, their expression levels are too low to permit subsequent functional and structural characterisation. Furthermore, a large proportion of the expressed CfTX proteins form insoluble inclusion bodies that require solubilisation and refolding prior to purification as soluble, potentially active recombinant proteins. Attempts to optimise cell growth and expression conditions for CfTX-1 or -2 by changing experimental parameters such as cell density prior to induction, final IPTG concentration, incubation temperature and length of expression time, produced no definitive improvement in the expression level or solubility of either protein within the limited timeframe of the research project. Monitoring cell density during small-scale expression studies revealed that induction of CfTX-1 and -2 expression caused a decline in bacterial cell density over time, thus indicating that the expression products are toxic to *E. coli*. Given that CfTX-1 and -2 are predicted to contain transmembrane spanning regions and act as pore-forming toxins (Brinkman and Burnell, 2007), it is possible that the recombinant proteins cause damage to vital bacterial membrane systems, resulting in cell death and/or impaired cell function.

Despite this set-back to the research project, several other strategies are available to potentially improve the yield and/or solubility of CfTX-1 and -2 in future expression studies (for a selection of comprehensive reviews see Peti and Page, 2007; Jana and Deb, 2005; Hunt, 2005; Swartz, 2001; Baynex, 1999). For example, a change in the growth medium (e.g. 2× YT or Super Broth) and/or the addition of supplements (e.g. glycerol, glucose etc.) can have a marked effect on levels of protein expression (Moore et al., 1993; Broedel et al., 2001). Similarly, the solubility and functional activity of the expressed proteins may be dependent on the presence of metal ions, potential cofactors or prosthetic groups which can be added to the growth medium (Stevens, 2000). While the BL21(DE3)-RIPL CodonPlus *E. coli* strain used in this study is usually suitable for the expression of toxic genes, other strains with enhanced toxicity tolerance are also available and may be more suitable for the expression of CfTX-1 and -2 (e.g. OverExpress C41(DE3) and C43(DE3); Lucigen Corp., USA).

Other options to potentially improve yields of recombinant CfTX-1 and -2 include changing the expression constructs and/or the expression system. In the case of the expression constructs, the type of fusion tag (e.g. $(His)_6$, GST, *Strep*, MBP, etc.) and its position relative to

the target gene (N- or C-terminal) can have a dramatic impact on protein yield and solubility (Hunt, 2005). Alternatively, the cDNA encoding CfTX-1 and -2 could be modified by deleting the hydrophobic transmembrane spanning regions of the genes. Although such deletions would inactivate or impair the functionality of the proteins, the modified expression products may still be suitable for antibody production. In relation to the expression system, while bacterial systems are considered a good starting point for economical, large-scale recombinant protein expression, other expression systems such as yeast, insect or mammalian cells may be necessary to obtain higher protein activity, correct disulphide bonds formation or eukaryotic-specific post-translational modifications. Another promising alternative is the use of a cell-free expression system to potentially overcome the difficulties of synthesising CfTX-1 and -2 in cell-based systems. While cell-free expression kits are generally more expensive and relatively small-scale compared to bacterial expression systems, sufficient amounts of the CfTX proteins could be synthesised to allow a number of crucial functional assays and structural studies.

The discovery that CfTX-1 and -2 expression in *E. coli* is extremely low at least helps to explain why the immunoscreening experiment detailed in Chapter 3 failed to detect these two major nematocyst toxins. With such low expression levels, it is highly unlikely that the antibodies could detect expressed CfTX proteins using standard cDNA expression library immunoscreening protocols. To investigate this hypothesis further, replicate library phage were screened sequentially with α-CfTX antibodies and a CfTX-specific radiolabelled oligonucleotide probe, using methods previously described (Sections 2.4.13 and 2.4.14). Results of the immunoscreening and DNA screening experiments differed dramatically. Out of ~5×10^4 random library plaques, nine antigenic plaques were detected using α-CfTX antibodies (~0.02% total library clones), whereas ~100 radiolabelled plaques (~0.2% total library clones) were detected by autoradiography (see Figure 7.8a and b, respectively). In theory, approximately one third of the total CfTX-1 and -2 clones present in the unidirectional library would be correctly in-frame for expression in pBSK-. However, the proportion of antigenic plaques detected was much lower than expected. More importantly, none of the antigenic plaques coincided with any of the radiolabelled plaques, thus suggesting that the former are "false positives", similar to those obtained in the immunodetection experiment (Chapter 3). These findings highlight the increased sensitivity, reliability and usefulness of screening *C. fleckeri* cDNA libraries with DNA probes rather than antibodies, particularly when searching for clones that have relatively low representation in the library and/or are poorly expressed.

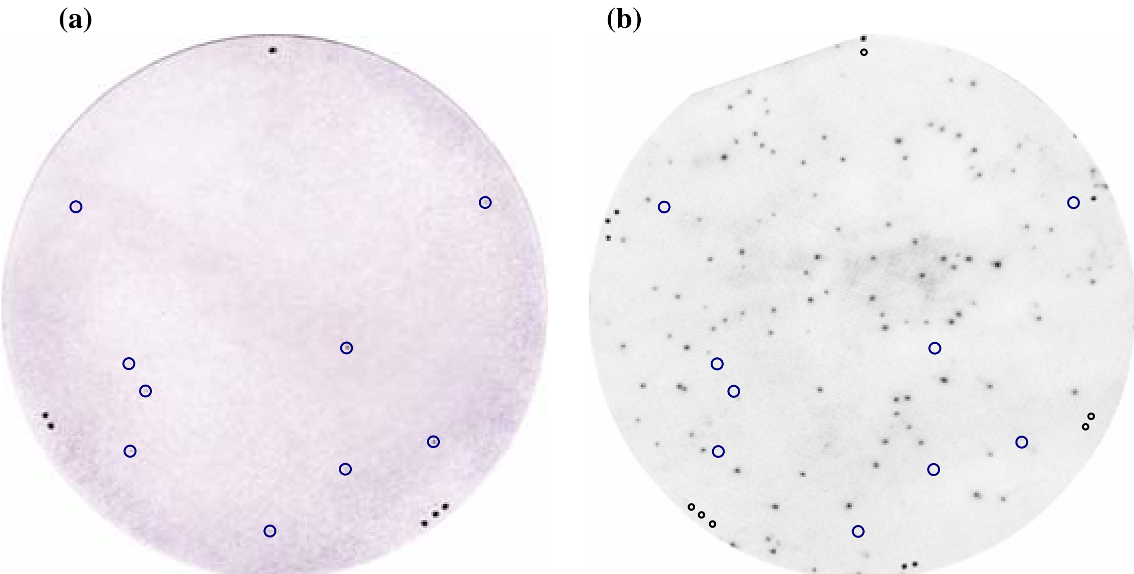


Figure 7.8 A comparison of two cDNA library screening strategies for the detection of CfTX-1 and -2 clones from ~5 × 10^4 replicate phage plaques. (a) Immunodetection of expressed cDNA library clones using α-CfTX antibodies; (b) Autoradiographic detection of cDNA library clones using a radiolabelled CfTX-specific oligonucleotide probe. Blue circles indicate antigenic plaques in (a) and their corresponding positions in (b). Black dots (•, •• and •••) indicate pin holes in the membranes used to mark the orientation of the membranes on the agar plate. Black circles (◦, ◦◦ and ◦◦◦) indicate the transposed pin hole positions of (a) onto (b).

CHAPTER 8

General Conclusions and Future Research

8.1 Research Overview

Cubozoan jellyfish produce an array of toxic proteins which are stored and delivered by their nematocysts. Although these toxins are primarily used for prey capture and possibly the deterrence of predators, they are also clinically important due to their painful and damaging effects in envenomed humans. Consequently, significant research effort has focussed on the characterisation of venom proteins from the large and extremely venomous box jellyfish, *Chironex fleckeri*, and to a lesser extent, from species such as *Chiropsalmus quadrigatus* and *Carybdea marsupialis*. However, despite several decades of biochemical, toxicological and pharmacological research, few toxic cubozoan proteins have been isolated and characterised, and their mechanisms of action remain unclear. In particular, progress in the biochemical characterisation of bioactive box jellyfish proteins has been hindered by issues relating to protein instability and hydrophobicity, as well as variability in protein sources, extraction methods and analytical techniques (see Chapter 1).

To narrow this gap in our current knowledge of cubozoan toxins, this thesis describes the molecular and biochemical characterisation of venom proteins from the harmful and potentially life-threatening box jellyfish, *C. fleckeri*. The aims of this study were to identify the major protein components of the venom, provide insight into their possible structures, functions and mechanisms of action, and based on this information, explore the use of biotechnology to advance current box jellyfish research and potentially improve contemporary treatments for *C. fleckeri* stings in humans.

8.2 Is There an Easy Way to Isolate and Identify *C. fleckeri* Venom Proteins?

An initial strategy to isolate and identify putative *C. fleckeri* venom proteins involved the construction and screening of a novel *C. fleckeri* tentacle cDNA expression library using antibodies (Chapter 3). The library was screened with three types of antibodies raised against

C. fleckeri milked or nematocyst venoms to detect random cDNA clones that expressed antigenic venom proteins. In principle, the strategy was sound and a number of previous studies successfully used cDNA expression libraries to isolate toxic proteins from venomous organisms such as snakes (e.g. Welton and Burnell, 2005; Assakura et al., 2003) and insects (e.g. Fang et al., 1988; Jones et al., 1992). However, no putative venom clones were isolated in this study using this immunoscreening approach. Instead, cDNA clones encoding 44 different non-venom proteins were collectively isolated. As detailed in Chapter 3, several inherent problems are likely to explain the ineffectiveness of the immunodetection strategy to isolate *C. fleckeri* venom clones. In particular, and as also demonstrated in Chapter 7, issues relating to poor venom protein expression and/or the toxicity of the expressed venom proteins to their bacterial host are likely to have contributed significantly to the unsuccessful detection of these target proteins. Nevertheless, this study has provided the first preliminary survey of the transcriptome of *C. fleckeri* and greatly enhances the number of *C. fleckeri* nucleotide and deduced amino acid sequences that are currently available in the publicly accessible database, GenBank (www.ncbi.nlm.nih.gov/Genbank/index.html). Publication of the box jellyfish sequences adds to the diversity of cnidarian species represented in the database and hence the sequences will be particularly useful for future phylogenetic studies investigating the early evolution of metazoans.

8.3 Biochemical Characterisation of *C. fleckeri* Venom Proteins

The failure to isolate putative *C. fleckeri* venom protein clones by immunoscreening the expressed *C. fleckeri* cDNA library (Chapter 3) prompted an immediate shift in research strategy towards the use of alternative biochemical and proteomic methods. Firstly, nematocyst-derived venom proteins were extracted and analysed using a variety of biochemical methods including SDS-PAGE, 2DGE and western blot (Chapter 4). Nematocysts were isolated from *C. fleckeri* tentacles prior to venom extraction to exclude extraneous tentacle proteins that would not normally be injected into envenomed animals. Nematocyst venom was subsequently prepared from the isolated nematocysts using a rapid bead mill homogenisation method previously adopted by several researchers (e.g. Bailey et al., 2005; Carrette and Seymour, 2004, 2005; Kitner et al., 2005; Ramasamy et al., 2003, 2004; Winter et al., 2007a).

To ensure that the nematocyst venom extracts used in this study were biologically active, lethality studies were conducted on crayfish and haemolytic activity was tested on sheep erythrocytes (Chapters 4 and 6). Bioassays confirmed that *C. fleckeri* nematocyst venom was

lethal to crayfish by intramuscular injection, but the time to death was dependent on the ionic strength of the buffer used for venom extraction (Chapter 4). For example, nematocyst venom (10ng total protein/g crayfish) extracted with high-salt buffer was lethal to the majority of crayfish within 3 minutes (172 ± 25s; n = 4; one survived). In contrast, the concentration of nematocyst venom extracted with low-salt buffer required to kill the majority of crayfish within a similar time-frame (187 ± 20s; n = 5; one survived) was significantly higher (429 ng total protein/g crayfish). Haemolytic assays confirmed that the nematocyst venom extracts were potently haemolytic to suspensions of sheep erythrocytes (HU_{50} = 5 ng/mL; Chapter 6) and caused the formation of large cleared zones of haemolysis in 5% sheep blood agar (Chapter 4).

SDS-PAGE analyses revealed that *C. fleckeri* nematocyst venom contains a variety of proteins ranging in molecular mass from 10–180 kDa, but the overall protein composition is dominated by the presence of two major proteins, CfTX-1 (~43 kDa) and CfTX-2 (~45 kDa). Western blot analyses demonstrated that several of these venom proteins, including the CfTX proteins, are strongly antigenic to CSL box jellyfish antivenom and rabbit antibodies raised against bead-mill nematocyst venom.

As described in Chapter 4, additional biochemical and immunological investigations, included:

- an examination of the efficiency of bead mill homogenisation to extract all the soluble venom proteins contained in *C. fleckeri* nematocysts
- a study to determine the effects of salt concentration on nematocyst venom protein extraction
- a comparison of protein composition between nematocyst and milked venoms
- an interspecies comparison of nematocyst venom protein composition between two species of chirodropid jellyfish: one harmful to humans (*C. fleckeri*) and one relatively harmless to humans (*C. bronzie*)
- a comparison of antibody specificity between commercially available box jellyfish antivenom (CSL Ltd), which is raised in sheep against milked venom, and rabbit and mouse polyclonal antibodies raised during this study against *C. fleckeri* nematocyst venom

Several interesting conclusions are based on the results of these investigations. Firstly, bead mill homogenisation represents an efficient method for the extraction of nematocyst venom proteins, however, the protein composition (and lethality) of nematocyst venom is

enhanced by the presence of salts in the extraction buffer. The composition of nematocyst venom is generally consistent from batch to batch of non-lyophilised nematocysts (using the same buffer), with the exception of two relatively major proteins (~39 and 41 kDa) and a minor protein (~12 kDa) that are not always observed. As discussed in Chapter 6, factors such as ontogeny, gender, spatial distribution of jellyfish populations and environmental factors may explain some of the variations observed in nematocyst venom protein composition. Comparatively, the protein composition of nematocyst venom is also reasonably consistent when extracted from lyophilised or non-lyophilised nematocysts, with the exception of a few minor proteins that vary in abundance and antigenicity to CSL box jellyfish antivenom.

In contrast, and somewhat surprisingly, the protein composition of reconstituted milked venom appears to be more unpredictable than that of nematocyst venom and depends upon the solubility of its lyophilised proteins at the time of reconstitution. The preliminary findings of this study indicate that the solubility of lyophilised milked venom proteins decreases with increased storage time, possibly due to oxidation of the proteins. These findings could have important implications for the efficacy of CSL box jellyfish antivenom, which is used for the treatment of *C. fleckeri* stings in humans, however, further research is required to investigate this unexpected phenomenon.

A comparison of the protein composition of *C. fleckeri* and *C. bronzie* nematocyst venoms was limited due to the small number of *C. bronzie* nematocysts available for this study. However, preliminary results indicate that the nematocyst venoms of the two species vary significantly. Previous pharmacological studies suggest that the cardiotoxic and neurotoxic effects of *C. fleckeri* and *C. bronzie* nematocyst venoms are similar, but *C. fleckeri* nematocyst venom is more potent than *C. bronzie* nematocyst venom (Ramasamy et al., 2003, 2005a; Winter et al., 2007a). Therefore, the differences in protein composition observed between species in this study could explain the previously observed differences in venom potency. Most notably, homologues of the major CfTX proteins in *C. fleckeri* venom appear to be significantly reduced in abundance in *C. bronzie* venom. Hence, CfTX-1 and -2 could be responsible for the more potent lethal, neurotoxic and/or cardiotoxic effects observed in envenomed animals. Furthermore, as found in this study and corroborated in another (Bailey et al., 2005), *C. fleckeri* nematocyst venom is significantly more haemolytic to sheep erythrocytes than *C. bronzie*, thus suggesting that the CfTX proteins may also significantly contribute to the potent cytolytic activity ascribed to *C. fleckeri* venom.

Despite the differences in venom protein composition between the two species of chirodropid jellyfish, significant cross-reactivity of rabbit antibodies raised against *C. fleckeri* nematocyst venom was observed in a western blot of *C. bronzie* nematocyst venom. The results suggest that antibodies raised against *C. fleckeri* nematocyst venom may have the potential to reduce the biological effects of *C. bronzie* envenoming. In contrast, less specific binding of CSL box jellyfish antivenom to *C. bronzie* venom proteins (amidst a high background of non-specific binding) was observed in a western blot of *C. bronzie* nematocyst venom, thus suggesting that the efficacy of CSL box jellyfish antivenom to counter the effects of *C. bronzie* stings may vary to that of antibodies raised against *C. fleckeri* nematocyst venom.

Administration of CSL box jellyfish antivenom is currently recommended for the treatment of severe and life-threatening *C. fleckeri* stings in humans (Tibballs, 2006). However, a recurrent finding in this study is that CSL box jellyfish antivenom, when used in western blot analyses of *C. fleckeri* nematocyst venom, detects several antigenic proteins but not the full complement of venom proteins detected by SDS-PAGE analyses. Tentatively, this finding suggests that the milked venom used to prepare CSL box jellyfish antivenom may lack some of the proteins that are present in *C. fleckeri* nematocyst venom. The results also support the assertions of several researchers that CSL box jellyfish antivenom is only partially effective in counteracting the *in vivo* and *in vitro* effects of *C. fleckeri* nematocyst venom, possibly due to differences in the active constituents of milked and nematocyst venoms (Endean and Sizemore, 1988; Ramasamy et al., 2003, 2004; Winter et al., 2007a).

To further examine the antigenicity of *C. fleckeri* venom proteins, western blot analyses of nematocyst venom proteins were also performed using mouse and rabbit antibodies raised against nematocyst venom itself. The western blot results indicated that, similar to CSL box jellyfish antivenom, antibodies raised against nematocyst venom bind to several but not all of the proteins contained in *C. fleckeri* nematocyst venom. Moreover, different patterns of antigenic proteins are observed when the same nematocyst venom is blotted with antibodies raised in different host species. The results of the study suggest that the immunological response of mammals to *C. fleckeri* nematocyst venom (and possibly milked venom) is species-specific and variable, due to the presence of venom proteins that are not significantly immunogenic. The clinical implication of these findings is that, similar to the criticisms of CSL box jellyfish antivenom, antivenoms prepared from nematocyst venom may be only partially effective in neutralising the biological effects of *C. fleckeri* envenoming. Therefore, further

pharmacological studies are necessary to compare the efficacy of antibodies raised against *C. fleckeri* nematocyst venom with that of CSL box jellyfish antivenom.

8.4 Molecular and Biochemical Characterisation of CfTX-1 and -2

Chapter 5 describes for the first time the identification, cloning and sequencing of venom proteins derived from the nematocysts of *C. fleckeri*. The two newly discovered proteins, CfTX-1 and -2 (~43 and 45 kDa, respectively), were targeted for further molecular and biochemical characterisation due to their dominant abundance in *C. fleckeri* nematocyst venom (Chapter 4) and their similar molecular masses to other previously characterised box jellyfish toxins (Nagai et al., 2000a, b, 2002).

A combination of N-terminal amino acid sequencing, peptide mass fingerprinting, RT-PCR and cDNA library screening enabled the cloning of two distinct cDNA sequences (1789 and 1624 bp, respectively) encoding CfTX-1 and -2 (Chapter 5). The N-terminal amino acid sequence of each translated clone is preceded by a signal peptide, which presumably directs the post-translational transport of each protein to the nematocyst. A comparison of the deduced amino acid sequences of mature CfTX-1 and -2 confirmed that the proteins are quite similar (73% identical) and protein database searches indicated that the CfTX proteins are related to lethal, haemolytic box jellyfish toxins, CqTX-A (>72%), CrTXs (26%) and CaTX-A (>24%) previously isolated from *C. quadrigatus, C. rastoni* and *C. alata*, respectively (Nagai et al., 2002, 2000a, b).

Biochemical studies revealed that CfTX-1 and -2 are difficult to separate by electrophoretic or chromatographic methods, however, the two proteins can be significantly co-purified using either size-exclusion or cation-exchange chromatography (Chapter 6; Brinkman and Burnell, 2008). Under native conditions, the molecular mass of the co-purified CfTX proteins was 370 kDa, indicating that an aggregated form of these proteins is present in the native venom, perhaps an octamer or mixture of oligomers. The co-purified CfTX proteins were also potently haemolytic to sheep erythrocyte suspensions (HU_{50} = 14 ng/mL) and caused the formation of large cleared zones of haemolysis in 5% sheep blood agar. The haemolytic activity of the CfTX proteins exhibited a sigmoidal dose-response curve (Chapter 6), characteristic of stoichiometric cytolysins (Hessinger et al., 1988). Although CfTX-1 and -2 were not significantly degraded during purification, the proteins were prone to adsorption to the size-exclusion and cation-exchange media and their haemolytic activity relative to crude nematocyst venom was reduced.

8.5 Additional Cytolytic Proteins in *C. fleckeri* Nematocyst Venom

During the fractionation studies of *C. fleckeri* nematocyst venom described in Chapter 6, two additional proteins exhibiting potent haemolytic activity were also partially purified. The proteins (~39 and 41 kDa) were twice as haemolytic to sheep erythrocytes (HU_{50} = 7 ng/mL) as co-purified CfTX-1 and -2, but were not always observed in different batches of nematocysts. Under native conditions, the estimated molecular mass of the co-purified proteins was 145 kDa, suggesting the formation of tetrameric (or possibly trimeric) quaternary structures.

Due the relatively high abundance of the 39 and 41 kDa proteins in some batches of *C. fleckeri* nematocysts compared to others, the variable concentrations of the proteins could have a major impact on the potency and variety of biological activities elicited by *C. fleckeri* nematocyst venom. Furthermore, the 39 and 41 kDa proteins were not antigenic to CSL box jellyfish antivenom or rabbit antibodies raised against *C. fleckeri* nematocyst venom, suggesting that neither antibodies could attenuate the potential biological effects of the two proteins. However, the 39 and 41 kDa proteins may not have been antigenic in this study because they were absent or not significantly abundant in the milked or nematocyst venoms that were used for antibody production. Conversely, the findings may simply imply that the two proteins are not immunogenic and would therefore require chemical modification (e.g. conjugation to an immunogenic carrier protein) before a strong immunological response is elicited in host animals. Nevertheless, the discovery of the potently haemolytic 39 and 41 kDa proteins is important because they may represent a novel class of cytolytic proteins that are produced by cubozoan jellyfish. Further biochemical and molecular studies are therefore necessary to identify and characterise the two proteins and determine their clinical relevance (if any) in human envenoming.

8.6 The Emergence of a Novel Family of Bioactive Cubozoan Proteins

Progress in the molecular characterisation of cubozoan protein and peptide toxins using proteomic and molecular cloning techniques is considerably less advanced than for toxins from other cnidarians such as sea anemones and hydrae, however in recent years, this field of research has gathered momentum. The first reported cubozoan protein toxins to be sequenced were CrTX-A and CrTX-B isolated from the carybdeid box jellyfish, *C. rastoni* (Nagai et al., 2000a). Following this break-through in cubozoan venom research, homologous protein toxins, CaTX-A and CqTX-A, were successively isolated from *C. alata* and *C. quadrigatus* (Nagai et

al., 2000b, 2002). In a parallel study, a basic 42 kDa haemolytic protein (CAH1) was also isolated from the nematocyst venom of *C. alata* venom (Chung et al., 2001). The partial N-terminal amino acid sequence of CAH1 is very similar (if not identical) to the deduced sequence reported for CaTX-A (Nagai et al., 2000b), thus suggesting that CAH1 and CaTX-A are probably the same protein. Finally, the discovery of CfTX-1 and -2 in *C. fleckeri* nematocyst venom in this study (Chapters 4 and 5) augments this newly emerging family of box jellyfish toxins, as summarised in Table 8.1.

Table 8.1
Novel family of bioactive box jellyfish proteins.

Species	Toxin Name	Exp. M_r[a] (kDa)	Calc. M_r[b] (kDa)	Calc. pI[c]	GenBank Accession No.	Reference
C. rastoni	CrTX-A	43	47.3	9.1	AB015878	Nagai et al. (2000a)
	CrTX-B	46	-	-	-	
C. alata	CaTX-A	43	49.0	8.8	AB036714	Nagai et al. (2000b)
	CaTX-B	45	-	-	-	
	CAH1	42	-	-	-	Chung et al. (2001)
C. quadrigatus	CqTX-A	44	49.6	8.6	AB045319	Nagai et al. (2002)
C. fleckeri	CfTX-1	43	49.1	8.3	EF636902	Brinkman and Burnell (2007)
	CfTX-2	45	49.9	7.8	EF636903	(see also Chapter 5)

[a] Exp. Mr is the experimental molecular mass of protein toxins determined by SDS-PAGE, expressed in kilodaltons (kDa).
[b] Calc. M_r is the calculated molecular mass of deduced amino acid sequences corresponding to the mature toxins.
[c] Calc. pI is the calculated iso-electric point of each mature toxin; Calc. M_r and pI-values were calculated using ProtParam (Gasteiger et al., 2005).

No significant amino acid sequence homology exists between the box jellyfish toxins and other known proteins, indicating that the toxins belong to a novel family of bioactive proteins (Nagai et al., 2002; Chapter 5). Characteristically, the box jellyfish toxins are labile, basic proteins (42-46 kDa) that exhibit potent haemolytic activity. Bioactivity assay data obtained for selected toxins also suggest that the proteins are lethal and cause localised pain, inflammation and dermonecrosis (Nagai et al., 2000a, b, 2002). The researchers found that CrTX-A, CaTX-A and CqTX A are lethal to crayfish, CrTX-A and -B are lethal to mice, and CrTX-A causes pain, inflammation and necrosis of the skin similar to that observed in envenomed humans (Nagai et al., 2000a, b, 2002).

A multiple sequence alignment of the five similar box jellyfish amino acid sequences is shown in Chapter 5. Although 63 amino acids are highly conserved throughout the five homologous cubozoan sequences, phylogenetic analysis (Chapter 5) suggests that the chirodropid proteins (CfTX-1, CfTX-2 and CqTX-A) have diverged during evolution from the carybdeid proteins (CrTXs and CaTX-A). Such divergence in protein sequences also implies that the structures, specificities and functions of the toxins could vary to some extent between Classes Chirodropida and Carybdeida. Furthermore, genes encoding similar proteins have not been identified in the genome of *Nematostella vectensis* (sea anemone) nor isolated from other Cnidaria, such as hydra and scyphozoans (true jellyfish), suggesting that the box jellyfish proteins may have evolved as highly specialised cubozoan toxins.

The structural features of the box jellyfish toxins have yet to be elucidated. However, several amino acid sequence analysis methods predict that the five cubozoan toxins have a similar pattern of secondary structures (Chapter 5). The consensus of prediction methods suggest that the first 300 amino acids of the toxins are dominated by α-helices, loop structures and possibly one β-strand, whereas the C-terminal regions are comprised exclusively of β-strands and loop structures. Therefore, the two regions of significantly different topology may correspond to distinct domains.

Based on secondary structure predictions for the five box jellyfish toxins, at least two common structural features have been identified that may relate to the potent cytolytic activity of the proteins. Firstly, an amphiphilic α-helix is predicted in the N-terminal region of each protein (Nagai et al., 2000a, b, 2002; Chapter 5) and secondly, a common transmembrane-spanning region (TSR1) containing 1 or 2 hydrophobic α-helices is predicted in the five cubozoan sequences, which coincides with several highly conserved amino acids (Chapter 5). In the absence of 3-D structural information, no consistent pattern(s) of highly conserved cationic residues flanking either the amphiphilic α-helices or the hydrophobic regions of TSR1 is clearly evident, as otherwise described for several types of cytolytic toxins (Kini and Evans, 1989). However, the potential involvement of α-helices suggests that the family of box jellyfish toxins may act as α-pore-forming toxins, with their pore-forming domains consisting of anything from 1 or 2 helices to a bundle of several helices (Parker and Feil, 2005). Moreover, remote protein homology analyses predicted that the cubozoan proteins share weak structural similarity to α-pore-forming insecticidal δ-endotoxins produced by the bacterium, *Bacillus thuringiensis* (Chapter 5).

8.7 Future Research

The molecular characterisation of toxic cubozoan proteins is currently in its infancy and to date only a small group of homologous box jellyfish toxins has been successfully cloned and sequenced. Undoubtedly, there are other clinically-relevant bioactive proteins contained in cubozoan nematocysts that will be identified and characterised as progress in this field of research continues.

The proteins in the newly discovered family of box jellyfish toxins are particularly interesting because they are all potent cytolysins with the potential to be lethal and cause cutaneous pain, inflammation and necrosis. The toxins are also the most abundant proteins in the nematocyst venoms and tentacle extracts from which they are isolated, suggesting that they could be the primary cause of similar harmful effects in envenomed humans. Presently, only the lethal and localised effects of CrTX-A have been established in a mammalian system (Nagai et al., 2000a). Therefore, more research is necessary to determine if the other homologous toxins elicit similar effects in mammals. Furthermore, a comparative study of the properties and expression levels of the box jellyfish toxins could help to explain the variation in biological effects caused by the venoms of different box jellyfish species, particularly between Orders Carybdeida and Chirodropida.

The mechanisms underlying the biological activities of the box jellyfish toxins are unknown, however, secondary structure predictions suggest that the family of proteins may act as α-pore-forming toxins by altering cell membrane permeability and disrupting normal transmembrane ion concentration gradients. Nevertheless, the validity of this hypothesis cannot be determined without initiating critical biochemical, pharmacological and structural studies. The molecular mechanism of interaction between the toxins and lipid membranes has yet to be studied and the specificities of toxin-receptor binding require investigation. As more sequence data becomes available for homologous toxins in other box jellyfish species, the identification of highly conserved amino acid residues and/or secondary structures/motifs that are crucial to the mechanism of pore-forming action may be refined to enable targeted cDNA mutagenesis or chemical modifications in functional (expression) studies.

Clearly, the difficulties encountered during the purification of bioactive cubozoan proteins are a major concern, as exemplified by the inconsistent results of studies on *C. fleckeri* (see Chapter 1 for a review). These problems also limit the ability of researchers to undertake functional and structural studies that are necessary to elucidate the molecular mechanisms of

isolated bioactive proteins. The continuing trend towards the use of isolated nematocysts rather than intact tentacle tissue is likely to overcome some variability in future research data. However, problems relating to the lability, non-specific binding and poor resolution of proteins during purification are more difficult to overcome and commonly result in reduced yields of pure bioactive proteins (Chapter 6). Obtaining sufficient amounts of crude nematocyst venom is also an issue, particularly for studies involving small chirodropid jellyfish (e.g. Ramasamy et al., 2005b; Winkel et al., 2005). For these reasons, recombinant expression technology may offer an alternative to the isolation of native cubozoan proteins. However, as demonstrated in Chapter 7, the expression of major cubozoan toxins in sufficient quantities for subsequent functional and structural studies has yet to be accomplished. In contrast, commendable progress has been achieved in the characterisation of toxins from other cnidarians using recombinant expression analysis. In particular, several expression studies have accelerated the elucidation of the structures, functions and molecular mechanisms of cytolytic toxins in sea anemones (see Anderluh and Maček, 2002 for a review). Therefore, with further research effort directed at overcoming poor levels of recombinant venom protein expression, similar outcomes could also be achieved for the characterisation of cubozoan toxins.

The versatility of cDNA libraries cannot be underestimated in the discovery of novel or homologous toxins from venomous organisms. In the case of box jellyfish toxins, CfTX-1 and -2, full-length cDNAs encoding the proteins were successfully obtained from a *C. fleckeri* tentacle cDNA expression library using available peptide data to selectively screen the library (Chapter 5). Although immunoscreening the cDNA expression library to detect putative venom protein clones was unsuccessful in this study (Chapter 3), random sequencing of the library may also offer an opportunity to discover new toxins produced by cubozoan jellyfish. Similar strategies have been successfully used for the detection of toxins expressed by the scyphozoan jellyfish, *Cyanea capillata* (Yang et al., 2003), and other venomous organisms such as snakes (e.g. Ching et al., 2006) and spiders (e.g. Satake et al., 2004).

Other proteomic methods currently under-utilised for the characterisation of cubozoan venom proteins include 2-D gel technology, peptide mass fingerprinting using single-stage mass spectrometry (MS), more sophisticated peptide analyses using tandem MS (MS/MS) and optional interfacing of MS with liquid chromatography (LC-MS, LC-MS/MS). In contrast, several reported studies have used a combination of these techniques to characterise the venom proteins and peptides of organisms such as sea anemones (Zaharenko et al., 2008), snakes (e.g. Kulkeaw et al., 2007), scorpions (e.g. Miyashita et al., 2007), spiders (e.g. Herzig et al., 2008)

and cone snails (Jakubowski et al., 2006). The adoption of any one of these different molecular approaches could benefit future cubozoan venom research by assisting in the identification and characterisation of new and existing families of bioactive proteins.

REFERENCES

ABC News, 2006. Jellyfish sting kills 7-year-old. Australian Broadcasting Corporation, www.abc.net.au/news/stories/2006/01/08/1543650.htm

ABC News, 2007. Boy's death prompts stinger warning. Australian Broadcasting Corporation, www.abc.net.au/news/stories/2007/11/13/2089350.htm

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25, 3389-3402.

Anderluh, G., Maček, P., 2002. Cytolytic peptide and protein toxins from sea anemones (Anthozoa: Actinaria). *Toxicon* 40, 111-124.

Assakura, M.,T., Silva, C.A., Mentele, R., Camargo, A.C.M., Serrano, S.M.T., 2003. Molecular cloning and expression of structural domains of bothropasin, a P-III metalloproteinase from the venom of *Bothrops jararaca. Toxicon* 41, 217-227.

Azuma, H., Ishikawa, M., Nakajima, T., Satoh, A., Sekizaki, S., 1986a. Calcium-dependent contractile response of arterial smooth muscle to a jellyfish toxin (pCrTX: *Carybdea rastonii*). *British Journal of Pharmacology* 88, 549-559.

Azuma, H., Sekizaki, S., Satoh, A., Nakajima, T., Ishikawa, M., 1986b. Platelet aggregation caused by a partially purified jellyfish toxin from *Carybdea rastonii*. *Toxicon* 24, 489-499.

Azuma, H., Sekizaki, S., Satoh, A., Nakajima, T., 1986c. Platelet aggregation caused by Carybdea rastonii toxins (CrTX-I, II and III) obtained from a jellyfish, *Carybdea rastonii* (42305). *Proceedings of the Society for Experimental Biology and Medicine* 182, 34-42.

Bailey, P.M., Bakker, A.J., Seymour, J.E., Wilce, J.A., 2005. A functional comparison of the venom of three Australian jellyfish – *Chironex fleckeri*, *Chiropsalmus sp.* and *Carybdea xaymacana* – on cytosolic Ca^{2+}, haemolysis and *Artemia sp.* lethality. *Toxicon* 45, 233-242.

Bailey, P.M., Little, M., Jelinek, G.A., Wilce, J.A., 2003. Jellyfish envenoming syndromes: unknown toxic mechanisms and unproven therapies. *Medical Journal of Australia* 178, 34-37.

Balashova, N.V., Crosby, J.A., Ghofaily, L.A., Kachlany, S.C., 2006. Leukotoxin confers beta-hemolytic activity to *Actinobacillus actinomycetemcomitans*. *Infection and Immunity* 74, 2015-2021.

Baneyx, F., 1999. Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology* 10, 411-421.

Barnes, J.H., 1967. Extraction of cnidarian venom from living tentacle. In: Russell, F.E., Saunders, P.R. (Eds), Animal Toxins: A collection of papers presented at the First International Symposium on Animal Toxins, Atlantic City, New Jersey, 9-11 April 1966. Pergamon Press, Oxford. pp. 115-129.

Baxter, E.H., Marr, A.G.M., 1969. Sea wasp (*Chironex fleckeri*) venom: lethal, haemolytic and dermonecrotic properties. *Toxicon* 7, 195-210.

Baxter, E.H., Marr, A.G.M., 1975. Sea wasp toxoid: an immunizing agent against the venom of the box jellyfish, *Chironex fleckeri*. *Toxicon* 13, 423-435.

Baxter, E.H., Marr, A.G.M., Lane, W.R., 1968. Immunity to the venom of the sea wasp *Chironex fleckeri. Toxicon* 6, 45-50.

Beadnell, C.E, Rider, T.A., Williamson, J.A., Fenner, P.J., 1992. Management of a major box jellyfish (*Chironex fleckeri*) sting. Lessons from the first minutes and hours. *Medical Journal of Australia* 156, 665-668.

Beck, G., Habicht, G.S., 1986. Isolation and characterization of a primitive interleukin-1-like protein from an invertebrate, *Asteria forbesi*. *Proceedings of the National Academy of Sciences of the United States of America* 83, 7429-7433.

Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* 340, 783-795.

Berney, C., Pawlowski, J., Zaninetti, L., 2000. Elongation factor 1-alpha sequences do not support an early divergence of the Aceola. *Molecular Biology and Evolution* 17, 1032-1039.

Bloom, D.A., Burnett, J.W., Alderslade, P., 1998. Partial purification of box jellyfish (*Chironex fleckeri*) nematocyst venom isolated at the beachside. *Toxicon* 36, 1075-1085.

Bloom, D.A., Burnett, J.W., Hebel, J.R., Alderslade, P., 1999. Effects of Verapamil and CSL antivenom on *Chironex fleckeri* (box-jellyfish) induced mortality. *Toxicon* 37, 1621-1626.

Bloom, D.A., Radwan, F.F., Burnett, J.W., 2001. Toxinological and immunological studies of capillary electrophoresis fractionated *Chrysaora quinquecirrha* (Desor) fishing tentacle and *Chironex fleckeri* nematocyst venoms. *Comparative Biochemistry and Physiology* 128**C**, 75-90.

Boulware, D.R., 2006. A randomized, controlled field trial for the prevention of jellyfish stings with a topical sting inhibitor. *Journal of Travel Medicine* 13, 166-167.

Boyd, W., 1984. Sea-wasp antivenom in a toddler. *Medical Journal of Australia* 140, 504.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72, 248-254.

Brinkman, D., Burnell, J., 2007. Identification, cloning and sequencing of two major venom proteins from the box jellyfish, *Chironex fleckeri*. *Toxicon* 50, 850-860.

Brinkman, D., Burnell, J., 2008. Partial purification of cytolytic venom proteins from the box jellyfish, *Chironex fleckeri*. *Toxicon* 51, 853-863.

Broedel, S.E., Papciak, S.M., Jones, W.R., 2001. The selection of optimal media formulations for improved expression of recombinant proteins in *E. coli*. Technical Bulletin 2. Athena Enzyme Systems Group. www.athenaes.com.

Burnett, J.W., 1991. Clinical manifestations of jellyfish envenomation. *Hydrobiologia* 216/217, 629-635.

Burnett, J.W., 1998. The case for use of verapamil in alarming *Chironex* stings. *Anaesthesia and Intensive Care* 26, 461-462.

Burnett, J.W., Calton, G.J., 1973. Purification of sea nettle nematocyst toxins by gel diffusion. *Toxicon* 11, 243-248.

Burnett, J.W., Calton, G.J., 1983. Response of the box-jellyfish (*Chironex fleckeri*) cardiotoxin to intravenous administration of verapamil. *Medical Journal of Australia* 2, 192-194.

Burnett, J.W., Calton, G.J., 1986. Pharmacological effects of various venoms on cutaneous capillary leakage. *Toxicon* 24, 614-617.

Burnett, J.W., Calton, G.J., 1987. Venomous pelagic coelenterates: chemistry, toxicology, immunology and treatment of their stings. *Toxicon* 25, 581-602.

Burnett, J.W., Calton, G.J., 2004. The case for use of verapamil in alarming *Chironex* stings remains. *Toxicon* 44, 817-818.

Burnett, J.W., Currie, B., Fenner, P., Rifkin, J., Williamson, J., 1996a. Cubozoans (‘Box Jellyfish’). In: Williamson, J.A., Fenner, P.J., Burnett, J.W., Rifkin, J.F. (Eds), Venomous and Poisonous Marine Animals: A Medical and Biological Handbook. Surf Life Saving Queensland Incorporated and University of New South Wales Press, Sydney, Australia, pp. 236-283.

Burnett, J.W., Othman, I.B., Endean, R., Fenner. P.J., Callanan, V.I., Williamson, J.A., 1990. Verapamil potentiation of *Chironex* (box-jellyfish) antivenom. *Toxicon* 28, 242-244.

Burnett, J.W., Rifkin, J., Williamson, J., 1996b. Cnidarian envenomation. In: Williamson, J.A., Fenner, P.J., Burnett, J.W., Rifkin, J.F. (Eds), Venomous and Poisonous Marine Animals: A

Medical and Biological Handbook. Surf Life Saving Queensland Incorporated and University of New South Wales Press, Sydney, Australia, pp. 174-179.

Calton, G.J., Burnett, J.W., 1986. Partial purification of *Chironex fleckeri* (sea wasp) venom by immunochromatography with antivenom. *Toxicon* 24, 416-420.

Carrette, T.J., 2002. An investigation into the extraction and ecology if venom in two species of cubozoan, *Chironex fleckeri* and *Chiropsalmus* sp. Thesis submitted for the research degree of Master of Science, James Cook University, Townsville, Australia.

Carrette, T., Alderslade, P., Seymour, J., 2002a. Nematocyst ratio and prey in two Australian cubomedusans, *Chironex fleckeri* and *Chiropsalmus* sp. *Toxicon* 40, 1547-1551.

Carrette, T., Cullen, P., Little, M., Peiera, P.L., Seymour, J.E., 2002b. Temperature effects on box jellyfish venom: a possible treatment for envenomated patients? *Medical Journal of Australia* 177, 654-655.

Carrette, T., Seymour, J., 2004. A rapid and repeatable method for venom extraction from Cubozoan nematocysts. *Toxicon* 44, 135-139.

Carrette, T., Seymour, J., 2005. Vascular dopplers: a new way of recording cardiac parameters in envenomed organisms. *Toxicon* 45, 541-544.

Ceroni, A., Frasconi, P., Passerini, A., Vullo, A., 2006. DISULFIND: a disulfide bonding state and cysteine connectivity prediction server, *Nucleic Acids Research* 34 (Web Server issue), W177-W181.

Ching, A.T.C., Rocha, M.M.T., Paes Leme, A.F., Pimenta, D.C., Furtado, M. de F.D., Serrano, S.M.T., Ho, P.L., Junqueira-de-Azevedo, I.L.M., 2006. Some aspects of the venom proteome of the Colubridae snake *Philodryas olfersii* revealed from a Duvernoy's (venom) gland transcriptome. *FEBS Letters* 580, 4417-4422.

Cho, H.M., Jun, D.Y., Bae, M.A., Ahn, J.D., Kim, Y.H., 2000. Nucleotide sequence and differential expression of the human 3-phosphoglycerate dehydrogenase gene. *Gene* 245, 193-201.

Chung, J.J., Ratnapala, L.A., Cooke, I.M., Yanagihara, A.A., 2001. Partial purification and characterization of a haemolysin (CAH1) from Hawaiian box jellyfish (*Carybdea alata*) venom. *Toxicon* 39, 981-990.

Claros, M.G., von Heijne, G., 1994. TopPred II: an improved software for membrane protein structure predictions. *Computer Applications in the Biosciences* 10, 685-686.

Clauser, K.R., Baker, P.R., Burlingame, A.L., 1999. Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Analytical Chemistry* 71, 2871-2882.

Collins, S.P., Comis, A., Marshall, M., Hartwick, R.F., Howden, M.E.H., 1993. Monoclonal antibodies neutralising the haemolytic activity of box-jellyfish (*Chironex fleckeri*) tentacle extracts. *Comparative Biochemistry and Physiology* 106B, 67-70.

Collins, A.G., Schuchert, P., Marques, A.C., Jankowski, T., Medina, M., Schierwater, B., 2006. Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models. *Systematic Biology* 55, 97-115.

Comis, A., Hartwick, R.F., Howden, E.H., 1989. Stabilization of lethal and hemolytic activities of box jellyfish *(Chironex fleckeri*) venom. *Toxicon* 27, 439-447.

Cooper, L.Z., Madoff, M.A., Weinstein, L., 1964. Hemolysis of rabbit erythrocytes by purified staphylococcal alpha-toxin. *Journal of Bacteriology* 87, 127-135.

Coughlan, J.P., Seymour, J., Cross, T.F., 2006. Isolation and characterisation of seven polymorphic microsatellite loci in the box jellyfish (*Chironex fleckeri*, Cubozoa, Cnidaria). *Molecular Ecology Notes* 6, 41-43.

Cribb, A.B., Cribb, J.W., 1981. In: Wild Medicine in Australia. William Collins Pty Ltd, Sydney. pp. 228.

Crone, H.D., 1976. Chemical modification of the haemolytic activity of extracts from the box jellyfish *Chironex fleckeri. Toxicon* 14, 97-107.

Crone, H.D., Keen, T.E.B., 1969. Chromatographic properties of the hemolysin from the cnidarian *Chironex fleckeri*. *Toxicon* 7, 79-87.

Crone, H.D., Keen, T.E.B., 1971. Further studies on the biochemistry of the toxins from the sea wasp *Chironex fleckeri*. *Toxicon* 9, 145-151.

Cuff, J.A., Barton, G.J., 1999. Evaluation and improvement of multiple sequence methods for protein secondary structure prediction. *Proteins* 34, 508-519.

Currie, B., 1994. Clinical implications of research on the box-jellyfish *Chironex fleckeri*, *Toxicon* 32, 1305-1313.

Currie, B.J., Wood, Y.K., 1995. Identification of *Chironex fleckeri* envenomation by nematocyst recovery from skin. *Medical Journal of Australia* 162, 478-480.

Cuypers, E., Yanagihara, A., Karlsson, E., Tytgat, J., 2006. Jellyfish and other cnidarian envenomations cause pain affecting TRPV1 channels. *FEBS Letters* 580, 5728-5732.

da Silveira, R.B., Wille, A.C.M., Chaim, O.M., Appel, M.H., Silva, D.T., Franco, C.R.C., Toma, L., Mangili, O.C., Gremski, W., Dietrich, C.P., Nada, H.B., Veiga, S.S., 2007. Identification, cloning, expression and functional characterisation of an astacin-like metalloprotease toxin from *Loxosceles intermedia* (brown spider) venom. *The Biochemical Journal* 406, 355-363.

de Hostos, E.L., 1999. The corinin family of actin-associated proteins. *Trends in Cell Biology* 9, 345-350.

Dufton, M.J., Hider, R.C., 1988. Structure and pharmacology of elapid cytotoxins. *Pharmacology and Therapeutics* 36, 1-40.

Endean, R., 1987. Separation of two myotoxins from nematocysts of the box jellyfish (*Chironex fleckeri*). *Toxicon* 25, 483-492.

Endean, R., 1988a. How *Chironex* injects its venom. In: Pearn, J., Covacevich, J. (Eds), Venoms and Victims. The Queensland Museum and Amphion Press, Brisbane, Australia. pp. 9-14.

Endean, R., 1988b. Venom of *Chironex*, the world's most venomous animal. In: Pearn, J., Covacevich, J. (Eds), Venoms and Victims. The Queensland Museum and Amphion Press, Brisbane, Australia. pp. 15-24.

Endean, R., Duchemin, C., McColm, D., Fraser, E.H., 1969. A study of the biological activity of toxic material derived from nematocysts of the Cubomedusan *Chironex fleckeri*. *Toxicon* 6, 179-204.

Endean, R., Henderson, L., 1969. Further studies of toxic material from nematocysts of the cubomedusan *Chironex fleckeri* Southcott. *Toxicon* 7, 303-314.

Endean, R., Henderson, L., 1974. Some aspects of the biological activity of crude nematocyst toxin from *Chironex fleckeri.* In: Humm, H.J., Lane, C.E. (Eds), Bioactive Compounds from the Sea. Marcel Dekker, New York. pp. 1-13.

Endean, R., Monks, S.A., Cameron, A.M., 1993. Toxins from the box-jellyfish *Chironex fleckeri*. *Toxicon* 31, 397-410.

Endean, R., Noble, M., 1971. Toxic material from the tentacles of the cubomedusan *Chironex fleckeri*. *Toxicon* 9, 255-264.

Endean, R., Rifkin, J., 1975. Isolation of the different types of nematocyst from the cubomedusan *Chironex fleckeri*. *Toxicon* 13, 375-376.

Endean, R., Rifkin, J.F., Daddow, L.Y.M., 1991. Envenomation by the box-jellyfish *Chironex fleckeri*: how nematocysts discharge. *Hydrobiologia* 216/217, 641-648.

Endean, R., Sizemore, D.J., 1987. The use of verapamil to counter the effects of myotoxins from nematocysts of the box jellyfish *Chironex fleckeri*. In: Gopalakrishnakone, P., Tan, C.K. (Eds), Progress in Venom and Toxin Research. Pergamon Press, Oxford, pp. 499-507.

Endean, R., Sizemore, D.J., 1988. The effectiveness of antivenom in countering the actions of box jellyfish (*Chironex fleckeri*) nematocyst toxins in mice. *Toxicon* 26, 425-431.

Engel, U., Özbek, S., Streitwolf-Engel, R., Petri, B., Lottspeich, F., Holstein, T.W., 2002. Nowa, a novel protein with minicollagen Cys-rich domains, is involved in nematocyst formation in *Hydra*. *Journal of Cell Science* 115, 3923-3934.

Evans, R.J., Bray, J., Childs, J.D., Vigers, G.P.A., Brandhuber, B.J., Skalicky, J.J., Thompson, R.C., Eisenberg, S.P., 1995. Mapping receptor binding sites in interleukin (IL)-1 receptor antagonist and IL-1β by site-directed mutagenesis. *The Journal of Biological Chemistry* 270, 11477-11483.

Fang, K.S.Y., Vitale, M., Fehlner, P., King, T.P., 1988. cDNA cloning and primary structure of a white-face hornet allergen, antigen 5. *Proceedings of the National Academy of Sciences of the United States of America* 85, 895-899.

Fenner, P.J., 1991. Cubozoan jellyfish envenomation syndromes and their medical treatment in northern Australia. *Hydrobiologia* 216/217, 637-640.

Fenner, P.J., Harrison, S. L., 2000. Irukandji and *Chironex fleckeri* envenomation in tropical Australia. *Wilderness and Environmental Medicine* 11, 233-240.

Fenner, P.J., Williamson, J.A., 1996. Worldwide deaths and severe envenomation from jellyfish stings. *Medical Journal of Australia* 165, 658-661.

Fenner, P.J., Williamson, J.A., Burnett, J.W, Rifkin, J., 1993. First aid treatment of jellyfish stings in Australia: Response to a newly differentiated species. *Medical Journal of Australia* 158, 498-501.

Freeman, S.E., Turner, R.J., 1969. A pharmacological study of the toxin of a Cnidarian *Chironex fleckeri*. *British Journal of Pharmacology* 35, 510-522.

Freeman, S.E., Turner, R.J., 1971. Cardiovascular effects of toxins isolated from the cnidarian *Chironex fleckeri* Southcott. *British Journal of Pharmacology* 41, 154-166.

Freeman, S.E., 1974. Actions of *Chironex fleckeri* toxins on cardiac transmembrane potentials. *Toxicon* 12, 395-404.

Galitsky, N., Cody, V., Wojtczak, A., Ghosh, D., Luft, J.R., Pangborn, W., English, L., 2001. Structure of the insecticidal bacterial δ-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. *Acta Crystallographica* D57, 1101-1109.

Garnier, P., Goudey-Perrière, F., Breton, P., Dewulf, C., Petek, F., Perrière, C., 1995. Enzymatic properties of the stonefish (*Synanceia verrucosa* Bloch and Schneider, 1801) venom and purification of a lethal, hypotensive and cytolytic factor. *Toxicon* 33, 143-155.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch A., 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker, J.M. (Ed.), The Proteomics Protocols Handbook. Humana Press, Totowa, USA, pp. 571-607.

Gaucher, E.A., Miyamoto, M.M., Benner, S.A., 2001. Function-structure analysis of proteins using covarion-based evolutionary approaches: elongation factors. *PNAS* 98, 548-552.

Geer, L.Y., Domrachev, M., Lipman, D.J., Bryant, S.H., 2002. CDART: protein homology by domain architecture. *Genome Research* 12, 1619-1623.

Gershwin, L., 2005. *Carybdea alata* auct. and *Manokia stiasnyi*, reclassification to a new family with description of a new genus and two new species. *Memoirs of the Queensland Museum* 51, 501-523.

Gershwin, L., 2006. Comments of *Chiropsalmus* (Cnidaria: Cubozoa: Chirodropida): a preliminary revision of the Chiropsalmidae, with descriptions of two genera and two new species. *Zootaxa* 1231, 1-42.

Ghadessy, F.J., Chen, D., Kini, M., Chung, M.C.M., Jeyaseelan, K., Khoo, H.E., Yuen, R., 1996. Stonustoxin is a novel lethal factor from stonefish (*Synanceja horrida*) venom. *The Journal of Biological Chemistry* 271, 25575-25581.

Grochulski, P., Masson, L., Borisova, S., Pusztai-Carey, M., Schwartz, J-L., Brousseau, R., Cygler, M., 1995. *Bacillus thuringiensis* CryIA(a) insectical toxin: crystal structure and channel formation. *Journal of Molecular Biology* 254, 447-464.

Gross, S.R., Kinzy, T.G., 2005. Translation elongation factor 1A is essential for regulation of the actin cytoskeleton and cell morphology. *Nature Structural and Molecular Biology* 12, 772-778.

Gupta, R.S., Golding, G.B., 1993. Evolution of HSP70 gene and its implications regarding relationships between archeabacteria, eubacteria, and eukaryotes. *Journal of Molecular Evolution* 37, 573-582.

Gusmani, L., Avian, M., Galil, B., Patriarca, P., Rottini, G., 1997. Biologically active polypeptides in the venom of the jellyfish *Rhopilema nomadica. Toxicon* 35, 637-648.

Hannappel, E., 2007. β-Thymosins. *Annals of the New York Academy of Sciences* 1112, 21-37.

Harry, J.L., Wilkins, M.R., Herbert, B.R., Packer, N.H., Gooley, A.A., Williams, K.L., 2000. Proteomics: capacity versus utility. *Electrophoresis* 21, 1071-1081.

Hartwick, R., 1987. The box-jellyfish. In: Covacevich, J., Davie, P., Pearn, J. (Eds), Toxic Plants and Animals: A Guide For Australia. The Queensland Museum, Brisbane, Australia, pp. 99-105.

Hartwick, R.F., 1991. Distributional ecology and behaviour of the early life stages of the box-jellyfish *Chironex fleckeri*. *Hydrobiologia* 216/217, 181-188.

Hartwick, R., Callanan, V., Williamson, J., 1980. Disarming the box jellyfish: nematocyst inhibition in *Chironex fleckeri*. *Medical Journal of Australia* 1, 15-20.

Hawkes, R., 1982. Identification of concanavalin A-binding proteins after sodium dodecyl sulfate-gel electrophoresis and protein blotting. *Analytical Biochemistry* 123, 143-146.

Hellstern, S., Stetefeld, J., Fauser, C., Lustig, A., Engel, J., Holstein, T.W., Özbek, S., 2006. Structure/function analysis of spinalin, a spine protein of *Hydra* nematocysts. *FEBS Journal* 273, 3230-3237.

Herrmann, D., Hatta, M., Ullerich-Hoffmeister, S.A.H., 2005. Thypedin, the multicopy precursor for the hydra peptide pedin, is a β-thymosin repeat-like domain containing protein. *Mechanisms of Development* 122, 1183-1193.

Herzig, V., Khalife, A.A., Chong, Y., Isbister, G.K., Currie, B.J., Churchill, T.B., Horner, S., Escoubas, P., Nicholson, G.M., Hodgson, W.C., 2008. Intersexual variations in Northern (*Missulena pruinosa*) and Eastern (*M. bradleyi*) mouse spider venom. *Toxicon* 51, 1167-1177.

Herzig, V., Ward, R.J., Ferreira dos Santos, W., 2004. Ontogenetic changes in *Phoneutria nigriventer* (Araneae, Ctenidae) spider venom. *Toxicon* 44, 635-640.

Hessinger, D.A., 1988. Nematocyst venoms and toxins. In: Hessinger, D.A., Lenhoff, H.M. (Eds), *The Biology of Nematocysts*. Academic Press Incorporated, New York. pp. 333-368.

Hoffmeister, S.A.H., 1996. Isolation and characterization of two new morphogenetically active peptides from *Hydra vulgaris. Development* 122, 1941-1948.

Hofmann, K., Stoffel, W., 1993. TMbase – a database of membrane spanning protein segments. *Biological Chemistry Hoppe-Seyler* 374, 166-170.

Holstein, T.W., Benoit, M., van Herder, G., Wanner, G., David, C.N., Gaub, H.E., 1994. Fibrous mini-collagens in *Hydra* nematocysts. *Science* 265, 402-404.

Horton, P., Nakai, K., 1997. Better prediction of protein cellular localization sites with the k nearest neighbors classifier. In: Proceedings of the Fifth International Conference on Intelligence Systems for Molecular Biology. AAAI Press, Menlo Park, CA, pp. 147-152.

Hunt, I., 2005. From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Protein Expression and Purification* 40, 1-22.

Jakubowski, J.A., Kelley, W.P., Sweedler, J.V., 2006. Screening for post-translational modifications in conotoxins using liquid chromatography/mass spectrometry: an important component of conotoxin discovery. *Toxicon* 47, 688-699.

Jana, S., Deb, J.K., 2005. Strategies for efficient production of heterologous proteins in *Escherichia coli. Applied Microbiology and Biotechnology* 67, 289-298.

Jones, D.T., 1999. Protein secondary structure prediction based on position-specific scoring matrices. Journal of Molecular Biology 292, 195-202.

Jones, D., Sawicki, G., Wozniak, M., 1992. Sequence, structure, and expression of a wasp venom protein with a negatively charged signal peptide and a novel repeating internal structure. *The Journal of Biological Chemistry* 267, 14871-14878.

Karplus, K., Barrett, C., Hughey, R., 1998. Hidden Markov models for detecting remote protein homologies. *Bioinformatics* 14, 846-856.

Karplus, K., Katzman, S., Shackleford, G., Koeva, M., Draper, J., Barnes, B., Soriano, M., Hughey, R., 2005. SAM-T04: what's new in protein-structure prediction for CASP6. *Proteins: Structure, Function, and Bioinformatics* 61(S7), 135-142.

Kass-Simon, G., Scappaticci, A.A., 2002. The behavioral and developmental physiology of nematocysts. *Canadian Journal of Zoology* 80, 1772-1794.

Keen, T.E.B., Crone, H.D., 1969a. Haemolytic properties of extracts of tentacles from the cnidarian *Chironex fleckeri*. *Toxicon* 7, 55-63.

Keen, T.E.B., Crone, H.D., 1969b. Dermonecrotic properties of extracts of tentacles from the cnidarian *Chironex fleckeri*. *Toxicon* 7, 173-180.

Kim, E., Lee, S., Kim, J.-S., Yoon, W.D., Lim, D., Hart, A.J., Hodgson, W.C., 2006. Cardiovascular effects of *Nemopilema nomurai* (Schyphozoa: Rhizostomeae) jellyfish venom in rats. *Toxicology Letters* 167, 205-211.

Kimball, A.B., Arambula, K.Z., Stauffer, A.R., Levy, V., Davis, V.W., Liu, M., Rehmus, W.E., Lotan, A., Auerbach, P.S., 2004. Efficacy of a jellyfish sting inhibitor in preventing jellyfish stings in normal volunteers. *Wilderness and Environmental Medicine* 15, 102-108.

Kini, R.M., Evans, H.J., 1989. A common cytolytic region in myotoxins, haemolysins, cardiotoxins and antibacterial peptides. *International Journal of Peptide and Protein Research* 34, 277-286.

Kinsey, B.E., 1986. In: Barnes on box jellyfish. James Cook University of North Queensland, Australia, pp. 29.

Kintner, A.H., Seymour, J.E., Edwards, S.L., 2005. Variation in lethality of two Australian chirodropid jellyfish venoms in fish. *Toxicon* 46, 699-708.

Kobayashi, M., Wada, H., Satoh, N., 1996. Early evolution of the metozoa and phylogenetic status as inferred from amino acid sequence of elongation factor-1α. *Molecular Phylogenetics and Evolution* 5, 414-422.

Koyama, J., 1965. Kinetic study on streptolysin S'. *The Journal of Biochemistry* 57, 103-108.

Koyama, T., Noguchi, K., Matsuzaki, T., Sakanashi, M., Nakasone, J., Sakanashi, M., Sakanashi, M., 2003. Haemodynamic effects of the crude venom from nematocysts of the box jellyfish *Chiropsalmus quadrigatus* (Habu-kurage) in anaesthetised rabbits. *Toxicon* 41, 621-631.

Kulkeaw, K., Chaicumpa, W., Sakolvaree, Y., Tongtawe, P., Tapchaisri, P., 2007. Proteome and immunome of the venom of the Thai cobra, *Naja kaouthia. Toxicon* 49, 1026-1041.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680-685.

Li, J., Carroll, J., Ellar, D.J., 1991. Crystal structure of insecticidal δ-endotoxin from *Bacillus thuringiensis*, at 2.5Å resolution. *Nature* 353, 815-821.

Liu, G., Grant, W.M., Persky, D., Latham, V.M., Jr, Singer, R.H., Condeelis, J., 2002. Interactions of elongation factor 1α with F-actin and β-actin mRNA: implications for anchoring mRNA in cell protrusions. *Molecular Biology of the Cell* 13, 579-592.

Lotan, A., Fishman, L., Zlotkin, E., 1996. Toxin compartmentation and delivery in the Cnidaria: the nematocyst's tubule as a multi-headed poisonous arrow. *The Journal of Experimental Zoology* 275, 444-451.

Lumley, J., Williamson, J.A., Fenner, P.J., Burnett, J.W., Colquhoun, D.M., 1988. Fatal envenomation by *Chironex fleckeri*, the north Australian box jellyfish: the continuing search for lethal mechanisms. *Medical Journal of Australia* 148, 527-533.

Lupas, A., Van Dyke, M., Stock, J., 1991. Predicting Coiled Coils from Protein Sequences, *Science* 252, 1162-1164.

Mackessy, S.P., Sixberry, N.M., Heybourne, W.H., Fritts, T., 2006. Venom of the brown treesnake, *Boiga irregularis*: ontogenetic shifts and taxa-specific toxicity. *Toxicon* 47, 537-548.

Madera, M., Vogel, C., Kummerfield, S.K., Chothia, C., Gough, J., 2004. The SUPERFAMILY database in 2004: additions and improvements. *Nucleic Acids Research* 32, D235-239.

Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales, N.R., Gwadz, M., Hao, L., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Krylov, D., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Thanki, N., Yamashita, R.A., Yin, J.J., Zhang, D., Bryant, S.H, 2007. CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Research* D35, 237-240.

Marchler-Bauer, A., Bryant, S.H., 2004 CD-Search: protein domain annotations on the fly. *Nucleic Acids Research* W32, 327-331.

Marsh, L.M., Slack-Smith, S.M., 1986. In: Sea Stingers, Western Australian Museum, Perth Australia, pp. 133.

Mastro, R., Hall, M., Llewellyn, L., 1998. Tackling proteinaceous venom with 2-dimensional electrophoresis. Poster presented at From Venoms to Drugs Conference, Heron Island, Australia, July 1998.

Mayer, M.P., Bukau, B., 2005. Hsp70 chaperones: cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences* 62, 670-684.

Meier, S., Jensen, P.R., Adamczyk, P., Bächinger, H.P., Holstein, T.W., Engel, J., Özbek, S., Grzesiek, S., 2007. Sequence-structure and structure-function analysis in cysteine-rich domains forming the ultrastable nematocyst wall. *Journal of Molecular Biology* 368, 718-728.

Miyashita, M., Otsuki, J., Hanai, Y., Nakagawa, Y., Miyagawa, H., 2007. Characterisation of peptide components in the venom of the scorpion *Liocheles australasiae* (Hemiscorpiidae). *Toxicon* 50, 428-437.

Moore, J.T., Uppal, A., Maley, F., Maley, G.F., 1993. Overcoming inclusion body formation in a high-level expression system. *Protein Expression and Purification* 4, 160-163.

Moore, R.C., Cyr, R.J., 2000. Association between elongation factor-1α and microtubules in vivo is domain dependent and conditional. *Cell Motility and the Cytoskeleton* 45, 279-292.

Moreira, D., Le Guyader, H., Phillipe, H., 1999. Unusually high evolutionary rate of the elongation factor 1α genes from the Ciliophora and its impact on the phylogeny of eukaryotes. *Molecular Biology and Evolution* 16, 234-245.

Mustafa, M.R., White, E., Hongo, K., Othman, I., Orchard, C.H., 1995. The mechanism underlying the cardiotoxic effect of the toxin from the jellyfish *Chironex fleckeri*. *Toxicology and Applied Pharmacology* 133, 196-206.

Nakai, K., Kanehisa, M., 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* 14, 897-911.

Nagai, H., 2003. Recent progress in jellyfish toxin study. *Journal of Health Science* 49, 337-340.

Nagai, H., Takuwa, K., Nakao, M., Ito, E., Miyake, M., Noda, M., Nakajima, T., 2000a. Novel proteinaceous toxins from the box jellyfish (sea wasp) *Carybdea rastoni*. *Biochemical and Biophysical Research Communications* 275, 582-588.

Nagai, H., Takuwa, K., Nakao, M., Sakamoto, B., Crow, G.L., Nakajima, T., 2000b. Isolation and characterization of a novel protein from the Hawaiian box jellyfish (sea wasp) *Carybdea alata*. *Biochemical and Biophysical Research Communications* 275, 589-594.

Nagai, H., Takuwa-Kuroda, K., Nakao, M., Oshiro, N., Iwanaga, S., Nakajima, T., 2002. A novel protein toxin from the deadly box jellyfish (sea wasp, habu-kurage) *Chiropsalmus quadrigatus*. *Bioscience, Biotechnology and Biochemistry* 66, 97-102.

Naguib, A.M.F., Bansal, J., Calton, G.J., Burnett, J.W., 1988. Purification of *Chironex fleckeri* venom components using *Chironex* immunoaffinity chromatography. *Toxicon* 26, 387-394.

Nevalainen, T.J., Peuravuori, H.J., Quinn, R.J., Llewellyn, L.E., Benzie, J.A.H., Fenner, P.J., Winkel, K.D., 2004. Phospholipase A2 in Cnidaria. *Comparative Biochemistry and Physiology B* 139, 731-735.

Noguchi, K., Sakanashi, M., Matsuzaki, T., Nakasone, J., Sakanashi, M., Koyama, T., Hamadate, N., Sakanashi, M., 2005. Cardiovascular effects and lethality of venom from nematocysts of the box-jellyfish *Chiropsalmus quadrigatus*. *Toxicon* 45, 519-526.

Nüchter, T., Benoit, M., Engel, U., Özbek, S., Holstein, T.W., 2006. Nanosecond-scale kinetics of nematocyst discharge. *Current Biology* 16, R316-318.

Olson, C.E., Pockl, E.E., Calton, G.J., Burnett, J.W., 1984. Immunochromatographic purification of a nematocyst toxin from the cnidarian *Chironex fleckeri* (sea wasp). *Toxicon* 22, 733-742.

O'Reilly, G.M., Isbister, G.K., Lawrie, P.M., Treston, G.T., Currie, B.J., 2001. Prospective study of jellyfish stings from tropical Australia, including the major box jellyfish, *Chironex fleckeri*. *Medical Journal of Australia* 175, 652-655.

Othman, I., Burnett, J.W., 1990. Techniques applicable for purifying *Chironex fleckeri* (box-jellyfish) venom. *Toxicon* 28, 821-835.

Othman, I., Eldila, M.T., Mustaffa, M.R., Musa, M.Y., Norazila, 1996. Studies on the jellyfish *Carybdea rastoni*. *Toxicon* 34, 167-168.

Pardo-López, L., Gómez, I., Rausell, C., Sánchez, J., Soberón, M., Bravo, A., 2006. Structural changes of the Cry1Ac oligomeric pre-pore from *Bacillus thuringiensis* induced by *N*-acetylgalactosamine facilitates toxin membrane insertion. *Biochemistry* 45, 10329-10336.

Parker, M.W., Feil, S.C., 2005. Pore-forming protein toxins: from structure to function. *Progress in Biophysics and Molecular Biology* 88, 91-142.

Pelham H., 1988. Heat-shock proteins. Coming in from the cold. *Natur*e 332 776-777.

Pestarino, M., De Anna, E., Masini, M.A., Sturla, M., 1997. Localization of interleukin-1β mRNA in the cerebral ganglion of the protochordate, *Styela plicata*. *Neuroscience Letters* 222, 151-154.

Peti, W., Page, R., 2007. Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expression and Purification* 51, 1-10.

Pineda, M., Vilaseca, M.A., Artuch, R., Santos, S., García-González, M.M., Sau, I., Aracil, A., van Schaftingen, E., Jaekin, J., 2000. 3-phosphoglycerate dehydrogenase deficiency in a patient with West syndrome. *Developmental Medicine and Child Neurology* 42, 629-633.

Pi, C., Liu, J., Wang, L., Jiang, X., Liu, Y., Peng, C., Chen, S., Xu, A., 2007. Soluble expression, purification and functional identification of a disulfide-rich conotoxin derived from *Conus litteratus*. *Journal of Biotechnology* 128, 184–193.

Ramasamy, S., Isbister, G.K., Seymour, J.E., Hodgson, W.C., 2003. The *in vitro* effects of two chirodropid (*Chironex fleckeri* and *Chiropsalmus sp.*) venoms: efficacy of box jellyfish antivenom. *Toxicon* 41, 703-711.

Ramasamy, S., Isbister, G.K., Seymour, J.E., Hodgson, W.C., 2004. The *in vivo* cardiovascular effects of box jellyfish *Chironex fleckeri* venom in rats: efficacy of pre-treatment with antivenom, verapamil and magnesium sulphate. *Toxicon* 43, 685-690.

Ramasamy, S., Isbister, G.K., Seymour, J.E., Hodgson, W.C., 2005a. The *in vivo* cardiovascular effects of an Australasian box jellyfish (*Chiropsalmus sp.*) venom in rats. *Toxicon* 45, 321-327.

Ramasamy, S., Isbister, G.K., Seymour, J.E., Hodgson, W.C., 2005b. The *in vivo* cardiovascular effects of the Irukandji jellyfish (*Carukia barnesi*) nematocyst venom and a tentacle extract in rats. *Toxicology Letters* 155, 135-141.

Rang, H.P., Dale, M.M., Ritter, J.M., 1999. In: Pharmacology. Churchill Livingstone, Edinburgh, 4th ed, pp. 830.

Rausell, C., García-Robles, I., Sánchez, J., Muñoz-Garay, C. Martínez-Ramírez, A.C., Real, M.D., Bravo, A., 2004. Role of toxin activation on binding and pore formation activity of the *Bacillus thuringiensis* Cry3 toxins in membranes of *Leptinotarsa decemlineata* (Say). *Biochimica et Biophysica Acta* 1660, 99-105.

Renner, T., Waters, E.R., 2007. Comparative genomic analysis of the Hsp70s from five diverse photosynthetic eukaryotes. *Cell Stress and Chaperones* 12, 172-185.

Rifkin, J., 1996a. Animal taxonomy. In: Williamson, J.A., Fenner, P.J., Burnett, J.W., Rifkin, J.F. (Eds), Venomous and Poisonous Marine Animals: A Medical and Biological Handbook. Surf Life Saving Queensland Incorporated and University of New South Wales Press, Sydney, Australia, pp. 58-62.

Rifkin, J., 1996b. Jellyfish mechanisms. In: Williamson, J.A., Fenner, P.J., Burnett, J.W., Rifkin, J.F. (Eds), Venomous and Poisonous Marine Animals: A Medical and Biological Handbook. Surf Life Saving Queensland Incorporated and University of New South Wales Press, Sydney, Australia, pp. 121-143.

Rifkin, J., Endean, R., 1983. The structure and function of the nematocysts of *Chironex fleckeri*. *Cell and Tissue Research* 233, 563-577.

Rifkin, J., Williamson, J., Fenner, P., 1996. Anthozoans, hydrozoans and scyphozoans. In: Williamson, J.A., Fenner, P.J., Burnett, J.W., Rifkin, J.F. (Eds), Venomous and Poisonous Marine Animals: A Medical and Biological Handbook. Surf Life Saving Queensland Incorporated and University of New South Wales Press, Sydney, Australia, pp. 180-235.

Rost, B., Yachdav, G., Liu, J., 2004. The PredictProtein Server. *Nucleic Acids Research* 32, W321-W32.

Rost, B., Fariselli, P., Casadio, R., 1996. Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Science* 5, 1704-1718.

Rottini, G., Gusmani, L., Parovel, E., Avian, M., Patriarca, P., 1995. Purification and properties of a cytolytic toxin in venom of the jellyfish *Carybdea marsupialis*. *Toxicon* 33, 315-326.

Rybakin, V., Clemen, C.S., 2005. Coronin proteins as multifunctional regulators of the cytoskeleton and membrane trafficking. *BioEssays* 27, 625-632.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, 2nd Edn, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sánchez-Rodríguez, J., Torrens, E., Segura-Puertas, L., 2006. Partial purification and characterisation of a novel neurotoxin and three cytolysins from the box jellyfish (*Carybdea marsupialis*) nematocyst venom. *Archives of Toxicology* 80, 163-168.

Satake, H., Villegas, E., Oshiro, N., Terada, K., Shinada, T., Corzo, G., 2004. Rapid and efficient identification of cysteines-rich peptides by random screening of a venom gland cDNA library from the hexathelid spider *Macrothele gigas*. *Toxicon* 44, 149-156.

Schaefer, H., Chamrad, D.C., Herrmann, M., Stuwe, J., Becker, G., Klose, J., Blueggel, M., Meyer, H.E., Marcus, K., 2006. Study of posttranslational modifications in lenticular αA-Crystallin of mice using proteomic analysis techniques. *Biochimica et Biophysica Acta* 1764, 1948-1962.

Seymour, J.E., Carrette, T.J., Sutherland, P.A., 2004. Do box jellyfish sleep at night? *Medical Journal of Australia* 181, 707.

Shorten, M., Davenport, J., Seymour, J.E., Cross, M.C., Carrette, T.J., Woodward, G., Cross, T.F., 2005. Kinematic analysis of swimming in Australian box jellyfish, *Chiropsalmus sp.* and *Chironex fleckeri* (Cubozoa, Cnidaria: Chirodropidae). *Journal of Zoology* 267, 371-380.

Singhamatr, P., Rojnuckarin, P., 2007. Molecular cloning of albolatin, a novel snake venom metalloprotease from green pit viper (*Trimeresurus albolabris*), and expression of its disintegrin domain. *Toxicon* 50, 1192-1200.

Söding, J., Biegert, A., Lupas, A.N., 2005. The HHPred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* 33, W244-248.

Stefano, G.B., Smith, E.M., Hughes, T.K., 1991. Opioid induction of immunoreactive interleukin-1 in *Mytilus edulis* and human immunocytes: an interleukin-1-like substance in invertebrate neural tissue. *Journal of Neuroimmunology* 32, 29-34.

Stevens, R.C., 2000. Design of high-throughput methods of protein production for structural biology. *Structure* 8, R177-185.

Swartz, J.R., 2001. Advances in *Escherichia coli* production of therapeutic proteins. *Current Opinion in Biotechnology* 12, 195-201.

Szczepanek, S. Cikala, M., David, C.N., 2002. Poly-γ-glutamate synthesis during formation of nematocyst capsules in *Hydra*. *Journal of Cell Science* 115, 745-751.

Tamkun, M.M., Hessinger, D.A., 1981. Isolation and partial characterisation of a haemolytic and toxic protein from the nematocyst venom of the Portuguese Man-of-War, *Physalia physalis. Biochimica et Biophysica Acta* 667, 87-98.

Thompson J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22, 4673-4680.

Tibballs, J., 2006. Australian venomous jellyfish, envenomation syndromes, toxins and therapy. *Toxicon* 48, 830-859.

Tibballs, J., Williams, D., Sutherland, S.K., 1998. The effects of antivenom and verapamil on the haemodynamic actions of *Chironex fleckeri* (box jellyfish) venom. Anaesthesia and Intensive Care 26, 40-45.

Trapane, T.L., Lattman, E.E., 2007. Seventh meeting on the critical assessment of techniques for protein structure prediction. *Proteins: Structure, Function, and Bioinformatics* 69(S8), 1-207.

Tusnády, G.E., Simon, I., 1998. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *Journal of Molecular Biology* 283, 489-506.

Underwood, A.H., Seymour, J.E., 2007. Venom ontogeny, diet and morphology in *Carukia barnesi*, a species of Australian box jellyfish that causes Irukandji syndrome. *Toxicon* 49, 1073-1082.

Welton, R.E., Burnell, J.N., 2005. Full length nucleotide sequence of a Factor V-like subunit of oscutarin from *Oxyuranus scutellatus scutellatus* (coastal Taipan). *Toxicon* 46, 328-336.

Williamson, J.A.H., 1985. In: Exton, D. (Ed.), The Marine Stinger Book. Surf Lifesaving Association of Australia, Queensland State Centre Incorporated, Brisbane, Australia, 3rd ed.

Williamson, J.A., Fenner, P.J., Burnett, J.W., Rifkin, J.F. (Eds), 1996. Venomous and Poisonous Marine Animals: A Medical and Biological Handbook. Surf Life Saving Queensland Incorporated and University of New South Wales Press, Sydney, Australia, pp. 800.

Williamson, J.A., Le Ray, L.E., Wohlfahrt, M., Fenner, P.J., 1984. Acute management of serious envenomation by the northern Australian box-jellyfish (*Chironex fleckeri*). *Medical Journal of Australia* 141, 851-853.

Wiltshire, C.J., Sutherland, S.K., Fenner, P.J., Young, A.R., 2000. Optimization and preliminary characterisation of venom isolated from 3 medically important jellyfish: the box (*Chironex fleckeri*), Irukandji (*Carukia barnesi*), and blubber (*Catostylus mosaicus*) jellyfish. *Wilderness and Environmental Medicine* 11, 241-250.

Winkel, K.D., Tibballs, J., Molenaar, P., Lambert, G., Coles, P., Ross-Smith, M., Wiltshire, C., Fenner, P.J., Gershwin, L.-A., Hawdon, G.M., Wright, C.E., Angus, J.A., 2005.

Cardiovascular actions of the venom from the Irukandji (*Carukia barnesi*) jellyfish: effects in human, rat and guinea-pug tissues *in vitro* and in pig *in vivo. Clinical and Experimental Pharmacology and Physiology* 32, 777-788.

Winter, K.L., Fernando, R., Ramasamy, S., Seymour, J., Isbister, G.K., Hodgson, W.C., 2007a. The *in vitro* vascular effects of two chirodropid (*Chironex fleckeri and Chiropsella bronzie*) venoms. *Toxicology Letters* 168, 13-20.

Winter, K.L., Isbister, G.K., Seymour, J.E., Hodgson, W.C., 2007b. An *in vivo* examination of the stability of venom from the Australian box jellyfish *Chironex fleckeri. Toxicon* 49, 804-809.

Winter, K.L., Isbister, G.K., Schneider, J.J.., Konstantakopoulos, N., Seymour, J.E., Hodgson, W.C., 2008. An examination of the cardiovascular effects of an 'Irukandji' jellyfish, *Alatina nr mordens*. *Toxicology Letters* 179, 118-123.

Yang, Y., Cun, S., Xie, X., Lin, J.,Wei, J., Yang, W., Mou, C., Yu, C., Ye, L., Lu, Y., Fu, Z., Xu, A., 2003. EST analysis of gene expression in the tentacle of *Cyanea capillata. FEBS Letters* 538, 183-191.

Zaharenko, A.J., Ferreira Jr, W.A., Oliviera, J.S., Richardson, M., Pimenta, D.C., Konno, K., Portaro, F.C.V., de Freitas, J.C., 2008. Proteomics of the neurotoxic fraction from the sea anemone *Bundosoma cangicum* venom: novel peptides belonging to new classes of toxins. *Comparative Biochemistry and Physiology* D, 219-225.

Zhu, K., Zhao, J., Lubman, D.M., Miller, F.R., Barder, T.J., 2005. Protein p*I* shifts due to posttranslational modifications in the separation and characterisation of proteins. *Analytical Chemistry* 77, 2745-2755.

APPENDIX 1

Peptide Mass Fingerprinting

Lewis Adler, APAF

Protein samples were destained and subjected to a 16 hour in-gel tryptic digest at 37ºC. The resulting peptides were extracted using $H_2O:CH_3CN$ (9:1, + 0.1% formic acid) for 15 minutes. The digested peptides were separated by nano-LC using a CapLC system (Agilent 1100 Series, Agilent Technologies, Germany). Sample (15µL) was injected onto a peptide trap (Michrome peptide Captrap) for preconcentration and desalted with 0.1% formic acid at 10µL/min The peptide trap was then switched into line with the analytical column containing C18 RP silica (SGE ProteCol C18, 300A, 3µm, 150µm × 10 cm). Peptides were eluted from the column using a linear solvent gradient, with steps, from $H_2O:CH_3CN$ (95:5; + 0.1% formic acid) to $H_2O:CH_3CN$ (20:80, + 0.1% formic acid) at 500nL/min over a 45 min period. The LC eluent was subjected to positive ion nanoflow electrospray analysis on an Applied Biosystems QSTAR XL mass spectrometer (ABI, CA, USA). The QSTAR was operated in an information dependant acquisition mode (IDA). In IDA mode a TOFMS survey scan was acquired (m/z 400-2000, 1.0s), with the four largest multiply charged ions (counts >25) in the survey scan sequentially subjected to MS/MS analysis. MS/MS spectra were accumulated for 1 s (m/z 50-2000).

The LC/MS/MS data was processed into peak lists for database searching using Mascot Ion/Ion Search (Matrix Sciences, http://www.matrixscience.com/). The sequences that were determined were subjected to a BLAST search on NCBI (http://www.ncbi.nlm.nih.gov/BLAST/). It is noted that leucine and isoleucine are isomers (reported as '[L/I]') that cannot be distinguished by mass spectrometry. Phenylalanine and methionine sulfoxide (a common artefact of 2D electrophoresis), reported as '[F/M_o]', together with glutamine (Q) and lysine (K), differ in mass by only 0.04 Da. Sometimes it can be difficult to distinguish these amino acids and unequivocal identifications may require further analysis.

APPENDIX 2

Functional Classification of Antigenic *C. fleckeri* Proteins

Supplementary Table 2.1

Classification of expressed *C. fleckeri* proteins according to their function.

Cluster No.	No. Clones	Homologous Protein (Best BLAST Match)
General Metabolism		
2	16	3-PGDH
6	2	trans-homoaconitate synthase
18	1	hypoxia-induced gene 1 (HIG1)-like
22	1	TDC-1 (tyrosine decarboxylase family)
26	1	short chain oxidoreductase
28	1	ornithine aminotransferase
Cell replication, signalling and motility		
7	1	replication factor C
17	1	chromosome segregation ATPase-like
19	1	B-cell translocation gene 1 (BTG family)
21	1	KIF27A homologue (kinesan family)
23	1	memo-like protein
Protein synthesis and post-translational processing		
1	29	EF1α
14	1	proprotein convertase-like protein
20	1	polyA binding protein
Molecular chaperones and regulation of protein activity		
3	7	HSP70
9	1	cyclophilin A
12	1	DnaJ (HSP40) b11 homologue
27	1	calreticulin
32	1	calponin
35	1	serine/threonine protein kinase
Structural Proteins		
4	5	filamin
5	3	ABP, coronin
11	1	nonmuscle myosin heavy chain-like
24	1	viral A-type inclusion protein or myosin heavy chain-like
Small molecule and protein/peptide transport		
10	1	ATPase-like protein
15	1	peroxisomal biogenesis factor 14
16	1	metaxin-like
25	1	mitochondrial inorganic phosphate carrier protein
30	1	transportin-3 homologue

Cluster No.	No. Clones	Homologous Protein (Best BLAST Match)
Unknown Function		
8	1	alkylhydroperoxidase (AhpD) core
13	1	thypedin
29	1	cell wall anchor family protein
31	1	translationally-controlled tumour protein
33	1	embryonic-1
34	1	folliculin
36	1	predicted protein
37	1	predicted protein
38	1	predicted protein
39	1	nil
40	1	nil
41	1	nil
42	1	nil
43	1	nil
44	1	nil

Supplementary Table 2.1 *(cont')*

APPENDIX 3

C. fleckeri Nucleotide Sequence Clusters Obtained Using CSL Box Jellyfish Antivenom Immunodetection

Supplementary Table 3.1

Assembled clusters that contain more than two independent *C. fleckeri* nucleotide sequences.

Cluster No.	Protein Match	Clone ID	Comment
1	EF1α	1-11-3	full-length
		1-12-1	
		1-13-1	
		2-4-1	full-length
		2-5-1	full-length
		2-6-3	full-length
		2-7-1	
		2-8-3	
		2-9-1	full-length
		2-10-1	full-length
		2-11-1	
		2-12-1	
		2-13-1	
		2-14-3	
		3-1-1	
		3-4-1	full-length
		3-5-1	
		3-8-1	full-length
		3-10-1	
		3-11-1	
		3-12-1	
		3-17-1	full-length
		3-20-1	full-length
		3-22-1	full-length
		3-24-1	
		4-4-1	full-length
		4-5-1	full-length
		4-6-1	full-length
		4-7-1	
2	3-PGDH	1-3-1	full-length
		2-2-1	
		3-2-3	
		3-9-1	full-length
		3-19-1	
		4-3	
		5-4	
		6,7-6-1/2	full-length
		11-2-1	
		13-1	
		21-3	

Cluster No.	Protein Match	Clone ID	Comments
2 *(cont')*	3-PGDH	22-3	
		25-1	
		36-1	
		40-1	
		45-1-1	full-length
3	HSP70	2-3-1	full-length
		3-6-1	full-length
		3-13-1	full-length
		1-1-1	
		27-2	
		42-1	
		43-1	
4	filamin	9-2	
		6,7-6-3	
		15-1	
		23-3	
		33-1	
5	ABP	3-16-1	
		4-3-3	
		19-1-1	

Supplementary Table 3.1 *(cont')*

APPENDIX 4

Nucleotide and Deduced Amino Acid Sequences of Expressed *C. fleckeri* Clones Obtained by Immunoscreening a Tentacle cDNA Library

4.1 Background

Expressed *C. fleckeri* cDNA library clones were detected using CSL box jellyfish antivenom, and mouse and rabbit anti-*C. fleckeri* nematocyst venom antibodies (see Chapter 3). Nucleotide sequences were processed and aligned using Sequencher (GeneCodes Corp.). Nucleotide and translated nucleotide sequences of isolated immunogenic cDNA clones, shown below in order of cluster number, were subjected to BLASTN, BLASTX and BLASTP analysis (Altschul et al., 1997), respectively. Putative start codons (ATG) and initiating methionines (M_i) are indicated in bold. During BLASTP analysis, putative conserved domains were detected by searches (CD-search; Marchler-Bauer and Bryant, 2004) of the Conserved Domain Database (CDD; Marchler-Bauer et al., 2007). Unless otherwise referenced, brief descriptions of detected conserved domains (CD) were sourced from the CDD and linked protein databases, including Pfam (pfam.sanger.ac.uk), COGs (www.ncbi.nlm.nih.gov/COG/grace/uni.html), PRotein K(c)lusters (www.ncbi.nlm.nih.gov/sites/entrez?db=proteinclusters), Smart (smart.embl-heidelberg.de), or InterPro (www.ebi.ac.uk/interpro/).

4.2 Antigenic clones detected with CSL box jellyfish antivenom

4) Filamin A - Consensus of Clones 9-2, 6/7-6-3, 15-1, 23-3 and 33-1

The translated clone contains 5 filamin-type immunoglobulin domains (Accession No. smart00557) spanning deduced residues 1-94, 97-150, 176-283, 284-371 and 387-476 of the 5'-end of the clone (a). Each domain forms a rod-like structure in the **actin-binding** cytoskeleton protein, filamin. The C-terminal repeats of filamin bind β1-integrin. No clear ORF was detected in the 3'-end of the clone (b).

(a)

```
  1 - GACGTCGAATCGGCTCCCCCAGATGGATACAGTGACATTCTTGCCTTCCCCAGCGCCAAACCACAACAGGAT - 72
  1 -  D  V  E  S  A  P  P  D  G  Y  S  D  I  L  A  F  P  S  A  K  P  Q  Q  D  - 24

 73 - TTTGTCATCAAGTTTGCSGGCTCCAGCGATCTCAGTGCATCGGTGACGAGACCGTCAGGGAAAGAAGATGAT - 144
 25 -  F  V  I  K  F  A  G  S  S  D  L  S  A  S  V  T  R  P  S  G  K  E  D  D  - 48

145 - GCTGAAGTCGTCGAAACTGGRAAGGATACGTACACGGTTCGTTTTGTGCCTCGGGAAACCGGAGAGCATCTG - 216
 49 -  A  E  V  V  E  T  G  K  D  T  Y  T  V  R  F  V  P  R  E  T  G  E  H  L  - 72

217 - GTCCATGTTAAACATAGGAAACGAGACATTCCTGGAAGTCCTTTCAAGGTCCTTGTGGAAACCCCAACTGGT - 288
 73 -  V  H  V  K  H  R  K  R  D  I  P  G  S  P  F  K  V  L  V  E  T  P  T  G  - 96
```

```
 289 - GGAGCTGCGGCCTGCCGTGCATTCGGACCAGGGTTGGAGAAAGGTGTTTCTAACCAGCCATGCAAATTCACT - 360
  97 -  G  A  A  A  C  R  A  F  G  P  G  L  E  K  G  V  S  N  Q  P  C  K  F  T  - 120

 361 - GTTGTCACAAGAGATGCAGGGCCAGGAGGATTGGCAGTAGCTGTCGAAGGACCAGCAAAGGCTGAAATTCAG - 432
 121 -  V  V  T  R  D  A  G  P  G  G  L  A  V  A  V  E  G  P  A  K  A  E  I  Q  - 144

 433 - TGCCATGATAACGGGGATGGATCCTGTGATATTACGTGGTTCCCAGTAGAACCTGGCGAGTACACAATCCAC - 504
 145 -  C  H  D  N  G  D  G  S  C  D  I  T  W  F  P  V  E  P  G  E  Y  T  I  H  - 168

 505 - ATTCGATTTGCTGACGAACCAATCCCAGATAGTCCATTTAATGTGTACGTTGTCCCAGAAGATGAGAGCTCA - 576
 169 -  I  R  F  A  D  E  P  I  P  D  S  P  F  N  V  Y  V  V  P  E  D  E  S  S  - 192

 577 - GCTATGAGCATGGCACAGTTTAAAGAACAGGTTTTGAGGGTTGGGCAACAAGCCAGTTTTGCAGTGCAGATG - 648
 193 -  A  M  S  M  A  Q  F  K  E  Q  V  L  R  V  G  Q  Q  A  S  F  A  V  Q  M  - 216

 649 - AAAGGCAAGAAGGGTAAAATCAGTGCGTCCGTCGTAGCACCATCATCCAGTGCCGAGATTGAATGCACCGTA - 720
 217 -  K  G  K  K  G  K  I  S  A  S  V  V  A  P  S  S  S  A  E  I  E  C  T  V  - 240

 721 - GTTGAATTAGATGCCGGAAATTATGCAGTACGATTTGTGCCGAGGGAACTTGGTGATCATATGGTCAATGTT - 792
 241 -  V  E  L  D  A  G  N  Y  A  V  R  F  V  P  R  E  L  G  D  H  M  V  N  V  - 264

 793 - TTCTTAGATAAACAACATATTCCTGGAAGTCCTTTCAAAGTACGCGTTGGTGGTATTGAAGGAGATCCAAGC - 864
 265 -  F  L  D  K  Q  H  I  P  G  S  P  F  K  V  R  V  G  G  I  E  G  D  P  S  - 288

 865 - AAAGTGCGAGCGTATGGACCCGGCCTACGTGGTGGAGTCATTCATCGAGAAGCTGAGTTTACCGTGAATGCA - 936
 289 -  K  V  R  A  Y  G  P  G  L  R  G  G  V  I  H  R  E  A  E  F  T  V  N  A  - 312

 937 - CTCGAAGCTGGGTCTGGTGCATTGGCACTGTCAATCGATGGACCCGCTAAAGTCAAGATGAACTGTGTGGAA - 1008
 313 -  L  E  A  G  S  G  A  L  A  L  S  I  D  G  P  A  K  V  K  M  N  C  V  E  - 336

1009 - CTAGACGATGGCACGTATCAGGTTAGTTACACCCCAACAGTAGCTGGTAGCTACGAGATAAGTATTCGTTTC - 1080
 337 -  L  D  D  G  T  Y  Q  V  S  Y  T  P  T  V  A  G  S  Y  E  I  S  I  R  F  - 360

1081 - GCAGGACAACACATCCCAGGAAGTCCCTACAAAGTTGTCATTTCGAGTACAGGTGATGGCGTTGATGGTGCC - 1152
 361 -  A  G  Q  H  I  P  G  S  P  Y  K  V  V  I  S  S  T  G  D  G  V  D  G  A  - 384

1153 - TATGGTCCGTCAGATGCAAGCAAATGTACATCCAAGGGACTCGGTCTGAAACGTGCAACAGTCGGAGAACCC - 1224
 385 -  Y  G  P  S  D  A  S  K  C  T  S  K  G  L  G  L  K  R  A  T  V  G  E  P  - 408

1225 - GCAACGTTTACAGTTAATGCCAGCAATGCGGGCCGAGGTATGATTCTTGTCGGCGTTGAAGGTCCAGTCATA - 1296
 409 -  A  T  F  T  V  N  A  S  N  A  G  R  G  M  I  L  V  G  V  E  G  P  V  I  - 432

1297 - CCAGCGAAAGAAATTTTGGTTAGGCACACCGGTAACAACGTTTATTCAGTCAATTACGTCTTAGAGGAGCCT - 1368
 433 -  P  A  K  E  I  L  V  R  H  T  G  N  N  V  Y  S  V  N  Y  V  L  E  E  P  - 456

1369 - GGAGAATACGTCCTACAAGTGATGTGGGGTGATTCGCACATTCCTGGCTCGCCTTTCCACGGTAACAGTCTA - 1440
 457 -  G  E  Y  V  L  Q  V  M  W  G  D  S  H  I  P  G  S  P  F  H  G  N  S  L  - 480

1441 - ATTTCGCATCCGATCGAGTCTCTAGTTGAGAGACTTAGTGGGAGGAAAACGGGAAGTGGACTTTCTAGTTTT - 1512
 481 -  I  S  H  P  I  E  S  L  V  E  R  L  S  G  R  K  T  G  S  G  L  S  S  F  - 504

1513 - CAAATCGTCCGC                                                             - 1524
 505 -  Q  I  V  R                                                              - 509
```

(b)

```
   1 - GTTGTTGTTGTTGTTGTTGACCTGAATAACCGTCGCGATGTCTGTCGATTCTGAACTCTTATACGCAGTAGA - 72
  73 - TTTTTCAATCTTAACGCTGAAGCACAAAATTATTTATGTCCCCACCATTTGTATCGTGTTGAGGTACAATGT - 144
 145 - TATATTCAGGAGCAATACAGTAGGCAATGTTGTGTGATGTAGTTAGTGTTGTCATAATGAATATCCCAGCTC - 216
 217 - TACCTCAATCAGCTTAATTTGTTTAGAAATTGAAGCTATCTTTGGGTTACAAGAGGAAAAGGGTTTCTTAAT - 288
 289 - TTTTTTGTACTTGCAGATTTGTGATAGTGCGATTCGCCTCAAATTGTGGGTGGGTGGTTGTTGACAATGTTA - 360
 361 - AGTTGTCATAAATGTGCATAATATGACCCAAGTATTGAAGGCAGACTTTCCATTGCATTGGACAATAGTTTC - 432
 433 - CCTCTCTTTTCATTCATCGTTTGCTTCGTTTTGATACAAGCCCGATGTACATAGCTTAAATTTACCCTTTTT - 504
 505 - TACGTCATTACTGTAGCTGCTGAAAAATTTCTGAATAATTTGGCATTGAAAAAAAAAAAAAAAAAAAAAAAA - 576
 577 - AAAAAAAAAAAAAAAAA                                                        - 593
```

6) Trans-homoaconitate synthase - Consensus of Clones 8-3 and 32-1

No putative conserved domain was detected in translated nucleotide sequence. In archaea, trans-homoaconitate synthase catalyses the condensation of α-ketoglutarate and acetyl-CoA to form trans-homoaconitate. The enzyme also catalyses the condensation of α-ketoadipate or α-ketopimelate with acetyl-CoA to form (R)-2-hydroxy-1,2,5-pentanetricarboxylic acid and (R)-2-hydroxy-1,2, 6-hexanetricarboxylic acid, respectively (Howell et al., 1998).

```
   1 - AATGAGCGCCCATTTGGAATCCTTTACGAAGAGTCCGGTAAGTACTTGCCTGAGCAACTTGGTGTGTGGACG - 72
   1 -  N  E  R  P  F  G  I  L  Y  E  E  S  G  K  Y  L  P  E  Q  L  G  V  W  T  - 24

  73 - GCGGCAGTTCGGAAAGAAATGGACAGATGTGGATTTCAGTCTGGTCATTTGCTAGTTCATGTCCATCAGCAA - 144
  25 -  A  A  V  R  K  E  M  D  R  C  G  F  Q  S  G  H  L  L  V  H  V  H  Q  Q  - 48

 145 - TGGGGATTGCAAGACACAACACAAATGGAATGTTTGGCAAATGGAGCAAATGGCATTTGGGCTGGTCTGTGC - 216
  49 -  W  G  L  Q  D  T  T  Q  M  E  C  L  A  N  G  A  N  G  I  W  A  G  L  C  - 72

 217 - GAGGAGGGAGCAGCCATGGGTCATTCTTGTTCTGCTGTCACTCTAATGAATTTAGTACGTCTGGGGAACAAA - 288
  73 -  E  E  G  A  A  M  G  H  S  C  S  A  V  T  L  M  N  L  V  R  L  G  N  K  - 96

 289 - AAAGTTTTAAAGCGCTACAATTGCACTGAGCTGCGACGGGCAGCACAGAATGTGACACGGATTACCACAGGA - 360
  97 -  K  V  L  K  R  Y  N  C  T  E  L  R  R  A  A  Q  N  V  T  R  I  T  T  G  - 120

 361 - GTTGGCCCACATCCAAAACAGCCGATTTATGGCGAACGTGCTCTGGATATGGTATTTGGCATGGATCAGTTC - 432
 121 -  V  G  P  H  P  K  Q  P  I  Y  G  E  R  A  L  D  M  V  F  G  M  D  Q  F  - 144

 433 - TTACCAAATGAAAAAGAATTCAGTTTAGCGCATTTCTTTGGAGAAAAGCCGGTGATGCGAATGACAACTTTA - 504
 145 -  L  P  N  E  K  E  F  S  L  A  H  F  F  G  E  K  P  V  M  R  M  T  T  L  - 168

 505 - GCCTCACCTGCTATGGTTGTTCAAAGGTTAAAGAATTTATTTGGGGAAGACTCACAGTTTACAGAGGCACTA - 576
 169 -  A  S  P  A  M  V  V  Q  R  L  K  N  L  F  G  E  D  S  Q  F  T  E  A  L  - 192

 577 - GCTAAGAAAATGTTGGCAGTCATGCTTGAAGATTTGCATAAGAATCGTAAAGAGGAATACATGAGCGCAGTT - 648
 193 -  A  K  K  M  L  A  V  M  L  E  D  L  H  K  N  R  K  E  E  Y  M  S  A  V  - 216

 649 - GGTCTCGCTCTCCTTTTTGACCGTGCTGGAGGCCAAATAACAGAAAATATGAGCAAAGTCATCGCATCAGTT - 720
 217 -  G  L  A  L  L  F  D  R  A  G  G  Q  I  T  E  N  M  S  K  V  I  A  S  V  - 240

 721 - GAATCTAAAAATATTCACGTCAAAGAACTCATTGCCGAGATTCGCCAAATGTGGGACGAATGGGACTTGAGA - 792
 241 -  E  S  K  N  I  H  V  K  E  L  I  A  E  I  R  Q  M  W  D  E  W  D  L  R  - 264

 793 - GATGGACAACTCGATAACTATCTAGAGTTTGATGCCTTCTACAATGGCTTTATGGCACCATATTTTGGTTGT - 864
 265 -  D  G  Q  L  D  N  Y  L  E  F  D  A  F  Y  N  G  F  M  A  P  Y  F  G  C  - 288

 865 - AATCGTTGTGATGAGACTAAGAAAAGCTTGAAAGCTATTGATATGGATAGTGATGGGCGCGTGGATTGGGAT - 936
 289 -  N  R  C  D  E  T  K  K  S  L  K  A  I  D  M  D  S  D  G  R  V  D  W  D  - 312

 937 - GAATTTGGCTCTATTGAAATGGGCTGGAAGACAATACCCAAACGTCAAGGATGCTGAGGAGCTATTGTCAAT - 1008
 313 -  E  F  G  S  I  E  M  G  W  K  T  I  P  K  R  Q  G  C  *                 - 330

1009 - TGCATTCCGCAAAGGTTTGGTGCCTGCTATGCAGGACGAAGTTGTCAAGGAGACAGACAAGTTGCCAGCTAG - 1080
1081 - GAAGCAAATTCCACCAGAGGATGATGATGATGATTTGCATTTGTACTTTGGCGATGACGCAGAAGACGAGGA - 1152
1153 - ATTTTTTGGAATGCCTTAACTATGTAGACTTGTTTTGGATGGACAGTTTGACGCTGGCGATAATTATATCAC - 1224
1225 - CATTATGACTAGAAATTGGCCAGCAGCAAAATGGTCTGGTTAGTGGGCTCTCAAGGGGTTCTTTGTTTCCTT - 1296
1297 - TTCTAGTACAATAGTGATATGTAGAGTCGCATGCAAATAATGCTTGTCCTTTTAACAAGTGCTCCCGTAAAA - 1368
1369 - ATTCCATTATGCTATAGTTGTAGTAACGTAAGTACAGAAGCTACCAAATAATATCTTGCATCCCAGAAGTGA - 1440
1441 - CAGAATAACATATCTTACGCCTGATTTTTATTAAAATTTCTGCGGCTGGTAAAAAAAAAAAAAAAAAA     - 1508
```

7) Replication Factor C (RFC) – Clone 1-5-1

The translated clone contains a conserved replication factor C small subunit 2 (Accession No. PRK12402) between residues 1-290. This subunit forms part of the RFC clamp loader complex which loads the DNA polymerase processivity clamp, proliferating cell nuclear antigen (PCNA), onto DNA (Johnson et al., 2006).

```
   1 - TCGTTTGAAACGATAAATAGGACCATGCTTAAAAATCTAAAAGACAAAAAACTGGAAGGAAAACTCAAAGCC - 72
   1 -  S  F  E  T  I  N  R  T  M  L  K  N  L  K  D  K  K  L  E  G  K  L  K  A  - 24

  73 - GAGCTGTCAAGTCTTGCACACCGTTATGTCTCCTCTCATCGTCCATCGGAAAAAGATCTACAACACTTCAAT - 144
  25 -  E  L  S  S  L  A  H  R  Y  V  S  S  H  R  P  S  E  K  D  L  Q  H  F  N  - 48

 145 - GTCCTTAAGAAACTGCGCCAAAATAAAGATTTAGTTATAACGAGACCAGACAAGGGAAACGGAGTTGTTGTT - 216
  49 -  V  L  K  K  L  R  Q  N  K  D  L  V  I  T  R  P  D  K  G  N  G  V  V  V  - 72

 217 - ATGAATCACAAGGACTATGTCAAAGTCGTAGTTCTGATGGAAGTTGATAAATTAACTAAAGATGCACAGCAT - 288
  73 -  M  N  H  K  D  Y  V  K  V  V  V  L  M  E  V  D  K  L  T  K  D  A  Q  H  - 96

 289 - GCACTTCGACGAACAATGGAAAAGTATGTATCAACATGTAGGATTGTGCTGTGCTGCCAATCAACGAGTAAA - 360
  97 -  A  L  R  R  T  M  E  K  Y  V  S  T  C  R  I  V  L  C  C  Q  S  T  S  K  - 120

 361 - GTGATTCCTGCCATCCGAAGCAGGTGTCTTGGCATCAGAGTCGCAGCACCATCTATTGAAGAGATTTGTCAA - 432
 121 -  V  I  P  A  I  R  S  R  C  L  G  I  R  V  A  A  P  S  I  E  E  I  C  Q  - 144

 433 - GTTCTGAATTACGTCTGCAAGAAGGAATGCCTATCCCTACCACCAGTTTTATGTAAGAGGATTGCAGTGACG - 504
 145 -  V  L  N  Y  V  C  K  K  E  C  L  S  L  P  P  V  L  C  K  R  I  A  V  T  - 168

 505 - TCAGAAAGAAATTTACGAAGGGCACTTCTTATCTGTGAAGCCTGCCGTGTACAACAGTATCCATTTAGCGAT - 576
 169 -  S  E  R  N  L  R  R  A  L  L  I  C  E  A  C  R  V  Q  Q  Y  P  F  S  D  - 192

 577 - GAGCAAAATATACCTGTCCCTGCCTGGGAAGTATTTTTAAGACAAACAGCATCAAATATCATTCAACAGCAA - 648
 193 -  E  Q  N  I  P  V  P  A  W  E  V  F  L  R  Q  T  A  S  N  I  I  Q  Q  Q  - 216

 649 - AGTCCAAAACAACTTCTTGCTATTCGTTCACAATTGTACAAATTGCTGACACATTGTATACCTCCCGATATT - 720
 217 -  S  P  K  Q  L  L  A  I  R  S  Q  L  Y  K  L  L  T  H  C  I  P  P  D  I  - 240

 721 - ATCATGAAGGGTCTGTTAACAGAGCTGCTTAAAAATTGTGATGGGACACTTAAGACTGAGGTGACGCAGCTC - 792
 241 -  I  M  K  G  L  L  T  E  L  L  K  N  C  D  G  T  L  K  T  E  V  T  Q  L  - 264

 793 - GCAGCACATTACGAACATCGATTGCAGACTGGAAGCAAGGCCATTTTTCATTTGGAAGCGTTTGTAGCCAAA - 864
 265 -  A  A  H  Y  E  H  R  L  Q  T  G  S  K  A  I  F  H  L  E  A  F  V  A  K  - 288

 865 - TTCATGAGTATTTATAAGCAATTCTTGGAGGAAGGATTTGCTGACATGTTTTAACTAAATAGCTGACTTGAA - 936
 289 -  F  M  S  I  Y  K  Q  F  L  E  E  G  F  A  D  M  F  *                    - 305

 937 - CAAAGTTTAAACATTTACTTACCTGTTACATTGGCTTAAGTCTGTAATAGCTGAGGAGAGGAGATTATGGTT - 1008
1009 - GGTTGCTTACAGTAAGTTGCTTTCGGATAATTTAAAAGTAGACAGGGTGGCCTTATGTGAGAAACTGAAGTG - 1080
1081 - ATCTTATGATAATTATTATCACAACTTTAGTAANCGCGTGAAATTATTATATCATCTAATAATAAAAATGCC - 1152
1153 - TGACCTGCTGAAAAAAAAAAAAAA                                                 - 1176
```

8) AhpD Core – Clone 1-6–1

Uncharacterized conserved protein domain (Accession No. COG2128) detected (unknown function) in translated clone between residues 1-153. Although the translated clone appears to be full-length, homologues generally have ~30 extra residues in the N-terminal region.

```
   1 - AAAAACTTTGTTTGAAAAGATTTAATGTTAATAACCTATTCTTTCCATTTCCTTTCTTTTTCTACATGAATG - 72
   1 -              *                                                      *  M  - 1

  73 - CAGTCTGGTTTTATTCCAAATGTATTCCTTGCACTTGCTCATCGGCCTGACGAGTTCCGTGCATTCTTTTCC - 144
   2 -  Q  S  G  F  I  P  N  V  F  L  A  L  A  H  R  P  D  E  F  R  A  F  F  S  - 25

 145 - TATTATGATGCACTGATGTTAAAAGAGGATGGCTCTCTGACAAAAGCTGAAAAAGAAATGATTGTTGTTGCT - 216
  26 -  Y  Y  D  A  L  M  L  K  E  D  G  S  L  T  K  A  E  K  E  M  I  V  V  A  - 49

 217 - ACAAGTGCAGTCAATAATTGTTTATATTGTGTTGTTGCCCATGGTGCTTTATTACGGATTTATTCTAAGAAA - 288
  50 -  T  S  A  V  N  N  C  L  Y  C  V  V  A  H  G  A  L  L  R  I  Y  S  K  K  - 73

 289 - CCCTTGCTAGGAGATCAAGTGGCCATCAACTGGAAACATGCAGATTTAAGTAGTCGAGAAAAAGCAATATTA - 360
  74 -  P  L  L  G  D  Q  V  A  I  N  W  K  H  A  D  L  S  S  R  E  K  A  I  L  - 97

 361 - GAATTTGCTTTGACCATTGGATCAGGGAAACCAGTTCTTGATGCGGAGATCGAAAAATTGCGAGAACATGAC - 432
  98 -  E  F  A  L  T  I  G  S  G  K  P  V  L  D  A  E  I  E  K  L  R  E  H  D  - 121

 433 - CTCAGTTCAGACGACATCTGGGATATAGGAGCAATTGCTGCATTCTTTGCTATGTCCAATAGGTTGGCACAT - 504
 122 -  L  S  S  D  D  I  W  D  I  G  A  I  A  A  F  F  A  M  S  N  R  L  A  H  - 145

 505 - TTGTCCAACATGCGACCGAATGATGAATTTTATTTATTGGGAAGATTGCGAAAGACAAAGCAAAAATAAAAA - 576
 146 -  L  S  N  M  R  P  N  D  E  F  Y  L  L  G  R  L  R  K  T  K  Q  K  *     - 167

 577 - GTTACTAACCAACTTTGTAATGATTTTTTAGTTGTTGTCTAAGGACAATTACAATTTATTTAATATGTCGTG - 648
 649 - TGAGGTGGACATCCCAAACATTTTGACAGGGTCTGGAATGCGAGAAAAAATTTATTTGATGTAGGAAACTTCA - 720
 721 - ACCCTAGATTAAGAAATGTTTTGACGTCAATAAAGACTAACTTCATAAGAGGAAATTCTTGTATGCCCCAAA - 792
 793 - TTATTCAAAAATGATTATGTAAGACCTTTTTTATGGTAAAAATAGCTTGTAGGATTAAGTATTTAAGGAGAA - 864
 865 - GATGTAGATGTAAAGTAGTTGTAGTGATGCAGGTAATTGGTGATTGATTGGATACACAGTCAGATGGTGAAT - 936
 937 - TATTCTGCTATCCCTTAGGATTATTGCCGTAAGGATAATCGTATTTGTGACACATGGTGATGTTTATTATTT - 1008
1009 - GCATCATTAACTACCATAAATTGCCAGAAATCAGGGAAAAA                                - 1049
```

9) Cyclophilin A - Clone 1-9-1

Translated clone appears to be full-length, but no clear poly(A)-tail is present. Cyclophilin A is a highly abundant, cytosolic protein that catalyses the isomerisation of peptide bonds from *cis* to *trans* at proline residues. Cyclophilin A plays important roles in protein folding, regulation of protein activity and cell-signalling (Wang and Hietman, 2005).

```
   1 - TGGAATTTGAACAGCCTGGCTAAATTTCTTTCCGTTGGCTTCGTATTCGCAAGAAACTTTGCAAGAAGAGGG - 72

  73 - ATGGCTTCGTCTAATCCAAGGTGCTATTTTGATTTGACGGCAGATGGCGCAGACGTGGGCCGAGTCGTCTTC - 144
   1 -  M  A  S  S  N  P  R  C  Y  F  D  L  T  A  D  G  A  D  V  G  R  V  V  F  - 24

 145 - GAGCTCCGAAGTGATGTTGTTCCAAAGACAGCAGAGAATTTCAGAGCTCTTTGCACAGGAGAGAAGGGTTTT - 216
  25 -  E  L  R  S  D  V  V  P  K  T  A  E  N  F  R  A  L  C  T  G  E  K  G  F  - 48

 217 - GGGTATGAAGGCTCAACCTTCCACAGGATCATTCCAGAGTTCATGTGCCAGGGAGGAGATTTTACAAGAGGT - 288
  49 -  G  Y  E  G  S  T  F  H  R  I  I  P  E  F  M  C  Q  G  G  D  F  T  R  G  - 72

 289 - AATGGGACAGGAGGAAAATCCATCTATGGTGAAAAGTTTCCTGACGAGAATTTCAAACTTAGGCACACTGGC - 360
  73 -  N  G  T  G  G  K  S  I  Y  G  E  K  F  P  D  E  N  F  K  L  R  H  T  G  - 96

 361 - CCAGGGATTCTTTCAATGGCCAATGCTGGTCCAAACACTAATGGATCTCAGTTCTTCATTTGCACTGTTAAG - 432
  97 -  P  G  I  L  S  M  A  N  A  G  P  N  T  N  G  S  Q  F  F  I  C  T  V  K  - 120
```

```
 433 - ACTGAGTGGCTTGACAACAAACATGTCGTATTTGGGTCAGTAGTGGAAGGGCTTGGACATATCAAAACAGTT - 504
 121 -  T  E  W  L  D  N  K  H  V  V  F  G  S  V  V  E  G  L  G  H  I  K  T  V  - 144

 505 - GAATCCTATGGCAGCAGAAATGGGAAGACTTCAAAGAAGGTTGTCATAAAGAAGTGTGGACAACTATAAACA - 576
 145 -  E  S  Y  G  S  R  N  G  K  T  S  K  K  V  V  I  K  K  C  G  Q  L  *     - 168

 577 - TTTGTTTGGCTTTTAGCAAGTTCCTGCCCTTTAAGGAATGAAAGCAAGTACGGTCTAGGGAACGTGGTCCTG - 648
 649 - CTGAGGCATTTGGTTTTGCAGATGCCCTGGCTTGAAATGTGGTCTTTGAAAAATTAAGTAACTGATGTTTGC - 720
 721 - AGGGTTATTTGGAGTAATATGAAGACATACATAACATACGTATTCACCTTCACTGACTTAGCCATACGTTCT - 792
 793 - GCTTTGTAATGTAGAAAAATGTGGGTTAAGGTACATTTGTGTAAGGCTCTGGTGAGAAGCATGAACATGCCT - 864
 865 - TTGTATGTTTTTGTCTGAGGCACTGCTTGTGTCAGTTTCTGTTAGGCTGCAGCCTTGCAGGTAGTGATCAGT - 936
 937 - TGCCACTTGTGAAAGTATAATTGTTCTCTTCTGTTCAAGCAGTTGATATTTGTTGGGAACTGAATACATAAG - 1008
1009 - ATGAAGG                                                                  - 1015
```

10) ATPase-like protein – Clone 1-10-1 - ~280 residues missing from N-terminus

Residues 1-336 of translated clone contain a MgtA domain (Accession No. OG0474). MgtA is a cation transport ATPase domain involved in inorganic ion transport and metabolism.

```
   1 - AAGCTGAAGTCTCGAGGATTCATCGTATTGATGTGTGGCGATGGCACAAATGATGTAGGCGCCCTTAAACAT - 72
   1 -  K  L  K  S  R  G  F  I  V  L  M  C  G  D  G  T  N  D  V  G  A  L  K  H  - 24

  73 - TCACACGTGGGTGTCGCTTTACTGCCTGGCTCACCTGAGAAGTGGCAAGATCCACAGAAGAAAGAGAAAATT - 144
  25 -  S  H  V  G  V  A  L  L  P  G  S  P  E  K  W  Q  D  P  Q  K  K  E  K  I  - 48

 145 - AAAAGAGAGCTGGCTGATGCTGAAGATGCCAAAAAATCTAAAGGAAGTGCACCAAGTCGAGCTACGTCAAAG - 216
  49 -  K  R  E  L  A  D  A  E  D  A  K  K  S  K  G  S  A  P  S  R  A  T  S  K  - 72

 217 - GCTGCAAGAATGAGGGCCGTTACAAGAGGTGATACAGACTCACCTGCATTGAAAGCGGCAAACGATCCACAG - 288
  73 -  A  A  R  M  R  A  V  T  R  G  D  T  D  S  P  A  L  K  A  A  N  D  P  Q  - 96

 289 - AAAAAACTTCGTGAGCTGCTCAAGGAGTTGGAGGAGATCGATCAAGCTTCAGTTGTTAAGTTAGGAGATGCC - 360
  97 -  K  K  L  R  E  L  L  K  E  L  E  E  I  D  Q  A  S  V  V  K  L  G  D  A  - 120

 361 - TCCATTGCATCTCCATTCACTTCAAAATTTGCCAGCATCACCTGCATTCTTCACGTCATCAGACAAGGCCGC - 432
 121 -  S  I  A  S  P  F  T  S  K  F  A  S  I  T  C  I  L  H  V  I  R  Q  G  R  - 144


 433 - TGTACGTTGGTGACTACGCTACAGATGTTTAAAATCCTCGCTCTAAATGCACTTATTTTAGCATACAGTCAA - 504
 145 -  C  T  L  V  T  T  L  Q  M  F  K  I  L  A  L  N  A  L  I  L  A  Y  S  Q  - 168

 505 - AGTGTGTTGTACCTGAAAGGAGTTAAATTCAGTGATGGCCAGGCAACGCTGCAGGGAATACTTCTTGCTGGC - 576
 169 -  S  V  L  Y  L  K  G  V  K  F  S  D  G  Q  A  T  L  Q  G  I  L  L  A  G  - 192

 577 - TGTTTTCTTTTCATATCACGATCCAAGCCATTGGATATTCTTGCTAAAAGAAGACCATTGCCTAATATTTTT - 648
 193 -  C  F  L  F  I  S  R  S  K  P  L  D  I  L  A  K  R  R  P  L  P  N  I  F  - 216

 649 - AACCTCTATACAATTTTGACTGTGCTTGGGCAATTCGGTGTTCATTTCTGCGCAATGTTGTATATTGTACAG - 720
 217 -  N  L  Y  T  I  L  T  V  L  G  Q  F  G  V  H  F  C  A  M  L  Y  I  V  Q  - 240

 721 - GAGGCATTCGCCTTGACGCCACCCAAGGAAGAATTTGTTGACTTAGAGAAGAAATTTGAGCCAAATTTATTG - 792
 241 -  E  A  F  A  L  T  P  P  K  E  E  F  V  D  L  E  K  K  F  E  P  N  L  L  - 264

 793 - AACAGCGCAGTATACATAATAGCCATGTCCTTACAAATGGCAACATTTGCAGTTAATTACAGAGGTAGACCA - 864
 265 -  N  S  A  V  Y  I  I  A  M  S  L  Q  M  A  T  F  A  V  N  Y  R  G  R  P  - 288

 865 - TTTATGGAAAGCCTCCACGAAAATCGGCCCTTGCTACTTAGTTTATTAGGTTCTTGTATTGCACTTATCACT - 936
 289 -  F  M  E  S  L  H  E  N  R  P  L  L  L  S  L  L  G  S  C  I  A  L  I  T  - 312

 937 - CTGGTTGTTAATGGAATACCGGAATTTACACAACAGTTTGAGATTGTCGAGTTTCCACCCGAGTTCCGATCC - 1008
 313 -  L  V  V  N  G  I  P  E  F  T  Q  Q  F  E  I  V  E  F  P  P  E  F  R  S  - 336


1009 - ATGCTGCTTCAAATACTTGCTGCTGATATCATTTTAGCTTTTATTGTTGACAGATTTCTCCGTGTTTTGTTT - 1080
 337 -  M  L  L  Q  I  L  A  A  D  I  I  L  A  F  I  V  D  R  F  L  R  V  L  F  - 360
```

```
1081 - GGAAACTCGAAACTTCGAACTTAGTGTCATCATTATTAGGATTTTATTTGAGTATGATGACTGTAAGAGTTA - 1152
 361 -  G  N  S  K  L  R  T  *                                                 - 367

1153 - TACATATATTTACACTTGGCAAATTTGGAAAACAAAATGGTATGCACAAAAACGTACATGTCAATGGGCAAG - 1224
1225 - TGTACTTGCAGATGGTTATTTGGATTGTCATTAATTGGATAGATCTTTGAGATGGAATCATGCTCGTAAATC - 1296
1297 - CTGTGAAGTATTTTCTGCCTTGTTTTTCCGATGCTAGAGAAAGAGTTATACGTGCAAATCAAAAAAA      - 1363
```

11) Nonmuscle myosin heavy chain – Clone 2-1-1

Approx. 1500 residues are missing from the N-terminus of the translated clone. Translated residues 1-260 of the 5'-end of the clone (a) and residues 1-125 of the 3'-end of the clone (b) contain the conserved domain of the myosin tail 1 family (Accession No. pfam01576). The myosin molecule is a multi-subunit complex made up of two heavy chains and four light chains it is a fundamental contractile protein found in all eukaryote cell types. This family consists of the coiled-coil myosin heavy chain tail region. The coiled-coil is composed of the tail from two molecules of myosin. These can then assemble into the macromolecular thick filament. The coiled-coil region provides the structural backbone the thick filament.

(a)

```
  1 - CAGGTTGAACTTCAAGAACTTATCGACAGCAAAGATGATGTTGGTAAAAATGTGCATGAACTCGAAAAGGCA - 72
  1 -  Q  V  E  L  Q  E  L  I  D  S  K  D  D  V  G  K  N  V  H  E  L  E  K  A  - 24

 73 - AAACGTCTGCTGGAGCAACAAGTACAAGAGCAGAAAACGCAAATCGAAGAACTCGAAGATGAATTACAAGCC - 144
 25 -  K  R  L  L  E  Q  Q  V  Q  E  Q  K  T  Q  I  E  E  L  E  D  E  L  Q  A  - 48

145 - ACTGAGGATGCCAAATTGCGGCTCGAGGTCAACATGCAGGCCCAAAAAGCCACTTTTGAGCGTGACCTCGCC - 216
 49 -  T  E  D  A  K  L  R  L  E  V  N  M  Q  A  Q  K  A  T  F  E  R  D  L  A  - 72

217 - GCTAAAGATGAACAGATCGAAGAAGCCAGGAGAGGGCTTGTCAAACAGTTACGAGACATGGAGAATGAACTG - 288
 73 -  A  K  D  E  Q  I  E  E  A  R  R  G  L  V  K  Q  L  R  D  M  E  N  E  L  - 96

289 - GAGGAAGAAAGGAAAGCAAAGACCAATGCGGTTAGTGCTAAACGCAAATTGGAAGGAGATCTTAAAGACATG - 360
 97 -  E  E  E  R  K  A  K  T  N  A  V  S  A  K  R  K  L  E  G  D  L  K  D  M  - 120

361 - GAAACACAACTTGAGCAAGCCAATAGAATCAAGGAAGAGGGTGTGAAGCAGCTGAAGAAGTACCAGGGTCAA - 432
121 -  E  T  Q  L  E  Q  A  N  R  I  K  E  E  G  V  K  Q  L  K  K  Y  Q  G  Q  - 144

433 - ATGAAAGATGTCCAGCGCGACTTGGAAGACGCTCGACAATCACGAGACGAGTTAGCACAGCAAGTCAAGGAC - 504
145 -  M  K  D  V  Q  R  D  L  E  D  A  R  Q  S  R  D  E  L  A  Q  Q  V  K  D  - 168

505 - AATGAAAAGAAATTAAAGCTATTGGAGGCAGATTTCTTACAGATGCAAGAGGATCTGAGTGCTGCTGAACGA - 576
169 -  N  E  K  K  L  K  L  L  E  A  D  F  L  Q  M  Q  E  D  L  S  A  A  E  R  - 192

577 - GCCAGACGAACTGTTGAAGCTGAAAGGGATGAACTGCTAGAAGAGATCAATAACAACACAACTAGTAGGAAT - 648
193 -  A  R  R  T  V  E  A  E  R  D  E  L  L  E  E  I  N  N  N  T  T  S  R  N  - 216

649 - TCTATGGCTGAAGAAAAAAAAAGGTTAGAGGCAGCTATCGCGCAATTAGAGGAGGATCTGGAAGAGGAACGA - 720
217 -  S  M  A  E  E  K  K  R  L  E  A  A  I  A  Q  L  E  E  D  L  E  E  E  R  - 240

721 - ACTCAGAACGAGTTGATGCGTGAGAAAGTGAAACGGGCAAATCTTCAAATGGACCAAATGGG            - 782
241 -  T  Q  N  E  L  M  R  E  K  V  K  R  A  N  L  Q  M  D  Q  M              - 260
```

(b)

```
  1 - TGGCAAAACTTGCTGATATGGAAACATCTTTTGCGATCAAGGAATAAAAACACCATTCAGGCCTTGGAGACG - 72
  1 -  W  Q  N  L  L  I  W  K  H  L  L  R  S  R  N  K  N  T  I  Q  A  L  E  T  - 24

 73 - AAGATTATTAATCTCGAGGAACAGTTGGATCATGAAGCAAAAGAAAAAGCCCAGCTTAGCAAAGCTAGTCGC - 144
 25 -  K  I  I  N  L  E  E  Q  L  D  H  E  A  K  E  K  A  Q  L  S  K  A  S  R  - 48

145 - AGACTTGAGAAGAAGATAAAAGAAATGACCTTGAGTATTGAAGAAGAGCGGCGCCATACTGATCAATACAAA - 216
 49 -  R  L  E  K  K  I  K  E  M  T  L  S  I  E  E  E  R  R  H  T  D  Q  Y  K  - 72
```

```
217 - GACCAGGCTGAAAAGGTCAATCAACGTATGAAGAGCTTAAAACGACAATTGGAAGAAACTGAGGAAGAGGTC - 288
 73 -  D  Q  A  E  K  V  N  Q  R  M  K  S  L  K  R  Q  L  E  E  T  E  E  E  V  - 96

289 - ACCCGCCTTAACGCTGCCAAGAGAAAACTGCAGCGTGATCTTGACGAACAAGTGGAACAGGTGGAGTCCACG - 360
 97 -  T  R  L  N  A  A  K  R  K  L  Q  R  D  L  D  E  Q  V  E  Q  V  E  S  T  - 120

361 - CAACGAGAGCTCAATCAAGCTCGCGCAAGAGCCAGACGTGGGACAGGTGGAGGAAGCATGTCTAGATCTGGG - 432
121 -  Q  R  E  L  N  Q  A  R  A  R  A  R  R  G  T  G  G  G  S  M  S  R  S  G  - 144

433 - TACGACACTCCAACACGCAGGCGTGGAAAAGACGACGACGATGTGCCTGAAGAAGCTGGCGAAGATTGAGAT - 504
145 -  Y  D  T  P  T  R  R  R  G  K  D  D  D  D  V  P  E  E  A  G  E  D  *     - 166

505 - CTTGCCACGGGGAAACGGTTGCAATGGAATTTTCAGGCCCACAGATTATAGTGAAGAAATATATTTTAGGAC - 576
577 - TTGGGAATAACAATTTTACATTTTCTGTAGCCATAGTCCTAATCTTGGTTATTTGGAATAATTTTACTAATT - 648
649 - GAAATATGGGAAGAGTTGAGTCCAGACGTTGAATATGTTGTGGTGTCGTGAAAATTATTTCAATTCATGACA - 720
721 - ACGCCTGCATTACTGATCACATCGCATTTTCCAATAAACATGAACATTCCTTGTTATTGTCAACAACCTTTA - 792
793 - TTACAAATAAATTTTTGTAGCAACTGAAAAAAAAAAAAAAAAAA                             - 836
```

12) Heat Shock Protein 40 (DnaJ) – Clone 3'-2

Deduced residues 1-192 of the 5'-end of the clone (a) and deduced residues 1-51 of the 3'-end of the clone (b) belong to the DnaJ-class of molecular chaperones with C-terminal Zn finger domains (Accession No. COG0484). Family members are involved in posttranslational modification, protein turnover and chaperone activity.

(a)

```
  1 - GTTAAAAAGATGGGTGGATTTGATGATGGGGGCTTTGATCCATTTTCAACGTTCTTTGGTGGTTTTGGATTC - 72
  1 -  V  K  K  M  G  G  F  D  D  G  G  F  D  P  F  S  T  F  F  G  G  F  G  F  - 24

 73 - AGTTTTGGTGGTGATAATGGTAAAAAGCGAGGAGAAATTCCTAGGGGACCTGATGTTATTATGGATTTAGAA - 144
 25 -  S  F  G  G  D  N  G  K  K  R  G  E  I  P  R  G  P  D  V  I  M  D  L  E  - 48

145 - GTTAGTTTAGAAGAATTGTACACTGGAGATTTTGTTGAGGTAATGAGAGCAAAACCTGTTGCAAAGCAAGTG - 216
 49 -  V  S  L  E  E  L  Y  T  G  D  F  V  E  V  M  R  A  K  P  V  A  K  Q  V  - 72

217 - CCTGGAACAAGGAAATGTAATTGCAGAATGGAAATGAGAACAACACAGCTTGGTCCCGGGAGATTCCAGATG - 288
 73 -  P  G  T  R  K  C  N  C  R  M  E  M  R  T  T  Q  L  G  P  G  R  F  Q  M  - 96

289 - TCACAGGAGGAAGTTTGTGATGAGTGTCCACAAATAAAATATGTCACACAGGAGAAACTTCTTGAAATAGAA - 360
 97 -  S  Q  E  E  V  C  D  E  C  P  Q  I  K  Y  V  T  Q  E  K  L  L  E  I  E  - 120

361 - ATTGAACCAGGCATGGTTGATGGACAAGAATATCCTTTCCATGCTGAAGGCGAACCTCATATTGATGGTGAA - 432
121 -  I  E  P  G  M  V  D  G  Q  E  Y  P  F  H  A  E  G  E  P  H  I  D  G  E  - 144

433 - CCTGGTGACCTCAAGTTTAAAATTAAGGAGCAAAAACACGATCGGTTTGAACGTCGTGGTGAAGATTTGTAC - 504
145 -  P  G  D  L  K  F  K  I  K  E  Q  K  H  D  R  F  E  R  R  G  E  D  L  Y  - 168

505 - ACAAATGTAACAATCTCATTGACAGATGCTTTGACTGGATTTGAAATGGCCGTTCCTCATCTTGATGGTCAT - 577
169 -  T  N  V  T  I  S  L  T  D  A  L  T  G  F  E  M  A  V  P  H  L  D  G  H  - 192
```

(b)

```
  1 - AGGATGCCAAACTATGATAATAATCATCAGAAAGGTGTATTGCATGTAACGTTTGATGTAGAATTTCCGAAG - 72
  1 -  R  M  P  N  Y  D  N  N  H  Q  K  G  V  L  H  V  T  F  D  V  E  F  P  K  - 24

 73 - GGTGAACTTTCTGAGGAAGAGAAAACAGGTATTAAGGAAATATTAAAGCAGGCTTCAAAACAGACCGTCTAT - 144
 25 -  G  E  L  S  E  E  E  K  T  G  I  K  E  I  L  K  Q  A  S  K  Q  T  V  Y  - 48

145 - AATGGGTTATGACGCCGTTCATAGAATATTAATATAGGCTGCCATTTATTAAAATCTGTCAAATTGTGACCA - 216
 49 -  N  G  L  *                                                              - 51

217 - TTTGTGATGTACTCGACGAGCTTAGTGTATTTCAATCTGCTATTAGCTACAACAAATCTCTTTCTTCATAAA - 288
289 - TTGTGTTAGTCCTGCTTTTATCCATTCTAATGAATTATAAATATATTCTTTGTGGCATCCCAAGTGTAGATA - 360
361 - CGTAGGTATTATATAATGTAATACGAAGATTTGTCTTCCTCGTGCTTGTCTTTATGTCATCCATTTACGTTA - 432
```

```
433 - AGAACTGTCTCACAGTACTATACCGTGTTTGTTGTTCGTCTGCGAGTAGGCAATAACAGGTCTCCACCTAAA - 504
505 - TTTGGATTATTCACTCTGTATACTGTTGGTTTTAGTAGTTTTAATTATTCACTTCCAATTGATGTTCAGGGG - 576
577 - CATTGTATTATTAATTGTTGTCAATACTGCTGCGATTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA    - 645
```

14) Proprotein convertase-like protein – Clone 3-18-1

Residues 1-83 of the translated clone correspond to a conserved domain of the subtilase family (Accession No. pfam00082). Subtilases are a family of serine proteases that appear to have independently and convergently evolved an Asp/Ser/His catalytic triad, like that found in the trypsin serine proteases (see pfam00089). Residues 142-211 of the translated clone also contain a protease-associated (PA) domain (CD Accession No. cl08238). The PA domain is an insert domain in a diverse fraction of proteases. The significance of the PA domain to many of the proteins in which it is inserted is undetermined. It may be a protein-protein interaction domain. At peptidase active sites, the PA domain may participate in substrate binding and/or promoting conformational changes, which influence the stability and accessibility of the site to substrate.

```
   1 - GAAAAGGAGGAAAACAAAATGATTACAACTGATCTTCACAATAAATGCACACAGGAGTTTAAAGGCACGTCA - 72
   1 -  E  K  E  E  N  K  M  I  T  T  D  L  H  N  K  C  T  Q  E  F  K  G  T  S  - 24


  73 - TCTGCTGCACCGCTAGCCGCTGGGATGTTTGCTCTGGTTTTACAAGCAAACCCCAATTTAACTTGGAGGGAT - 144
  25 -  S  A  A  P  L  A  A  G  M  F  A  L  V  L  Q  A  N  P  N  L  T  W  R  D  - 48

 145 - CTTCAACATCTCATTGTAAACACGGCACAAGTAACAAGTCCAGTGGATGAAGGATGGATGACGAATGGAGCT - 216
  49 -  L  Q  H  L  I  V  N  T  A  Q  V  T  S  P  V  D  E  G  W  M  T  N  G  A  - 72

 217 - GGCTTCCATTTCAATCACAAGTTTGGATTTGGTCGTTTGGACTCAGACAAAATGGTAGACCTTGCCAAACGC - 288
  73 -  G  F  H  F  N  H  K  F  G  F  G  R  L  D  S  D  K  M  V  D  L  A  K  R  - 96

 289 - TGGAAAACTGTACCCAAGCAGAGAAAATGTCAAGGGCCATCGAGCAGTTCTGCCCAGGAAATTCCTGCCGGA - 360
  97 -  W  K  T  V  P  K  Q  R  K  C  Q  G  P  S  S  S  S  A  Q  E  I  P  A  G  - 120

 361 - GGTACACTCTCTATTACCATAGATACCATAGCATGTGAAGGCACCGATCAGGAAATCACAAAACTAGAGCAC - 432
 121 -  G  T  L  S  I  T  I  D  T  I  A  C  E  G  T  D  Q  E  I  T  K  L  E  H  - 144

 433 - GTCACACTAACAGTATCATTTCAGCATCGTCGTCGTGGCGACGTGAGTGTGGATTTGTTCTCTCCATTGGGA - 504
 145 -  V  T  L  T  V  S  F  Q  H  R  R  R  G  D  V  S  V  D  L  F  S  P  L  G  - 168

 505 - ACGAGAAACGAAATGTTGTCTACTCGAAGGTATGATGATAGTGACAAAGGCTTACATGATTGGACATTTATG - 576
 169 -  T  R  N  E  M  L  S  T  R  R  Y  D  D  S  D  K  G  L  H  D  W  T  F  M  - 192

 577 - ACGGTGCATTGCTGGGGAGAAAATCCCAAAGGAGTCTGGACGTTGAACCTAACGGACAATATTCCCGCTTTA - 648
 193 -  T  V  H  C  W  G  E  N  P  K  G  V  W  T  L  N  L  T  D  N  I  P  A  L  - 216

 649 - ACTAGCAAAGGTTTTAACGTGAATGGGTATTTACAACATGATACAAGTAGACAGGAACCAGATGTTGAAGAT - 720
 217 -  T  S  K  G  F  N  V  N  G  Y  L  Q  H  D  T  S  R  Q  E  P  D  V  E  D  - 240

 721 - CTTGAAGAGGAGGTGATAGACGACGAAAAGAAAATGAAGGAGATCGAATCAAAGAGAAAAGGAAAAAAGGCC - 792
 241 -  L  E  E  E  V  I  D  D  E  K  K  M  K  E  I  E  S  K  R  K  G  K  K  A  - 264

 793 - CACCCAGATGATGGATCTGACGGCTTTCCGTACCCAGAGGGCGTTCGCAGGGCCAACGTAGAGGGTGTTAAA - 864
 265 -  H  P  D  D  G  S  D  G  F  P  Y  P  E  G  V  R  R  A  N  V  E  G  V  K  - 288

 865 - GGGGGTGGAGCGATACGAGACGACCTAGACGAGGATAGAAATCTGTTAACGACAGATGAGGATATCGATGCT - 936
 289 -  G  G  G  A  I  R  D  D  L  D  E  D  R  N  L  L  T  T  D  E  D  I  D  A  - 312

 937 - CTGTTGAGAGATACTGAAAGACGAACAGTTGTCACGGGCAATAATGAGAACGAAATATCACTCGACAGCATT - 1008
 313 -  L  L  R  D  T  E  R  R  T  V  V  T  G  N  N  E  N  E  I  S  L  D  S  I  - 336

1009 - CAAAGTTATGATGATGCGCCTGGTATGATGGGCGCTCTTGAGCAAGGTCAAGCTACCTTTGATGGATTTGAT - 1080
 337 -  Q  S  Y  D  D  A  P  G  M  M  G  A  L  E  Q  G  Q  A  T  F  D  G  F  D  - 360
```

```
1081 - AGAAATTTTATGAGTGACAGATCTCTTACAGGAAAACTACCGGCTGATGAGTCGGAAAGTTCCGAAGTTTAC - 1152
 361 -  R  N  F  M  S  D  R  S  L  T  G  K  L  P  A  D  E  S  E  S  S  E  V  Y  - 384

1153 - GGTTATGATCGAAACATTGAGAATGTACGTAAGATAATGCAAGCTGATTCTAATAAGAGAATGGGTCTTTCA - 1224
 385 -  G  Y  D  R  N  I  E  N  V  R  K  I  M  Q  A  D  S  N  K  R  M  G  L  S  - 408
1225 - TACCAGGATCCAAGAGTACAAGTGCAGGAGGAAACTGGCTCCCAGAGAGTGCAAGTTAGTAATGGTTATCAG - 1296
 409 -  Y  Q  D  P  R  V  Q  V  Q  E  E  T  G  S  Q  R  V  Q  V  S  N  G  Y  Q  - 432

1297 - GAACCTTTCATTCCAGGACTAAATCCCTCAACAGGCAGAGTGAGCTCAGGACAATTACTGGAGTGGGATTTA - 1368
 433 -  E  P  F  I  P  G  L  N  P  S  T  G  R  V  S  S  G  Q  L  L  E  W  D  L  - 456

1369 - GTATTTTACGGAACCGGGCCAGAAGAAGAAGACAATTGACTCAAGGAGATTGATGGCTTAGGACTATCTATT - 1440
 457 -  V  F  Y  G  T  G  P  E  E  E  D  N  *                                   - 468

1441 - TTGTACTGTGTTTTATTGCTATTTCCAGACTTGCTTACATGTACCTCTTATTATCAAAA              - 1499
```

15) Peroxisomal biogenesis factor 14 – Clone 3-21-1

Translated residues 15-131 at the 5'-end of the clone (a) contains a peroxisomal membrane anchor protein (Pex14p) conserved region, as part of the Pex14N superfamily (Accession No. pfam04695). This is a family of peroxisomal membrane anchor proteins that bind the PTS1 (peroxisomal targeting signal) receptor and are required for the import of PTS1-containing proteins into peroxisomes. Loss of functional Pex14p results in defects in both the PTS1 and PTS2-dependent import pathways. Deletion analysis of this conserved region implicates it in selective peroxisome degradation. In the majority of members this region is situated at the N-terminus of the protein. No putative conserved domain was detected in the translated residues of the 3'-end of the clone (b).

(a)

```
   1 - GCAGAAGGAGAAGAGATCTTAGAAGAAAAGAAAAATCAGAATGTTGATCTTCTCAAGGTCGACACAGCTGTA - 72
   1 -  A  E  G  E  E  I  L  E  E  K  K  N  Q  N  V  D  L  L  K  V  D  T  A  V  - 24

  73 - AAATTCTTGCAAAATCCAAAGGTGGCTGCAAGTCTACTGTCAAGGCGGATTGATTTTCTTCGAGACAAGGGA - 144
  25 -  K  F  L  Q  N  P  K  V  A  A  S  L  L  S  R  R  I  D  F  L  R  D  K  G  - 48

 145 - TTAACAGAGACAGAAATTCAGCTTGCAGTTGAGAGATCAGGTGCAACAGAATTTACAGGAACGGCACAATTG - 216
  49 -  L  T  E  T  E  I  Q  L  A  V  E  R  S  G  A  T  E  F  T  G  T  A  Q  L  - 72

 217 - ACACAGGGTCGCAATGAAGGGGATAAAACAATTGAAAGACGAGACCAGCCAGTAACAACAGCACAGGCATCA - 288
  73 -  T  Q  G  R  N  E  G  D  K  T  I  E  R  R  D  Q  P  V  T  T  A  Q  A  S  - 96

 289 - ACAAGAGGGACACGATGGAAAGACTTATTTGGAGCTGCAGTTATTGGTGCTGGTTTCGTGTATAGTATGGGA - 360
  97 -  T  R  G  T  R  W  K  D  L  F  G  A  A  V  I  G  A  G  F  V  Y  S  M  G  - 120

 361 - TACTTGGTGAAGTCACATGTCATTCCTTCGATTTTTGGAAAACAAGGGAAGTCTGACGAAGAGGAACAAATA - 432
 121 -  Y  L  V  K  S  H  V  I  P  S  I  F  G  K  Q  G  K  S  D  E  E  E  Q  I  - 144

 433 - CAATTTAAATTAACACTACAAGAATTGCGTGGTGGATTTGGTGATATTGCCGTCGAAATGAGAGACACTTTC - 504
 145 -  Q  F  K  L  T  L  Q  E  L  R  G  G  F  G  D  I  A  V  E  M  R  D  T  F  - 168

 505 - CGAAGCATTCAGCAATTTCTTGATGAACAACAGAGAGTGTTGGGAGGTTTAAGAGGTGAAGTATCAGTCTGC - 576
 169 -  R  S  I  Q  Q  F  L  D  E  Q  Q  R  V  L  G  G  L  R  G  E  V  S  V  C  - 192

 577 - AAGTCCAAACTTGAGGGAATTCCATCGGATATGGCCAGTCTAAGTGAAATTAAAAATGAACTGTCGTCACTT - 648
 193 -  K  S  K  L  E  G  I  P  S  D  M  A  S  L  S  E  I  K  N  E  L  S  S  L  - 216

 649 - AAAGGATTGGTCTTGGGCAGGCGACAGTTTCCACCTTCGCCATTTGCAGCAAGCAGTACATCCCCGTCTATC - 720
 217 -  K  G  L  V  L  G  R  R  Q  F  P  P  S  P  F  A  A  S  S  T  S  P  S  I  - 240

 721 - CCGGCATGGCAAAAGGCAGCTCCTGTTTTACGTGGAGAAAGCAAAGTATCGTCATCAAGACAAGATTCTCCT - 792
 241 -  P  A  W  Q  K  A  A  P  V  L  R  G  E  S  K  V  S  S  S  R  Q  D  S  P  - 264
```

```
793 - GCCAGTTACATCTCGTCTTCT                                                    - 813
265 -  A  S  Y  I  S  S  S                                                     - 271
```

(b)

```
  1 - AAATCGACGAAGTCACAAGGCGTCCCGAAGGATAGGCCCACGTTGTGTTCAAACTTTGACCTGAAAATAACA - 72
  1 -  K  S  T  K  S  Q  G  V  P  K  D  R  P  T  L  C  S  N  F  D  L  K  I  T  - 24

 73 - TACAGTTCACGACAAACATACTCCGTTCCATTTGCGATTAAGGGAGTAATGACCGTGAAGTTCAGATATATT - 144
 25 -  Y  S  S  R  Q  T  Y  S  V  P  F  A  I  K  G  V  M  T  V  K  F  R  Y  I  - 48

145 - AACTTAGCTCGCAGTAATTGCCTGGGCAAGCGCTTTCGACTATTTGACTAGTCGATACCCAGTAAATATTGA - 216
 49 -  N  L  A  R  S  N  C  L  G  K  R  F  R  L  F  D  *                       - 64

217 - TAGATTTCCAGAAAATTCATAGAAGAGTAAACGCCGTGGGCTTCTGAATCAACGATTGTGAGACATCGATAT - 288
289 - TTACCTTGCTTTTCTCGCAATCTTAGTTGTGTGCGTTTGAGAAAAGAAATGCCAGGCGTTTAGACTTGCATT - 360
361 - GGCTCCTTACCTGCCAGCGAACATTATCAACTGCTCATGCGCATATTATGTGTTGTTTATTGTGAGTGACCT - 432
433 - AAGCATGGAACAGTCTCCGAACGGAAAACTTTCTCGATCACTCAAGCATTGAAAGTTTTATATGTGGAGAAC - 504
505 - AAATATGTATCATATTTTGCAGCGACATAATACTTAAGAATGTTAAAAAATCTCATGGGGGAAATTGTTTCA - 576
577 - AGTGTAAACACAGAAGTCCTGCGTTGTAGTGAAAGTTTTGAGCAGCGTTAAGACATAATTTTGTACTGCTTT - 648
649 - GTAAAAAGTTTCAGTGTTTTGGACAGTATTGTGACACCGTTGCTTGTTGTGGAGATTTACAGCTATAAACAA - 720
721 - GTTTAGAGACAAAATTAAAAAAAAAAAAAAAAAAAAA                                    - 757
```

16) Metaxin-like protein – Clone 3-23-1

Deduced residues 32-104 of the 5'-end of the clone (a) contain a conserved GST-N metaxin-like domain (Accession No. cd03080); a heterogenous group of proteins, predominantly uncharacterised, with similarity to metaxins and glutathione S-transferases (GSTs). Metaxin 1 is a component of a preprotein import complex of the mitochondrial outer membrane. It extends to the cytosol and is anchored to the mitochondrial membrane through its C-terminal domain. In mice, metaxin is required for embryonic development. In humans, alterations in the metaxin gene may be associated with Gaucher disease. One characterized member of this subgroup is a novel GST from *Rhodococcus* with toluene o-monooxygenase and γ-glutamylcysteine synthetase activities. Other members include the cadmium-inducible lysosomal protein CDR-1 and its homologs from *C. elegans*, and the failed axon connections (fax) protein from *Drosophila*. CDR-1 is an integral membrane protein that functions to protect against cadmium toxicity and may also have a role in osmoregulation to maintain salt balance in *C. elegans*. The *fax* gene of *Drosophila* was identified as a genetic modifier of Abelson (Abl) tyrosine kinase. The fax protein is localized in cellular membranes and is expressed in embryonic mesoderm and axons of the central nervous system. No clear ORF was detected in the 3'-end of the clone (b).

(a)

```
  1 - GGTGCTGCTGTTGAAGCTGCTCCTGAGAAAACTGAGGATACAGAGAAAGTGGAACAGGTCGAAGAGGAAAAG - 72
  1 -  G  A  A  V  E  A  A  P  E  K  T  E  D  T  E  K  V  E  Q  V  E  E  E  K  - 24

 73 - AAGGACGCTGACACGCCGCCCAAAGTTCTATTGCATCAGTTCCCGCCAGGTAGAAGTGTTGTAAGTTACTCT - 144
 25 -  K  D  A  D  T  P  P  K  V  L  L  H  Q  F  P  P  G  R  S  V  V  S  Y  S  - 48

145 - CTGTTCTGCTTGAAGTTGGAGACACTATTACGAATGCATAAAATTCCTTACAGCAATATTTACAGTTATAAA - 216
 49 -  L  F  C  L  K  L  E  T  L  L  R  M  H  K  I  P  Y  S  N  I  Y  S  Y  K  - 72

217 - ATGTCAAAGAAGGGACTTATGCCATGGATCGAGTATAAGGGAGAGAAAATTGCCGATTCTAGCTTTGCCTGG - 288
 73 -  M  S  K  K  G  L  M  P  W  I  E  Y  K  G  E  K  I  A  D  S  S  F  A  W  - 96
```

```
289 - GACCACTTGGTCAAGGCCTTTGAATTGGAAGTCGATGGACATTTGAATGATGAACAAAAAGCAACGAGTAGA - 360
 97 -  D  H  L  V  K  A  F  E  L  E  V  D  G  H  L  N  D  E  Q  K  A  T  S  R  - 120

361 - GCTTTGCAAGTTATGTTGGAGGAAAATACAAGTTGGACGTTGTCATATCATCGCTGGATCAATGACTACGGT - 432
121 -  A  L  Q  V  M  L  E  E  N  T  S  W  T  L  S  Y  H  R  W  I  N  D  Y  G  - 144

433 - GAATGGAAGAAGTTGATACAACCCAATTCTGGAGGAGGGATAGCGTTCACTGTTGGGATGAAGATGTCACAA - 504
145 -  E  W  K  K  L  I  Q  P  N  S  G  G  G  I  A  F  T  V  G  M  K  M  S  Q  - 168
505 - CGTAAGATGAAAAATCTCTTGGAGGGGCACGGTATTGGTAGGCATACCAAGGAAGAAATTTTCACTATCGCC - 576
169 -  R  K  M  K  N  L  L  E  G  H  G  I  G  R  H  T  K  E  E  I  F  T  I  A  - 192

577 - GAAAAAGATTTACGGGCTTTGTCAGGCTTCCTCGATGAAAAGAAATTCTTGATGGGCGAACAACCAAGCTGC - 648
193 -  E  K  D  L  R  A  L  S  G  F  L  D  E  K  K  F  L  M  G  E  Q  P  S  C  - 216

649 - GTGGATGCAGTCGCATTTGGAATACTCAGTAACATATTTGTATGTAGGAACAGAAGTCTCCTCATCCGGA   - 718
217 -  V  D  A  V  A  F  G  I  L  S  N  I  F  V  C  R  N  R  S  L  L  I  R     - 239
```

(b)

```
  1 - AAGACTGAAGATCATGCAGCGAGTCGAGGAGGAACATCTCTAGACAGCAACTGTTAGTTTACATTTTTACCG - 72
 73 - TTGTGTCACTCGGATCTAAGAGTGATTAATTAGTACGCGTAACGCATCAAGGCATTCTGCCTTAAAAACTTT - 144
145 - ACTGTTCCAAATAAAGGAATATTGCTTAAGAATAGATCCTACTAATTTCTTACAAGTAGTAAGTAGACTAAC - 216
217 - TCTTATTGCGTTCGTAAAGATTTGGGGCAGTTTCTCCCGTGCACCTCAACTAAGCGTTTGTTATATGCACTG - 288
289 - TCAGCAAATATGGACGTTCTACTGCGATGCTCATCAGTTGCTTTACAAACTCTGTCCCAGCGGGACTTAAGT - 360
361 - GGATGAACTTAGTATTGAATATTCGAGTCGTCAGAAAGACGAGTCGGAGCTTCGAGCTTAATTTTCGTGACA - 432
433 - ACTACGAACCATGCATTCAGTATCGTCATAATAGATAATAGATAACATTTTCAGGGGTTAGCGTTCCGAGGA - 504
505 - GCAGACAACTGACTATGATTCCTTGGGAATTTCCATGTCCATGGTGATCGCTGAGCATGGCTGAATTGGACC - 576
577 - AGACCATAATGGAGTCTGGTGCCTACCGGGCCCCAGACCTTGTCAGCGGTTCGCTTGCTCCTGACGATGTTC - 648
649 - CAGATAATAGTAAGACTTGTTCCCCATGAGTTTGGCAAAATTGTTTTTAATTGCAAATATGATACTTGTGAA - 720
721 - TTGTTAGCTTTTCTTAAACATTCATTATATTTTCTGATTAAAGTGGAGTTGTTCAGTTAACTTTTTTGTTTG - 792
793 - TTATGCGAAAACTTCGACTTTCCTATCAAAAAAAAAAAAAAAAAA                            - 837
```

17) Chromosome segregation ATPase-like protein – Clone 4-1-1

Translated clone is a member of the chromosome segregation ATPase family (Accession No. COG1196); proteins involved in cell division and chromosome partitioning. Approx 500 residues are missing from the N-terminus of the translated clone and the 3'-end of the clone appears to be prematurely truncated.

```
   1 - GAGATTCAAGTCAGCAAAACAGTGGATGAAAATGCTTACCTACGTAAGGAGCTTACAACAACACAAAATAAG - 72
   1 -  E  I  Q  V  S  K  T  V  D  E  N  A  Y  L  R  K  E  L  T  T  T  Q  N  K  - 24

  73 - CTTGAAAATTTGTCATCAATATTGGAGAAAAGTGTGGTGGAGAACTCAAGACTACAAAATGAGGCAACATCC - 144
  25 -  L  E  N  L  S  S  I  L  E  K  S  V  V  E  N  S  R  L  Q  N  E  A  T  S  - 48

 145 - TACAAATCCAATGTTACCAAACTTGAAAGGGAGATCTTGGTGACTAAAGAGGAATCTTCTCAACTGCAGACT - 216
  49 -  Y  K  S  N  V  T  K  L  E  R  E  I  L  V  T  K  E  E  S  S  Q  L  Q  T  - 72

 217 - GCAGTCGTGACACATGAAGACCGAAATGAGAATCTCTTAACGGAGCTCGATTCGTCTCGTCGTCGAGTCTTT - 288
  73 -  A  V  V  T  H  E  D  R  N  E  N  L  L  T  E  L  D  S  S  R  R  R  V  F  - 96

 289 - GAACTCGAAGCGATGCATGACATGATTGTCAAGGAAAATCAAACAATCAAATATGAACTTGAAGTGAAGATC - 360
  97 -  E  L  E  A  M  H  D  M  I  V  K  E  N  Q  T  I  K  Y  E  L  E  V  K  I  - 120

 361 - AAGCAACTTGAAGAGCAAATAAGTATGACCATGAAACAGTCGTCTGACTACAAGAATCAATGTGAAGATTAT - 432
 121 -  K  Q  L  E  E  Q  I  S  M  T  M  K  Q  S  S  D  Y  K  N  Q  C  E  D  Y  - 144

 433 - AAACTTCGCATTAAAAAGGTTCAAGAGGAATACACGCGAGTGCAACAGGAGTGCGATGACTGGAGAATTAAA - 504
 145 -  K  L  R  I  K  K  V  Q  E  E  Y  T  R  V  Q  Q  E  C  D  D  W  R  I  K  - 168

 505 - CATGATGAAGTGTACAAAGCACAGACAGCAAAAGATCGCATCATTGAGACACTGAAGCAGGAAATTACTACC - 576
 169 -  H  D  E  V  Y  K  A  Q  T  A  K  D  R  I  I  E  T  L  K  Q  E  I  T  T  - 192

 577 - ATAACTGGAGAATTGGACAGCGTTATTTTTGATGCTGAATCAAAGAAGACACTCATAGAGAAGATCACCGAG - 648
 193 -  I  T  G  E  L  D  S  V  I  F  D  A  E  S  K  K  T  L  I  E  K  I  T  E  - 216

 649 - GAAAAGAATGAGCTAACGGCTGATTTGACATCTACTCAGCAACGTCTGTATGAACTAGAGAGTATCTACAAA - 720
 217 -  E  K  N  E  L  T  A  D  L  T  S  T  Q  Q  R  L  Y  E  L  E  S  I  Y  K  - 240

 721 - TCCAGCGAAGAAGAGAAAGATAAATTGCAAATTGAGATCAGGGAGATACACAAGAGACTCATGGAGGTAGAG - 792
 241 -  S  S  E  E  E  K  D  K  L  Q  I  E  I  R  E  I  H  K  R  L  M  E  V  E  - 264

 793 - GCCGAACTTTCAAATGCCCAAAGAGATTCAGACAAGGAGACACAGAGTCTAAAGCAAAAACTTGCTCGACTG - 864
 265 -  A  E  L  S  N  A  Q  R  D  S  D  K  E  T  Q  S  L  K  Q  K  L  A  R  L  - 288

 865 - GAACAACAAATTGAGATTGTGATGAAAGACAGAGATTATGTGGCTGAGCAGCTTGAAACGGAAAGAAGCACA - 936
 289 -  E  Q  Q  I  E  I  V  M  K  D  R  D  Y  V  A  E  Q  L  E  T  E  R  S  T  - 312

 937 - TCCCACAAGTACAAAGAAGACTTTGAGAGAATTCAACGGGAGTATCTTGAGCATCAGAAAATTTGTGAAAGG - 1008
 313 -  S  H  K  Y  K  E  D  F  E  R  I  Q  R  E  Y  L  E  H  Q  K  I  C  E  R  - 336

1009 - TACAGCTCGCAGCTTATGGAGAAGGACAGAGTGATTGAGTCGTTGAAAAACGACATTTACCGAATAGAATCT - 1080
 337 -  Y  S  S  Q  L  M  E  K  D  R  V  I  E  S  L  K  N  D  I  Y  R  I  E  S  - 360

1081 - GAAATCGAGACTCTGAGAGAGCAATTAAGAAAAGCACACATGCAAACTGAAGTTGCAGAGCAGCAAAGAGAC - 1152
 361 -  E  I  E  T  L  R  E  Q  L  R  K  A  H  M  Q  T  E  V  A  E  Q  Q  R  D  - 384

1153 - AAAGCTCAGGAAAAACTTGCAACTGTTGAAGCGCGAATTCATGATTTGCAAGTTGAAAATGATCAACTGAAG - 1224
 385 -  K  A  Q  E  K  L  A  T  V  E  A  R  I  H  D  L  Q  V  E  N  D  Q  L  K  - 408

1225 - AAGAGACTCAACAGTCTTGAGAAGGATCTAAATACCATGAGTAACAAACTAGCTGATTCTGAAGCCGATAAT - 1296
 409 -  K  R  L  N  S  L  E  K  D  L  N  T  M  S  N  K  L  A  D  S  E  A  D  N  - 432

1297 - CAGAAATTAAAGTCAAGGATTGTAGAATTGGAGGAGGTAGTAAGGACTCAAAAAAATGTGATAAGTTCACTT - 1368
 433 -  Q  K  L  K  S  R  I  V  E  L  E  E  V  V  R  T  Q  K  N  V  I  S  S  L  - 456

1369 - GAGAGTGATGTTAACCAGTGCTAAGATAGTAAAAAAAAAAAAAAAA                           - 1414
 457 -  E  S  D  V  N  Q  C  *                                                  - 463
```

18) Hypoxia induced gene 1-like – Clone 4-2-1

Deduce amino acids 3-76 contain a conserved hypoxia induced protein region, HIG 1 N (Accession No. pfam04588). This family is found in proteins thought to be involved in the response to hypoxia. Family members mostly come from diverse eukaryotic organisms however bacterial members have been identified. This region is found at the N-terminus of the member proteins which are predicted to be transmembrane. <10 residues appear to be missing from the N-terminal end of the translated clone.

```
  1 - AGTACTGAGTCAGAGGAGTATGAATCACAAACAGCGAAGTTGATAAGAAAGAGTAAGAAAGAGCCTTTCATA - 72
  1 -  S  T  E  S  E  E  Y  E  S  Q  T  A  K  L  I  R  K  S  K  K  E  P  F  I  - 24

 73 - CCGATAGGATTCCTTGGACTTATTGGTGCGCTTGCTTATGGAGCTTACGCATATAAAACCAGAGGACCTGCT - 144
 25 -  P  I  G  F  L  G  L  I  G  A  L  A  Y  G  A  Y  A  Y  K  T  R  G  P  A  - 48

145 - ATGACTACAAGTAGATATTTGATGCGGCTCAGAGTACTAGCTCAAAGCTGTGTTGTTGGCTCCATTGTTCTT - 216
 49 -  M  T  T  S  R  Y  L  M  R  L  R  V  L  A  Q  S  C  V  V  G  S  I  V  L  - 72

217 - GGTGTTGCAGTGACATCAATGCAAGACGATACTGATGTCAAGTCAGAAAAATGATAAAGTATGGTGTATGCA - 288
 73 -  G  V  A  V  T  S  M  Q  D  D  T  D  V  K  S  E  K  *                    - 89

289 - AAGAACCTTTGGTGCAGGATTTTGATTAAAGAAGGACCATGGACGAAATTTCACTTGTTTTTCAAACAATGA - 360
361 - GAAAGATATCTCACTAATACTTCTGTAGTATAGATAGTCATATCAACCGATCTATTTATGTTGCTTGAAATA - 432
433 - TTTATTGATATTTGTTTGTTTTTTGTACGTATTTATTATTTTTGTTTGTTTGCTTTATTTTTATGCATACTT - 504
505 - CTATTTATTTACACATGTTCGATTGTATAAATGGGAGGTAAGCATTTTGAATGTTTGCCTAAAAGGCAGTGT - 576
577 - TTGATGGTTTAACAGCAGTTAATTTGGAAGTCCATAATTCAATAATTGAGGACTCTGTGATCGTTGTGCTCA - 648
649 - TGTGACTATGCAATTTACAACAACAAGATTTTCAGTATTTTGTGAAATGCAAATGTACCTTAATAAAATTTA - 720
721 - AGATGGAAAAAAAAAAAAAAAAAAAA                                               - 746
```

19) B-cell translocation gene 1 – Clone 6/7-6-4

Deduced residues 1-118 of the 5'-end of the clone corresponds to the B-cell translocation gene (BTG) superfamily of proteins (Accession No. pfam00742). Members of the BTG1 family are implicated as antiproliferative factors. They are characterized by their rapid, but transient, expression in response to factors that induce growth arrest and subsequent differentiation (Raburn et al., 1995).

```
  1 - TCAAATTGGACGTTGACACTAAATAATGTGAACTTTCTTACTTATCTCTCTCTCTCACCGCTTATACACTCT - 72
 73 - CCACCCTGGATCAGAAGCACGAACTGAATCGAAATACACACCAGCCTTCTTTAGTACGGACATTATTATCAC - 144
                               *                          *                    -

145 - ACAGTGGCGATACCCGCGAGAATTGTTAGCATCGGCCACCATTACGAAATGTTGTTAGAAGTACGTTCAGCG - 216
  1 -                                                   M  L  L  E  V  R  S  A  - 8

217 - ACCACCTTTTTAAAGGATCTTATTCACAGAAAAGGACACGTTGCGGTGGAAAAACTGGATAAATTCTCACAG - 288
  9 -  T  T  F  L  K  D  L  I  H  R  K  G  H  V  A  V  E  K  L  D  K  F  S  Q  - 96

289 - AGTCTAGAGAAACAGTTATGTGAAAAGTATAGAAATCACTGGCATCCAGAAAAGCCAAATTTAGGATCAGGA - 360
 33 -  S  L  E  K  Q  L  C  E  K  Y  R  N  H  W  H  P  E  K  P  N  L  G  S  G  - 120

361 - TTCCGATGTATTCGGTATTTTCACACCAACTTAGATCCATTGTTGGCTAATGCAGTCCGAGAAAGTGGCATG - 432
 57 -  F  R  C  I  R  Y  F  H  T  N  L  D  P  L  L  A  N  A  V  R  E  S  G  M  - 144

433 - ACTGCTCAAGATCTCAACGTATGTTTACCGTCTGATCTTACGCTATGGATTGATCCAGATGATGTTGCCTAC - 504
 81 -  T  A  Q  D  L  N  V  C  L  P  S  D  L  T  L  W  I  D  P  D  D  V  A  Y  - 168

505 - AGAATTGGGGAGGATGGCTCAATTTGTTCACTTTTAGATACATGTTTTCTAGAGGAGGAGGTTGTAAGCAAA - 576
105 -  R  I  G  E  D  G  S  I  C  S  L  L  D  T  C  F  L  E  E  E  V  V  S  K  - 192
```

```
577 - GCCAACAGCATGTCAGCGTTGAGTGAAGGCCAATGTTGGTATTCGCATAAACAATTGTACAAGTGGAATCCA - 648
129 -  A  N  S  M  S  A  L  S  E  G  Q  C  W  Y  S  H  K  Q  L  Y  K  W  N  P  - 216

649 - ATGTCGTTCAAACCATACCAACATTATCCAGGAAATTTCAAATCCAGAGAATGTATAAGACACATAGACATA - 720
153 -  M  S  F  K  P  Y  Q  H  Y  P  G  N  F  K  S  R  E  C  I  R  H  I  D  I  - 240

721 - TTATTAGAGAATGGGTTGAGCGATTGCTTAGTCTTTACTA                                 - 760
177 -  L  L  E  N  G  L  S  D  C  L  V  F  T                                   - 189
```

20) Polyadenylate binding protein - Clone 10-1-1

Deduced residues 188-201 of the 5'-end of the clone (a) contain a glycine-rich region (underlined). Also, deduced residues 90-162 contain a conserved region known as the RNA recognition motif (RRM; Accession No. cl02586). Many eukaryotic proteins that are known or supposed to bind single-stranded RNA contain one or more copies of the RNA recognition motif (a.k.a. RRM, RBD, or RNP domain), which spans ~90 amino acids. RRMs are found in a variety of RNA binding proteins, including heterogeneous nuclear ribonucleoproteins (hnRNPs), proteins implicated in regulation of alternative splicing, and protein components of small nuclear ribonucleoproteins (snRNPs). The motif also appears in a few single stranded DNA binding proteins. The RRM structure consists of four strands and two helices arranged in an alpha/beta sandwich, with a third helix present during RNA binding in some cases. Unable to confirm the correct reading frame for the 3'-end of the clone (b).

(a)

```
  1 - CGACGTCGATCACTGCAATCCTGTGGAGTACGTAGTTTTATTTTTGACTGAATTAATCATAGCATAATGTCT - 72
  1 -                                                  *  I  N  H  S  I  M  S  - 2

 73 - GACGAGGAACCTGGTAATGATACGTTATTGGACGCGGCCCTTGATAGTTCAAAAGACGATTACTCACTGGAT - 144
  3 -  D  E  E  P  G  N  D  T  L  L  D  A  A  L  D  S  S  K  D  D  Y  S  L  D  - 26

145 - GGAGATACGAGCTTGAATGCTGATGATGCTACTTACGATCCGGAGCTGGAGGCCATTAAGGCAAGAGTTAAA - 216
 27 -  G  D  T  S  L  N  A  D  D  A  T  Y  D  P  E  L  E  A  I  K  A  R  V  K  - 50

217 - GAGATGGAGGAAGAGGCACAAAAGCTTAGTGAAATGCAGAAAGAAGTTGAAGAATCTTTAATGAGTCCTACA - 288
 51 -  E  M  E  E  E  A  Q  K  L  S  E  M  Q  K  E  V  E  E  S  L  M  S  P  T  - 74

289 - GCTTCACAACCAGTAATGGAAGATAAAAAGGAAGTCGATGCTCGTTCAATATATGTAGGAAATGTTGATTAC - 360
 75 -  A  S  Q  P  V  M  E  D  K  K  E  V  D  A  R  S  I  Y  V  G  N  V  D  Y  - 98

361 - TCAGCAACAGCTGCTGAATTAGAGCAACATTTTCATGGTTGTGGCGCAGTAAATCGCGTTACCATCTTGTGT - 432
 99 -  S  A  T  A  A  E  L  E  Q  H  F  H  G  C  G  A  V  N  R  V  T  I  L  C  - 122

433 - GACAAATTCAGTGGACACCCCAAAGGGTTTGCTTATGTGGAATTTTCAGACAAAGACAGCATACAAAATGCA - 504
123 -  D  K  F  S  G  H  P  K  G  F  A  Y  V  E  F  S  D  K  D  S  I  Q  N  A  - 146

505 - GTTTCACTTAATGATTCGCTTTTCAAAGGAAGGCAAATAAAGGTTGTCCCAAAGAGAACCAATGTACCTGGG - 576
147 -  V  S  L  N  D  S  L  F  K  G  R  Q  I  K  V  V  P  K  R  T  N  V  P  G  - 170

577 - CTTACATCAACAGACCGTGGATATCGAGCACGTGGACGTGGATATTCACGTGGCTATCGGGGCTATGGTGGG - 648
171 -  L  T  S  T  D  R  G  Y  R  A  R  G  R  G  Y  S  R  G  Y  R  G  Y  G  G  - 194

649 - TTTGGTGGTGGGGGTTATGGATTTTCACCCCGACCTATAGAGGTCGTGGTGCATAAGA               - 706
195 -  F  G  G  G  G  Y  G  F  S  P  R  P  I  E  V  V  V  H  K                 - 213
```

(b)

```
  1 - GTTTTTGTTGGCCTATGTTTGACGGTATTGCTTCTTTGCTTCCCTTTGTAAATCGTCATGTGCATTTTGTTT - 72
 73 - AAACGTACATCTGGCAATTTGCTATTGTTCACTGTATCGTGACCTCAAAACTGTTCCCTGCTCAATGTAAAA - 144
145 - ATTTACTGTGCGTGCTTGGATTGTGTGGGGCGAGAGATAAAAAAGAAATTAAAAGATAAAACACTGGAGAAA - 216
217 - GGAAAATAAAGAACATGGATCAGTAGACGCGAGAGAAATCAAAATAATCAGAAAAAAGAACAAAAAGACATA - 288
289 - AAAAGACAACAATAACAAGATCTGTGAGCGGAAGTGGAGCATTTCATGTAGATATATTGGATAATATCCGCG - 360
361 - TTACCATTGCGTGGAGTGCTAATGGAAATAGTTGAAAAACCTTGCTGAAGCAGTGAAACCATTTTCAAAAAA - 432
433 - AAAAAAAAA                                                                - 441
```

21) KIF27A homologue (kinesan family) - Clone 12-2

No putative conserved domain was detected in the translated nucleotide sequences of the 5'- (a) or 3'- (b) ends of the clone; however, the translated clone is homologous with proteins that contain a conserved kinesin motor domain characteristic of the KIF4-like subfamily (Accession No. cd01372). Members of this group seem to perform a variety of functions, and have been implicated in neuronal organelle transport and chromosome segregation during mitosis. Kinesins are microtubule-dependent molecular motors that play important roles in intracellular transport and in cell division. In most kinesins, the motor domain is found at the N-terminus (N-type).

(a)

```
  1 - GAATTGGAGAAACTTGTAGGTCATAACAGAAAGCTTAGTGATCTTGAGCAAGATTTGAAAAGACGCCAAGAG - 72
  1 -  E  L  E  K  L  V  G  H  N  R  K  L  S  D  L  E  Q  D  L  K  R  R  Q  E  - 24

 73 - ATTCTTGAACGAAAAGAAGCTCTGATTGCTGAAAGGGTCGCATTAGAAAACAAAAAGACGAGAGCAAGTCAG - 144
 25 -  I  L  E  R  K  E  A  L  I  A  E  R  V  A  L  E  N  K  K  T  R  A  S  Q  - 48

145 - TTACTATCACAGGATACAGTTCGATTGTCCAGCCGATTGGAAGAAATCAGGCAACAGTTAAATGAACTACAA - 216
 49 -  L  L  S  Q  D  T  V  R  L  S  S  R  L  E  E  I  R  Q  Q  L  N  E  L  Q  - 72

217 - AGTCAAGGCAGAACCGAGAAAAAGAAGAATCAGACTGATGTTGATAATCTCATAGAAACCAAAACAGAATTG - 288
 73 -  S  Q  G  R  T  E  K  K  K  N  Q  T  D  V  D  N  L  I  E  T  K  T  E  L  - 96

289 - GAGAAACAGAAAAAAAGCCTTGAAGCCAGGCTTGAATATGATACTGTTATTGATATAAAAGAAGAGCAAAGA - 360
 97 -  E  K  Q  K  K  S  L  E  A  R  L  E  Y  D  T  V  I  D  I  K  E  E  Q  R  - 120

361 - TTACAGGATCTGATTCAGGGAATTGAAACACTTGAGACAACGATACAGTATAAATCTGATTTAATTCAAGAG - 432
121 -  L  Q  D  L  I  Q  G  I  E  T  L  E  T  T  I  Q  Y  K  S  D  L  I  Q  E  - 144

433 - AAGGAACATGTAGTAAGAAACTGTCAGCCTGCTGATGTATGTATTGCTAAACTGGAGTGCAAATTAGGTGAG - 504
145 -  K  E  H  V  V  R  N  C  Q  P  A  D  V  C  I  A  K  L  E  C  K  L  G  E  - 168

505 - CTGAGCAAAACCGAGATCAAATCTTTGTTAAGTGTCTATCTGAAGAAAGTTGTAGCGTCGAAAGATGCTGAA - 576
169 -  L  S  K  T  E  I  K  S  L  L  S  V  Y  L  K  K  V  V  A  S  K  D  A  E  - 192

577 - TATCAAGCAAAAGTAGACCGAGAAATGATGGAAGTACGTGTGGAGGAACAAGCACAAACGATTCAAGAGTTG - 648
193 -  Y  Q  A  K  V  D  R  E  M  M  E  V  R  V  E  E  Q  A  Q  T  I  Q  E  L  - 216

649 - GAACATGTCCTTGCA                                                          - 663
217 -  E  H  V  L  A                                                           - 221
```

(b)

```
  1 - GTAGTGGAGCCACTTCGACTTCCTCCAGTGCAAATAACTCCGAAAGACGTTAAGTCCTATCGGGAGAGATCG - 72
  1 -  V  V  E  P  L  R  L  P  P  V  Q  I  T  P  K  D  V  K  S  Y  R  E  R  S  - 24

 73 - CGAAATCAAAGTACCTTAGAAGGACTTGCGAGTCAGTTAGAAAAACAAGGTTCTGATCAACATCAATTTCTT - 144
 25 -  R  N  Q  S  T  L  E  G  L  A  S  Q  L  E  K  Q  G  S  D  Q  H  Q  F  L  - 48

145 - AAGGAAATTTTCGAATCGCAGCCTTCATTAAGATTAACAAAGAAGGCAGGCGACAATTAAGTTATATAAACG - 216
 49 -  K  E  I  F  E  S  Q  P  S  L  R  L  T  K  K  A  G  D  N  *              - 67

217 - TGGTTGAAGAGGAAATTTTTGAAAAACAAGGAAAATTGTTTGAACAAGACCGTCCAGGATTTCCGCATGCTG - 288
289 - AATTCGCATTCGCACTATGAGTGTGTCAAAAGGAATTTCTGAAGCAGTTCTCAAATGTCCTGTTTGCGCTAA - 360
361 - TACAGACCACGTTTCTTTAGGAAACAAGACACGTTTGCATGGTAGACTCGGTGTATCATGGTTCAAACAATA - 432
433 - GTGATCCTCTAAAACAGTCAATTCCTTAAGAGCCGTGTATGAACCTTAGCATGATGCTGAATCGACTACTGG - 504
505 - AAAGAAGCTTAAGGGAGTAGATAGAATGGGTCTATTCCCAACACACTCGCCAATCCTGTAGTATTCCCAAAT - 576
577 - TGTAGTGTAGGATTGTACAAAGAAGAAATAGCAATGAAGAAAAATCAAGACAATTTCTTAACACATATTCTG - 648
649 - TTGTGTAGAGAAATGACCGTAAATAACTTTCCTGATTACTGAAATATATAACAAGTTATTGACTTTTATATA - 720
721 - AAAAAAAAAAAA                                                             - 732
```

22) Tyrosine decarboxylase 1 - Clone 12'-2

In comparison to homologues, translated clone has ~250 residues missing from the N-terminus and the C-terminus appears to have been prematurely truncated by 100-200 residues. The quality of DNA sequencing is relatively poor in this region. Deduced residues 1-241 contain a conserved pyridoxal-dependent decarboxylase domain (Accession No. pfam00282). Members of this family are involved in amino acid metabolism and transport.

```
   1 - GATAAATACAGCTTATTATCTCGATTCGTTGCATATTGTTCCGATGAGACTCATAGCGCTTTCGAGAAAGCC - 72
   1 -  D  K  Y  S  L  L  S  R  F  V  A  Y  C  S  D  E  T  H  S  A  F  E  K  A  - 24

  73 - TGTAAAATTGCTGCAGTGACTTGCAGAGCCTTGGAATCTAATGCAGATAATGAGCTTGATGGTGATGTGGTG - 144
  25 -  C  K  I  A  A  V  T  C  R  A  L  E  S  N  A  D  N  E  L  D  G  D  V  V  - 48

 145 - GATAAAGCTATTCAGGAAGATAGAGAGAAAGGCTTAATTCCTTTTATTGTCTGTGCCACTTTGGGAACAACA - 216
  49 -  D  K  A  I  Q  E  D  R  E  K  G  L  I  P  F  I  V  C  A  T  L  G  T  T  - 72

 217 - TCAACTTGTTCGTTTGACAAAGTGAAGGAGATAGGCCCTATTTGCAAAGAAAATAAGATTTGGTTTCACATT - 288
  73 -  S  T  C  S  F  D  K  V  K  E  I  G  P  I  C  K  E  N  K  I  W  F  H  I  - 96

 289 - GATGCAGCCTACGCTGGTTCTGCCTTCATATGTCCAGAATATCGGCCTCTGTTTGATGGTGTTGAGTATGCC - 360
  97 -  D  A  A  Y  A  G  S  A  F  I  C  P  E  Y  R  P  L  F  D  G  V  E  Y  A  - 120

 361 - GACTCATGCAATATAAGTTGCAACAAACTAATGCTTGTCAATTTCGACTGCTCTCCGCTTTGGGTGAAAGAT - 432
 121 -  D  S  C  N  I  S  C  N  K  L  M  L  V  N  F  D  C  S  P  L  W  V  K  D  - 144

 433 - AAACAGAAAGTAATAAATTCTCTGAGTACTGATTTTGTACTCTTTAGCAAGAATGGGGAAGACTTGATAGTG - 504
 145 -  K  Q  K  V  I  N  S  L  S  T  D  F  V  L  F  S  K  N  G  E  D  L  I  V  - 168

 505 - GACTTCAAGGACATGCAAGTCCAGCTTGGCCGAAGATTCCGTGCGTTGAAGCTCTGGTTCGTTGTACGTACG - 576
 169 -  D  F  K  D  M  Q  V  Q  L  G  R  R  F  R  A  L  K  L  W  F  V  V  R  T  - 192

 577 - TATGGAGTGAAGGGGTTACAGCAGCATGTTTATCGGCACTGCAAATTAGCAAAATATTTTGCTGGACTAGTA - 648
 193 -  Y  G  V  K  G  L  Q  Q  H  V  Y  R  H  C  K  L  A  K  Y  F  A  G  L  V  - 216

 649 - GAGGTTGATGAAAGATTTGAAGTTTTGGGTTGTCGCTTGGGCTTATTCCTCTTCCGCCTTAAGGGAGACAAT - 720
 217 -  E  V  D  E  R  F  E  V  L  G  C  R  L  G  L  F  L  F  R  L  K  G  D  N  - 240

 721 - GAACTAAACGAAGCTTTGATCCAGAAAATACAAGCAGATGGCACAATTCATTTTTCTCCAACATTTGTGAAA - 792
 241 -  E  L  N  E  A  L  I  Q  K  I  Q  A  D  G  T  I  H  F  S  P  T  F  V  K  - 264

 793 - GGTCACGCGCCGTCACTCACGGATCGTAATTTGTTCCTAAGAATGCCACACCGGGAAAACCACCTTAGATAA - 864
 265 -  G  H  A  P  S  L  T  D  R  N  L  F  L  R  M  P  H  R  E  N  H  L  R  *  - 287

 865 - TTCCACTGGAATTACCACTTCCACGGAAAGTTGCCTCACTGATATTTTGGAGAAAAGATCACCCAGGAAATG - 936
 937 - ATAAGTGAAATTATTTCAGTTCACAATGTGGACTATTTTGTGTTTGCAAGTTGTTCGTATGAGTGACTTTAC - 1008
1009 - TGTATGTTTAATGTTATAATTTTGGTGAGTGTATTTTTAGTAATACTTTATGGTTATTCAACTTTTATATTT - 1080
1081 - TAACTCTCGTTACAGTTAAGAAGATGCACGTAAGAGGGCTGGTGTCCACCAAGTTTAAAAAAAAAAAAAAAA - 1152
1153 - AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA     - 1220
```

23) Memo-like protein - Clone 1-14-1

Sections of the translated clone (underlined) are homologous to a protein known as Memo (mediator of ErbB2-driven cell motility) between deduced residues 160-217 of the 5'-end of the clone (a) and 1-34 of the 3'-end of the clone (b). However, no clear ORF including an initiating methionine was detected prior to this region of significant homology. Therefore clone requires re-sequencing.

Memo is a 297-amino-acid protein that co-precipitates with the C-terminus of ErbB2 and is required for ErbB2-driven cell motility (Qiu et al., 2008). A Memo-like protein domain (Accession No. pfam01875) was also detected between deduced residues 162-208 of the 5'-end of the clone. It has been suggested that Memo controls cell migration by relaying extracellular chemotactic signals to the microtubule cytoskeleton.

(a)

```
  1 - CTTTGATAGTGTTCATTTAAGATGGAAGATGTTTAAGACTGTAGTTTTGCTTTGCAATGATCATGGTGGACC - 72
  1 -  L  *  *  C  S  F  K  M  E  D  V  *  D  C  S  F  A  L  Q  *  S  W  W  T  - 24

 73 - CCATGGTACTCCTAGTCATCCCTCAAGTAGGCTGCATTGGCATGTTGTATCTTGGTATTATAGACAATGCTA - 144
 25 -  P  W  Y  S  *  S  S  L  K  *  A  A  L  A  C  C  I  L  V  L  *  T  M  L  - 48

145 - CTAAAAGTGAAAGAGTATAGTAGAGAATTACCCATCTTAAAGTGGCTCCATATGTGTAAAGTGATTTGGCAG - 216
 49 -  L  K  V  K  E  Y  S  R  E  L  P  I  L  K  W  L  H  M  C  K  V  I  W  Q  - 72

217 - TGTGCTACATTTAACCTTTCCACAAGAAGGAAGTAATTCACCTTGAAAATTTTGTACCACTAAAAGATTGTT - 288
 73 -  C  A  T  F  N  L  S  T  R  R  K  *  F  T  L  K  I  L  Y  H  *  K  I  V  - 96

289 - CAAGGAAGTAATTGTCAGGAAAATAAAGAACTTTCTCCAACTGCAGTGACAAGCTGCCTTCTTAAGTAAGAA - 360
 97 -  Q  G  S  N  C  Q  E  N  K  E  L  S  P  T  A  V  T  S  C  L  L  K  *  E  - 120

361 - GGCTATAATTTGATCATATGTTACTTTTTTTACTGTACTATGCATGATATGTTAAGCAGTAAGGGGGAAGGC - 432
121 -  G  Y  N  L  I  I  C  Y  F  F  Y  C  T  M  H  D  M  L  S  S  K  G  E  G  - 144

433 - TGGTTTACAATGCTTTGCTTCTGCTGATCTTTGTTTTTATATTGTAGTGCAAAAGGGAACTTCAAAGTAGTT - 504
145 -  W  F  T  M  L  C  F  C  *  S  L  F  L  Y  C  S  A  K  G  N  F  K  V  V  - 168

505 - CCAATCTTAGTAGGAAGTACTAGTCAAGAAAAAGAGGAAATGTATGGAAAACTTTTTAGCAAGTATTTGGAA - 576
169 -  P  I  L  V  G  S  T  S  Q  E  K  E  E  M  Y  G  K  L  F  S  K  Y  L  E  - 192

577 - GATCCTAGTAATTTCTTTGTTATTTCATCAGATTTTTGTCATTGGGGTAAAAGATTTCGTTATATGTATTAT - 648
193 -  D  P  S  N  F  F  V  I  S  S  D  F  C  H  W  G  K  R  F  R  Y  M  Y  Y  - 216

649 - GAC                                                                      - 651
217 -  D                                                                       - 217
```

(b)

```
  1 - TCAACAGTTGCAAATTTTCAGTTCCACTCATATGCACAGTCTAATAGATGCCGTGATTTTAATGACAGTTCA - 72
  1 -  S  T  V  A  N  F  Q  F  H  S  Y  A  Q  S  N  R  C  R  D  F  N  D  S  S  - 24

 73 - GTGAGTTATGCTGCTGGTTCCTTGACATTAGGTTAACATTATTGCCAGCAGTGCCCAAGAACATTAGATATT - 144
 25 -  V  S  Y  A  A  G  S  L  T  L  G  *                                     - 35

145 - GAAGACACATGCATCTTAACATAAACACAATGGAGGTCTTTCTTATACAAGCTGAGTCAAGTGGTTTATGTG - 216
217 - GATACTCTAAGGCCATTTCACAAGTTTTTACAGCACATGAACACATTATTGTATACACTTATGATAATTGCC - 288
289 - ATGGGTTTGTGGCCACTTGAGGAGACTTTCCACAGGTTAAAAAATGTCAGCACCTAATATTTGTCTAAATAA - 360
361 - TGGTATACACTGTGCCCTTCAAACTGTGTTTAATATATTTAGTTTGAACTTGGGTTCCAGGCTCTGAAAAAG - 432
433 - GATTGATGTTATAGCTGAAAGGGTTTAAGAATGAGAACATGTAGTCCAAACCAGTAGATGATCCAAGGACTG - 504
505 - GTGAGTCTATATGGAATGATGTAGGCCAATTAAAATTGCTGAGTATTCTAGATTCATTTTTTGCTAATGTAA - 576
577 - CATGCATATGTTGGTTTTATGAAGACTGTATTGGCATTTTTTCTGTAATCTAAATATAAGTATGCATGTTGT - 648
649 - TTGTTGTAGTGTCACTATTAGGTCCAAAAACTAATTACATTTGCACAGTCAATCCAGCATGCAAGGTAATTA - 720
721 - CAGATTGGAAAAATGTAAATTTTGATTTTGTGAGAGAATTTGGTGGTCGTTAAAAAAAAAAAAAAAAAAAAA - 792
793 - AAAA                                                                     - 796
```

24) Viral A-type inclusion protein or myosin heavy chain - Clone 1-16-1

BLAST analysis indicates that the translated 5'-end of the clone (a) is homologous to *Trichomonas vaginalis* viral A-type inclusion protein (Accession No EAX99974; 3748 residues; 26% match; E-value = 6E-14) and more weakly homologous to *Trypanosoma cruzi* myosin heavy chain (Accession No. EAN86336; 3543 residues; 24% match; E-value = 3E-12). In contrast, the best match to the 3'-end of the clone (b) is *Caenorhabditis elegans* myosin heavy chain (Accession No. AAA28124; 1966 residues; 23% match; E-value = 9E-08), whereas the match to *Trichomonas vaginalis* viral A-type inclusion protein (Accession No EAX99974; 3748 residues; 22% match; E-value = 7E-06) is weaker. Therefore, complete clone sequencing is required to improve protein homology predictions. Also, no clear poly(A)-tail is present in 3'-end of the clone, suggesting that clone may be prematurely truncated. No significant conserved domain was detected in the translated clone.

(a)

```
  1 - GAAAACAGCAAGTTGAAGACAGAACTGCAAAGTTTGAAAGAAATGGAGAAAAATTATTTGGAAAAAATGGAT - 72
  1 -  E  N  S  K  L  K  T  E  L  Q  S  L  K  E  M  E  K  N  Y  L  E  K  M  D  - 24

 73 - GAAAGTGCTCAAATGATTACCAAATTTGAAGACGATTTGAAGGTAAAAGAGGCAGAAGTACAGCAAATGAGA - 144
 25 -  E  S  A  Q  M  I  T  K  F  E  D  D  L  K  V  K  E  A  E  V  Q  Q  M  R  - 48

145 - GGCGAATTAGCTTTGAAAGAGGAAGAGATTAAACATTTTGAAATAAAGCTGAAGCAGCTAGAGGACAAGTAC - 216
 49 -  G  E  L  A  L  K  E  E  E  I  K  H  F  E  I  K  L  K  Q  L  E  D  K  Y  - 72

217 - AATAGTAAGGACTTAGAAGAAAAGCTAGAAGATAACGGGAAACTAACAGAAAAATTACAGATAGATTTGTCT - 288
 73 -  N  S  K  D  L  E  E  K  L  E  D  N  G  K  L  T  E  K  L  Q  I  D  L  S  - 96

289 - GAATCGACCAAAGATTTAGAGGAGTTGAGGCTGGAGAATGATATTTTACAAGACCAGATAACTGAACTTGAA - 360
 97 -  E  S  T  K  D  L  E  E  L  R  L  E  N  D  I  L  Q  D  Q  I  T  E  L  E  - 120

361 - GAATCAAATACTACTTTGCACAAACAAGTTGATGAAAAGGACAAGTATGAACAAATGTTGAGAGATTTAGAG - 432
121 -  E  S  N  T  T  L  H  K  Q  V  D  E  K  D  K  Y  E  Q  M  L  R  D  L  E  - 144

433 - ATGGAAAATACTGCACTGAAACAGGATATTGTAACCATAACAGGTGAAACAGAGAGATTGAAAGAGGACTTG - 504
145 -  M  E  N  T  A  L  K  Q  D  I  V  T  I  T  G  E  T  E  R  L  K  E  D  L  - 168

505 - GATGAAATGAATTCTTCGCGTCAGGAGGAATTATTTAAGCTGAATCAGGAAAACAAGTCCCTCACTGAAACA - 576
169 -  D  E  M  N  S  S  R  Q  E  E  L  F  K  L  N  Q  E  N  K  S  L  T  E  T  - 192

577 - ATTGCAAGTTTGCAAGGCAATAGAGAACTTAACGTAGACGTGTCT                            - 621
193 -  I  A  S  L  Q  G  N  R  E  L  N  V  D  V  S                             - 207
```

(b)

```
  1 - GAGGAGAAATATAGACACGAGGTGATGAAGAGACAAGGTTGGATGAAGGAGCAAGAAGACATGACCGGATAC - 72
  1 -  E  E  K  Y  R  H  E  V  M  K  R  Q  G  W  M  K  E  Q  E  D  M  T  G  Y  - 24

 73 - AAGGATGACAAGATTAAAGAACTGCAGCATGACATACAACAGGCAGGCAGAGAAATAATTCAGCTGCGTCAA - 144
 25 -  K  D  D  K  I  K  E  L  Q  H  D  I  Q  Q  A  G  R  E  I  I  Q  L  R  Q  - 48

145 - GAAACGAGAAAATATGCAGAAGAAAACCAAGAGCTACAGCAAGAATTGTATGTACTTAGGATAAATGCTCAG - 216
 49 -  E  T  R  K  Y  A  E  E  N  Q  E  L  Q  Q  E  L  Y  V  L  R  I  N  A  Q  - 72

217 - TCACAGGATGACTCCAAGACTGATTCACATGCAGCACAAGCTCTTCAAGGTGAAAATACAGCGAAAATTCAA - 288
 73 -  S  Q  D  D  S  K  T  D  S  H  A  A  Q  A  L  Q  G  E  N  T  A  K  I  Q  - 96

289 - CAGTTAACTGACGAGATTGAAATGCAGCGACAAAGATATGAGCTGGAGATCAAAAGCTTAAAGTCTTCTCTC - 360
 97 -  Q  L  T  D  E  I  E  M  Q  R  Q  R  Y  E  L  E  I  K  S  L  K  S  S  L  - 120

361 - GGTGCGTTGAAATCGGAACATGTCGGGTCTCTTAAGTCAGATTACGATCGGATTGCATTAGAGCGTGTGGAC - 432
121 -  G  A  L  K  S  E  H  V  G  S  L  K  S  D  Y  D  R  I  A  L  E  R  V  D  - 144

433 - TTACAAGTTCAGATGGCAGAAATGGAAGAAGAACTCTCCTCCAAACAACTTACGTTAGATGAAATGTTTACT - 504
145 -  L  Q  V  Q  M  A  E  M  E  E  E  L  S  S  K  Q  L  T  L  D  E  M  F  T  - 168
```

```
505 - AAGAATCGCTTTTTAGAAATCGAGATGAAAAATGCCAAATCAAGAATTTCTGACCTAACGGAAGAACTTGAA - 576
169 -  K  N  R  F  L  E  I  E  M  K  N  A  K  S  R  I  S  D  L  T  E  E  L  E  - 192

577 - AGAGTAAGAGAGACAGGTGGACACTAAGTTCAAATAAAGCAATAATACTAATCGGGCCATTTAATTAAGCGT - 648
193 -  R  V  R  E  T  G  G  H  *                                                - 200

649 - GTAGTTTTACGTAGCTAATGACTGTATAATTAATCCTGTTTGAATATTTTTATAAAT                - 705
```

25) Mitochondrial Carrier Protein - Clone 17-2

~260 residues missing from N-terminus of translated clone. Deduced residues 1-67 contain a conserved mitochondrial carrier protein domain (Accession No. pfam00153).

```
  1 - GCTGTTGTGTCTCACCCAGCTGATACTGTAGTATCAAAATTGAACAATGATGTTGGAAGCACTCCTGTCCAG - 72
  1 -  A  V  V  S  H  P  A  D  T  V  V  S  K  L  N  N  D  V  G  S  T  P  V  Q  - 24

 73 - GCTGCAAGAGAACTTGGGATGCGAGGTTTATGGAAAGGACTGGGACCTCGTATTTTGATGATCGGTACCCTT - 144
 25 -  A  A  R  E  L  G  M  R  G  L  W  K  G  L  G  P  R  I  L  M  I  G  T  L  - 48

145 - ACTGCCTTGCAATGGTTTATTTATGACACAGTGAAAGTCGTCTTCAAACTTCCAAGGCCACCACCACCAGAA - 216
 49 -  T  A  L  Q  W  F  I  Y  D  T  V  K  V  V  F  K  L  P  R  P  P  P  P  E  - 72

217 - ATGCCAGAGAGTTTAAAGAGAAAATTGGCATTGAAACAGGCAGCTTAAAAAGAACCGCAATAAAATTATTTG - 288
 73 -  M  P  E  S  L  K  R  K  L  A  L  K  Q  A  A  *                          - 87

289 - GATTGACTGTGTTCTGCCATTGGTAGACAAGAATTTTTGGATTGCAACTGAAAATTGTACAGAATAGATGGC - 360
361 - AAGATGAAATGATGTTTAAGTTATGTGACATTTCGGCATATTTTGAGGGGTTTTGTGTGAAAAGAATCCAGA - 432
433 - ATACATGTGTGATATGCCGTCTTTTACTCATGAGTAGTTGAGACTTTATTGGCTGTGTCTCCTGTAGGTTAA - 504
505 - TTGCTCTTATTTGGTAGGTGAAATGTATATGGTAAGACTCACCATTGTACCAATGAAGTTATTTGGCAATAT - 576
577 - TGATGGATGTTTTGTCTTTCATTTAACTTGTTTGGAAGAGTCGTATTTATTCATAGAAAATGCCATAACTAT - 648
649 - AAAAAAAAAA                                                               - 658
```

26) Short chain oxidoreductase homologue - Clone 18-3

Deduced residues 34-20 in the 5'-end of the translated clone (a) contain a conserved domain characteristic of 3-ketoacyl-(acyl-carrier-protein) reductase (*FabG* gene; CD Accession No. PRK12825). This protein catalyses the first of the two reduction steps in the elongation cycle of fatty acid synthesis.

(a)

```
  1 - GTCTATCATAGTTCGAGTCAGTCTTTCAACAAACTGAACGGGAACTGGAAATTCATAGGTGAAACACTTTTA - 72

 73 - ATAATGGCACTCAACACGGGAAGCTGGCAGAAGAACAGATTGATAACAGGAGCAAGCAGAGGTATTGGAAAA - 144
  1 -     M  A  L  N  T  G  S  W  Q  K  N  R  L  I  T  G  A  S  R  G  I  G  K  - 23

145 - GCAATTGCAGTGAAGTTGGCCAGGGATGGTGCAAATGTTATCATTGCTGCCAAAACAGCAGAACCACACCCC - 216
 24 -  A  I  A  V  K  L  A  R  D  G  A  N  V  I  I  A  A  K  T  A  E  P  H  P  - 47

217 - AAGTTGCCTGGAACTATTTTCACTGCTGCAGAAGAGATTGAAAAAGCTGGTGGCAAGTGTTTACCCTGTATT - 288
 48 -  K  L  P  G  T  I  F  T  A  A  E  E  I  E  K  A  G  G  K  C  L  P  C  I  - 71

289 - GTTGACATAAGACATGAAAATGAAGTACAAGATGCTGTTGAAAAAGGAGTAAAACTATTTGGTGGTATTGAC - 360
 72 -  V  D  I  R  H  E  N  E  V  Q  D  A  V  E  K  G  V  K  L  F  G  G  I  D  - 97

361 - ATTCTTGTGAACAATGCCAGTGCTATTAGTTTAACTGGGACTCTGGAGACTCCAATGAAAAGATATGATCTT - 432
 96 -  I  L  V  N  N  A  S  A  I  S  L  T  G  T  L  E  T  P  M  K  R  Y  D  L  - 119

433 - ATGAATGGTGTAAATGCAAGAGGAACGTATCTTTGCTCTCAAGCATGTATCCCATATCTTAAAAATGGCAAA - 504
120 -  M  N  G  V  N  A  R  G  T  Y  L  C  S  Q  A  C  I  P  Y  L  K  N  G  K  - 143
```

```
505 - AACCCACACATATTGACCATAAGCCCTCCACTTAATTTGAACCCTAAGTGGTTTAAGAATCATGTTGCTTAC - 576
144 -  N  P  H  I  L  T  I  S  P  P  L  N  L  N  P  K  W  F  K  N  H  V  A  Y  - 167

577 - ACAATGGCAAAATACGGGATGTCAATGTGTGTACTTGGAATG                               - 618
168 -  T  M  A  K  Y  G  M  S  M  C  V  L  G  M                                - 181
```

(b)

```
  1 - AAGAATGGAGCAGGATCTGCTGGACAAGGACCATACCAAGATGGTGCTAATGTAACGATGACGCTGGACAGT - 72
  1 -  K  N  G  A  G  S  A  G  Q  G  P  Y  Q  D  G  A  N  V  T  M  T  L  D  S  - 24

 73 - GACGACTTTGTCAAAATGTTCTCTGGTAAACTTCAGCCAACGCAAGCCTTTATGAGTGGGAAACTGAAAATC - 144
 25 -  D  D  F  V  K  M  F  S  G  K  L  Q  P  T  Q  A  F  M  S  G  K  L  K  I  - 48

145 - AAGGGTGATATGATGATGGCGATGAAGCTAGAGAAACTTATGAAACAAATGACGCCCAAACTGTAGACCACT - 216
 49 -  K  G  D  M  M  M  A  M  K  L  E  K  L  M  K  Q  M  T  P  K  L  *        - 69

217 - ATTAATGACTAGTTAGATTAACAACATATGCATTTAAGCAAGGTTAGCAATAACTGTAAACTGTTTAGCTTC - 288
289 - GGCAAGCGATTGTTTACATTGATCTGATGAATATCATTTGCCCCATATTTTAAACCCGTTCGCAAAGGAGGC - 360
361 - ATCAGGTAAAGTCCAGTTATAACTTGACCGAACACATTATCGAGGAATCGTATTCGTATACATTCGCATGCA - 432
433 - TCCTAAGTGAACGTTATGATTGGTGTTTGTAATGCATGGTTTTGGCAAAAATTATGTCGTCGCAACAGTGAG - 504
505 - TGAAAAAAAAAAAAAAAAAA                                                     - 524
```

27) Calreticulin - Clone 26-2

Deduced residues 1-260 of the 5'-end of the clone (a) and residues 1-48 of the 3'-end of the clone (b) contain the conserved domain of calreticulin (Accession No. pfam00262). Clone appears to be full-length but no stop codon is detected in the 3'-end. Calreticulin is a ubiquitous multifunctional protein found in a wide range of species and in all nucleated cell types. Calreticulin is a major **Ca^{2+}-binding**/storage chaperone residing in the ER lumen (Michalak et al., 1992). Calreticulin binds (buffers) Ca^{2+} with high capacity and participates in folding newly synthesized proteins and glycoproteins. It is an important component of the calreticulin/calnexin cycle and quality control pathways in the ER (Kopito, 1997).

(a)

```
  1 - CGGTGGAACCCATGTGTTCCCTGTTAAACTGACTTTAGATTCTCCTGATTTCACTATACACACACCAGAAGT - 72
                               *                    *
 73 - TGGCATAACTATAATCAACTAGAAATTATCACTATGAATTATATAGGTTTACTGCTAATCGCCTTTCTCGCG - 144
  1 -                                   M  N  Y  I  G  L  L  L  I  A  F  L  A  - 13

145 - TCTTGTCAAGCGGAGGGAGAGAAAAAGAAAGAAGATGAGAATGTGTTAAAGTATAAAAGGCCAGAGGTATCT - 216
 14 -  S  C  Q  A  E  G  E  K  K  K  E  D  E  N  V  L  K  Y  K  R  P  E  V  S  - 72

217 - GGGTTTGCATATATTGCAGAACCATTCCACGAGAAATCAGAATTTGATCAGCACTGGATTGCATCACAAGCA - 288
 38 -  G  F  A  Y  I  A  E  P  F  H  E  K  S  E  F  D  Q  H  W  I  A  S  Q  A  - 96

289 - AAAAAAGACGGTGTTGATTCTGCCATTTCCAAGTATGATGGACAGTGGGCAATTGAAGAAGCTTCTGAAAAC - 360
 62 -  K  K  D  G  V  D  S  A  I  S  K  Y  D  G  Q  W  A  I  E  E  A  S  E  N  - 120

361 - CCATTAACGGGAGACCAAGGTTTAGTACTTAAGACGCAAGCAAAGCATCATGCTGTAGCAACACTCTTACCT - 432
 86 -  P  L  T  G  D  Q  G  L  V  L  K  T  Q  A  K  H  H  A  V  A  T  L  L  P  - 144

433 - ACCAAGTTTGAATTTAAAGATCAACCTCTCATTTTACAGTATGAAGTAAAATTTCAGAATGCTCTTGAATGT - 504
110 -  T  K  F  E  F  K  D  Q  P  L  I  L  Q  Y  E  V  K  F  Q  N  A  L  E  C  - 168

505 - GGTGGTGCCTATGTCAAACTTCTCTCTGATTCACCTGATTTAGATTTAGAACAATTTACTGACAAGACTGGG - 576
134 -  G  G  A  Y  V  K  L  L  S  D  S  P  D  L  D  L  E  Q  F  T  D  K  T  G  - 192

577 - TACACCATTATGTTTGGTCCAGACAAATGTGGTGAAAGTAGTAAGTTACATTTTATATTCCGGCATAAAAAT - 648
158 -  Y  T  I  M  F  G  P  D  K  C  G  E  S  S  K  L  H  F  I  F  R  H  K  N  - 216

649 - CCAGTGACGGG                                                              - 659
182 -  P  V  T                                                                 - 184
```

(b)

```
  1 - TATAAGGGTATCTGGAAGCCAAAGAGAATACCAAATCCTGATTTCTTTGAGGATAAAGAGCCTTTCAAAATG - 72
  1 -  Y  K  G  I  W  K  P  K  R  I  P  N  P  D  F  F  E  D  K  E  P  F  K  M  - 24

 73 - ACACCTATCGGCGGACTTGGATTTGAACTCTGGTCAATGACTTCGGATATTCTTTTTGATAATATCATTTTA - 144
 25 -  T  P  I  G  G  L  G  F  E  L  W  S  M  T  S  D  I  L  F  D  N  I  I  L  - 48

145 - ACAAGCGACCAAGCTGTTGTTGATCAATGGACCGCAGACACGTGGGATTTGAAACGCACAAAAGAATTCGAG - 216
 49 -  T  S  D  Q  A  V  V  D  Q  W  T  A  D  T  W  D  L  K  R  T  K  E  F  E  - 72

217 - ACACTGTCAAGTAGTGAGGGTGGCCTATGGAATCAACTAATGGAAGCAACACATGAGCGCCCATGGCTTTGG - 288
 73 -  T  L  S  S  S  E  G  G  L  W  N  Q  L  M  E  A  T  H  E  R  P  W  L  W  - 96

289 - ATTCTTTATGCTGCTGTTGTGTTCCTTCCGTTTGTTCTCATTGCCGCATTTTGCTTTCCAAGTAAAAAGGAT - 360
 97 -  I  L  Y  A  A  V  V  F  L  P  F  V  L  I  A  A  F  C  F  P  S  K  K  D  - 120

361 - AAGGCTGCTGAAAGAAAGAAGACCGACGAACCCACCCCGGATGATGAGCCTGAAAGTAAGGAAGAAGGCAAA - 432
121 -  K  A  A  E  R  K  K  T  D  E  P  T  P  D  D  E  P  E  S  K  E  E  G  K  - 144

433 - GAGGACGAAGAACAAGCTGGGGAAAGCAAAGACGACGCTGAAGAGGAGAAGAAAGAGGGCAATGAAGAAGAG - 504
145 -  E  D  E  E  Q  A  G  E  S  K  D  D  A  E  E  E  K  K  E  G  N  E  E  E  - 168

505 - GAAGAAGCAGAGGAAGATAGTAATAAAAAAAAAAAAAAAAA                                - 545
169 -  E  E  A  E  E  D  S  N  K  K  K  K  K                                   - 181
```

28) Ornithine aminotransferase - Clone 28-2

Deduced residues 1-258 contain the conserved domain of ornithine-oxo-acid transaminase (Accession No. PRK04073), the bacterial homologue of eukaryotic ornithine aminotransferase. Ornithine aminotransferase is a key enzyme in the pathway that converts arginine and ornithine into the major excitatory and inhibitory neurotransmitters glutamate and GABA. ~170-180 residues are missing from the N-terminus of the translated clone.

```
  1 - CGTACTTTAGCTGCCGTTTCATCGTCCAGCGATCCCGACTGTTATGGTGGATTTGGACCTTACATGCCAGGA - 72
  1 - R  T  L  A  A  V  S  S  S  S  D  P  D  C  Y  G  G  F  G  P  Y  M  P  G   - 24

 73 - TTTGAAATGGTGCCTTACAATGACCTTGAAGCGTTAGAGAATTCAATAAAAGATCCTAATACAGCTGCATTT - 144
 25 - F  E  M  V  P  Y  N  D  L  E  A  L  E  N  S  I  K  D  P  N  T  A  A  F   - 48

145 - ATGGTAGAGCCCATCCAAGGAGAAGCTGGAGTTATGGTTCCAGATGCTGGTTACCTTAAAGGTGTCAGAAAT - 216
 49 - M  V  E  P  I  Q  G  E  A  G  V  M  V  P  D  A  G  Y  L  K  G  V  R  N   - 72

217 - CTATGTGACAAGTATAATGTATTATTCATTGCAGATGAAGTACAAACTGGACTTGCTAGAACTGGACGAATG - 288
 73 - L  C  D  K  Y  N  V  L  F  I  A  D  E  V  Q  T  G  L  A  R  T  G  R  M   - 96

289 - TTATGTGTTGACCATGATGAAGTGCGACCAGATCTTGTCATACTTGGGAAAGCTTTATCAGGTGGAGTTATG - 360
 97 - L  C  V  D  H  D  E  V  R  P  D  L  V  I  L  G  K  A  L  S  G  G  V  M   - 120

361 - CCAGTATCTGCTGTCCTTGCAGATGATGATGTCATGCTTTGCATTCAACCAGGAGAGCATGGCTCCACTTTT - 432
121 - P  V  S  A  V  L  A  D  D  D  V  M  L  C  I  Q  P  G  E  H  G  S  T  F   - 144

433 - GGAGGAAATCCTCTTGGATGCAAAGTAGCTATTGCTGCTTTGGAGGTTATACAGGAGGAAAATCTATGTGAT - 504
145 - G  G  N  P  L  G  C  K  V  A  I  A  A  L  E  V  I  Q  E  E  N  L  C  D   - 168

505 - AATGCTGAAAGGCTTGGGAAAATCTTGAGAGAAGAACTATCCCAACTGCCTGCAGAGATTGTGCAAACTGTG - 576
169 - N  A  E  R  L  G  K  I  L  R  E  E  L  S  Q  L  P  A  E  I  V  Q  T  V   - 192

577 - CGAGGAAAAGGGTTGCTCAATGCGGTCGTCATCAAAGAAACTCCAGAATACGATGCTTGGAAGGTTTGCTTA - 648
193 - R  G  K  G  L  L  N  A  V  V  I  K  E  T  P  E  Y  D  A  W  K  V  C  L   - 216

649 - AAATTACGAGATAACGGCTTGCTTGCAAAACCGACACATGGTGACATTATAAGATTTGCACCACCGCTTTGC - 720
217 - K  L  R  D  N  G  L  L  A  K  P  T  H  G  D  I  I  R  F  A  P  P  L  C   - 240
```

```
721 - ATCACGGAAGAACAAATTCGAGAATGTATTGCAATCATCAAAGATACCATCAAGTCATTTATGTAACATTCT - 792
241 - I  T  E  E  Q  I  R  E  C  I  A  I  I  K  D  T  I  K  S  F  M  *        - 261

793 - TTTTGTATTGTTATTAAGAATAGAAATTGGCTGTTGATGTGACCTGTTGTGGTATTGCCTTATTTATAGATG - 864
865 - ACTGAAGTAGTTGATGCATCATAGTGGGAATTTTAGTGATTCTTGGACTGAGTATTGTATAGCAACATTTCA - 936
937 - TAGCAAGTAAAA                                                             - 948
```

29) Cell wall anchor protein - Clone 29-1

No putative conserved domain was detected in translated nucleotide sequence of the 5'- (a) or 3'- (b) ends of the clone; however, BLAST homologues contain conserved domains similar to members of the nuclear pore complex protein-like family (nsp1-like; Accession No. pfam05064).

(a)

```
  1 - AATGTGTCGGGCTTTAACATCTACGATGGTATTTATGCTAGGAATCGACCCGATGTGCGACCAGTTGTGACT - 72
  1 -  N  V  S  G  F  N  I  Y  D  G  I  Y  A  R  N  R  P  D  V  R  P  V  V  T  - 24

 73 - GTTCCGCCTTTTATCCAGACACAATCTTCTTCTGTACTTCAGCAACATGTGTCAACATTTGATCGACTAGCA - 144
 25 -  V  P  P  F  I  Q  T  Q  S  S  S  V  L  Q  Q  H  V  S  T  F  D  R  L  A  - 48

145 - TTGGTCAATTCGTATTTACGTCAAGTTTACGTCATGAACAAGAGTGCTCAACAGACAGGTGTGACATTTAGT - 216
 49 -  L  V  N  S  Y  L  R  Q  V  Y  V  M  N  K  S  A  Q  Q  T  G  V  T  F  S  - 72

217 - AGTAATGTACTGAACTCGAATGCTCCAGGAGCTGCTAATTTTGCCAGTTGTATAAACTCCGAGTTCACTGCT - 288
 73 -  S  N  V  L  N  S  N  A  P  G  A  A  N  F  A  S  C  I  N  S  E  F  T  A  - 96

289 - ACGTCATCGAGTACGTTATGGTCAATTACATCTCCAGCAGTCAATGGAATATCCTCGTTTACGCAGAGTGTG - 360
 97 -  T  S  S  S  T  L  W  S  I  T  S  P  A  V  N  G  I  S  S  F  T  Q  S  V  - 120

361 - GCATCATATGGTAGCAGGGAAGATACTGCGGTTTTAGGTGGTTCTACAGTGCACAGAGTTGGGGGCTATGAA - 432
121 -  A  S  Y  G  S  R  E  D  T  A  V  L  G  G  S  T  V  H  R  V  G  G  Y  E  - 144

433 - CCATATCAACCACTAACAAGTCACTTGTTTAGTGTAAGCTCAGCAAATCAGGGTCAGCCGGACGAAACATCA - 504
145 -  P  Y  Q  P  L  T  S  H  L  F  S  V  S  S  A  N  Q  G  Q  P  D  E  T  S  - 168

505 - GACTCATTGGACGTGCAGTCTGGAAGTTATCTTACTGTTGGTCAGTTACCTAGATATCCGATTGCTAGT    - 573
169 -  D  S  L  D  V  Q  S  G  S  Y  L  T  V  G  Q  L  P  R  Y  P  I  A  S     - 191
```

(b)

```
  1 - ACGACGTTCTTCCGACTAGGTCGAATTTTCCCTCCTCTTGTGCTATGCGACAAGAGACTTCTGTGTCGACAA - 72
  1 -  T  T  F  F  R  L  G  R  I  F  P  P  L  V  L  C  D  K  R  L  L  C  R  Q  - 24

 73 - ATCAGAGTAAGGAGAACAGTGATGTCATCGTTATCGACCTGGACGATTGATTTTAGTTGCGATTACTTAAGA - 144
 25 -  I  R  V  R  R  T  V  M  S  S  L  S  T  W  T  I  D  F  S  C  D  Y  L  R  - 48

145 - TCGAAACCACGGCGAAGCAAGGATAAGGCATTTTATAAAGAAATACCCCTTATGTATAATTGTATTTTTTAT - 216
 49 -  S  K  P  R  R  S  K  D  K  A  F  Y  K  E  I  P  L  M  Y  N  C  I  F  Y  - 72

217 - TCCTTGGCTTGTCGGAAGACGACAAGCTACGGGGTATTATTGGTAGATTCAAAGCACAACTTTTAATTAACT - 288
 73 -  S  L  A  C  R  K  T  T  S  Y  G  V  L  L  V  D  S  K  H  N  F  *        - 93

289 - CGTGCTGTGTTCGTTAGATCGTTAGCTAGTTAGTGTTCGCTTTAGCAAATGAATACATTGTGTAGGTTGTTC - 360
361 - TAGAGTGATAACGTATTTAGGACGTCTCCTGCTATAGAAACATTGTATATTGAAAGGACATCATAGCCAAAT - 432
433 - GTGGTTAGAAGTAGTTATAGCATTCATGAGTCATTCTGTTAGAATGTCAGCTAACTATAATAGTGTATTGTA - 504
505 - CATGGGTAGTAACATTTTATGTTTTCACAAAAAAAAAAAAAAAAA                            - 549
```

30) Transportin-3 - Clone 31-1

~700 residue missing from the N-terminus of the translated 5'-end of the clone (a). No putative conserved domain was detected; however, closest BLAST match contains an exportin 1-like protein domain (Xpo1; Accession No. pfam08389). Exportin 1 is a nuclear export receptor that interacts with leucine-rich nuclear export signal (NES) sequences, and Ran-GTP, and is involved in translocation of proteins out of the nucleus. Unable to confirm the correct reading frame for the 3'-end of the clone (b).

(a)

```
  1 - GATGAATATGGAAGTAACCCACAGTGTGTTCCAGGACTCATTGAAATGCTTAAGGCATTTTGTGATCCGACA - 72
  1 -  D  E  Y  G  S  N  P  Q  C  V  P  G  L  I  E  M  L  K  A  F  C  D  P  T  - 24

 73 - TTTCGTAGACTGTCACAACAGGGTGGCCTTGTACAGCATCCTGACACAGTAGATGACTTCTTTAGATTGTGT - 144
 25 -  F  R  R  L  S  Q  Q  G  G  L  V  Q  H  P  D  T  V  D  D  F  F  R  L  C  - 48

145 - ATAAGGTTTTTACAACACTGCCCTAGAGAATTTTTCCAAAATGATGCCATGGACACTGTTCTTTTGCTTGCA - 216
 49 -  I  R  F  L  Q  H  C  P  R  E  F  F  Q  N  D  A  M  D  T  V  L  L  L  A  - 72

217 - ATTGCTGCCATTGCTTTGGACCACAGAGAAGCAAATGCTTCCGTTACCAAGTTTTAGTTGATACCATCAGA - 288
 73 -  I  A  A  I  A  L  D  H  R  E  A  N  A  S  V  T  K  F  L  V  D  T  I  R  - 96

289 - TGTTCCGCGGGATCTCGACATCCTGCACAGGAAAAGATCGCACGACAATTGTTAAGCAAACATGGCTTAGCC - 360
 97 -  C  S  A  G  S  R  H  P  A  Q  E  K  I  A  R  Q  L  L  S  K  H  G  L  A  - 120

361 - ATCACCAAAGGGATTGTGAAGGCCTGCACCGGTGGTATACAAACGTTTATGTTACCTGAAGTTGGCGATGTC - 432
121 -  I  T  K  G  I  V  K  A  C  T  G  G  I  Q  T  F  M  L  P  E  V  G  D  V  - 144

433 - CTTTGGGAATTATTAGTATTTGAGAAACAGAAAACTCTTCAGTGGTTGCAAGAGGCATTAAATGATCTTCCA - 504
145 -  L  W  E  L  L  V  F  E  K  Q  K  T  L  Q  W  L  Q  E  A  L  N  D  L  P  - 168

505 - TCACATGGCCCCACTGGTACTATAGTTGCTTCACAGGAGCAAGTTGGACAATTTTACTTGACTGTGTCAAGT - 576
169 -  S  H  G  P  T  G  T  I  V  A  S  Q  E  Q  V  G  Q  F  Y  L  T  V  S  S  - 192

577 - GCTGATTCTGTCAAAGGGATTTGGCGAGCGTTCCGGGACTTCTCC                             - 621
193 -  A  D  S  V  K  G  I  W  R  A  F  R  D  F  S                                - 207
```

(b)

```
  1 - TCATCAATCTCACGTCTCAGCTTACCGTGAAGAAGAGAAGGATGACTCACTTCAAGCTATGAAGCACAAAAC - 72
 73 - AATTTCAACTTATTTTGTTGAGAAGGAAGCTGCACTATTAAGACCAAGGAGGTCAACAAGGACAGAGGAGAG - 144
145 - AGAACGGAATTTATCAGAGACATTTTTGTAATGTTTTATATAATTATGATAAGAGCATCATTCTGACGACAT - 216
217 - TGCTTAACATATTGATTCATAAACATGAAAAAGGAACAGTTTCAGTTTTCAGTGCTCCTAAAACATTTGTTT - 288
289 - GTCACCTGTACTCCATGATACGTTTTTAATTTTGCTTGCCACAGTGCTGTGGAAATATCGATTGTTCTACAC - 360
361 - AATCATCCAATTTCTGGCTTCAGTGAATCTCATTCGAATTTTATTTTACCAAGATTTCATTTAAACATGCCA - 432
433 - AGGTTAGCATACAGGAGACATTGATCAGCCCATTTTTAAAGGAGCAGTTGTATTTTAAATGAGCTAGACTCT - 504
505 - ACCGTTGTTATACTAGGGTTATTTCTCAGGTTGAGGACGTCAGACGTTTGCAGTGAAATGCCCCATGTATCA - 576
577 - TACTGTGGGTTATATAATCTGTTTTGCCATCCAAAAAAAAAAAAAAAA                          - 624
```

31) Translationally-controlled tumour protein (TCTP) - Clone 37-1

Only ~10 residue missing from the N-terminus of the translated clone. Deduced residues 1-156 contain a conserved domain corresponding to the TCTP family (Accession No. pfam00838). TCTP is a protein which has been found to be preferentially synthesized in mammalian cells during the early growth phase of some types of tumour, but which is also expressed in normal cells. The physiological function of TCTP is still not known. It was first identified as a histamine-releasing factor, acting in IgE+- dependent allergic reactions. In addition, TCTP has been shown to bind to tubulin in the cytoskeleton, has **a high affinity for calcium**, is the binding target for the antimalarial compound artemisinin, and is induced in vitamin D-dependent apoptosis. TCTP production is thought to be controlled at the translational as well as the transcriptional level. Close homologues have been found in yeast, plants, nematodes and hydra (see InterPro: IPR001983).

```
  1 - GATGAACTATTTAGTGACAGCTATAAAGTGGAGACTGTGGATGATATCTATTACAAAGTGAAAGGGAAATGG - 72
  1 -  D  E  L  F  S  D  S  Y  K  V  E  T  V  D  D  I  Y  Y  K  V  K  G  K  W  - 24

 73 - GTAACAGAAGTTACAGATTGTAGCAATGTAGATATTGGAGGAAATAAATCAGCTGAAGAACCAACAGAGGAG - 144
 25 -  V  T  E  V  T  D  C  S  N  V  D  I  G  G  N  K  S  A  E  E  P  T  E  E  - 48

145 - GCAGATGAAGTTACCAAAACGTCTGGTTGCAATATTGCCCTGGCAAATAGAATTTCTGAGTTGGGATACACA - 216
 49 -  A  D  E  V  T  K  T  S  G  C  N  I  A  L  A  N  R  I  S  E  L  G  Y  T  - 72

217 - AAAACTGAATATAAGAAGCATATAAAGAAATACTGCCTAAGAATTGTTGAGCACTTACAAGAAAAGGACCCC - 288
 73 -  K  T  E  Y  K  K  H  I  K  K  Y  C  L  R  I  V  E  H  L  Q  E  K  D  P  - 96

289 - AGTAGGGTATCTGTTTTTAAGGCAAAAATGCAAGGTGCTGTCAAGACTTTTCTTGAGAAATATGATGATCTT - 360
 97 -  S  R  V  S  V  F  K  A  K  M  Q  G  A  V  K  T  F  L  E  K  Y  D  D  L  - 120

361 - ACATCAATGTGGGGAGAAAGTTATGACGATACTGCAATGATTCCGTTATTGGAAGTAGAGGATGATGGTACA - 432
121 -  T  S  M  W  G  E  S  Y  D  D  T  A  M  I  P  L  L  E  V  E  D  D  G  T  - 144

433 - TGCTACATGTTCTTTTGGAAGGATGGAGTGGCTGTTGAAAAATATTAATTGGACTGAGCCTGTGCAGCAGTT - 504
145 -  C  Y  M  F  F  W  K  D  G  V  A  V  E  K  Y  *                          - 159

505 - GTCAGATTGGATAAGTGACCGTGCACAATTTTCTGGTGGAACAATTTTTCGTTATTGTAATCTTCTGTCCAA - 576
577 - GAATCATCTTGTGTCAGTTAGGAAATTGTTTAGTTAATACACCCAAATACTTGTCATATTTGTGCTTGAGCC - 648
649 - TTAACGAAATATTGTGACATCCAAATGTATGTTGTTTTAAATTTTCTATATTTTCTTGTTTGCTCCAATTCA - 720
721 - TGCTACCCAAGAGGTACGTTCTTACCTCGACGTGTTAGCACATTGAATCCACAACTTGGCAGTTTGTTTCAG - 792
793 - TAATGCCCAAATACCTTGAACGCTCGTTTTGTAATAAGATATGAACGTTGAATAAGATACAAATCCTTTTGA - 864
865 - TGACATAAAAAAAAAAAAAAAA                                                   - 886
```

32) Calponin - Clone 38-1

BLAST alignments suggest that the translated clone is full-length; however, no initiating methionine is present where expected, near deduced residue 16 of the 5'-end of the clone (a). Deduced residues 40-108 of the 5'-end of the clone and 1-44 of the 3-end of the clone (b) contain a conserved calponin homology (CH) domain (Accession No cd00014). A CH domain is an actin-binding domain which may be present as a single copy or in tandem repeats (which increases binding affinity). The CH domain is found in cytoskeletal and signal transduction proteins, including **actin-binding proteins** like spectrin, alpha-actinin, dystrophin, utrophin, and fimbrin, proteins essential for regulation of cell shape (cortexillins), and signaling proteins (Vav).

(a)

```
  1 - GAAAAGCTGCGTTTGTTGAACAGAACAAGCACATTTATTCCTAAAGATGGCTCAGAGACCAAAGGATTCGGC - 72
  1 -  E  K  L  R  L  L  N  R  T  S  T  F  I  P  K  D  G  S  E  T  K  G  F  G  - 24

 73 - CTTTCGGCAGAAACATCCAAAAAGATTATGGGAAAATATAATCCAGAGCTAGAAGCACAAGTTATTGACTGG - 144
 25 -  L  S  A  E  T  S  K  K  I  M  G  K  Y  N  P  E  L  E  A  Q  V  I  D  W  - 48

145 - ATCAACACGATTTTACAAAAAGACATCACTAATGGAAAATCGGGACCAGATGAGGTTCACAAAGCATTAAAA - 216
 49 -  I  N  T  I  L  Q  K  D  I  T  N  G  K  S  G  P  D  E  V  H  K  A  L  K  - 72

217 - GATGGATCCATACTGTGCTCACTGATGAATGCTGTGAATCCAGACGCAAACCTTAAGATTAATAAGGGCAAA - 288
 73 -  D  G  S  I  L  C  S  L  M  N  A  V  N  P  D  A  N  L  K  I  N  K  G  K  - 96

289 - ATGCCTTTTCATCAGATGGAAAATATTGGCAACTTT                                     - 324
 97 -  M  P  F  H  Q  M  E  N  I  G  N  F                                      - 108
```

(b)

```
  1 - AAGGCACAAAGCAATGGATATGCTGGACCAACTCTCGGACCAAAGGAGGCCGAGCAGAATGTCCGAGAGTTT - 72
  1 -  K  A  Q  S  N  G  Y  A  G  P  T  L  G  P  K  E  A  E  Q  N  V  R  E  F  - 24

 73 - GACGAAGAAACTATTGCAGAAAGTAAAGCACATATCGGTTTACAGATGGGTACCAACAAGTTGGCATCCCAG - 144
 25 -  D  E  E  T  I  A  E  S  K  A  H  I  G  L  Q  M  G  T  N  K  L  A  S  Q  - 48

145 - GCCGGTCAAAATTTTGGTAAAACGAGAGCAATCATCGATTGAAGGGACGAAATCGAAAGTAAATATCTAAGG - 216
 49 -  A  G  Q  N  F  G  K  T  R  A  I  I  D  *                               - 61

217 - AACAATTAAAGAAACATTTCCAACAATCTACAGATCTGAATTTTGCCACACGGAATTTGAACTGTCTTACCG - 288
289 - TCACTACAATTGCACATTTCGAAAATTTTTTAATGTTACTTTGTTAATACAAATGATTTCTCATCTGTGTTA - 360
361 - GTAGTTGTAGTAGATTTGCAGTTTTATGTGCAAAGTTTTCGCCTAATCTAAAACAGTATCTCAGTTACATAC - 432
433 - ACCATGTCTATTCATCGACCTCTTCATTTTGTGCCAGACCAAAGCAGACAACTTTTAGTTGCCGACAGTACT - 504
505 - ATCAAGAATCGATGTTGTGTTAAAGTTGATCGTTTGATATTAATAAACGCACATTCGACTAATTAGTGTAGT - 576
577 - GTCACTTGGTAACGTAAGCAAATTTGATCCACTTGTAGAAAACGGCGACAGCTTGTTATGATTGCCATCCCC - 648
649 - TAATCTGTAAAGTACTTGCCTGATACAAAGGTGTGATGTCTCAAAAAAA                        - 697
```

33) Embryonic-1 - Clone 39-1

A putative signal peptide (20 residues) is underlined (SignalP; Emanuelsson et al., 2007). No putative conserved domain was detected; however, BLAST analysis revealed that the translated clone is similar to Hydra embryonic-1 (HyEMB-1). HyEMB-1 represents a maternal gene that is strongly expressed in the interstitial cell derivatives accumulating in the developing ovary and in early embryos (up to the gastrula stage); however, expression is not detectable in adult polyps (Genikhovich et al., 2006). The role of HyEMB-1 in *Hydra* embryogenesis is unknown.

```
   1 - AAGTTGGTGAGGGTAGGAAACACATTGAATGAAATCTGCTTCTAAACACAAGCTGCTGCCATGTTGCTCAGT - 72
   1 -                                            *                 M  L  L  S  - 4

  73 - CTTATCCTACTAGGAGTTGTATTCCCGTCATATTATTTGGTCAACGCCAAATGCGCATTCTATCATCACACT - 144
   5 -  L  I  L  L  G  V  V  F  P  S  Y  Y  L  V  N  A  K  C  A  F  Y  H  H  T  - 28

 145 - TGGTACAAACCTATGGGATGGTGTCGCCACTCGGGTTCGCAACTTTTTGTGGGAGATTTCAATGGCGACAAA - 216
  29 -  W  Y  K  P  M  G  W  C  R  H  S  G  S  Q  L  F  V  G  D  F  N  G  D  K  - 52

 217 - AGAGATGACATGCTTTGTCATGATACAGCTGGAAGGAAGTACATAGCCTTTGCGAGAAAGGCCGGAACCTTC - 268
  53 -  R  D  D  M  L  C  H  D  T  A  G  R  K  Y  I  A  F  A  R  K  A  G  T  F  - 76

 289 - TTTAGAAATGATTGGTATCGAAATTTGGGCTGGTGCAGACACAAAAATGGTCGTTTGTTCATCGGTGACTTC - 360
  77 -  F  R  N  D  W  Y  R  N  L  G  W  C  R  H  K  N  G  R  L  F  I  G  D  F  - 100

 361 - AATGGAGATAAAAGAAAAGATATGTTGTGTCATGATGTTAAAACTGGCTACAAATGGATTGCATTTGCAAAT - 432
 101 -  N  G  D  K  R  K  D  M  L  C  H  D  V  K  T  G  Y  K  W  I  A  F  A  N  - 124

 433 - AAAAGAGGTCACTTTACTGGGACAGGATGGCATCGCAATCTAAAGTGGTGCTTTCATGCTGGTTCTAAGTTA - 504
 125 -  K  R  G  H  F  T  G  T  G  W  H  R  N  L  K  W  C  F  H  A  G  S  K  L  - 148

 505 - TTTATTGGTGACTTCAATGGGGATGGCAAAGACGATATGTTATGTCATGACACTCATGGACGGAAGTGGATT - 576
 149 -  F  I  G  D  F  N  G  D  G  K  D  D  M  L  C  H  D  T  H  G  R  K  W  I  - 172

 577 - GCTTTTGCTAACAGATACGGACACTTCACAGGTACTAATTGGTATAGAAATATGGGCTGGTGTAGACATGGT - 648
 173 -  A  F  A  N  R  Y  G  H  F  T  G  T  N  W  Y  R  N  M  G  W  C  R  H  G  - 196

 649 - GGCTCACGTCTTTACATCGGCGATTTCAATGGAGATGGAAGATCTGATATGTTGTGCCACGATACTGCAGCA - 720
 197 -  G  S  R  L  Y  I  G  D  F  N  G  D  G  R  S  D  M  L  C  H  D  T  A  A  - 220

 721 - CGAATGTGGATTGCTTTATCAGGATGTCAAGGACATTTTACAACAACCACTTGGTATCGAAGTAACTTCCGA - 792
 221 -  R  M  W  I  A  L  S  G  C  Q  G  H  F  T  T  T  T  W  Y  R  S  N  F  R  - 244

 793 - TGGTGCAGTACAAGTGGTAAGCTCTTCGTTGGCGATTTTAATGCAGATGGCAAATCAGACGTTCTTTGCCAC - 864
 245 -  W  C  S  T  S  G  K  L  F  V  G  D  F  N  A  D  G  K  S  D  V  L  C  H  - 268

 865 - CACAAAAAGACTGGACAAAAGTGGATTGCATACTCCAAGGGAAGAGGACATTTCACTGGTACCGACTGGTAT - 936
 269 -  H  K  K  T  G  Q  K  W  I  A  Y  S  K  G  R  G  H  F  T  G  T  D  W  Y  - 292

 937 - CGTGCCATGGGATGGTGTGCACACTCTGGAGCAAAGATATTTATTGGAGATTACAACGGAGACAAATTGGAC - 1008
 293 -  R  A  M  G  W  C  A  H  S  G  A  K  I  F  I  G  D  Y  N  G  D  K  L  D  - 316


1009 - GACATGCTGTGTCATGACAGCATAGGAAGGAAATGGATAGCCTATGGAGAATGTTATTGAGTTGCGTGAAAA - 1080
 317 -  D  M  L  C  H  D  S  I  G  R  K  W  I  A  Y  G  E  C  Y  *              - 335

1081 - ATCCTGTAGATATTGGCAGATTTTGTTAAGCACGTTTGAGCCGGGCGTCATTTTGTAACATCTGGTTTTGAT - 1152
1153 - GGCAGTTTTATCAAGTTTGTAGAAGCTCTTGCCGTTGCTCAACGAAGAAAATGAGAATCTGGTTAACTGTAA - 1224
1225 - AAAAAAAAAAAAAAAAA                                                        - 1241
```

34) Folliculin - Clone 41-1

~340 residues missing from the N-terminus of the translated clone (a). Unable to confirm the correct reading frame for the 3'-end of the clone (b). No putative conserved domain was detected in translated nucleotide sequences of the 5'- or 3'- ends of the clone. The function(s) of folliculin is not known but it may act as a tumour suppressor to help control the growth and division of cells. It may also be involved in energy and/or nutrient sensing through the AMPK and mTOR signaling pathways (see UniProtKB/Swiss-Prot Q3B7L5).

(a)

```
  1 - CTGCGTCATATGAACATGGTAATTGGAGCTGATAAGATGCGGTTGATTTCCTTCCATGTTGTCCGTGGCGAT - 72
  1 -  L  R  H  M  N  M  V  I  G  A  D  K  M  R  L  I  S  F  H  V  V  R  G  D  - 24

 73 - CAACTCATAGTAAGAGGTAAACAAAAAGGAGCTGTTTTATCGGCATTGCGCGTACTTAAGGAACTGATTCCA - 144
 25 -  Q  L  I  V  R  G  K  Q  K  G  A  V  L  S  A  L  R  V  L  K  E  L  I  P  - 48

145 - GGAGGCTGCTGCAAATGTTGTCCGTACTCGACGTATTATCTGGATTCGTGGAAATTTAATTTCCTTGGCTTG - 216
 49 -  G  G  C  C  K  C  C  P  Y  S  T  Y  Y  L  D  S  W  K  F  N  F  L  G  L  - 72

217 - GTAGATACAGCCAAAATACCTATGCATGTCTTGACATCAGAGCTCTTCGTGCTAATGGACATTGATTGTTTA - 288
 73 -  V  D  T  A  K  I  P  M  H  V  L  T  S  E  L  F  V  L  M  D  I  D  C  L  - 96

289 - CATGAAGAGGTTGACCGGGAAGGATCCCCAGCGACTTCAAGTGAGACGGAAAATAGTTTTGATAACTATCAA - 360
 97 -  H  E  E  V  D  R  E  G  S  P  A  T  S  S  E  T  E  N  S  F  D  N  Y  Q  - 120

361 - ATTCATATCCATGGCTGTCCTGTCAGTGATTGTCCAACGTATCTTTTACGAGTACTGAGCGCCCTGACAGAC - 432
121 -  I  H  I  H  G  C  P  V  S  D  C  P  T  Y  L  L  R  V  L  S  A  L  T  D  - 144

433 - GAAGAATTTACTGATTCGGTTTTTGACGTAATGTTATCAGCCTTAAAGGAGGAATGGATCAACAAAGTGA    - 502
145 -  E  E  F  T  D  S  V  F  D  V  M  L  S  A  L  K  E  E  W  I  N  K  V     - 167
```

(b)

```
  1 - AAGTAGAGATTCTCTAGTTCTGCTTCCAGTAACTTCCAGAGATGCAGTGTCGTGGGGCTAATATAGGACCAG - 72
 73 - TGGATTATATAAAGTTATACTTTTACCATCCATTTGACGAAGAGTTGTTATTGTGCCACCCATTTGATGAAG - 144
145 - AGAAAGAGAAGTAGAGGTGTCCAAAAATTGAAATGCAATGGACGCCGCCTGAGTCTTAATTACAGAATTGCA - 216
217 - AGACCAAGATTTTTTAGATAGAAAACTCCTGGATTGATCGGTTTTTGGTTAGAAAAGGTTGTTAATTGATAA - 288
289 - CATTTCTGTTGTTTGCTGCATATTCGTACACCATCTCATTACAAATGGATTGGTCATGTTTGATGCAGTGCA - 360
361 - TTTGGTGATAATGAGTAGTGATATTCGCACTAGGGGCCCATTTGAGAATGTCGTGGGTATAATAAAGGTGTG - 432
433 - AAGCCTCCACCGGCAATATGCGAATCACATTTTATGTAATTGTTTCAGAAGTAAAGTTTGCTTCCATTTTTA - 504
505 - TCTTTGTACTACTGATATGGCTAAGAAAAAAAAAAAAAAAAAAAAAAAAAA                       - 554
```

35) Serine/threonine protein kinase - Clone 46-1

More than 250 residues missing from the N-terminus of the translated 5'-end of the clone (a). Deduced residues 4-195 of the 5'-end of the clone feature a conserved serine/threonine protein kinase, catalytic domain (S_Tkc; Accession No. cd00180). Members include phosphotransferases of the serine or threonine-specific kinase subfamily. The enzymatic activity of these protein kinases is controlled by phosphorylation of specific residues in the activation segment of the catalytic domain, sometimes combined with reversible conformational changes in the C-terminal autoregulatory tail. Unable to confirm the correct reading frame for the 3'-end of the clone (b).

(a)

```
  1 - AACAAAACTCTCTTTCTTGTGATGAAGAAGTACTCCTGCACTCTCAGTCAGCTGCTACGCTCATACGCAGAC - 72
  1 -  N  K  T  L  F  L  V  M  K  K  Y  S  C  T  L  S  Q  L  L  R  S  Y  A  D  - 24

 73 - ACAAACTATGATTATTCACTCCCTTTACTTGTGCAACTGCTGAACAGTCTTGCTTTCTTACGCTCACACAAC - 144
 25 -  T  N  Y  D  Y  S  L  P  L  L  V  Q  L  L  N  S  L  A  F  L  R  S  H  N  - 48

145 - GTAGCACATAGAGATCTCAAAAGTGACAATATTCTTGTGGAAATGAGCAATAATTCCTATCCTGTGCTTGCG - 216
 49 -  V  A  H  R  D  L  K  S  D  N  I  L  V  E  M  S  N  N  S  Y  P  V  L  A  - 72

217 - TTGACGGATTTTGGGTGTTCTTTAAGGGACGATAACTATGGCCTAAAAATTCCGTACTACACAAGTGAGATG - 288
 73 -  L  T  D  F  G  C  S  L  R  D  D  N  Y  G  L  K  I  P  Y  Y  T  S  E  M  - 96

289 - TCAAAAGGGGGCAATCCCGCCTTAATGGCACCTGAGATTGCAAATACTTCAGCTGGTCCAGGAAAATTTCTT - 360
 97 -  S  K  G  G  N  P  A  L  M  A  P  E  I  A  N  T  S  A  G  P  G  K  F  L  - 120

361 - GATTATGACAAAGCTGATGTGTGGGCTGCTGGTGCAATCGCATACGAAATCTTTGGGCGAGAAAATCCACTG - 432
121 -  D  Y  D  K  A  D  V  W  A  A  G  A  I  A  Y  E  I  F  G  R  E  N  P  L  - 144

433 - ACTGGCACTGCCAAGAATGAGTTTCAACTAGTAACAGGAGGTAAGAACTCATTTGTGGATAAACTTATTAGC - 504
145 -  T  G  T  A  K  N  E  F  Q  L  V  T  G  G  K  N  S  F  V  D  K  L  I  S  - 168

505 - TTAATGCTCCAGCCCAATCCTAATCGAAGAATATCGGCATCAGTTGCAGCAACTTACGCAAATCTGCTGCTC - 576
169 -  L  M  L  Q  P  N  P  N  R  R  I  S  A  S  V  A  A  T  Y  A  N  L  L  L  - 192

577 - TTTTCTCCCTGGTATAACAATACGATCGAAATGCCAACGGAAGGACAGGTTAGGAGGTGGCTGGAGCTTGCC - 648
193 -  F  S  P  W  Y  N  N  T  I  E  M  P  T  E  G  Q  V  R  R  W  L  E  L  A  - 216

649 - AAACGGCTGACACTGTTTGAAT                                                   - 670
217 -  K  R  L  T  L  F  E                                                     - 223
```

(b)

```
  1 - AGTAGCATTTACAAAGTGCAAGACTTTTATTATGTTGTTGATTAGCAAAAAGTCAGGTCGAGCAATGTTGAC - 72
 73 - TGTAAACCTTTACGTTTCTACCATATCCACGTCAACACCAAGGCCAATGTTGAAATTCTAGGCCAAGTGGAA - 144
145 - GGACCAGTGTCTCAATACTAAGGTTTGCCATAGGCTATTCATTGGACCCCCAACACTGAATTGCAAAGCTCA - 216
217 - AGTGTATATGGCAATGTATCACTCAAACAAATTGATCACAAATTAATCAATCAATTAATCAATCAGCCAACC - 288
289 - AATAAATTACTGAATCAATAACAGTGGGTACTCGGAATAATAAACAACAAACACATCAATGTCAACAATAAC - 360
361 - AGCAACAACAGAAACAACAATAGCAAATTAACAAGCATCGAATGGTCAAATACAATAACTACCGGAGTGAGC - 432
433 - AAGCATACAATAAATTTGTAATATATGCTCATCTGTTTTAGCTTAGATAATAGTAGTGACGATGTAAGTCAT - 504
505 - TTAAATACGTACGTGCACAAATTAACAAAAACATAATAAAAGCATAGTCAAAACATAATAAGCTTATTAGTC - 576
577 - AATAATAAATCATACACCAAATATCATCTACATCAGAGAAAATCGATGTATCATATATTCGATACAAAA    - 645
```

36) Predicted Protein - Clone 3-3

No putative conserved domain was detected in the translated clone.

```
  1 - CTCACAAGAAGGAAAAGCGATTGGCAGATTGAATCAACCAATTTAACACCATGTCCTGAAACTACTCATTTC - 72
  1 -  L  T  R  R  K  S  D  W  Q  I  E  S  T  N  L  T  P  C  P  E  T  T  H  F  - 24

 73 - TTTACCATAGTGCAATTTGCAGTGGTGTTCAGTGCTGAAAATGAATTAGAAGAGGTCACCGAAGATCCAGCT - 144
 25 -  F  T  I  V  Q  F  A  V  V  F  S  A  E  N  E  L  E  E  V  T  E  D  P  A  - 48

145 - GTTGTTAAGAAGATCATCAGGGTCCATACGACAACAGTTATCCACAAGAGTGGGTGGCTCAAAATAAATTCC - 216
 49 -  V  V  K  K  I  I  R  V  H  T  T  T  V  I  H  K  S  G  W  L  K  I  N  S  - 72

217 - AGACCAGTATGTTTTGGTGCAAAAGGAAACACCTATGGATCTTTCAAATCCCTGTACACTAAGCGCGCCCAT - 288
 73 -  R  P  V  C  F  G  A  K  G  N  T  Y  G  S  F  K  S  L  Y  T  K  R  A  H  - 96

289 - CAGGTGAGACTGGTTCGTGTCTCTGGATATTTGACATGTAACTCCAACCACAACAAAGCTTTCTTCGACTGT - 360
 97 -  Q  V  R  L  V  R  V  S  G  Y  L  T  C  N  S  N  H  N  K  A  F  F  D  C  - 120

361 - GGGTCAGCTGCTGCTCATTACTTTTCGGTCTTTCTCACCACGACTAGGAATGCTAAGGGAATCATCATTCCC - 432
121 -  G  S  A  A  A  H  Y  F  S  V  F  L  T  T  T  R  N  A  K  G  I  I  I  P  - 144

433 - CCCGCACCTATTCACTACCATGCCAAATACCATGGATGGTATTACCTTAACGGACACAAACCAGGTACGTCC - 504
145 -  P  A  P  I  H  Y  H  A  K  Y  H  G  W  Y  Y  L  N  G  H  K  P  G  T  S  - 168

505 - GTCATCTCTCAAACTACCGGTGCCCATGGCTACACTTTCGTCAAGGGACATGCCTACCACATCTGGTACGGT - 576
169 -  V  I  S  Q  T  T  G  A  H  G  Y  T  F  V  K  G  H  A  Y  H  I  W  Y  G  - 192

577 - GAAGATTTGTTCAACTGGACCGAAGGAGATAACGGAGGAAGAGTCTGCGTTAACGTGTACCTTTTGTAATGG - 648
193 -  E  D  L  F  N  W  T  E  G  D  N  G  G  R  V  C  V  N  V  Y  L  L  *     - 214

649 - ATGACCACGAAGGACTCGTAATTGGCTTAAGACTGGCTCCACTAATATACGAATCTATCTTATCTGGAAGCA - 720
721 - ATGTGTCAGGCATAAAGCTTAGGGATCATGTAAACTTAAACTGAACTCCAGTTGACAGGACACTTTTCTAAA - 792
793 - CAGCATTTTTACAAGTTTTGGCATTACTCCATGTTAAAAATAGCATCTGTCCAATTTCTCATTTTAAGTCCT - 864
865 - GACACAGTCATATTTTCGACGCATTGCACTTTTCTAAATTGTCAGTTCTCTTTAGAATTTCTTAAACCAGCA - 936
937 - ATGAATAAAGCAAAAGCTTTGAGATAAAAAAAAAA                                      - 971
```

37) Predicted Protein (function unknown) – Clone 3-7-1

No putative conserved domain was detected in translated nucleotide sequence.

```
  1 - AAAGTTTTAAAGCGCTACAATTGTACTGAGCTGCGACGGGCAGCACAGAATGTGACACGGATTACCACAGGA - 72
  1 -  K  V  L  K  R  Y  N  C  T  E  L  R  R  A  A  Q  N  V  T  R  I  T  T  G  - 24

 73 - GTAGGACCACATCCAAAACAGCCGATTTATGGCGAACGTGCTCTGGATATGGTATTTGGCATGGATCAGTTC - 144
 25 -  V  G  P  H  P  K  Q  P  I  Y  G  E  R  A  L  D  M  V  F  G  M  D  Q  F  - 48

145 - TTACCAAATGAAAAAGAATTCAGTTTAGCGCATTTCTTTGGAGAAAAGCCGGTGATGCGAATGACAACTTTA - 216
 49 -  L  P  N  E  K  E  F  S  L  A  H  F  F  G  E  K  P  V  M  R  M  T  T  L  - 72

217 - GCCTCACCTGCTATGGTTGTTCAAAGGTTAAAGAATTTATTTGGGGAAGACTCACAGTTTACAGAGGCACTA - 288
 73 -  A  S  P  A  M  V  V  Q  R  L  K  N  L  F  G  E  D  S  Q  F  T  E  A  L  - 96

289 - GCTAAGAAAATGTTGGCAGTCATGCTTGAAGATTTGCATAAGAATCGTAAAGAGGAATACATGAGCGCAGTT - 360
 97 -  A  K  K  M  L  A  V  M  L  E  D  L  H  K  N  R  K  E  E  Y  M  S  A  V  - 120

361 - GGTCTCGCTCTCCTTTTTGACCGTGCTGGAGGCCAAATTACAGAAAATATGAGCAAAGTCATCGCATCAGTT - 432
121 -  G  L  A  L  L  F  D  R  A  G  G  Q  I  T  E  N  M  S  K  V  I  A  S  V  - 144

433 - GAATCTAAAAATGTTCACATCAAAGAACTCATTGCCGAGATTCGCCAAATGTGGGACGAATGGGACTTGAGA - 504
145 -  E  S  K  N  V  H  I  K  E  L  I  A  E  I  R  Q  M  W  D  E  W  D  L  R  - 168

505 - GATGGACAACTCGATAACTATCTAGAGTTTGATGCCTTCTACAATGGCTTTATGGCGCCATATTTTGGCTGT - 576
169 -  D  G  Q  L  D  N  Y  L  E  F  D  A  F  Y  N  G  F  M  A  P  Y  F  G  C  - 192
```

```
 577 - TATCGTTGTGATGAGACTAAGAAAGCCTTGAAAGCTATTGATATGGATAGTGATGGGCGTGTGGATTGGGAT - 648
 193 -  Y  R  C  D  E  T  K  K  A  L  K  A  I  D  M  D  S  D  G  R  V  D  W  D  - 216

 649 - GAATTTGCCCTCTACTTGAAATGGGCTGGAAGACAATACCCAAACGTCAAGGATGCTGAGGAGCTATTGTCA - 720
 217 -  E  F  A  L  Y  L  K  W  A  G  R  Q  Y  P  N  V  K  D  A  E  E  L  L  S  - 240

 721 - ATTGCATTCCGCAAAGGTTTGGTGCCTGCTATGCAGGACGAAGTTGTTAAGGAGACAGACAAGTTGCCAGCT - 792
 241 -  I  A  F  R  K  G  L  V  P  A  M  Q  D  E  V  V  K  E  T  D  K  L  P  A  - 264

 793 - AGGAAGCAAATTCCACCAGAGGATGATGATGATGATTTGCATTTGTACTTTGGCGATGACGCAGAAGACGAG - 864
 265 -  R  K  Q  I  P  P  E  D  D  D  D  D  L  H  L  Y  F  G  D  D  A  E  D  E  - 288

 865 - GAATTTTTTGGAATGCCTTAACTATGTAGACTTGATTTGGATGGACAGTTTGACGCTGGTGATAATTATATC - 936
 289 -  E  F  F  G  M  P  *                                                      - 294

 937 - ACCATTATGACTAGAAATTGGCCAGCAGCAAAATGGTCTGGTTAGTGGGCTCTCTAGGGGTTCTTTGTTTCC - 1008
1009 - TTTTCTAGTATAATAGTGATATGTAGAGTCGCATGCAAATAATGCTTGTCCTTTTAACAAGTGCTCCCGTAA - 1080
1081 - AAATTCCATTATGCAATAGTTGTAGTAACGTAAGTACAGAAGCTACAAAATAATATCTTGCATCCCAGAAGT - 1152
1153 - GACAGAATACAAAAAA                                                         - 1168
```

38) Predicted protein - Clone 20-1

Translated residues 58-101 coincide with a neurotransmitter-gated ion channel transmembrane region (Accession No. pfam02932). This family includes the four transmembrane helices that form the ion channel. At least 30 residues are missing from the N-terminus of the translated clone. The 3'-UTR seems abnormally long, but the 3'-end of the coding region matches the C-terminus of *N. vectensis* and *H. sapiens* homologues.

```
   1 - GAGTATATGCATCACAATCAAGAGCCTGGTGGGAGCGTTTTGTTAACAACAACAACAGAAACATCAACGGAT - 72
   1 -  E  Y  M  H  H  N  Q  E  P  G  G  S  V  L  L  T  T  T  T  E  T  S  T  D  - 24

  73 - AGTAGTGACGTAAGCGATGCCATAGTTTGTTCTTTAGATGATACGGCAATATGTGAAGATATGTTTAAGAAA - 144
  25 -  S  S  D  V  S  D  A  I  V  C  S  L  D  D  T  A  I  C  E  D  M  F  K  K  - 48

 145 - AAATTGTGGATTGAATTGAAAACCTTGAGCAACAGAGCTCAAGAGGAAACAAAACGGGAAAGAATTTCAGGT - 216
  49 -  K  L  W  I  E  L  K  T  L  S  N  R  A  Q  E  E  T  K  R  E  R  I  S  G  - 72

 217 - GAATGGAAAAGGACATCATCAATTATTGATCGCTTGTTTTTTCTCACATTTACGCTATCACTTTGTGCAGCA - 288
  73 -  E  W  K  R  T  S  S  I  I  D  R  L  F  F  L  T  F  T  L  S  L  C  A  A  - 96

 289 - ACAGCATGGCTTTTTTTGACTCCATCTAAAATCCAGATTTAAGTTTGAAGTAGTTTCTTTGTAGGTCTCCGT - 360
  97 -  T  A  W  L  F  L  T  P  S  K  I  Q  I  *                                - 109

 361 - CAGTAGCATATACACGGGCAGTCAGACACCTTTGTTTATCGGGGCATCAGCAGCGAATCCAGCTCAAATTGT - 432
 433 - GAGGGGTGCTGAAGTGTCTGATAATCCTTGTACACTGCTCAGTCCAGTTTCCAAAAGCGTAAAAATTTTCAA - 504
 505 - AAATTTTGAATATGCTAGACAATATGGGGTGCTTCCACGAAACGTAAGCATCCGCAAAATCGAAAGCAGATT - 576
 577 - GGCAACTTCCCATTCGGTGGTGCTCAAGCCCTGTTATTTGACTGGATAAAGTCGATTCTGAGGTAATTGTAA - 648
 649 - CTCGGACACGAGATGCTCGAAAGCTTTCATTTCTACAGAAATGTACAGTTTGATCCGTGGATTAATCAGATG - 720
 721 - TCAATGACCGGAAATTGCAGTGGAGATTTGCTGAGGAAAGACTACCAAGGGCTTTTCTTGTAGTGAGTGACA - 792
 793 - TTGATTGAAATTCACTACTTCTGGTTTGTAAGGTTATTATCGTATAACACAGACCACGGCCTCCCCCATCAT - 864
 865 - CTAGTAATTAACAACCCAAAACCACTGTTACTTTGAACCTCGGCCATCACTCTCTAACCGTGCTCTCAAAGC - 936
 937 - AAGATTGTTACTCAGTGGCATTCGACATTACATGTATCTTTGATAGCCAAGAAATACGTATCAATACTTTGG - 1008
1009 - CAATGATTGTAGTATTTCTGTCAACAATAAATTGTCTCTTAGTTAATGATTAGAGTTGAACTTGATGAACAT - 1080
1081 - TGACCTTCAAGTTGACGTAGCTCCATGAGTGATGAAATGTACTCCAGTCTCTTTTTTATGGAAAATTTGTAA - 1152
1153 - TTGAAACTGTAAATATAATGCAAGAAATAAAAAAACATATTGAAATAAAAAAAAAAAAAAAAAAAAA      - 1219
```

39) Clone 1-8-1

Potential initiating methionines are highlighted in bold. No putative conserved domain was detected in the translated clone, however, an unusual repeat pattern (consensus RRRRL/FY) is observed.

1 - TCGATCAAG**ATG**AAGCGATTTCGACTGACCGGCGTGTTCCTGTTCTTACTCTTGGCAGTTGCAGCTCATGCT - 72
1 - S I K **M** K R F R L T G V F L F L L L A V A A H A - 24

73 - GCCGTAGGACACAAAGAAGTAGTTGATGTTAATGATGGAGCGGTGGTTCCTGAGGAAGATGACGAGCACGCC - 144
25 - A V G H K E V V D V N D G A V V P E E D D E H A - 48

145 - CTCGACGCAGCAGATGAGGGCAATAATAAGGCAGAAGTG**ATG**AACGAGCCAAAAGTGTCATTTGATGCACTT - 216
49 - L D A A D E G N N K A E V **M** N E P K V S F D A L - 72

217 - GCAGGGCGGAGACGCCACTACAGACGCAGACGACTCTACAGACGCAGACGATTGTACAGACGCAGACGATTG - 288
73 - A G R R R H Y R R R R L Y R R R R L Y R R R R L - 96

289 - TACAGGCGCAGACGATTGTACAGACGTAGACGATTGTTCAGGCGCAGACGATTGTACAGACGTAGACGATTG - 360
97 - Y R R R R L Y R R R R L F R R R R L Y R R R R L - 120

361 - TTCAGGCGCAGACGATTGTACAGACGCAGACGATTGTTCAGGCGCAGACGATTGTTCAGACGCAGACGATTG - 432
121 - F R R R R L Y R R R R L F R R R R L F R R R R L - 144

433 - TACAGACGTAGACGATTGTTCAGGCGCAGACGATTGTACAGACGCAGACGATTGTTCAGGCGCAGACGATTG - 504
145 - Y R R R R L F R R R R L Y R R R R L F R R R R L - 168

505 - TTCAGGCGCAGACGATTGTACAGACGCAGACGATTGTTCAGGCGCAGACGATTGTTCAGGCGCAGACGATTG - 576
169 - F R R R R L Y R R R R L F R R R R L F R R R R L - 192

577 - TACAGACGCAGACGATTGTTCAGGCGCAGACGATTGTTCAGGCGCAGACGATTGTTCAGGCGCAGACGATTG - 648
193 - Y R R R R L F R R R R L F R R R R L F R R R R L - 216

649 - TACAGACGCAGACGATTGTTCAGGCGCAGACGATTGTTCAGGCGCAGACGATTGTACAGACGCAGACGATTG - 720
217 - Y R R R R L F R R R R L F R R R R L Y R R R R L - 240

721 - TTCAGGCGCAGACGATTGTTCAGGCGCAGACGATTGTTCAGGCGCAGACGATTGTTCAGGCGCAGACGATTG - 792
241 - F R R R R L F R R R R L F R R R R L F R R R R L - 264

793 - TACAGACGTAGACGATTGTTCAGGCGCAGACGATTGTACAGACGTAGACGATTGTTCAGGCGCAGACGATTG - 864
265 - Y R R R R L F R R R R L Y R R R R L F R R R R L - 288

865 - TACAGACGCAGACGATTGTTCAGGCGCAGACGATTGTTCAGGCGCAGACGATTGTACAGACGCAGACGATTG - 936
289 - Y R R R R L F R R R R L F R R R R L Y R R R R L - 312

937 - TACAGACGCAGACGATTGTACAGACGCAGACGATTGTACAGACGCAGACGATTCTACAGGCGCAGACGAGTC - 1008
313 - Y R R R R L Y R R R R L Y R R R R F Y R R R R V - 336

1009 - TAAAGGCGTTGATGATCATACAGGCGCCGAAGAGTTTTCGTCCAATGACAGCATTACAGATGCAGATGGTAT - 1080
*

1081 - TAAAGGCCAGGAACGTCCCTTGTAGACAGACAATATTTTTCCTTCTTAATCAATGACATTTTATTCTATTAC - 1152
1153 - ACTTACTGGTAGTTAAACTAGCGCGTCTGCACTAGGGTTTTGTAACCAATGATGCTTTATTCATGCTCTTAG - 1224
1225 - CTTGAAAGCATTATTTTGTATCGCATATAGACTTTGCTTATGAACTAGATGAGCTGATGCTGTTTTTGTTAC - 1296
1297 - AGTTGCTGTTTAGTCATTTGAGTTTCGGTGATCTAGGTTAATCGGTGTTTTTATAATTATTTGAGAAGGAAA - 1368
1369 - TATATTCTGAGTAAAAAAAAAAAAAAAAAAA - 1399

40) Clone 3-3-3

No putative conserved domain was detected in the translated clone.

```
   1 - GCAAAAGACATTGCAATTGAACTAAAAGCTCTTGACTTGTTAGAGATTATCAATGAACATGAAGAGAGACTG - 72
   1 -  A  K  D  I  A  I  E  L  K  A  L  D  L  L  E  I  I  N  E  H  E  E  R  L  - 24

  73 - AATCCTAATTACGAAAATCAGCCACAAACAGTTTTCTCTTTTCCAAGAGCAAGATTTGATCATGTACCAACC - 144
  25 -  N  P  N  Y  E  N  Q  P  Q  T  V  F  S  F  P  R  A  R  F  D  H  V  P  T  - 48

 145 - TTATTTAGAAGATCTTCTTCCCTTGTGTCTGGTAAATTAACCCTAAAACCAATTTTGGAAAACGCAAATAGC - 216
  49 -  L  F  R  R  S  S  S  L  V  S  G  K  L  T  L  K  P  I  L  E  N  A  N  S  - 72

 217 - GAACATGATACTGAATCAGAGGGGGAGGAAAATGGGTATGAATCACTCGATGAAGGAAATTTGAGTGATCAT - 288
  73 -  E  H  D  T  E  S  E  G  E  E  N  G  Y  E  S  L  D  E  G  N  L  S  D  H  - 96

 289 - AGCTCACAAGGATATACAGACAGTGGAAGTATTGATAGCGTGTTTACAGCCGTGACACTTAGCTGTGGATCT - 360
  97 -  S  S  Q  G  Y  T  D  S  G  S  I  D  S  V  F  T  A  V  T  L  S  C  G  S  - 120

 361 - ACACCTATGAATGGAAAGACCGTTGTAGGAGGATTCAGGAGAGAGCGGCACTTGTCACTAAGGAGTTTTAAT - 432
 121 -  T  P  M  N  G  K  T  V  V  G  G  F  R  R  E  R  H  L  S  L  R  S  F  N  - 144

 433 - AGTAATGACAGTTTGATTCGGATGACAAAAACGTTTAATCAAAATGGGCGTAGCAACAGTGGTTTTTCTGGA - 504
 145 -  S  N  D  S  L  I  R  M  T  K  T  F  N  Q  N  G  R  S  N  S  G  F  S  G  - 168

 505 - AAAGACGTATTTAACATTCCCAAAACAACTATTTGTCGTACAAGATCACAAGAATCCTCATCACTGAAGTCT - 576
 169 -  K  D  V  F  N  I  P  K  T  T  I  C  R  T  R  S  Q  E  S  S  S  L  K  S  - 192

 577 - ACTCTTGTCAGGAGGCTGTCGTTAGTTGATGACTCTCTCCTAGGCAATGGACTAGGCTCCAATGGCTTTTTT - 648
 193 -  T  L  V  R  R  L  S  L  V  D  D  S  L  L  G  N  G  L  G  S  N  G  F  F  - 216

 649 - GACGTCAAGGAGGAAGACGAATGTGTTAACAATGACCAACTGTCTTACGCAGTTGACCAAAATAGAAATATT - 720
 217 -  D  V  K  E  E  D  E  C  V  N  N  D  Q  L  S  Y  A  V  D  Q  N  R  N  I  - 240


 721 - TCCTTTACTCCAATAAAGGAGACTATAGAATATGATTTCATAGACAGGATAAGTCTTTCTTCTAATGACTCA - 792
 241 -  S  F  T  P  I  K  E  T  I  E  Y  D  F  I  D  R  I  S  L  S  S  N  D  S  - 264

 793 - AATGAAAATATTAAAATCCCTAGACTTCGAAAAATAGCATCGAGTGCATTTTCTGTCGGAAAAGCATTTAGC - 864
 265 -  N  E  N  I  K  I  P  R  L  R  K  I  A  S  S  A  F  S  V  G  K  A  F  S  - 288

 865 - AGTGGAGAAAACATATCGTGTGATGAAGATATTCTGTCTATTTTACGCCCGGAAGAAAGAGGCTCGCTTGAA - 936
 289 -  S  G  E  N  I  S  C  D  E  D  I  L  S  I  L  R  P  E  E  R  G  S  L  E  - 312

 937 - AAATTGCCGGTTCAAGGATCAACAAATTTACAGTGGTGGAGTGATTTACAGCCTTTACCAATATGTCCTCAC - 1008
 313 -  K  L  P  V  Q  G  S  T  N  L  Q  W  W  S  D  L  Q  P  L  P  I  C  P  H  - 336

1009 - TTAAACTCAATGACCAAATTCTGATGCTCACACTCCACAAACTATAAGTTTGATTCATTTCAGTGGCATATC - 1080
 337 -  L  N  S  M  T  K  F  *                                                   - 343

1081 - CACAGGAGGAAGGAGGGGGCGCTGACCCTCACTTTTGAAAGACGTGAGATCTGGAACTAACATTGACAAGAA - 1152
1153 - TAGCTGAAAGTACTTGGCCTTATAGCATCTTTCTTGAGATAAACCTCAGAACCCCCTACCCCCAGTTTTCTA - 1224
1225 - ATGCTCCATACGTCACGTAATAGAAGCTAGAGGAAGCAGAATATTTATTGTAATGTTTGGAGTAGGGTTTGT - 1296
1297 - AAATATTATATCATCGTCTTCTTTCCTTGCGAAGAAGGAACTATAATATATTTTTAGATAAATCAACATGTT - 1368
1369 - TATTTTGCTACTTTTAATTAAAATTAATAGAAGATCTTAAAGTCCATATTTTCCGTTTATTATGTTTATCAG - 1440
1441 - ATGGTTCTGTTAAAATTTCTAAGATGTAAATGTCATACTGAAATAACGAGGCAATGAAATGAAAAAAAAAAA - 1513
```

41) Clone 30-1

No putative conserved domain was detected in the translated clone.

```
   1 - TTGCATTCCATTGGTGCAGTTGGTAACATTAAAATGGGTCTTCGTGGTTCATTTGCATTGATTGGTTTCAAA - 72
   1 -  L  H  S  I  G  A  V  G  N  I  K  M  G  L  R  G  S  F  A  L  I  G  F  K  - 24

  73 - GGAACTGGAAAGAAACCAAATTGGGTCAAACAATTTGTTGCCAAATCAGGACATGGACCAGCAAGAATTATG - 144
  25 -  G  T  G  K  K  P  N  W  V  K  Q  F  V  A  K  S  G  H  G  P  A  R  I  M  - 48

 145 - TACCAATTGTGGTTATTTGGATTCAGACTAACAGGTCTCAAGCAGATTAAAATTGGCGGGCAGATTGGTGGA - 216
  49 -  Y  Q  L  W  L  F  G  F  R  L  T  G  L  K  Q  I  K  I  G  G  Q  I  G  G  - 72

 217 - CACATAAATGGTGGAACGATACTGAAGAAATTTAAAATTACTGGACCAATGGGTGGTGGACATTTTAAAGTC - 288
  73 -  H  I  N  G  G  T  I  L  K  K  F  K  I  T  G  P  M  G  G  G  H  F  K  V  - 96

 289 - GGTAGCCTGAATCTTGGAAAAATGCCAGCTGGATTCCACTTCAAGAAAATCAAGATTGGTGGACAAGGTAAA - 360
  97 -  G  S  L  N  L  G  K  M  P  A  G  F  H  F  K  K  I  K  I  G  G  Q  G  K  - 120

 361 - CTAGGGCATCTTGGAGTTGGTACCATTAAGCTTGGAAAATGCCAGCCGGATTCCACTTCAAGCAAATCAAGG - 432
 121 -  L  G  H  L  G  V  G  T  I  K  L  G  K  C  Q  P  D  S  T  S  S  K  S  R  - 144

 433 - GTGGTACCATTAAGCTTGGAAAAATGCCAGCCGGATTCCACTTCAAGCAAATCAAGGTTGGTGGGCAAATTA - 504
 145 -  V  V  P  L  S  L  E  K  C  Q  P  D  S  T  S  S  K  S  R  L  V  G  K  L  - 168

 505 - AAGTAGGCAGTCATGGAGTTGGTGGGAAAGTTAACGTTGGTCATCTTGGAGGCGCAATCAAACTCTTCGGTC - 576
 169 -  K  *  A  V  M  E  L  V  G  K  L  T  L  V  I  L  E  A  Q  S  N  S  S  V  - 192

 577 - GCAAAGACGTCCCGCATCCCGCGAGAGCCAATGACGTCGCTGATGGAACTGAGTACACTGACAATGACGTCG - 648
 193 -  A  K  T  S  R  I  P  R  E  P  M  T  S  L  M  E  L  S  T  L  T  M  T  S  - 216

 649 - CTGATACTACTGAATCCATTGACAACGCGGTGAACAAAATTGCTCATTAACCGGAAATACTAGTCGAACTAT - 720
 217 -  L  I  L  L  N  P  L  T  T  R  *                                          - 169

 721 - TTTCATGTTGTGTATTTTCTATTTAAACTAAGTATGGGGTTTCTTTAGCTGTCCAATCCTGTTTGTCCTTTA - 792
 793 - AAGTATTGTTTTGTCAACATGCACAATCGTACGTCTTCCACATGTTGGAGCTTAGGAAGTATTTTAATAAGC - 864
 865 - ACACACTACCAATGTATGTTTTTGCTGTTAATGTCTAGTTTAATGGCTATTAATGGCTCTTGTAAACAGCCA - 936
 937 - TGGACTCTGAAAAATATAATAAAAATGACCTTGTAAAACATTAATTATAGCTTTGGGAATAATTAAAGAATC - 1008
1009 - GTTTGTTTGAAAAAAAAAAAAAAAA                                                - 1033
```

42) Clone 1-1-1

No clear ORF was detected.

```
   1 - GCCGCCTTTTACATAAAGAGCTAACCTGTTCATCCTTTGGACTTTTTGGTTGACCTGACTTGGACATTTTTT - 72
  73 - CCAGGTCGGTAACATACCTCGTCCCGTCGTCGCTAACGCGATATGAATTCGAGGGACCACTAACACCGACGT - 144
 145 - TCGACAACCGAACGAGTGGAGCAAATGTCTGGTGCCGGCCCACCAGCGAAACAGTTGAAGGGTTACCTTTCA - 216
 217 - TCTGTGGCAGCTGCAGCTCCTGATTTTGTATTTTGAAGGAGGTGATGGAAACTGTTTTGCGTCTTTGCTTGA - 288
 289 - AACGTTTCATCCGTTCTGTTACGAAGATATTTTACGTTTTCGGAGCATGTTACTCGTCGCCCGCGTTCTCGA - 360
 361 - GCTGAAAATGTATGACCCGATGGCCTACCTGTAAAATGAATTTAGTCAAAATACCATGAAGTTTCGACAGTT - 432
 433 - AACAGCACCCGATGGTTTTTTTCGGTTGGCAGTTACAAAAGTTTAATATTTTTTTGCGATCATTGTATCTGT - 504
 505 - TGATTTTGTGAATAGCTAATGATATTAACGAACTGTGTACATAGCGCAATAAATCACTGAGGCTTAAAGCCA - 576
 577 - CGTAGATACAAAATGCGATCGGACATCTTTTCCGAGTCCTGTATATATTGAGAATTGTCAAAAGAGCAAGAG - 648
 649 - AAGCTATAGTGCAAACATTAAAAAAAAAAAAAAAAAAAAAAAA                              - 691
```

43) Clone 1-4-1

No clear ORF was detected.

```
  1 - CTCATTTTTAGCGCCCGCACCCGAAATATTCCCATTGACAGCGTTGGAAAAAATTAAACTAAGTTTAATATG - 72
 73 - ATGTATTTTATTAGAAAGTAGGCCTGATTAACACCAGTTGATATCACAATCCTTCCCTAGCCATCCGTAGGG - 144
145 - CTTAAAATGGATTTCGAATTTTGCAAAGAAACGAATTTATTGCCGCGCTCGAATCGTCGCGTCTGACGCGCA - 216
217 - CTTATGATGGTCGTGGCGTAATATTTTTAATATTTTAACGTTTTCTTTCGTTTTTTATTTGTTTGCACAGCT - 288
289 - TTCTGTAGGAGTGACAAGTTTATAATTTCCCAGGCATTTTTTATTTGCAGAATGTTTGGAAATAGTGGATTA - 360
361 - TTTGGTTGTGCAGAATATTTTGTTTTAGAATAGAAAGCTATCTGTAGAGTATTAAGGTCTTTTGGTTTGGCT - 432
433 - TTGCTTTCTTATTTCAAGTTTGTTGTGATCGAAAAGAATTCGAAAACAAAACCCTCGCAATTTCAGGGTTTT - 504
505 - CAAGTGTAATTTGACATGTACATATAAATAATTGCTTAGAGGTAGTCTTAAAAAGGATATACTGATTTAGAC - 576
577 - GTTTACTTTTACTGTGTATATAAAGTTTCTATTTGCATCTAGCACTGCTGATAGATTCGGGCTACATTTTCC - 648
649 - GACTGTCAGTCCGAGTTCTATTGCCTACGCCCAGTTTCAAACTTTACTTTCAAGTGTGCTGTCCAATTTAAA - 720
721 - GTGTTTGTAAATAGTAGACTACCGTTAACGTCCTGGCAGTATTTGAACAGAATTGCTTCAAGTGAAATAAAG - 792
793 - TTTGTCTTATGAATCAAAAAAAAAAAAAAAAAAAA                                      - 827
```

44) Clone 1-7-1

No clear ORF was detected.

```
  1 - CCGCTGTACTTGTCTTCATCTGAGAATTTTGTCCTGGTGACGTAGCAAGCATGCATCTACTCTTGCGAGAAG - 72
 73 - CTTTAATACGTTCGCAATAGGAATCTCAATCATGGACCGTACACTTGCATACAAGTTTATGATGTACTATAT - 144
145 - ATACTAATTGGCATTGACGAAATTTACTAATAAACTCGAAGTTCGAGAGCAATTTGGTAGCAGTGGATGTAA - 216
217 - GAAGAATTTGTTTAACAATAAGAATTTTATAAACGGGTTTCCTACCAATGCCTCATGATATATGAGGCTGAG - 288
289 - CTTCATCGTTTTGTTCCATGGGAAGGATCTTTTCTCCTCGTATAATGTTTTCTTTTTTATGTATCTTGGTAC - 360
361 - GTATCTCATGCTGAAGTTTGGAAGAATCAAATTGACGTACCTTTACTTGCCTGCTTCTTGGTTAAGAAAGAT - 432
433 - TTACAAATCAAAAGGGAATTTTTGTGATAAGAAATCTAACGAATTTATCATGTCTAAGATTTGAAAAAAAAA - 504
505 - AAAAAAAAAAAAAAAAAAAAAAA                                                  - 527
```

4.3 Antigenic clones detected with mouse antibodies

1) HSP70 – Consensus of Clones 14-1, 17-2 and 20-2

Clones were sequenced from the 5'-end only. Translated sequence similar (~80%) to corresponding region of HSP70 clone 7-1 obtained by CSL box jellyfish antivenom immunoscreening experiment and therefore is likely to be a second isoform. ~316 residues are missing from the N-terminus of the complete protein. Deduced residues 1-266 form part of a conserved multifunctional HSP70 domain (Accession No. pfam00012).

```
  1 - GTTATTCAGGCCATTAAAGACTCGGCAATTGACAAGGAAAAAATCCATGAAGTTGTCTTGGTTGGTGGTTCA - 72
  1 -  V  I  Q  A  I  K  D  S  A  I  D  K  E  K  I  H  E  V  V  L  V  G  G  S  - 24

 73 - ACAAGAATCCCAAAGATACAATCTTTGCTGCAAGACCATTTTGGGGGTAAAGAATTAAACAAGTCGATTAAC - 144
 25 -  T  R  I  P  K  I  Q  S  L  L  Q  D  H  F  G  G  K  E  L  N  K  S  I  N  - 48

145 - CCGGATGAAGCTGTTGCATACGGAGCAGCTGTACAAGCTGCCATCCTGGCTGGAGACAAAAGTGACGAGGTC - 216
 49 -  P  D  E  A  V  A  Y  G  A  A  V  Q  A  A  I  L  A  G  D  K  S  D  E  V  - 72

217 - CAAGATCTTCTTCTTCTCGATGTCGCGCCGTTGTCTCTTGGAATTGAAACTGCTGGTGGCGTGATGACGGCA - 288
 73 -  Q  D  L  L  L  L  D  V  A  P  L  S  L  G  I  E  T  A  G  G  V  M  T  A  - 96

289 - CTGATCAAGAGGAACACGACAATCCCATCGAAGTACAGTCAAAACTTCACTACATACGCAGATAATCAGCCT - 360
 97 -  L  I  K  R  N  T  T  I  P  S  K  Y  S  Q  N  F  T  T  Y  A  D  N  Q  P  - 120

361 - GGAGTGTTAATTCAAGTGTTTGAAGGAGAAAGAGCAATGACGAAAGACAACAACATGTTGGGAAAGTTCGAG - 432
121 -  G  V  L  I  Q  V  F  E  G  E  R  A  M  T  K  D  N  N  M  L  G  K  F  E  - 144

433 - TTAAGTGGAATCCCTCCAGCACCACGTGGTGTGCCCCAGATTGAAGTCACGTTTGATCTTGATGCGAATGGA - 504
145 -  L  S  G  I  P  P  A  P  R  G  V  P  Q  I  E  V  T  F  D  L  D  A  N  G  - 168

505 - ATATTAAATGTGTCGGCAGTAGACAAGAGTACTGGGAAGGAGAATAAAATTACTATCACGAACGACAAAGGA - 576
169 -  I  L  N  V  S  A  V  D  K  S  T  G  K  E  N  K  I  T  I  T  N  D  K  G  - 192

577 - CGTCTATCGAAAGAAGATATTGAGCGCATGGTTAATGAAGCCGAGAAATACAAAAGTGAAGACGATGCTCAA - 648
193 -  R  L  S  K  E  D  I  E  R  M  V  N  E  A  E  K  Y  K  S  E  D  D  A  Q  - 216

649 - CGGGAGAGGATTCAGGCAAAGAATGGTCTTGAAGGATATGCATACCAGATGAAGAGCACGATCGAAGATGAC - 720
217 -  R  E  R  I  Q  A  K  N  G  L  E  G  Y  A  Y  Q  M  K  S  T  I  E  D  D  - 240

721 - AAGGTAAAAGACAAGATTAAGCGAAGAGGACAAGAAAACCAATTGGTGACAAAATTGCCAAAGGAAAGTAAT - 792
241 -  K  V  K  D  K  I  K  R  R  G  Q  E  N  Q  L  V  T  K  L  P  K  E  S  N  - 264

793 - TGGGAATT                                                                 - 800
265 -  W  E                                                                    - 266
```

4.4 Antigenic clones detected with rabbit antibodies

1) Calreticulin – Consensus of Clones 9-1, 9-2, 10-3, 11-5, 11-7, 13-8 and 13-10

Clone appears to be full-length but no stop codon is detected in the 3'-end. Interestingly, no stop codon is present in the translated clone 26-2 from the CSL antivenom immunoscreening experiment either. However, the pairwise sequence analysis indicates that the translated clones are quite distinct (only 21% sequence homology). Deduced residues 19-331 contain the conserved domain of calreticulin (Accession No. pfam00262). Calreticulin is a ubiquitous multifunctional protein found in a wide range of species and in all nucleated cell types. Calreticulin is a major **Ca^{2+}-binding**/storage chaperone residing in the ER lumen (Michalak et al., 1992). Calreticulin binds (buffers) Ca^{2+} with high capacity and participates in folding newly synthesized proteins and glycoproteins. It is an important component of the calreticulin/calnexin cycle and quality control pathways in the ER (Kopito, 1997).

```
   1 - GTTGAGGTTCTGATTTGTCTAGTTGTGTGTAGAGGCTTTCCAAAAGGGAGCTCGTAGAAATCCTAAAATATG - 72
   1 -                                                        *        *     M  - 1

  73 - AGGCTTTTAGTAGCTTTAACGGCTGTGTTATTCACGGTAGTGGAATGTAAAATCTATTTTTTTGAAAAATTT - 144
   2 -  R  L  L  V  A  L  T  A  V  L  F  T  V  V  E  C  K  I  Y  F  F  E  K  F  - 25

 145 - GAAGATGATAGCTGGGAGAAACGATGGACATATTCAACTTTCAAAGGCCCAGAGGCTGGCAAATTTAAGCTG - 216
  26 -  E  D  D  S  W  E  K  R  W  T  Y  S  T  F  K  G  P  E  A  G  K  F  K  L  - 49

 217 - ACTGCTGGGAAGTTTTATGGTGATCCAGAAAAAGATAAGGGAATACAGACATCACAAGATGCAAAATTCTAC - 288
  50 -  T  A  G  K  F  Y  G  D  P  E  K  D  K  G  I  Q  T  S  Q  D  A  K  F  Y  - 73

 289 - CAGATTTCAGCCAAATTAGATGAAGTATTTACAAATAAAGATAAGCCACTTGTGTTGCAGTTTCAAGTAAAG - 360
  74 -  Q  I  S  A  K  L  D  E  V  F  T  N  K  D  K  P  L  V  L  Q  F  Q  V  K  - 97

 361 - CATGAGCAATCTATTGACTGTGGTGGTGGCTATATAAAACTCTTCCCCAAAGATTTGGAGCCAGAAAAAATG - 432
  98 -  H  E  Q  S  I  D  C  G  G  G  Y  I  K  L  F  P  K  D  L  E  P  E  K  M  - 121

 433 - AATGGTGATTCACCTTACTATGTAATGTTTGGTCCTGACATATGTGGACCTGGAACAAAGAAGGTCCATGTT - 504
 122 -  N  G  D  S  P  Y  Y  V  M  F  G  P  D  I  C  G  P  G  T  K  K  V  H  V  - 145

 505 - ATTTTTAACTATCAAGGAAAAAATCTTCTCACAAAGAAAGATATCAGGTGCAAAGATGATGAATTTACTCAC - 576
 146 -  I  F  N  Y  Q  G  K  N  L  L  T  K  K  D  I  R  C  K  D  D  E  F  T  H  - 169

 577 - CTTTACACTCTGATTGTCAAACCAGATAATACATATGAAGTTCGTATAGACAATGAGAAAGTAGAAGGTGGT - 648
 169 -  L  Y  T  L  I  V  K  P  D  N  T  Y  E  V  R  I  D  N  E  K  V  E  G  G  - 193

 649 - GACCTAGAAGCTGATTGGGACTTTTTACCACCCAAGAAAATTAAAGATCCTGAAGCTAAGAAGCCAGAAGAC - 720
 194 -  D  L  E  A  D  W  D  F  L  P  P  K  K  I  K  D  P  E  A  K  K  P  E  D  - 217

 721 - TGGGATGACAAAGCCAAGATTGATGACCCAGATGACAAGAAACCAGAGGATTGGGACAAACCAGAACTAATC - 792
 218 -  W  D  D  K  A  K  I  D  D  P  D  D  K  K  P  E  D  W  D  K  P  E  L  I  - 241

 793 - CCTGACCCAGATGCAAAGAAACCAGATGACTGGGATGATGAAGAAGATGGTGAATGGGAGCCACCAATGATT - 864
 242 -  P  D  P  D  A  K  K  P  D  D  W  D  D  E  E  D  G  E  W  E  P  P  M  I  - 265

 865 - AACAATCCTGAATATAAGGGAGAATGGAAGCCAAAACAGATTGATAATCCTAATTATAAAGGGGAGTGGATT - 936
 266 -  N  N  P  E  Y  K  G  E  W  K  P  K  Q  I  D  N  P  N  Y  K  G  E  W  I  - 289

 937 - CATCCAGAGATTGATAATCCAGAATATAAGCCTGACCCAGAATTGTACAAGTATGATGACATCTTTCATGTT - 1008
 290 -  H  P  E  I  D  N  P  E  Y  K  P  D  P  E  L  Y  K  Y  D  D  I  F  H  V  - 313


1009 - GGTTATGAAATTTGGCAGGTGAAGTCTGGTACAATCTTTGACAATATATTTGTTGGTGATGATGTTGAGGAA - 1080
 314 -  G  Y  E  I  W  Q  V  K  S  G  T  I  F  D  N  I  F  V  G  D  D  V  E  E  - 337
```

```
1081 - GCAGAGAAGGCTGCTAAGGAAACATTTGAAGTAACAAAGGAGGGAGAAAAGAAAATGAAAGATAAACAAGAT - 1152
 338 -  A  E  K  A  A  K  E  T  F  E  V  T  K  E  G  E  K  K  M  K  D  K  Q  D  - 361

1153 - GAGGAGGAAAGGAAAAAAAAAAAAAAAAAA                                           - 1182
 362 -  E  E  E  R  K  K  K  K  K  K                                            - 371
```

2) Excision Repair Protein – Consensus of Clones 7-3, 8-10 and 15-2

~740 residues are missing from the N-terminus of the translated 5'-end of the clone (a). Deduced residues 103-238 contain a conserved helicase superfamily C-terminal domain (Accession No. cd00079). This domain is found in a wide variety of helicases and helicase related proteins and is associated with DEXDc-, DEAD-, and DEAH-box proteins, yeast initiation factor 4A, Ski2p, and Hepatitis C virus NS3 helicases. All helicases share the ability to unwind nucleic acid duplexes with a distinct directional polarity. They utilise the free energy from nucleoside triphosphate hydrolysis to fuel their translocation along DNA, unwinding the duplex in the process. Specific features present within the domain include a conserved nucleotide binding region (residues highlighted in green) and an ATP-binding site (residues highlighted in red). Deduced residues 1-286 contain a conserved domain similar to superfamily II DNA/RNA helicases (SNF2 family). Members of the SNF2 family are involved in transcription / DNA replication, recombination and repair.

(a)

```
  1 - ATCAATCCATATCTATTGAGAAGACTGAAAAAGGATGTCAAAACAAACTTACAATTACCTACAAAGAATGAA - 72
  1 -  I  N  P  Y  L  L  R  R  L  K  K  D  V  K  T  N  L  Q  L  P  T  K  N  E  - 24

 73 - CAGGTTTTATTTTGTCGGCTTACTGACGAGCAAACTAACCTTTACCAATTGTACCTGAAGTCTAGAGAGGTT - 144
 25 -  Q  V  L  F  C  R  L  T  D  E  Q  T  N  L  Y  Q  L  Y  L  K  S  R  E  V  - 48

145 - AACGCTATTCTTGATGGCAGAATGAAGGTATTCCCTGGTTTAATCATGCTGCGAAAGATTTGCAATCATCCA - 216
 49 -  N  A  I  L  D  G  R  M  K  V  F  P  G  L  I  M  L  R  K  I  C  N  H  P  - 72

217 - GATCTTACTTCAGCGGCCGGATCTTTAGAGTGGATGAAAGTAAAAGAGGAAAAAGCAGCAACTAAATCAAAT - 288
 73 -  D  L  T  S  A  A  G  S  L  E  W  M  K  V  K  E  E  K  A  A  T  K  S  N  - 96

289 - AGGAAGGATCAACCAGAAGAAATTGAATACGAAGGTGGCTATGGCTTTTGGCAACGCTCTGGAAAGCTTATT - 360
 97 -  R  K  D  Q  P  E  E  I  E  Y  E  G  G  Y  G  F  W  Q  R  S  G  K  L  I  - 120

361 - GTAGTTGAAGCATTGTTGAGGTTATGGAAAAAGCAAGGACATAGAGTATTGCTTTTCTCACAAACCAGACAG - 432
121 -  V  V  E  A  L  L  R  L  W  K  K  Q  G  H  R  V  L  L  F  S  Q  T  R  Q  - 144

433 - ATGCTAGATATCCTAGAGTCTTTTGTTAAAGAACAAGGCTATACCTACATGAGAATGGATGGATCAACATCC - 504
145 -  M  L  D  I  L  E  S  F  V  K  E  Q  G  Y  T  Y  M  R  M  D  G  S  T  S  - 168

505 - ATCGCAAGTCGACAGCCAGCCGTGCAGAAATTCAACTCGGACACATCAATTTTTGTTTATCTCCTGACAACG - 576
169 -  I  A  S  R  Q  P  A  V  Q  K  F  N  S  D  T  S  I  F  V  Y  L  L  T  T  - 192

577 - CGAGTTGGTGGTTTAGGACTTAATCTTATTGGTGCAGATCGAATCGTCATTTATGATCCAGACTGGAATCCA - 648
193 -  R  V  G  G  L  G  L  N  L  I  G  A  D  R  I  V  I  Y  D  P  D  W  N  P  - 216

649 - AGCACTGACACTCAAGCCAGAGAAAGATCGTGGCGAATTGGTCAAAATAAACATGTTACAATTTATCGCCTG - 720
217 -  S  T  D  T  Q  A  R  E  R  S  W  R  I  G  Q  N  K  H  V  T  I  Y  R  L  - 240

721 - CTCACTACTGGCACAATTGAAGAAAAAATCTATCACAGGCAAATTTTCAAACAGTTCCTCACAAATCGGGTG - 792
241 -  L  T  T  G  T  I  E  E  K  I  Y  H  R  Q  I  F  K  Q  F  L  T  N  R  V  - 264

793 - CTTACAAATCCCATGCAGCGACGTTTCTTCAAATCAAATGATTTGTTTGAGTTATTTACGTTAGGAGATACA - 864
265 -  L  T  N  P  M  Q  R  R  F  F  K  S  N  D  L  F  E  L  F  T  L  G  D  T  - 288

865 - CGTTCAAATCATGGAACGGAAACGAGTGCCATATTTGCAGGGACAGGATCTGAAGTAAAAGTACGAAGGAAA - 936
```

```
 289 -  R  S  N  H  G  T  E  T  S  A  I  F  A  G  T  G  S  E  V  K  V  R  R  K  - 312

 937 - AAATCAGATATTGATGAAGGCAAGCAAATGAAAAAATTGGGAAAGCAAAAGAAAGCTAAATTGACAACTTCA - 1008
 313 -  K  S  D  I  D  E  G  K  Q  M  K  K  L  G  K  Q  K  K  A  K  L  T  T  S  - 336

1009 - GAAGAAAGTAAAAAAATTATTGTCAAGGAACTACCTGATCTAGAGTCAAGGAGCGGAGATGCCTGGACAATT - 1080
 337 -  E  E  S  K  K  I  I  V  K  E  L  P  D  L  E  S  R  S  G  D  A  W  T  I  - 360

1081 - ACAGCCAAAGATCTCGAGGAGCTTAGTAGTCGAGATTCTAATGACTCTAAGTCATTTCAAGAGGGCGTAAGC - 1152
 361 -  T  A  K  D  L  E  E  L  S  S  R  D  S  N  D  S  K  S  F  Q  E  G  V  S  - 384

1153 - TCAAGTACTATTGGCGATGTGAAATCAAGTGTACTAGGCAACCGGAGTGTCAATTCTCTTGGCGATGATATC - 1224
 385 -  S  S  T  I  G  D  V  K  S  S  V  L  G  N  R  S  V  N  S  L  G  D  D  I  - 408

1225 - TATGATCATAAGTCATCTTCACCATTACAAACTTCTTCTGTTGACAACAAGGTGGCTATAGATAAACTTACC - 1296
 409 -  Y  D  H  K  S  S  S  P  L  Q  T  S  S  V  D  N  K  V  A  I  D  K  L  T  - 432

1297 - ATGGCTTGTTCAACATGTAATCATTCAGCTACTGTCACTGGTTTGCCTAAAGTAAACGAATCAGCTGGTGTG - 1368
 433 -  M  A  C  S  T  C  N  H  S  A  T  V  T  G  L  P  K  V  N  E  S  A  G  V  - 456

1369 - AAAAACGAAATACCATCGACAAGTAGGTCATCGTTCGCTCTTAAAAGTGAAATCACAGACAGGCAGAAAGAT - 1440
 457 -  K  N  E  I  P  S  T  S  R  S  S  F  A  L  K  S  E  I  T  D  R  Q  K  D  - 480

1441 - AAGCGATTTCGAAAGAAGAAGCGTAAGAAGCATGTTAAGGTGGACGGTGTAGTTATTGACAATCTTGATCGT - 1512
 481 -  K  R  F  R  K  K  K  R  K  K  H  V  K  V  D  G  V  V  I  D  N  L  D  R  - 504

1513 - CG                                                                       - 1514
```

(b)

```
   1 - GCGCTCAACCATGACAAAATCGAAGATCCGAATAATCCTGATTTTGTTTTAGTGGAAAAAGAAGCCGAACGT - 72
   1 -  A  L  N  H  D  K  I  E  D  P  N  N  P  D  F  V  L  V  E  K  E  A  E  R  - 24

  73 - GTGGCAAAGCAGGCTGCAGAATCCTTACGCAGATCAAGGCATGAATGTAGGCAGAATGTAGTCGGTGTGCCC - 144
  25 -  V  A  K  Q  A  A  E  S  L  R  R  S  R  H  E  C  R  Q  N  V  V  G  V  P  - 48

 145 - ACATGGACGGGCTTTCAAGGGTCTGCTGGTCGACCAGGACTGAAACCTCGATTTGGGCCGAGAAAGGAGGCC - 216
  49 -  T  W  T  G  F  Q  G  S  A  G  R  P  G  L  K  P  R  F  G  P  R  K  E  A  - 72

 217 - ATTAAGACTCAAGTAAATGACATGAATGAAGACTCCTGTAGTGATGACAACAAAAAGCCTCTTTTTAGCGGG - 288
  73 -  I  K  T  Q  V  N  D  M  N  E  D  S  C  S  D  D  N  K  K  P  L  F  S  G  - 96

 289 - AAAGAACTGCTGAGGACTAATAAGGACGGAAAATGTGCTGCACCCTCCTCTTCGGATCTTCTAGCGAAAATG - 360
  97 -  K  E  L  L  R  T  N  K  D  G  K  C  A  A  P  S  S  S  D  L  L  A  K  M  - 120

 361 - AGAAATAGAAATACGACAGTAGATCTTCCCGCGACTTCAGTGAATGCTGATAAAACAAACGAAGATAAACTT - 432
 121 -  R  N  R  N  T  T  V  D  L  P  A  T  S  V  N  A  D  K  T  N  E  D  K  L  - 144

 433 - GTTGAAGAAATACGAGATTTTGTCGCATATCGTGCAATGACGTTTGGACAGGCGACGACACAGGAAATCTTA - 504
 145 -  V  E  E  I  R  D  F  V  A  Y  R  A  M  T  F  G  Q  A  T  T  Q  E  I  L  - 168

 505 - GCTGAATTTCAATCTCGAATTCAAAAGAACAAAAACGCGCTCTTTAGGGAATTACTGAAGCAGGTTTGTTAC - 576
 169 -  A  E  F  Q  S  R  I  Q  K  N  K  N  A  L  F  R  E  L  L  K  Q  V  C  Y  - 192

 577 - TTGAGAAAGGTTGAAGGGGTTGGTACATGGTTTTTAAAAGAAGAATACCGCTAGTTTGTAACTAATTTATAT - 648
 193 -  L  R  K  V  E  G  V  G  T  W  F  L  K  E  E  Y  R  *                    - 209

 649 - GCGTTTTTAAAACGTACAGCCTGTAAAACAATATAATCTATATATAAAGAGGTATATTGCTGCCAATGGTTG - 720
 721 - TGTAATTTATATAAACTTAATAAAGTGTATTTTATTACATGTATTTCTTTGTTCTTACCTTAAGAGATAAGA - 792
 793 - AGGAAACGTCTTCATGGACTGATAAAAAAAAAAAAAAAAAAAAAAAAA                         - 840
```

4.5 References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25, 3389-3402.

Emanuelsson, O., Brunak, S. von Heijne, G., Nielsen, H., 2007. Locating proteins in the cell using TargetP, SignalP, and related tools. *Nature Protocols* 2, 953-971.

Genikhovich, G., Kürn, U., Hemmrich, G., Bosh, T.C.G., 2006. Discovery of genes expressed in *Hydra* embryogenesis. *Developmental Biology* 289, 466-481.

Howell, D.M., Harich, K., Xu, H., White, R.H., 1998. Alpha-keto acid chain elongation reactions involved in the biosynthesis of coenzyme B (7-mercaptoheptanoyl threonine phosphate) in methanogenic Archaea. *Biochemistry* 37, 10108-10117.

Johnson, A., Yao, N.Y., Bowman, G.D., Kuriyan, J., O'Donnell, M., 2006. The replication factor C clamp loader requires arginine finger sensors to drive DNA binding and proliferating cell nuclear antigen loading. *Journal of Biological Chemistry* 281, 35531-35543.

Kopito, R.R., 1997. ER quality control: the cytoplasmic connection. *Cell* 88, 427-430.

Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales, N.R., Gwadz, M., Hao, L., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Krylov, D., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Thanki, N., Yamashita, R.A., Yin, J.J., Zhang, D., Bryant, S.H, 2007. CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Research* D35, 237-240.

Marchler-Bauer, A., Bryant, S.H., 2004 CD-Search: protein domain annotations on the fly. *Nucleic Acids Research* W32, 327-331.

Mikalak, M., Milner, R.E., Burns, K., Opas, M., 1992. Calreticulin. *Biochemistry Journal* 285, 681-692.

Qui, C., Lienhard, S., Hynes, N.E., Badache, A., Leahy, D.J., 2008. Memo is homologous to nonheme dioxygenases and binds an ErbB2-derived phosphopeptide in its vestigial active site. *Journal of Biological Chemistry* 283, 2734-2740.

Raburn, D.J., Hamil, K.G., Tsuruta, J.K., O'Brien, D.A., Hall, S.H., 1995. Stage-specific expression of B cell translocation gene 1 in rat testes. *Endocrinology* 136, 5769-5777.

Wang, P., Heitman, J., 2005. The cyclophilins. *Genome Biology* 6, 226.1-226.6.

APPENDIX 5

Eukaryote EF1α Protein Sequences

Supplementary Table 5.1

The biological origin and accession numbers of protein sequences used in the phylogenetic analysis of EF1α (Chapter 3, Figure 3.8).

Species	Accession No.
Entamoeba histolytica (enteric parasite)	AAA29096
Plasmodium falciparum (malaria parasite)	Q00080
Tetrahymena pyriformis (ciliate)	BAA01856
Stylonychia lemnae (ciliate)	P25166
Euglena gracilis (flagellate)	CAA34769
Trypanosoma brucei (flagellate)	AAA16602
Porphyra purpurea (red alga)	AAA61793
Dictyostelium discoideum (soil amoeba)	P18624
Triticum aestivum (wheat)	AAA34306
Arabidopsis thaliana (cress)	CAA34453
Saccharomyces cerevisiae (yeast)	P02994
Puccinia gramina (rust)	P32186
Trichoderma reesei (filamentous)	P34825
Podospora anserina (filamentous)	Q01520
Axinella verrucosa (sponge)	CAE45763
Geodia cydonium (sponge)	CAA70221
Chironex fleckeri (cubozoan)	
Hydra vulgaris (hydrazoan)	CAA92323
Podocoryne carnea (hydrazoan)	CAD70569
Cladonema radiatum (hydrozoan)	AAT11876
Nematostella vectensis (sea anemone)	BAD02195
Scleronephthya gracillimum (soft coral)	AAT81474
Pocillopora damicornis (stony coral)	BAE66714
Artemia salina (crustacean)	CAA27055
Drosophila melanogaster (insect)	P05303
Xenopus laevis (frog)	NP_001080856
Danio rerio (fish)	AAI28792
Mus musculus (mouse)	NP_031932
Homo sapiens (human)	P68104

```
Entamoeba histolytica         1 -MPKEK----THINIVVIGHVDSGKSTTTGHLIYKCGGIDQRTIEKFEKE  45
Plasmodium falciparum         1 -MGKEK----THINLVVIGHVDSGKSTTTGHIIYKLGGIDRRTIEKFEKE  45
Tetrahymena pyriformis        1 MARGDK----VHINLVVIGHVDSGKSTTTGHLIYKCGGIDKRVIEKFEKE  46
Stylonychia lemnae            1 -MPKEK----NHLNLVVIGHVDSGKSTSTGHLIYKCGGIDKRTIEKFEKE  45
Euglena gracilis              1 -MGKEK----VHISLVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Trypanosoma brucei            1 -MGKEK----VHMNLVVVGHVDAGKSTATGHLIYKCGGIDKRTIEKFEKE  45
Porphyra purpurea             1 -MGKEK----QHVSIVVIGHVDSGKSTTTGHLIYKCGGIDKRAIEKFEKE  45
Dictyostelium discoideum      1 -MGKEK----THINIVVIGHVDAGKSTTTGHLIYKCGGIDKRVIEKYEKE  45
Triticum aestivum             1 -MGKEK----THINIVVIGHVDSGKSTTTGHLIYKLGGIDKRVIERFEKE  45
Arabidopsis thaliana          1 -MGKEK----FHINIVVIGHVDSGKSTTTGHLIYKLGGIDKRVIERFEKE  45
Saccharomyces cerevisiae      1 -MGKEK----SHINVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Puccinia gramina              1 -MGKEK----NHVNVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Trichoderma reesei            1 MGKEDK----THINVVVIGHVDSGKSTTTGHLIYQCGGIDKRTIEKFEKE  46
Podospora anserina            1 MGKEDK----THINVVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  46
Artemia salina                1 -MGKEK----IHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Drosophila melanogaster       1 -MGKEK----IHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Homo sapiens                  1 -MGKEK----THINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Chironex fleckeri             1 -MPKEK----THINIVVIGHVDSGKSTSTGHLIYKCGGIDKRAIEKFEKE  45
Nematostella vectensis        1 -MPKEK----IHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Hydra vulgaris                1 MSNKVKP----HINIVVIGHVDSGKSTSTGHMIYKCGGIDKRQIEKFEKE  46
Podocoryne carnea             1 MAPKEK--AKIHINIVVIGHVDSGKSTSTGHLIYKCGGIDKRTIEKFEKE  48
Cladonema radiatum            1 MAPKDKKDQKIHVNIVVIGHVDSGKSTSTGHLIYKCGGIDKRTIEKFEKE  50
Scleronephthya gracillimum    1 -MGKEK----THINIVVIGHVDSGKSTTTGHLIYKCGGIDKRAIEKFEKE  45
Pocillopora damicornis        1 -MGKEK----THINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Axinella verrucosa            1 -MPKEK----IHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Geodia cydonium               1 -MPKEK----THINIVVIGHVDSGKSTTTGHLIYKCGGIDKRAIEKFEKE  45
Xenopus laevis                1 -MGKEK----THINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Danio rerio                   1 -MGKEK----IHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
Mus musculus                  1 -MGKEK----THINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKE  45
                              1      *     *. .**.****.****.***.**. ****.* **..***  50

Entamoeba histolytica        46 SAEMGKGSFKYAWVLDNLKAERERGITIDISLWKFETSKYYFTIIDAPGH  95
Plasmodium falciparum        46 SAEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETPRYFFTVIDAPGH  95
Tetrahymena pyriformis       47 SAEQGKGSFKYAWVLDKLKAERERGITIDISLWKFETAKYHFTIIDAPGH  96
Stylonychia lemnae           46 AAEMGKGSFKYAWVLDKLKAERERGITIDIALWNFETAKSVFTIIDAPGH  95
Euglena gracilis             46 ASEMGKGSFKYAWVLDKLKAERERCITIDIALWKFETAKSVFTIIDAPGH  95
Trypanosoma brucei           46 AADIGKASFKYAWVLDKLKAERERGITIDIALWKFESPKSVFTIIDAPGH  95
Porphyra purpurea            46 AAEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETDKYNFTIIDAPGH  95
Dictyostelium discoideum     46 ASEMGKQSFKYAWVMDKLKAERERGITIDIALWKFETSKYYFTIIDAPGH  95
Triticum aestivum            46 AAEMNKRSFKYAWVLDKLKAERERGITIDIALWKFETTKYYCTVIDAPGH  95
Arabidopsis thaliana         46 AAEMNKRSFKYAWVLDKLKAERERGITIDIALWKFETTKYYCTVIDAPGH  95
Saccharomyces cerevisiae     46 AAELGKGSFKYAWVLDKLKAERERGITIDIALWKFETPKYQVTVIDAPGH  95
Puccinia gramina             46 AAELGKGSFKYAWVLDKLKAERERGITIDIALWKFETPKYYVTVIDAPGH  95
Trichoderma reesei           47 AAELGKGSFKYAWVLDKLKAERERGITIDIALWKFETPKYYVTVIDAPGH  96
Podospora anserina           47 AAELGKGSFKYAWVLDKLKAERERGITIDIALWKFETPKYYVTVIDAPGH  96
Artemia salina               46 AQEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETAKYYVTIIDAPGH  95
Drosophila melanogaster      46 AQEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETSKYYVTIIDAPGH  95
Homo sapiens                 46 AAEMGKGSFKYAWVLDKLKAERERGITIDISLWKFETSKYYVTIIDAPGH  95
Chironex fleckeri            46 AQEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETIKYCVTVIDAPGH  95
Nematostella vectensis       46 SSEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETLKYYVTVIDAPGH  95
Hydra vulgaris               47 AQEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETTKYVVTIIDAPGH  96
Podocoryne carnea            49 AQEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETTKFYVTIIDAPGH  98
Cladonema radiatum           51 AQEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETTKYQVTIIDAPGH 100
Scleronephthya gracillimum   46 AAEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETVKYYVTVIDAPGH  95
Pocillopora damicornis       46 ANEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETEKYYVTVIDAPGH  95
Axinella verrucosa           46 AQEMGKGSFKYAWVLDKLKAERERGITIDIALWKFETTKYYITVIDAPGH  95
Geodia cydonium              46 AQEIGKGSFKYAWVLDKLKAERAG-ITIDIALWKFETTKYYITVIDAPGH  94
Xenopus laevis               46 AAEMGKGSFKYAWVLDKLKAERERGITIDISLWKFETTKYYITIIDAPGH  95
Danio rerio                  46 AAEMGKGSFKYAWVLDKLKAERERGITIDISLWKFETSKYYVTIIDAPGH  95
Mus musculus                 46 AAEMGKGSFKYAWVLDKLKAERERGITIDISLWKFETTKYYITIIDAPGH  95
                             51 . . .* *******.* *****   *****.** **. .   *.****** 100
```

Supplementary Figure 5.1 Multiple amino acid sequence alignment of 29 EF1α proteins using ClustalW (Thompson et al., 1994). Identical amino acids are marked by an asterisk and similar amino acids are marked with a stop. Dashes indicate gaps introduced for better alignment. This alignment was used in the phylogenetic analysis of eukaryote EF1α proteins (Chapter 3, Figure 3.8).

```
Entamoeba histolytica       96 RDFIKNMITGTSQADVAILIVAAGTGEFEAGISKNGQTREHILLSYTLGV 145
Plasmodium falciparum       96 KDFIKNMITGTSQADVALLVVPADVGGFDGAFSKEGQTKEHVLLAFTLGV 145
Tetrahymena pyriformis      97 RDFIKNMITGTSQADVAILMIASPQGEFEAGISKDGQTREHALLAFTLGV 146
Stylonychia lemnae          96 RDFIKNMITGTSQADAAILIIASGQGEFEAGISKEGQTREHALLAFTMGV 145
Euglena gracilis            96 RDFIKNMITGTSQADAAVLVIDSTTGGFEAGISKDGQTREHALLAYTLGV 145
Trypanosoma brucei          96 RDFIKNMITGTSQADAAILIIASAQGEFEAGISKDGQTREHALLAFTLGV 145
Porphyra purpurea           96 RDFIKNMITGTSQADLAILVIASPPGEFEAGISQNGQTREHALLAYTLGV 145
Dictyostelium discoideum    96 RDFIKNMITGTSQADCAVLVIASPTGEFEAGIAKNGQTREHALLAYTLGV 145
Triticum aestivum           96 RDFIKNMITGTSQADCAVLIIDSTTGGFEAGISKDGQTREHALLAFTLGV 145
Arabidopsis thaliana        96 RDFIKNMITGTSQADCAVLIIDSTTGGFEAGISKDGQTREHALLAFTLGV 145
Saccharomyces cerevisiae    96 RDFIKNMITGTSQADCAILIIAGGVGEFEAGISKDGQTREHALLAFTLGV 145
Puccinia gramina            96 RDFIKNMITGTSQADCAILIIAAGTGEFEAGISKDGQTREHALLAFTLGV 145
Trichoderma reesei          97 RDFIKNMITGTSQADCAILIIAAGTGEFEAGISKDGQTREHALLAYTLGV 146
Podospora anserina          97 RDFIKNMITGTSQADCAILIIAAGTGEFEAGISKDGQTREHALLAYTLGV 146
Artemia salina              96 RDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGV 145
Drosophila melanogaster     96 RDFIKNMITGTSQADCAVLIVAAGTGEFEAGISKNGQTREHALLAFTLGV 145
Homo sapiens                96 RDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGV 145
Chironex fleckeri           96 RDFIKNMITGTSQADCAVLIVAGSTGEFEAGISKEGQTREHALLAYTLGV 145
Nematostella vectensis      96 RDFIKNMITGTSQADCAVLIIASGTGEFEAGISKNGQTREHALLAYTLGV 145
Hydra vulgaris              97 RDFIKNMITGTSQADCAVLIVASSTGEFEAGISKNGQTREHALLAFTLGV 146
Podocoryne carnea           99 RDFIKNMITGTSQADCAVLIVASSTGEFEAGISKNGQTREHALLAYTLGV 148
Cladonema radiatum         101 RDFIKNMTTGTSQADCAVLIVASSTGEFEAGISKNGQTREHALLAFTLGV 150
Scleronephthya gracillimum  96 RDFIKNMITGTSQADCAVLIVASGVGEFEAGISANGQTREHALLAYTLGV 145
Pocillopora damicornis      96 RDFIKNMITGTSQADCAVLVVAAGTGEFEAGISKNGQTREHALLSYTLGV 145
Axinella verrucosa          96 RDFIKNMITGTSQADCAVLIVAAGTGEFEAGISKNGQTREHALLAYTLGV 145
Geodia cydonium             95 RDFIKNMITGTSQADCAVLIVASGTGEFEAGISKNGQTREHALLAYSLGV 144
Xenopus laevis              96 RDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGV 145
Danio rerio                 96 RDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGV 145
Mus musculus                96 RDFIKNMITGTSQADCAVLIVAAGVGEFEAGISKNGQTREHALLAYTLGV 145
                           101 .****** ******* *.*..    * *.   .  ***.** **....** 150

Entamoeba histolytica      146 KQMIVGVNKMDAIQ--YKQERYEEIKKEISAFLKKTGYNPDKIPFVPISG 193
Plasmodium falciparum      146 KQIVVGVNKMDTVK--YSEDRYEEIKKEVKDYLKKVGYQADKVDFIPISG 193
Tetrahymena pyriformis     147 KQMIVCLNKMDEKTVNFSEERYQEIKKELSDYLKKVGYKPDTIPFIPISG 196
Stylonychia lemnae         146 KQMIVAVNKMDDKSVNWDQGRFIEIKKELSDYLKKIWLQPRQDPFIPISG 195
Euglena gracilis           146 KQMIVATNKFDDKTVKYSQARYEEIKKEVSGYLKKVGYNPEKVPFIPISG 195
Trypanosoma brucei         146 KQMVVCCNKMDDKTVNYGQERYDEIVKEVSAYIKKVGYNVEKVRFVPISG 195
Porphyra purpurea          146 KQMIVACNKMDDKNVNWSKERYEEVSKEMDLYLKKVGYNPPKVPKVPTSG 195
Dictyostelium discoideum   146 KQMIVAINKMDEKSTNYSQARYDEIVKEVSSFIKKIGYNPEKVAFVPISG 195
Triticum aestivum          146 KQMICCCNKMDATTPKYSKARYEEIVKEVSSYLKKVGYNPDKVPFVPISG 195
Arabidopsis thaliana       146 KQMICCCNKMDATTPKYSKARYDEIIKEVSSYLKKVGYNPDKIPFVPISG 195
Saccharomyces cerevisiae   146 RQLIVAVNKMDSVK--WDESRFQEIVKETSNFIKKVGYNPKTVPFVPISG 193
Puccinia gramina           146 RQLIVAINKMDTTK--WSEQRF-EIVKETSNFVKKVGYNPKSIAFVPISG 192
Trichoderma reesei         147 KQLIVAINKMDTAN--WAEARYQEIIKETSNFIKKVGFNPKAVAFVPISG 194
Podospora anserina         147 KQLIVAINKMDTTK--WSEARFNEIIKETSNFIKKVGYNPKTVAFVPISG 194
Artemia salina             146 KQLIVGVNKMDSTEPPFSEARFEEIKKEVSAYIKKIGYNPAAVAFVPISG 195
Drosophila melanogaster    146 KQLIVGVNKMDSTEPPYSEARYEEIKKEVSSYIKKIGYNPASVAFVPISG 195
Homo sapiens               146 KQLIVGVNKMDSTEPPYSQKRYEEIVKEVSTYIKKIGYNPDTVAFVPISG 195
Chironex fleckeri          146 KQMIVAVNKMDNTEPPFSESRFQEIEKEVSAYLKKIGYNPKAVAFVPISG 195
Nematostella vectensis     146 KQMLVCANKMDSTEPPYSEARFKEIQKEVANFLKKVGFNPKAVPFIPISG 195
Hydra vulgaris             147 KQMIVAVNKIDNTEPPYSEARFNEIKKEISAYVKKVGYDPKTVPVLPVSG 196
Podocoryne carnea          149 KQLIIGVNKIDNTEPPYSEARFNEIKKEVEGYVKKVGYNPKAVAILPISG 198
Cladonema radiatum         151 KQLIVGVNKIDNTEPPYSEARFTEITKEVSNYIKKVGYNPKAVACVPISG 200
Scleronephthya gracillimum 146 KQMIVGVNKMDSSEPPYSEKRYEEIKKEVGSYLKKVGFNPKAVAFVPISG 195
Pocillopora damicornis     146 KQLIVAVNKMDTTEPKYHEGRFNEIQKEVSGYVKKVGYNPKAVVFVPISG 195
Axinella verrucosa         146 KQLIVGVNKMDSTEPPYSEPRFNEIIKEVGAYIKKIGYNPKAVAFVPISG 195
Geodia cydonium            145 KQLIVGVNKMDSTEPPYSQARYDEITKEVGTYIKKVGYNPKAVAFVPISG 194
Xenopus laevis             146 KQLIVGINKMDSTEPPYSEKRYDEIVKEVSAYIKKIGYNPATVPFVPISG 195
Danio rerio                146 KQLIVGVNKMDSTEPPYSQARFEEITKEVSAYIKKIGYNPATVAFVPISG 195
Mus musculus               146 KQLIVGVNKMDSTEPAYSEKRYDEIVKEVSAYIKKIGYNPATVPFVPISG 195
                           151 .*..   ** *         *. *. **   ..**          .* ** 200
```

Supplementary Figure 5.1 (*continued*)

```
Entamoeba histolytica      194 FQGDNMIEPSTNMPWYKG------------------PTLIGALDSVTPPE 225
Plasmodium falciparum      194 FEGDNLIEKSDKTPWYKG------------------RTLIEALDTMQPPK 225
Tetrahymena pyriformis     197 FNGDNMLERSTNAPWYKG------------------PILVEALDALEPPK 228
Stylonychia lemnae         196 WHGDNMLEKSPNMPWFTG------------------STLIDALDALDQPK 227
Euglena gracilis           196 WNGDNMIEASENMGWYKG------------------LTLIGALDNLEPPK 227
Trypanosoma brucei         196 WQGDNMIEKSEKMPWYKG------------------PTLLEALDMLEPPV 227
Porphyra purpurea          196 WTGENLFERTGGDHALGKWY--------------KGPCLLEALDACDPPK 231
Dictyostelium discoideum   196 WNGDNMLERSDKMEWYK------------------GPTLLEALDAIVEPK 227
Triticum aestivum          196 FEGDNMIERSTNLDWYK------------------GPTLLEALDQINEPK 227
Arabidopsis thaliana       196 FEGDNMIERSTNLDWYK------------------GPTLLEALDQINEPK 227
Saccharomyces cerevisiae   194 WNGDNMIEATTNAPWYKGWEKET------KAGVVKGKTLLEAIDAIEQPS 237
Puccinia gramina           193 WHGDNMLEESTNMGWFKGWTKET------KAGVSKGKTLLDAIDAIEPPS 236
Trichoderma reesei         195 FNGDNMLTPSTNCPWYKGWEKET------KAGKFTGKTLLEAIDSIEPPK 238
Podospora anserina         195 FNGDNMLEASTNCPWYKGWEKEV------KGGKATGKTLLEAIDSIEPPK 238
Artemia salina             196 WHGDNMLEASDRLPWYKGWNIER------KEGKADGKTLLDALDAILPPS 239
Drosophila melanogaster    196 WHGDNMLEPSEKMPWFKGWSVER------KEGKAEGKCLIDALDAILPPQ 239
Homo sapiens               196 WNGDNMLEPSANMPWFKGWKVTR------KDGNASGTTLLEALDCILPPT 239
Chironex fleckeri          196 WHGDNMLEPTERMPWYKGWSISR------KEGDAKGKTLFEALDSILPPQ 239
Nematostella vectensis     196 FHGDNMLEKSENMPWFKQWTIERVDPATKKEANASGVTLFEGLDSILPPS 245
Hydra vulgaris             197 WHGDNMIEPSPNMSWYKGWEVEYKDTG-----KHTGKTLLEALDNIPLPA 241
Podocoryne carnea          199 WHGDSMLEESPNTKWFKGWATERVDEDK-KVINSSGKTLFEALDAIVPPS 247
Cladonema radiatum         201 WHGDNMIEPSTNMGWYKGWNIERKEG------KASGKTLLEALDAIVPPS 244
Scleronephthya gracillimum 196 WHGDNMLEESDKMKWYKGWNVER------KEGNAAGKTLFEALDSILPPK 239
Pocillopora damicornis     196 FHGDNMLEASENMPWFKGWSIER------KEGNASGKTLFNALDAILPPE 239
Axinella verrucosa         196 WHGDNMLEPSDNMSWFKGWNVER------KEGNATGKTLFTCLDSILPPK 239
Geodia cydonium            195 WHGDNMLEESPNMKWFKGWNVER------KEGNASGKTLFNPLDSILPPT 238
Xenopus laevis             196 WHGDNMLEPSPNMPWFKGWKVER------KEGNANGVSLLEALDTILPPS 239
Danio rerio                196 WHGDNMLEPSSNMGWFKGWKIER------KEGGANGVTLLEALDSILPPS 239
Mus musculus               196 WHGDNMLEPSPNMPWFKGWKVER------KEGNASGVSLLEALDTILPPT 239
                           201   *. .   .                            *   .*    *  250

Entamoeba histolytica      226 RPVDKPLRLPLQDVYKISGIGTVPVGRVETGILKPGTIVQFAPSGVSSEC 275
Plasmodium falciparum      226 RPYDKPLRIPLQGVYKIGGIGTVPVGRVETGILKAGMVLNFAPSAVVSEC 275
Tetrahymena pyriformis     229 RPVDKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMSIQFAPNKVIAEC 278
Stylonychia lemnae         228 RPKDKPLRLPLQDVYKIGGIGTVPVGRVETGLLKPGMVLTFAPMNITTEC 277
Euglena gracilis           228 RPSDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGDVVTFAPNNLTTEV 277
Trypanosoma brucei         228 RPSDKPLRLPLQTCTKIGGIGTVPVGRVETGVMKPGDVVTFAPANVTTEV 277
Porphyra purpurea          232 RPVDKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPSGLSTEV 281
Dictyostelium discoideum   228 RPHDKPLRIPLQDVYKIGGIGTVPVGRVETGIIKPGMVVTFAPAGLSTEV 277
Triticum aestivum          228 RPSDKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFGPTGLTTEV 277
Arabidopsis thaliana       228 RPSDKPLRLPLQDVYKIGGIGTVPVGRVETGMIKPGMVVTFAPTGLTTEV 277
Saccharomyces cerevisiae   238 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVIKPGMVVTFAPAGVTTEV 287
Puccinia gramina           237 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGTIKAGMVVTFAPANVTTEV 286
Trichoderma reesei         239 RPTDKPLRLPLQDVYKIGGIGTVPVGRIETGVLKPGMVVTFAPSNVTTEV 288
Podospora anserina         239 RPTDKPLRLPLQDVYKIGGIGTVPVGRIETGILKPGMVVTFAPSNVTTEV 288
Artemia salina             240 RPTEKPLRLPLQDVYKIGGIGTVPVGRVETGIIKPGMIVTFAPANITTEV 289
Drosophila melanogaster    240 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGLLKPGMVVNFAPVNLVTEV 289
Homo sapiens               240 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMVVTFAPVNVTTEV 289
Chironex fleckeri          240 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGIIKPGMVVHFSPANITTEV 289
Nematostella vectensis     246 RPSGLPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGTVVTFSPSNITTEV 295
Hydra vulgaris             242 RPSSKPLRLPLQDVYKIGGIGTVPVGRVETGILKPGMVVTFCPANLSTEV 291
Podocoryne carnea          248 RPSNKPLRLPLQDVYKIGGIGTVPVGRVETGKIMPGMVVTFAPCGISTEV 297
Cladonema radiatum         245 RPSDKPLRLPLQDVYKIGGIGTVPVGRVETGIIKPGIVVTFAPANVTTEV 294
Scleronephthya gracillimum 240 RPTDKPLRLPLQDVYKIGGIGTVPVGRTETGILKPGMVVKFAPVNITTEV 289
Pocillopora damicornis     240 RPTKKALRLPLQDVYKIGGIGTVPVGRVETGILKPGMVVTFAPTGLSTEV 289
Axinella verrucosa         240 RPTEKPLRLPLQDVYKIGGIGTVPVGRVETGILKPGMVVTISPANITTEV 289
Geodia cydonium            239 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGILKPGTVVTFSPAGISTEV 288
Xenopus laevis             240 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGILKPGMVVTFAPVNITTEV 289
Danio rerio                240 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGTLKAGMIVTFAPANVTTEV 289
Mus musculus               240 RPTDKPLRLPLQDVYKIGGIGTVPVGRVETGILRPGMVVTFAPVNITTEV 289
                           251 **    **.***   ** ********* *** .  *  .   *  . .*  300
```

Supplementary Figure 5.1 (*continued*)

```
Entamoeba histolytica      276 KSIEMHHTALAQAIPGDNVGFNVRNLTVKDIKRGNVASDAKNQPAVGCED 325
Plasmodium falciparum      276 KSVEMHKEVLEEARPGDNIGFNVKNVSVKEIKRGYVASDTKNEPAKGCSK 325
Tetrahymena pyriformis     279 KSVEMHHEQLPEAVPGDNVGFNIKGVSVKDIRRGNVASDAKNDPAKEAAT 328
Stylonychia lemnae         278 KSVEMHHESLTEAEPGDNVGFTVKNLSVKDLRRGYVASDSKNDPAKDTTN 327
Euglena gracilis           278 KSVEMHHEALTEAVPGDNVGFNVKNVSVKDIRRGYVASNAKNDPAKEAAD 327
Trypanosoma brucei         278 KSIEMHHEQLAEATPGDNVGFNVKNVSVKDIRRGNVCGNTKNDPPKEAAD 327
Porphyra purpurea          282 KSVEMHHEALTQAGPGDNVGFNVKNVSVKDLKRGYVCGDSKNDPPKGCAS 331
Dictyostelium discoideum   278 KSVEMHHEQLPEARPGDNVGFNVKNVSVKEIKRGMVAGDSKNDPPQETEK 327
Triticum aestivum          278 KSVEMHHESLLEALPGDNVGFNVKNVAVKDLKRGFVASNSKDDPAKEAAN 327
Arabidopsis thaliana       278 KSVEMHHESLLEALPGDNVGFNVKNVAVKDLKRGYVASNSKDDPAKGAAN 327
Saccharomyces cerevisiae   288 KSVEMHHEQLEQGVPGDNVGFNVKNVSVKEIRRGNVCGDAKNDPPKGCAS 337
Puccinia gramina           287 KSVEMHHEQLEAGMPGDNVGFNVKNVSVKDIRRGNVCGDTKCDPPKEAAS 336
Trichoderma reesei         289 KSVEMHHEQLAEGQPGDNVGFNVKNVSVKEIRRGNVAGDSKNDPPMGAAS 338
Podospora anserina         289 KSVEMHHEQLAEGVPGDNVGFNVKNVSVKEIRRGNVAGDSKNDPPMGAAS 338
Artemia salina             290 KSVEMHHESLEQASPGDNVGFNVKNVSVKELRRGYVASDSKNNPARGSQD 339
Drosophila melanogaster    290 KSVEMHHEALTEAMPGDNVGFNVKNVSVKELRRGYVAGDSKNNPPRGAAD 339
Homo sapiens               290 KSVEMHHEALSEALPGDNVGFNVKNVSVKDVRRGNVAGDSKNDPPMEAAG 339
Chironex fleckeri          290 KSVEMHHEALPEALPGDNVGFNIKNVSVKEIKRGMVCGDSKNDPPKEAKN 339
Nematostella vectensis     296 KSVEMHHESLAEALPGDNVGFNVKNVSVKDIKRGNVAGDFKNNPPKPCKS 345
Hydra vulgaris             292 KSVEMHHESLPEALPGDNVGFNVKNVSIKDIRRGMVASDSKNDPAIEAAS 341
Podocoryne carnea          298 KSVEMHHTAMPEALPGDNVGFNVKNVSIKEIKRGMVASDSKNDPAKEAKT 347
Cladonema radiatum         295 KSVEMHHEALPEALPGDNVGFNVKNVSVKDIKRGMVASDSKNDPAKESKS 344
Scleronephthya gracillimum 290 KSVEMHHEALAEALPGDNVGFNVKNVSVKDIKRGNVAGDSKNDPPQQAKN 339
Pocillopora damicornis     290 KSVEMHHEALTEALPGDNVGFNVKNVSVKEIKRGNVAGDTKNDPPKPTNS 339
Axinella verrucosa         290 KSVEMHHESLTEALPGDNVGFNIKNVSVKEVRRGMVAGDSKNDPPKAAKS 339
Geodia cydonium            289 KSVEMHHEALTEALPGDNVGFNVKNISVKDIKRGMVAGDSKNDPPKEAKT 338
Xenopus laevis             290 KSVEMHHEALSEALPGDNVGFNVKNVSVKDIRRGNVCGDSKTDPPQEAAG 339
Danio rerio                290 KSVEMHHESLTEALPGDNVGFNVKNVSVKDIRRGNVAGDSKNDPPMEAAN 339
Mus musculus               290 KSVEMHHEALSEALPGDNVGFNVKNVSVKDIRRGNVCGDSKADPPQEAAQ 339
                           301 **.***.  .    ****.**.......*...** *.   *  *   .   350

Entamoeba histolytica      326 FTAQVIVMNHPGQIRKGYTPVLDCHTSHIACKFEELLSKIDRRTGKSMEG 375
Plasmodium falciparum      326 FTAQVIILNHPGEIKNGYTPLLDCHTSHISCKFLNIDSKIDKRSGKVVEE 375
Tetrahymena pyriformis     329 FYSQVIIMNHPGQIQAGYTPVLDCHTAHIACKFETIHDKIDRRTGKSQEE 378
Stylonychia lemnae         328 FLAQVIVLNHPGQIQKGYAPVLDCHTAHIACKFDEIESKVDRRSGKVLEE 377
Euglena gracilis           328 FTAQVIILNHPGQIGNGYAPVLDCHTCHIACKFATIQTKIDRRSGKELEA 377
Trypanosoma brucei         328 FTAQVIILNHPGQIGNGYAPVLDCHTSHIACKFAEIESKIDRRSGKELEK 377
Porphyra purpurea          332 FNAQVIILNHPGEIHAGYAPVLDCHTAHIACKFSELILKMDRRSGKKLED 381
Dictyostelium discoideum   328 FVAQVIVLNHPGQIHAGYSPVLDCHTAHIACKFTEIVDKVDRRTGAVVAK 377
Triticum aestivum          328 FTSQVIIMNHPGQIGNGYAPVLDCHTSHIAVKFAELVTKIDRRSGKELEA 377
Arabidopsis thaliana       328 FTSQVIIMNHPGQIGNGYAPVLDCHTSHIAVKFSEILTKIDRRSGKEIEK 377
Saccharomyces cerevisiae   338 FNATVIVLNHPGQISAGYSPVLDCHTAHIACRFDELLEKNDRRSGKKLED 387
Puccinia gramina           337 FVAQVIVLNHPGQIGNGYCPVLDCHTAHIACKFDTLQQKIDRRTGKALEE 386
Trichoderma reesei         339 FTAQVIVMNHPGQVGAGYAPVLDCHTAHIACKFAELLEKIDRRTGKATES 388
Podospora anserina         339 FDAQVIVLNHPGQVGAGYAPVLDCHTAHIACKFSELLQKIDRRTGKAVEE 388
Artemia salina             340 FFAQVIVLNHPGQISNGYTPVLDCHTAHIACKFAEIKEKCDRRTGKTTEA 389
Drosophila melanogaster    340 FTAQVIVLNHPGQIANGYTPVLDCHTAHIACKFSEIKEKCDRRTGKTTET 389
Homo sapiens               340 FTAQVIILNHPGQISAGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLED 389
Chironex fleckeri          340 FTAQVIILNHPGEIHAGYSPVLDCHTAHIACKFSELKEKNDRRSGKKLED 389
Nematostella vectensis     346 FTAQVIVMNHPGEIHAGYSPVLDCHTAHIACKFDKLLEKIDRRSGKKLED 395
Hydra vulgaris             342 FKAQVIILNHPGEIHAGYQPVLDCHTAHIACKFAELLEKIDRRSGKVIET 391
Podocoryne carnea          348 FYAQVIILNHPGEIHAGYQPVLDCHTAHVACKFTELKQKCDRRSGKVLEE 397
Cladonema radiatum         345 FLAQVIILNHPGEIHAGYQPVLDCHTAHVACKFSEIQQKIDRRSGKVLEE 394
Scleronephthya gracillimum 340 FTAQVIIMNHPGEIHAGYSPVLDCHTAHIACKFAELKQKIDRRSGKATEE 389
Pocillopora damicornis     340 FNAQVIVLNHPGEIHAGYSPVLDCHTAHIACKFAELLEKIDRRSGKKLED 389
Axinella verrucosa         340 FLAQVIILNHPGQIHAGYSPVLDCHTAHIACKFAELKEKCDRRSGKKLED 389
Geodia cydonium            339 FTAQVIVLNHPGQIHAGYAPVLDCHTAHIACKFTELKEKIDR-SGKKLED 387
Xenopus laevis             340 FTAQVIILNHPGQISAGYSPVIDCHTAHIACKFAELKEKIDRRSGKKLED 389
Danio rerio                340 FTSQVIILNHPGQISQGYAPVLDCHTAHIACKFAELKEKIDRRSGKKLED 389
Mus musculus               340 FTSQVIILNHPGQISAGYSPVIDCHTAHIACKFAELKEKIDRRSGKKLED 389
                           351 * . **..****..  ** *..****.*.. .*   .  * *. .*      400
```

Supplementary Figure 5.1 (*continued*)

```
Entamoeba histolytica      376 GEP--EYIKNGDSALVKIVPTKPLCVEEFAKFPPLGRFAVRDMKQTVAVG 423
Plasmodium falciparum      376 NP---KAIKSGDSALVSLEPKKPMVVETFTEYPPLGRFAIRDMRQTIAVG 422
Tetrahymena pyriformis     379 NP---KFIKNGDAALVTLIPTKALCVEVFQEYPPLGRYAVRDMKQTVAVG 425
Stylonychia lemnae         378 EP---KFIKSGEAALVRMVPQKPMCVEAFNQYPPLGRFAVRDMKQTVAVG 424
Euglena gracilis           378 EP---KFIKSGDAAIVLMKPQKPMCVESFTDYPPLG-VSCGDMRQTVAVG 423
Trypanosoma brucei         378 AP---KSIKSGDAAIVRMVPQKPMCVEVFNDYAPLGRFAVRDMRQTVAVG 424
Porphyra purpurea          382 SP---KMIKSGDAAMVKMVASKPMCVEAFTSYPPLGRFAVRDMRQTVAVG 428
Dictyostelium discoideum   378 EGTAAVVLKNGDAAMVELTPSRPMCVESFTEYPPLGRFAVRDMRQTVAVG 427
Triticum aestivum          378 LP---KFLKNGDAGIVKMIPTKPMVVETFATYPPLGRFAVRDMRQTVAVG 424
Arabidopsis thaliana       378 EP---KFLKNGDAGMVKMTPTKPMVVETFSEYPPLGRFAVRDMRQTVAVG 424
Saccharomyces cerevisiae   388 HP---KFLKSGDAALVKFVPSKPMCVEAFSEYPPLGRFAVRDMRQTVAVG 434
Puccinia gramina           387 AP---KFVKSGDACLVKMVPSKAMCVEPFADYPPLGRFAVRDMRQTVAVG 433
Trichoderma reesei         389 AP---KFIKSGDSAIVKMIPSKPMCVEAFTDYPPLGRFAVRDMRQTVAVG 435
Podospora anserina         389 SP---KFIKSGDAAIVKMVPSKPMCVEAFTEYPPLGRFAVRDMRQTVAVG 435
Artemia salina             390 EP---KFIKSGDAAMITLVPSKPLCVEAFSDFPPLGRFAVRDMRQTVAVG 436
Drosophila melanogaster    390 EP---KAIKSGDAAIIVLVPSKPLCVESFQEFPPLGRFAVRDMRQTVAVG 436
Homo sapiens               390 GP---KFLKSGDAAIVDMVPGKPMCVESFSDYPPLGRFAVRDMRQTVAVG 436
Chironex fleckeri          390 EP---KSVKSGDAALIVLVPSKPMCVEPYSAYPPLGRFAVRDMKQTVAVG 436
Nematostella vectensis     396 NP---KMIKTGDAAMVEMIPSKPMCVETFTEFPPLGRFAVRDMKQTVAVG 442
Hydra vulgaris             392 EP---KMVKSGDAAIINLIPSKGMCVESFSQYPPLGRFAVRDMRQTVAVG 438
Podocoryne carnea          398 NP---KLVKSGDAAMITLTPSKPMCVEAFSDYAPLGRFAVRDMRQTVAVG 444
Cladonema radiatum         395 NP---KMVKSGDAAMINLVPSKPMCVESFASYPPLGRFAVRDMRQTVAVG 441
Scleronephthya gracillimum 390 NP---KCLKKGDAGIVVMIPSKPMCVETFTEYPPLGRFAVRDMKQTVAVG 436
Pocillopora damicornis     390 NP---KMIKSGDASIVKMIPSKPMCVEKFSDYAPLGRFAVRDMRQTVAVG 436
Axinella verrucosa         390 NP---KSVKSGDAAIVELVPSKPMCVEAFSEYPPLGRFAVRDMKQTVAVG 436
Geodia cydonium            388 NP---KFVKSGDAAIVILTPSKPMCVESFSEYPPLGRFAVRDMKQTVAVG 434
Xenopus laevis             390 NP---KSLKSGDAAIVEMVPGKPMCVESFSQYPPLGRFAVRDMRQTVAVG 436
Danio rerio                390 NP---KALKSGDAAIILMIPGKPMCVESFSQYPPLGRFAVRDMRQTVAVG 436
Mus musculus               390 NP---KSLKSGDAAIVEMVPGKPMCVESFSQYPPLGRFAVRDMRQTVAVG 436
                           401        .* *.. ..     . . ** .  . ***  .  **.**.*** 450

Entamoeba histolytica      424 VVKAVTP----------------------- 430
Plasmodium falciparum      423 IINQLKRKNLGAVTAKAPAKK--------- 443
Tetrahymena pyriformis     426 VIKKVEKKDK-------------------- 435
Stylonychia lemnae         425 VIKEVVKKEQKG-MVTKAAQKKK------- 446
Euglena gracilis           424 VIKSVNKKENTG-KVTKAAQKKK------- 445
Trypanosoma brucei         425 IIKAVTKKDGSGGKVTKAAVKASKK----- 449
Porphyra purpurea          429 VIKSVEKKEVEG-KMTKSAAKK-------- 449
Dictyostelium discoideum   428 VIKSTVKKAPGKAGDKKGAAAPSKKK---- 453
Triticum aestivum          425 VIKGVEKKDPTGAKVTKAAIKKK------- 447
Arabidopsis thaliana       425 VIKSVDKKDPTGAKVTKAAVKKGAK----- 449
Saccharomyces cerevisiae   435 VIKSVDKTE-KAAKVTKAAQKAAKK----- 458
Puccinia gramina           434 VIKSVEKTDGKAGKVTKAAVKPVPRSSYAR 463
Trichoderma reesei         436 VIKAVEKSSAAAAKVTKSAAKAAKK----- 460
Podospora anserina         436 VIKKVEKAAAGSGKVTKSAAKAGKK----- 460
Artemia salina             437 VIKSVNFKDPTAGKVTKAAEKAGKKK---- 462
Drosophila melanogaster    437 VIKSVNFKETTSGKVTKAAEKAQKKK---- 462
Homo sapiens               437 VIKAVDKKAAGAGKVTKSAQKAQKAK---- 462
Chironex fleckeri          437 VIKAVEKSDKGGGKVTKAAQKAGNPKKK-- 464
Nematostella vectensis     443 VIKSVDKTEAAGGKTTKAATKVIKGGKK-- 470
Hydra vulgaris             439 VIKEVDKTVAVAGKVTKSAQKAGVAAGKKK 468
Podocoryne carnea          445 VIKSVEKAEAGG-KMTKSAQKA--TGGKKK 471
Cladonema radiatum         442 VIKSVEKLKEER------------------ 453
Scleronephthya gracillimum 437 VIKAVEKSDVAAGKVTKAAQKAAKK----- 461
Pocillopora damicornis     437 VIKDVDKSEVAG-KVTKAAQKATGKKK--- 462
Axinella verrucosa         437 VIKSVEKQEGKGGKVTKAAQKAQKKK---- 462
Geodia cydonium            435 VIKSVEKKEKEG-KATKSAQKKGK------ 457
Xenopus laevis             437 VIKNVEKKIGGAGKVTKSAQKAQKAAK--- 463
Danio rerio                437 VIKAVDKKASSGGKVTKSAQKAAKTK---- 462
Mus musculus               437 VIKNVEKKSGGAGKVTKSAQKAQKAGK--- 463
                           451 ..                             480
```

Supplementary Figure 5.1 (*continued*)

APPENDIX 6

CfTX-1 Mass Fingerprinting Data - APAF

6.1 Background

Trypsin-digested CfTX-1 was analysed by MALDI-TOF-TOF mass spectrometry at APAF. A peak list, containing a list of measured monoisotopic masses (MH^+) and intensities of ionised peptides, was generated from the mass spectrum. The peak list was subjected to a Mascot Ion/Ion Search (Matrix Sciences; http://www.matrixscience.com/) to compare the experimental mass data for the CfTX peptides with calculated peptide mass values of peptides obtained from theoretically-digested amino acid sequences contained in a comprehensive protein database (e.g. NCBInr, the NCBI non-redundant protein database). The top-scoring match of the search was CqTX-A, a venom protein isolated from the box jellyfish *C. quadrigatus*.

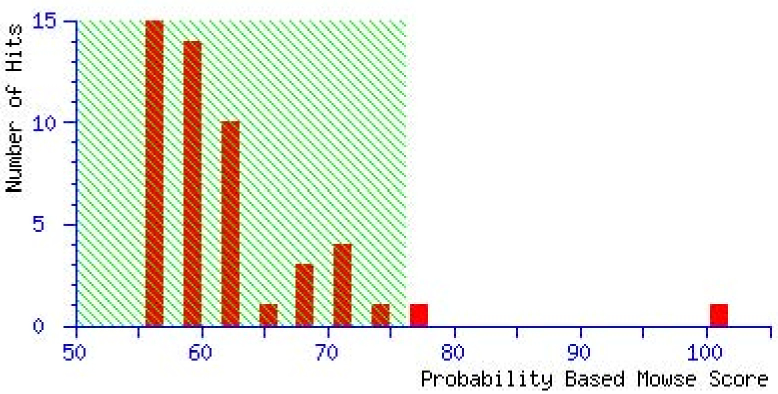
6.2 Mascot Ion/Ion Search Report

MATRIX SCIENCE Mascot Search Results

User : Mascot Daemon Email : aconnoll@proteome.org.au Search title : Project: Default, Spot Set: Default\050621jj_885397days_gs72-80, L MS data file : C:\ABMALDIData\ExportT2D\Default\050621jj_885397days_gs72-80\berni Database : NCBInr 20050526 (2471633 sequences; 837637098 residues) Timestamp : 24 Jun 2005 at 01:35:35 GMT Top Score : 101 for gi|20137750, Toxin A precursor (CqTX-A)

Probability Based Mowse Score

Score is -10*Log(P), where P is the probability that the observed match is a random event. **Protein scores greater than 76 are significant (p<0.05).**



Protein Summary Report

Switch to Peptide Summary Report

To create a bookmark for this report, right click this link: Protein Summary Report (Project: Default, Spot Set: Default\050621jj_885397days_gs72-80, Label: D1, Spot Id: 305133, Peak List Id: 519026, MS Job Run Id: 24254)

Index

	Accession	Mass	Score	Description
1.	gi|20137750	51769	101	Toxin A precursor (CqTX-A)
2.	gi|46100519	76281	77	hypothetical protein UM04979.1 [Ustilago maydi
3.	gi|15147837	60605	73	replicase [Soybean dwarf virus]
4.	gi|15147831	60613	72	replicase [Soybean dwarf virus]
5.	gi|436018	60542	72	RNA-dependent RNA polymerase [Soybean dwarf vi
6.	gi|62638861	18825	71	PREDICTED: similar to zinc finger protein 580
7.	gi|61556917	32947	71	four and a half LIM domains 5 (predicted) [Rat
8.	gi|48096814	779126	69	similar to Zinc finger protein 208 [Apis melli
9.	gi|294256	72151	68	major surface glycoprotein
10.	gi|29341421	128924	67	DNA helicase [Bacteroides thetaiotaomicron VPI
11.	gi|56786616	280543	65	P-512 [Borrelia hermsii]
12.	gi|29789189	98372	64	minichromosome maintenance protein 10 [Mus mus
13.	gi|63101617	98400	64	Minichromosome maintenance protein 10 [Mus mus
14.	gi|51228146	120232	64	Major Surface Glycoprotein (MSG), putative [Pn
15.	gi|49614454	15360	63	structural glycoprotein E2 [Bovine viral diarr
16.	gi|6606119	105083	63	DNA-dependent RNA polymerase II RPB140 [Dothid
17.	gi|48728671	16872	62	COG0789: Predicted transcriptional regulators
18.	gi|21228907	53887	61	Asparagine synthetase [glutamine-hydrolyzing]
19.	gi|439877	42516	61	reverse transcriptase
20.	gi|23619193	658551	61	hypothetical protein [Plasmodium falciparum 3D
21.	gi|29339896	32164	61	hypothetical protein [Bacteroides thetaiotaomi
22.	gi|31211211	31277	61	ENSANGP00000017908 [Anopheles gambiae str. PES
23.	gi|62650904	13811	61	PREDICTED: hypothetical protein XP_576055 [Rat
24.	gi|20451215	71546	60	Hypothetical protein C56C10.1 [Caenorhabditis
25.	gi|39934900	9546	60	hypothetical protein RPA1831 [Rhodopseudomonas
26.	gi|50740365	26639	60	PREDICTED: similar to SERTA domain containing
27.	gi|58039342	31193	60	hypothetical protein GOX0877 [Gluconobacter ox
28.	gi|31430304	88111	59	putative retroelement [Oryza sativa (japonica
29.	gi|6572987	8348	59	S-locus cysteine-rich protein [Brassica olerac
30.	gi|53804113	90090	59	fatty acid cis/trans isomerase, putative [Meth
31.	gi|61878917	29548	58	PREDICTED: similar to Hypothetical protein C6o
32.	gi|2314419	39465	58	H. pylori predicted coding region HP1247 [Heli
33.	gi|1519660	66969	58	Hypothetical protein F55A4.6 [Caenorhabditis e
34.	gi|56472336	140636	58	SMC3 protein, putative [Entamoeba histolytica
35.	gi|56419217	141182	58	ATP-dependent deoxyribonuclease subunit A [Geo
36.	gi|47223169	71554	58	unnamed protein product [Tetraodon nigroviridi
37.	gi|28629308	23063	57	putative protein [Escherichia coli]
38.	gi|12653257	52197	57	G patch domain and KOW motifs [Homo sapiens]
39.	gi|10946622	32884	57	four and a half LIM domains 5 [Mus musculus]
40.	gi|46242771	49454	57	Ctn003 [Bacteroides fragilis]
41.	gi|24429599	20446	57	hypothetical protein [Bradyrhizobium japonicum
42.	gi|58698322	32841	57	transposase-like [Wolbachia endosymbiont of Dr
43.	gi|42522429	26396	57	hypothetical protein Bd0857 [Bdellovibrio bact
44.	gi|13183552	399181	57	dystrophin-like protein DLP2 [Drosophila melan
45.	gi|23171723	399155	57	CG31175-PA, isoform A [Drosophila melanogaster

46. gi|39583795 86747 57 Hypothetical protein CBG22727 [Caenorhabditis
47. gi|19069618 32056 57 TRANSLATION INITIATION FACTOR 2 ALPHA SUBUNIT
48. gi|62088576 261846 57 low density lipoprotein-related protein 1 vari
49. gi|27383051 51084 56 putative transposase [Bradyrhizobium japonicum
50. gi|49526121 304286 56 unnamed protein product [Candida glabrata CBS1

Results List (truncated)

1. gi|20137750 **Mass:** 51769 **Score:** 101
Toxin A precursor (CqTX-A)

Observed	Mr(expt)	Mr(calc)	Delta	Start End	Miss	Peptide
777.41	776.40	776.42	-0.02	167 - 173	0	AFLDGVR
777.41	776.40	776.42	-0.02	167 - 173	0	AFLDGVR
963.52	962.51	962.54	-0.03	199 - 206	0	FIAMVVQR (M_o)
963.52	962.51	962.54	-0.03	199 - 206	0	FIAMVVQR (M_o)
979.51	978.50	978.53	-0.03	199 - 206	0	FIAMVVQR ($2M_o$)
979.51	978.50	978.53	-0.03	199 - 206	0	FIAMVVQR
1043.53	1042.52	1042.55	-0.03	307 - 314	1	FFDLMKVK (M_o)
1564.76	1563.75	1563.79	-0.04	393 - 405	1	NDDLEKFNFVPVK
1811.00	1809.99	1809.94	0.06	221 - 235	0	VLSMLELFTDLCSLR (C_a)
1826.98	1825.97	1825.93	0.04	221 - 235	0	VLSMLELFTDLCSLR (M_o)

No match to: 842.49, 843.41, 864.41, 917.56, 917.56, 937.42, 1010.52, 1050.56, 1187.62, 1187.62, 1255.61, 1364.67, 1377.53, 1466.69, 1764.89, 1888.86, 1934.97, 2034.95, 2036.94, 2036.94, 2048.96, 2163.05, 2256.09, 2287.14, 2711.87, 2728.34, 2728.34

MATRIX SCIENCE Mascot Search Results

Search Parameters

Type of search : MS/MS Ion Search
Enzyme : Trypsin
Variable modifications : Oxidation (M),Propionamide (C)
Mass values : Monoisotopic
Protein Mass : Unrestricted
Peptide Mass Tolerance : ± 100 ppm
Fragment Mass Tolerance: ± 0.3 Da
Max Missed Cleavages : 1
Instrument type : MALDI-TOF-TOF
Data File Name : C:\ABMALDIData\ExportT2D\Default\050621jj_885397days_gs72-80

Query1 (777.41,1+) : <no title>
Query2 (777.41,1+) : Label: D1, Spot_Id: 305133, Peak_List_Id: 519035, MSMS
Query3 (842.49,1+) : <no title>
Query4 (843.41,1+) : <no title>
Query5 (864.41,1+) : <no title>
Query6 (917.56,1+) : <no title>
Query7 (917.56,1+) : Label: D1, Spot_Id: 305133, Peak_List_Id: 519030, MSMS
Query8 (937.42,1+) : <no title>
Query9 (963.52,1+) : Label: D1, Spot_Id: 305133, Peak_List_Id: 519029, MSMS
Query10 (963.52,1+) : <no title>
Query11 (979.51,1+) : Label: D1, Spot_Id: 305133, Peak_List_Id: 519031, MSMS
Query12 (979.51,1+) : <no title>
Query13 (1010.52,1+): <no title>
Query14 (1043.53,1+): <no title>
Query15 (1050.56,1+): <no title>
Query16 (1187.62,1+): <no title>
Query17 (1187.62,1+): Label: D1, Spot_Id: 305133, Peak_List_Id: 519033, MSMS
Query18 (1255.61,1+): <no title>
Query19 (1364.67,1+): <no title>
Query20 (1377.53,1+): <no title>
Query21 (1466.69,1+): <no title>
Query22 (1564.76,1+): <no title>
Query23 (1764.89,1+): <no title>
Query24 (1811.00,1+): <no title>
Query25 (1826.98,1+): <no title>
Query26 (1888.86,1+): <no title>
Query27 (1934.97,1+): <no title>
Query28 (2034.95,1+): <no title>
Query29 (2036.94,1+): Label: D1, Spot_Id: 305133, Peak_List_Id: 519028, MSMS
Query30 (2036.94,1+): <no title>
Query31 (2048.96,1+): <no title>
Query32 (2163.05,1+): <no title>
Query33 (2256.09,1+): <no title>
Query34 (2287.14,1+): <no title>
Query35 (2287.14,1+): Label: D1, Spot_Id: 305133, Peak_List_Id: 519032, MSMS
Query36 (2711.87,1+): <no title>
Query37 (2728.34,1+): Label: D1, Spot_Id: 305133, Peak_List_Id: 519034, MSMS
Query38 (2728.34,1+): <no title>

{MATRIX SCIENCE} Mascot Search Results

Protein View

Match to: **gi|20137750**; Score: **101** **Toxin precursor (CqTX-A)**
Found in search of
C:\ABMALDIData\ExportT2D\Default\050621jj_885397days_gs72-80\bern

Nominal mass (M_r): **51769**; Calculated pI value: **8.63**
NCBI BLAST search of gi|20137750 against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Chiropsalmus quadrigatus
Links to retrieve other entries containing this sequence from NCBI Entrez:
gi|18146993 from Chiropsalmus quadrigatus

Variable modifications: Oxidation (M),Propionamide (C)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **11%**

Matched peptides shown in **Bold Red**

```
  1 MANMLYFSLL ALLFMTGIAS EGTISSGLAS LKAKIDAKRP SGKQLFDKVA
 51 NMQKQIEEKF SNDDERAKVM GAIGSLSTAV GKFQSGDPAK IASGCLDILV
101 GISSVLKDFA KFSPIFSILS MVVGLFSGTK AEESVGSVVK KVVQEQSDQE
151 LQEALYGVKR EYAVSKAFLD GVRNETSDLS PTEVSALGAN VPVYQGVRFI
201 AMVVQRIKNR KPRTESEIKR VLSMLELFTD LCSLRDLILL DLYQLVATPG
251 HSPNIASGIK EVSNLGREEY KKVFEDLLKT NDKETYLFLS YLYPRERNEQ
301 SQKIFKFFDL MKVKYDDRLK QDLTGIQVFS SLHWPNYFLC SSKDYLALIC
351 TKPYGSLRLD KLNDGFYSIK TTQSNPKVCH RYGEYILFTH DRNDDLEKFN
401 FVPVKLGERK IYLLSSKASP NKFAYVPKTA KGDLFFVDGI PSQLGYGNQG
451 YFTLATDENE QT
```

Mascot: http://www.matrixscience.com/

APPENDIX 7

Calculated Tryptic Peptide Monoisotopic Ion Masses of CfTX-1 and -2

7.1 MS-Digest Results for CfTX-1

Parameters

```
Database: User Protein
Considered modifications: | Oxidation of M | Acrylamide Modified Cys (Ca)|
Digest Used: Trypsin
Max. # Missed Cleavages: 3
User AA Formula 1: C2 H3 N1 O1
Cysteine Modification: acrylamide
Instrument Name: TOF-TOF
Minimum Digest Fragment Mass: 700
Maximum Digest Fragment Mass: 3000
Minimum Digest Fragment Length: 4

pI of Protein: 8.3
Protein MW: 49147
Amino Acid Composition: A23 C5 D29 E26 F24 G26 H5 I25 K44 L46 M8 N19 P16 Q16 R14
S39 T17 V32 W1 Y21

  1   ESTISSGLNSLKTKIDAKMPSGKQLFDKVVEMQKQIDAKFSNDDERAKVMGAIGSLSTAVGKFQSGDPAKIASGCLDILV
 81   GISSVLKDFAKFSPIFSILSLVVGLFSGTKAEESVGSVVKKAVQEQSDQELQEALYGVKREYAVSKAFLDGVRNETSDLS
161   PTEVSALAANVPIYQGVRFIAMVVQRIKYIKPKTESEIKRMLTMLELFTDLCSLRDLILLDLYQLVATPGHSPNIASGIK
241   EVSNLGREEYKKVFEDLLKNDDKETYLFLSYLYPREKNEQSRKIFNFFDLMKVKYDDRLKQDLTGVKIFSNVHWPNYFMC
321   SSNDYLALICTKPYGSLKLDKLNDGYYSIKTTQHDPKICHRYGNYILFTHKRNDDLEKFNFVPVKLEKREIYLLSSKESP
401   NKFAYVPQNADGALFFVDGIPSKVGYGNQGYFTLVE
```

Table 7.1

Calculated mass to charge ratio (m/z) values for monoisotopic peptide ions (mi) and average fragment masses (av) obtained by theoretical digestion of CfTX-1 amino acid sequence with trypsin. Considered modifications include oxidation of methionine (M_o) and alkylation of cysteine (C_a). Calculated monoisotopic peptide ions masses that are similar to experimental masses obtained by MALDI-TOF-TOF MS analysis (see Section 6.2) and their corresponding sequences are highlighted in green.

m/z(mi)	m/z(av)	Mod.	Start	End	Missed Cleavage	Sequence
706.3618	706.7752		194	199	0	(K)TESEIK(R)
733.3363	733.7574		373	378	0	(R)NDDLEK(F)
733.3913	733.9103		29	34	0	(K)VVEMQK(Q)
749.3862	749.9097	1M_o	29	34	0	(K)VVEMQK(Q)
760.4199	760.8705		301	307	0	(K)QDLTGVK(I)
761.3900	761.8174		278	283	1	(K)NEQSRK(I)
774.4104	774.8569		241	247	0	(K)EVSNLGR(E)
777.4254	**777.9038**		**147**	**153**	**0**	**(K)AFLDGVR(N)**
795.3995	795.8755		293	298	1	(K)VKYDDR(L)
809.4152	809.9025		295	300	1	(K)YDDRLK(Q)
826.4054	826.8896		351	357	0	(K)TTQHDPK(I)
849.4101	849.9241		63	70	0	(K)FQSGDPAK(I)
850.4822	851.0424		379	385	0	(K)FNFVPVK(L)
852.4574	852.9714		140	146	1	(K)REYAVSK(A)
862.4629	862.9638		194	200	1	(K)TESEIKR(M)
863.4873	864.0355		253	259	0	(K)VFEDLLK(N)
882.3588	882.8669		40	46	0	(K)FSNDDER(A)
889.4374	889.9460		372	378	1	(K)RNDDLEK(F)
889.5870	890.1639		187	193	1	(R)IKYIKPK(T)
890.4326	890.9337		276	282	1	(R)EKNEQSR(K)
946.5026	947.1482		15	23	1	(K)IDAKMPSGK(Q)
952.5350	953.1301		390	397	0	(R)EIYLLSSK(E)
962.4975	963.1476	1M_o	15	23	1	(K)IDAKMPSGK(Q)

m/z(mi)	m/z(av)	Mod.	Start	End	Missed Cleavage	Sequence
963.5444	**964.2252**		**179**	**186**	**0**	**(R)FIAMVVQR(I)**
979.5393	**980.2245**	**$1M_o$**	**179**	**186**	**0**	**(R)FIAMVVQR(I)**
991.5823	992.2107		252	259	1	(K)KVFEDLLK(N)
1001.5990	1002.2061		299	307	1	(R)LKQDLTGVK(I)
1004.5259	1005.1195		111	120	0	(K)AEESVGSVVK(K)
1018.5276	1019.1088		276	283	2	(R)EKNEQSRK(I)
1036.5786	1037.2111		293	300	2	(K)VKYDDRLK(Q)
1072.5310	1073.1978		342	350	0	(K)LNDGYYSIK(T)
1081.4909	1082.1213		40	48	1	(K)FSNDDERAK(V)
1108.6361	1109.3187		389	397	1	(K)REIYLLSSK(E)
1132.6208	1133.2946		111	121	1	(K)AEESVGSVVKK(A)
1150.5925	1151.3783		19	28	1	(K)MPSGKQLFDK(V)
1166.5874	1167.3777	$1M_o$	19	28	1	(K)MPSGKQLFDK(V)
1174.5965	1175.4438		284	292	0	(K)IFNFFDLMK(V)
1175.6453	1176.4291		13	23	2	(K)TKIDAKMPSGK(Q)
1190.5914	1191.4432	$1M_o$	284	292	0	(K)IFNFFDLMK(V)
1191.6402	1192.4285	$1M_o$	13	23	2	(K)TKIDAKMPSGK(Q)
1204.7235	1205.5608		179	188	1	(R)FIAMVVQRIK(Y)
1220.7038	1221.4943		379	388	1	(K)FNFVPVKLEK(R)
1220.7184	1221.5602	$1M_o$	179	188	1	(R)FIAMVVQRIK(Y)
1235.6478	1236.3727		1	12	0	(-)ESTISSGLNSLK(T)
1255.6470	**1256.4556**		**362**	**371**	**0**	**(R)YGNYILFTHK(R)**
1288.6929	1289.5459		29	39	1	(K)VVEMQKQIDAK(F)
1290.7086	1291.5619		49	62	0	(K)VMGAIGSLSTAVGK(F)
1302.6915	1303.6190		283	292	1	(R)KIFNFFDLMK(V)
1304.6879	1305.5453	$1M_o$	29	39	1	(K)VVEMQKQIDAK(F)
1306.7035	1307.5613	$1M_o$	49	62	0	(K)VMGAIGSLSTAVGK(F)
1318.6864	1319.6184	$1M_o$	283	292	1	(R)KIFNFFDLMK(V)
1323.6539	1324.4417		241	251	1	(K)EVSNLGREEYK(K)
1335.6586	1336.5244		351	361	1	(K)TTQHDPKICHR(Y)
1335.6791	1336.4934		253	263	1	(K)VFEDLLKNDDK(E)
1335.7518	1336.5807		189	199	1	(K)YIKPKTESEIK(R)
1364.7242	**1365.6445**		**24**	**34**	**1**	**(K)QLFDKVVEMQK(Q)**
1376.8049	1377.6829		379	389	2	(K)FNFVPVKLEKR(E)
1380.7192	1381.6439	$1M_o$	24	34	1	(K)QLFDKVVEMQK(Q)
1401.7599	1402.7524		284	294	1	(K)IFNFFDLMKVK(Y)
1406.6957	1407.7382	$1C_a$	351	361	1	(K)TTQHDPKICHR(Y)
1411.7481	1412.6443		362	372	1	(R)YGNYILFTHKR(N)
1417.7548	1418.7518	$1M_o$	284	294	1	(K)IFNFFDLMKVK(Y)
1428.7369	1429.6226		339	350	1	(K)LDKLNDGYYSIK(T)
1437.6605	1438.5025		35	46	1	(K)QIDAKFSNDDER(A)
1446.6900	1447.5970		424	436	0	(K)VGYGNQGYFTLVE(-)
1451.7489	1452.6168		241	252	2	(K)EVSNLGREEYKK(V)
1454.7638	1455.6638		141	153	1	(R)EYAVSKAFLDGVR(N)
1463.7740	1464.6686		252	263	2	(K)KVFEDLLKNDDK(E)
1464.7904	1465.6536		1	14	1	(-)ESTISSGLNSLKTK(I)
1478.8577	1479.7706		386	397	2	(K)LEKREIYLLSSK(E)
1489.8407	1490.8164		47	62	1	(R)AKVMGAIGSLSTAVGK(F)
1491.8530	1492.7694		189	200	2	(K)YIKPKTESEIKR(M)
1505.8356	1506.8157	$1M_o$	47	62	1	(R)AKVMGAIGSLSTAVGK(F)
1507.8003	1508.7221		390	402	1	(R)EIYLLSSKESPNK(F)
1529.8549	1530.9276		283	294	2	(R)KIFNFFDLMKVK(Y)
1540.8257	1541.7955		248	259	2	(R)EEYKKVFEDLLK(N)
1545.8498	1546.9270	$1M_o$	283	294	2	(R)KIFNFFDLMKVK(Y)
1550.8173	1551.7503		295	307	2	(K)YDDRLKQDLTGVK(I)
1564.8006	**1565.7770**		**373**	**385**	**1**	**(R)NDDLEKFNFVPVK(L)**
1564.8046	**1565.8204**		**264**	**275**	**0**	**(K)ETYLFLSYLYPR(E)**
1576.9309	1577.9164		187	199	2	(R)IKYIKPKTESEIK(R)
1577.8356	1578.8824		15	28	2	(K)IDAKMPSGKQLFDK(V)
1593.8305	1594.8818	$1M_o$	15	28	2	(K)IDAKMPSGKQLFDK(V)
1610.8649	1611.8524		140	153	2	(K)REYAVSKAFLDGVR(N)
1636.7925	1637.7570		35	48	2	(K)QIDAKFSNDDERAK(V)
1663.9014	1664.9107		389	402	2	(K)REIYLLSSKESPNK(F)
1687.9663	1689.0802		71	87	0	(K)IASGCLDILVGISSVLK(D)
1720.9017	1721.9657		372	385	2	(K)RNDDLEKFNFVPVK(L)
1759.0034	1760.2940	$1C_a$	71	87	0	(K)IASGCLDILVGISSVLK(D)
1764.9003	**1766.0904**		**358**	**371**	**1**	**(K)ICHRYGNYILFTHK(R)**

m/z(mi)	m/z(av)	Mod.	Start	End	Missed Cleavage	Sequence
1785.8948	1787.2264		201	215	0	(R)MLTMLELFTDLCSLR(D)
1801.8897	1803.2257	$1M_o$	201	215	0	(R)MLTMLELFTDLCSLR(D)
1817.8846	1819.2251	$2M_o$	201	215	0	(R)MLTMLELFTDLCSLR(D)
1821.9422	1823.1118		264	277	1	(K)ETYLFLSYLYPREK(N)
1834.1135	1835.3663		179	193	2	(R)FIAMVVQRIKYIKPK(T)
1835.9374	1837.3042	$1C_a$	358	371	1	(K)ICHRYGNYILFTHK(R)
1850.1085	1851.3657	$1M_o$	179	193	2	(R)FIAMVVQRIKYIKPK(T)
1856.9319	1858.4402	$1C_a$	201	215	0	(R)MLTMLELFTDLCSLR(D)
1864.9660	1866.2658		19	34	2	(K)MPSGKQLFDKVVEMQK(Q)
1872.9268	1874.4395	$1M_o$ $1C_a$	201	215	0	(R)MLTMLELFTDLCSLR(D)
1879.9185	1881.0647		342	357	1	(K)LNDGYYSIKTTQHDPK(I)
1880.9609	1882.2652	$1M_o$	19	34	2	(K)MPSGKQLFDKVVEMQK(Q)
1888.9217	**1890.4389**	**$2M_o$ $1C_a$**	**201**	**215**	**0**	**(R)MLTMLELFTDLCSLR(D)**
1892.0335	1893.1577		1	18	2	(-)ESTISSGLNSLKTKIDAK(M)
1896.9558	1898.2646	$2M_o$	19	34	2	(K)MPSGKQLFDKVVEMQK(Q)
1916.9687	1918.2385		278	292	2	(K)NEQSRKIFNFFDLMK(V)
1920.0259	1921.2801		24	39	2	(K)QLFDKVVEMQKQIDAK(F)
1921.0014	1922.2791		358	372	2	(K)ICHRYGNYILFTHKR(N)
1932.9636	1934.2379	$1M_o$	278	292	2	(K)NEQSRKIFNFFDLMK(V)
1935.0222	**1936.2289**		**373**	**388**	**2**	**(R)NDDLEKFNFVPVKLEK(R)**
1936.0208	1937.2795	$1M_o$	24	39	2	(K)QLFDKVVEMQKQIDAK(F)
1941.9959	1943.4150		200	215	1	(K)RMLTMLELFTDLCSLR(D)
1950.9782	1952.2966		284	298	2	(K)IFNFFDLMKVKYDDR(L)
1957.9908	1959.4144	$1M_o$	200	215	1	(K)RMLTMLELFTDLCSLR(D)
1966.9731	1968.2959	$1M_o$	284	298	2	(K)IFNFFDLMKVKYDDR(L)
1973.9857	1975.4137	$2M_o$	200	215	1	(K)RMLTMLELFTDLCSLR(D)
1992.0385	1993.4929	$1C_a$	358	372	2	(K)ICHRYGNYILFTHKR(N)
2012.1467	2013.4430		92	110	0	(K)FSPIFSILSLVVGLFSGTK(A)
2013.0330	2014.6288	$1C_a$	200	215	1	(K)RMLTMLELFTDLCSLR(D)
2029.0279	2030.6282	$1M_o$ $1C_a$	200	215	1	(K)RMLTMLELFTDLCSLR(D)
2034.9978	**2036.2163**		**122**	**139**	**0**	**(K)AVQEQSDQELQEALYGVK(R)**
2036.9964	**2038.2783**		**260**	**275**	**1**	**(K)NDDKETYLFLSYLYPR(E)**
2045.0228	2046.6275	$2M_o$ $1C_a$	200	215	1	(K)RMLTMLELFTDLCSLR(D)
2121.1009	2122.4633		49	70	1	(K)VMGAIGSLSTAVGKFQSGDPAK(I)
2126.0665	2127.3789		362	378	2	(R)YGNYILFTHKRNDDLEK(F)
2137.0958	2138.4626	$1M_o$	49	70	1	(K)VMGAIGSLSTAVGKFQSGDPAK(I)
2149.1937	2150.6017		71	91	1	(K)IASGCLDILVGISSVLKDFAK(F)
2152.0339	2153.3901		29	46	2	(K)VVEMQKQIDAKFSNDDER(A)
2163.0928	**2164.3915**		**121**	**139**	**1**	**(K)KAVQEQSDQELQEALYGVK(R)**
2168.0288	2169.3894	$1M_o$	29	46	2	(K)VVEMQKQIDAKFSNDDER(A)
2191.0990	2192.4050		122	140	1	(K)AVQEQSDQELQEALYGVKR(E)
2220.2308	2221.8155	$1C_a$	71	91	1	(K)IASGCLDILVGISSVLKDFAK(F)
2236.1244	2237.4895		339	357	2	(K)LDKLNDGYYSIKTTQHDPK(I)
2256.1335	**2257.5667**		**403**	**423**	**0**	**(K)FAYVPQNADGALFFVDGIPSK(V)**
2294.1339	2295.5697		260	277	2	(K)NDDKETYLFLSYLYPREK(N)
2319.1939	2320.5802		121	140	2	(K)KAVQEQSDQELQEALYGVKR(E)
2320.2329	2321.7178		47	70	2	(R)AKVMGAIGSLSTAVGKFQSGDPAK(I)
2336.2279	2337.7171	$1M_o$	47	70	2	(R)AKVMGAIGSLSTAVGKFQSGDPAK(I)
2353.1816	2354.6605		40	62	2	(K)FSNDDERAKVMGAIGSLSTAVGK(F)
2369.1765	2370.6599	$1M_o$	40	62	2	(K)FSNDDERAKVMGAIGSLSTAVGK(F)
2389.1717	2390.6995		342	361	2	(K)LNDGYYSIKTTQHDPKICHR(Y)
2436.2194	2437.7313		264	282	2	(K)ETYLFLSYLYPREKNEQSR(K)
2460.2089	2461.9133	$1C_a$	342	361	2	(K)LNDGYYSIKTTQHDPKICHR(Y)
2473.3741	2474.9645		88	110	1	(K)DFAKFSPIFSILSLVVGLFSGTK(A)
2518.3585	2519.9816		63	87	1	(K)FQSGDPAKIASGCLDILVGISSVLK(D)
2572.2878	2573.9572		351	371	2	(K)TTQHDPKICHRYGNYILFTHK(R)
2589.3957	2591.1954	$1C_a$	63	87	1	(K)FQSGDPAKIASGCLDILVGISSVLK(D)
2629.3398	2631.1675		194	215	2	(K)TESEIKRMLTMLELFTDLCSLR(D)
2631.3261	2632.9049		154	178	0	(R)NETSDLSPTEVSALAANVPIYQGVR(F)
2643.3249	2645.1710	$1C_a$	351	371	2	(K)TTQHDPKICHRYGNYILFTHK(R)
2645.3347	2647.1668	$1M_o$	194	215	2	(K)TESEIKRMLTMLELFTDLCSLR(D)
2648.4658	2650.1098		216	240	0	(R)DLILLDLYQLVATPGHSPNIASGIK(E)
2661.3296	2663.1662	$2M_o$	194	215	2	(K)TESEIKRMLTMLELFTDLCSLR(D)
2700.3769	2702.3813	$1C_a$	194	215	2	(K)TESEIKRMLTMLELFTDLCSLR(D)
2716.3718	2718.3806	$1M_o$ $1C_a$	194	215	2	(K)TESEIKRMLTMLELFTDLCSLR(D)
2732.3667	2734.3800	$2M_o$ $1C_a$	194	215	2	(K)TESEIKRMLTMLELFTDLCSLR(D)
2811.3988	2813.1587		398	423	1	(K)ESPNKFAYVPQNADGALFFVDGIPSK(V)

m/z(mi)	m/z(av)	Mod.	Start	End	Missed Cleavage	Sequence
2868.4374	2870.1650		122	146	2	(K)AVQEQSDQELQEALYGVKREYAVSK(A)
2881.4658	2883.2911		253	275	2	(K)VFEDLLKNDDKETYLFLSYLYPR(E)
2979.5860	2981.5031		63	91	2	(K)FQSGDPAKIASGCLDILVGISSVLKDFAK(F)
2997.6547	2999.5397		92	120	1	(K)FSPIFSILSLVVGLFSGTKAEESVGSVVK(K)

MS-Digest in ProteinProspector 4.27.2basic
© Copyright (1995-2008) The Regents of the University of California

7.2 MS-Digest Results for CfTX-2

Parameters

```
Database: User Protein
Considered modifications: | Oxidation of M | Acrylamide Modified Cys |
Digest Used: Trypsin
Max. # Missed Cleavages: 3
User AA Formula 1: C2 H3 N1 O1
Minimum Digest Fragment Mass: 700
Maximum Digest Fragment Mass: 3000
Minimum Digest Fragment Length: 5

pI of Protein: 7.7
Protein MW: 49883
Amino Acid Composition: A26 C5 D28 E28 F24 G28 H6 I26 K41 L49 M5 N15 P16 Q18 R16 S43 T23
V27 W3 Y18

  1 ESTISSGLASLKAKIDIKKPTGKQLFDKVKSMEQALENKFSDDDERAKVMGAIGSLGTAIGKFQSGDPASIASGCLDILV
 81 GISSVLKDFAKFSPVFSILSLVVGLFSGTKAEESVSSVVTKAIQEQSDQELQEALYGVKREFAVSKAFLDGVRNEESDLR
161 PTEVSALAANIPVYQGVRFIAMVVQRIKYIKPKTESEIKRMLTMLELFTDLCSIRDLILLDLHQLIATPGHSPNIASGIK
241 EVTSLGREEYQRVFEDLLKTDDEETFLFLSYLYPKEKNEQSRKIFKFFDLIEVKYDDRFKLDLSGGQALSTLQWPNYYLC
321 PHNDYLANNCHDLRVGLKLEKLSDGFYTIKTYGRDPRTCYWTDDYVKISSTSNGELEKFSFVPVQVKGQKAYLLSTKKWP
401 HNFAYSQKTANGLLSILKDVPSKLGYGNQGFFTISTYSNPKNRHA
```

Table 7.2

Calculated mass to charge ratio (m/z) values for monoisotopic peptide ions (mi) and average fragment masses (av) obtained by theoretical digestion of CfTX-2 amino acid sequence with trypsin. Considered modifications include oxidation of methionine (M_o) and alkylation of cysteine (C_a). Calculated monoisotopic peptide ions masses that are similar to experimental masses obtained by MALDI-TOF-TOF MS analysis (see Section 6.2) and their corresponding sequences are highlighted in blue.

m/z(mi)	m/z(av)	Mod.	Start	End	Missed Cleavage	Sequence
706.3618	706.7752		194	199	0	(K)TESEIK(R)
724.3260	724.7525		248	252	0	(R)EEYQR(V)
761.3900	761.8174		278	283	1	(K)NEQSRK(I)
761.4152	761.8581		241	247	0	(K)EVTSLGR(E)
777.4254	**777.9038**		**147**	**153**	**0**	**(K)AFLDGVR(N)**
786.5084	786.9959		335	341	1	(R)VGLKLEK(L)
795.4611	795.9598		391	397	0	(K)AYLLSTK(K)
836.4625	836.9720		140	146	1	(K)REFAVSK(A)
843.3995	**843.9199**		**295**	**300**	**1**	**(K)YDDRFK(L)**
862.4629	862.9638		194	200	1	(K)TESEIKR(M)
863.4873	864.0355		253	259	0	(R)VFEDLLK(T)
864.4322	**864.9418**		**351**	**357**	**1**	**(K)TYGRDPR(T)**
877.5142	878.0655		24	30	1	(K)QLFDKVK(S)
883.3428	883.8515		40	46	0	(K)FSDDDER(A)
889.5870	890.1639		187	193	1	(R)IKYIKPK(T)
890.4326	890.9337		276	282	1	(K)EKNEQSR(K)

m/z(mi)	m/z(av)	Mod.	Start	End	Missed Cleavage	Sequence
923.5560	924.1350		391	398	1	(K)AYLLSTKK(W)
963.5444	**964.2252**		**179**	**186**	**0**	**(R)FIAMVVQR(I)**
979.5393	**980.2245**	**1M$_o$**	**179**	**186**	**0**	**(R)FIAMVVQR(I)**
999.6197	1000.2338		15	23	1	(K)IDIKKPTGK(Q)
1010.5557	**1011.2133**		**287**	**294**	**0**	**(K)FFDLIEVK(Y)**
1018.5276	1019.1088		276	283	2	(K)EKNEQSRK(I)
1029.6303	1030.2602		409	418	0	(K)TANGLLSILK(D)
1043.5408	**1044.1997**		**342**	**350**	**0**	**(K)LSDGFYTIK(T)**
1049.4932	1050.1826		31	39	0	(K)SMEQALENK(F)
1050.5982	**1051.2815**		**379**	**387**	**0**	**(K)FSFVPVQVK(G)**
1065.4881	1066.1820	1M$_o$	31	39	0	(K)SMEQALENK(F)
1082.4749	1083.1060		40	48	1	(K)FSDDDERAK(V)
1108.6361	1109.3187		388	397	1	(K)GQKAYLLSTK(K)
1135.5841	1136.2516		111	121	0	(K)AEESVSSVVTK(A)
1149.6375	1150.3309		278	286	2	(K)NEQSRKIFK(F)
1161.6626	1162.3827		19	28	1	(K)KPTGKQLFDK(V)
1164.5743	1165.2497		368	378	0	(K)ISSTSNGELEK(F)
1192.6420	1193.3475		1	12	0	(-)ESTISSGLASLK(A)
1198.7518	1199.4883		13	23	2	(K)AKIDIKKPTGK(Q)
1204.7235	1205.5608		179	188	1	(R)FIAMVVQRIK(Y)
1220.7184	1221.5602	1M$_o$	179	188	1	(R)FIAMVVQRIK(Y)
1236.7310	1237.4939		388	398	2	(K)GQKAYLLSTKK(W)
1274.7137	1275.5625		49	62	0	(K)VMGAIGSLGTAIGK(F)
1276.6566	1277.4912		29	39	1	(K)VKSMEQALENK(F)
1277.6062	1278.4212		399	408	0	(K)WPHNFAYSQK(T)
1290.7086	1291.5619	1M$_o$	49	62	0	(K)VMGAIGSLGTAIGK(F)
1292.6515	1293.4906	1M$_o$	29	39	1	(K)VKSMEQALENK(F)
1293.5456	1294.4340		358	367	0	(R)TCYWTDDYVK(I)
1335.7518	1336.5807		189	199	1	(K)YIKPKTESEIK(R)
1363.7732	1364.6405		379	390	1	(K)FSFVPVQVKGQK(A)
1364.5827	1365.6478	1C$_a$	358	367	0	(R)TCYWTDDYVK(I)
1388.8260	1389.6913		19	30	2	(K)KPTGKQLFDKVK(S)
1391.7740	1392.6020		1	14	1	(-)ESTISSGLASLKAK(I)
1398.8032	1399.7268		284	294	1	(K)IFKFFDLIEVK(Y)
1405.7011	1406.5964		398	408	1	(K)KWPHNFAYSQK(T)
1413.7624	1414.6516		339	350	1	(K)LEKLSDGFYTIK(T)
1438.7689	1439.6644		141	153	1	(R)EFAVSKAFLDGVR(N)
1466.7234	**1467.5879**		**241**	**252**	**1**	**(K)EVTSLGREEYQR(V)**
1473.8458	1474.8170		47	62	1	(R)AKVMGAIGSLGTAIGK(F)
1489.8407	1490.8164	1M$_o$	47	62	1	(R)AKVMGAIGSLGTAIGK(F)
1491.8530	1492.7694		189	200	2	(K)YIKPKTESEIKR(M)
1520.7744	1521.7235		342	354	1	(K)LSDGFYTIKTYGR(D)
1526.8981	1527.9020		283	294	2	(R)KIFKFFDLIEVK(Y)
1555.9054	1556.8541		409	423	1	(K)TANGLLSILKDVPSK(L)
1559.7740	1560.7575		287	298	1	(K)FFDLIEVKYDDR(F)
1568.7955	1569.7653		248	259	1	(R)EEYQRVFEDLLK(T)
1576.9309	1577.9164		187	199	2	(R)IKYIKPKTESEIK(R)
1594.8700	1595.8530		140	153	2	(K)REFAVSKAFLDGVR(N)
1630.9527	1631.9680		15	28	2	(K)IDIKKPTGKQLFDK(V)
1661.7264	1662.8292		355	367	1	(R)DPRTCYWTDDYVK(I)
1732.7635	1734.0430	1C$_a$	355	367	1	(R)DPRTCYWTDDYVK(I)
1785.8948	1787.2264		201	215	0	(R)MLTMLELFTDLCSIR(D)
1801.8897	1803.2257	1M$_o$	201	215	0	(R)MLTMLELFTDLCSIR(D)
1811.0313	**1812.1729**		**335**	**350**	**2**	**(R)VGLKLEKLSDGFYTIK(T)**
1817.8846	1819.2251	2M$_o$	201	215	0	(R)MLTMLELFTDLCSIR(D)
1834.1135	1835.3663		179	193	2	(R)FIAMVVQRIKYIKPK(T)
1834.9374	1836.1105		287	300	2	(K)FFDLIEVKYDDRFK(L)
1850.1085	1851.3657	1M$_o$	179	193	2	(R)FIAMVVQRIKYIKPK(T)
1856.9319	1858.4402	1C$_a$	201	215	0	(R)MLTMLELFTDLCSIR(D)
1861.0641	1862.1873		1	18	2	(-)ESTISSGLASLKAKIDIK(K)
1872.9268	1874.4395	1M$_o$,1C$_a$	201	215	0	(R)MLTMLELFTDLCSIR(D)
1888.9217	**1890.4389**	**2M$_o$,1C$_a$**	**201**	**215**	**0**	**(R)MLTMLELFTDLCSIR(D)**
1888.9552	1890.1188		342	357	2	(K)LSDGFYTIKTYGRDPR(T)
1890.9960	1892.1754		339	354	2	(K)LEKLSDGFYTIKTYGR(D)

m/z(mi)	m/z(av)	Mod.	Start	End	Missed Cleavage	Sequence
1907.9895	1909.2254		24	39	2	(K)QLFDKVKSMEQALENK(F)
1913.8182	1915.0114		31	46	1	(K)SMEQALENKFSDDDER(A)
1923.9844	1925.2248	$1M_o$	24	39	2	(K)QLFDKVKSMEQALENK(F)
1929.8131	1931.0108	$1M_o$	31	46	1	(K)SMEQALENKFSDDDER(A)
1941.9959	1943.4150		200	215	1	(K)RMLTMLELFTDLCSIR(D)
1948.0215	1949.2710		284	298	2	(K)IFKFFDLIEVKYDDR(F)
1957.9908	1959.4144	$1M_o$	200	215	1	(K)RMLTMLELFTDLCSIR(D)
1973.9857	1975.4137	$2M_o$	200	215	1	(K)RMLTMLELFTDLCSIR(D)
1980.9477	1982.2078		260	275	0	(K)TDDEETFLFLSYLYPK(E)
1993.9654	1995.2125		424	441	0	(K)LGYGNQGFFTISTYSNPK(N)
1998.1310	1999.4160		92	110	0	(K)FSPVFSILSLVVGLFSGTK(A)
2013.0330	2014.6288	$1C_a$	200	215	1	(K)RMLTMLELFTDLCSIR(D)
2029.0279	2030.6282	$1M_o,1C_a$	200	215	1	(K)RMLTMLELFTDLCSIR(D)
2045.0228	2046.6275	$2M_o,1C_a$	200	215	1	(K)RMLTMLELFTDLCSIR(D)
2049.0135	**2050.2434**		**122**	**139**	**0**	**(K)AIQEQSDQELQEALYGVK(R)**
2112.9502	2114.2659		31	48	2	(K)SMEQALENKFSDDDERAK(V)
2128.9452	2130.2652	$1M_o$	31	48	2	(K)SMEQALENKFSDDDERAK(V)
2138.9600	2140.3530		351	367	2	(K)TYGRDPRTCYWTDDYVK(I)
2140.2165	2141.5775		379	397	2	(K)FSFVPVQVKGQKAYLLSTK(K)
2140.9815	2142.3200		29	46	2	(K)VKSMEQALENKFSDDDER(A)
2156.9765	2158.3194	$1M_o$	29	46	2	(K)VKSMEQALENKFSDDDER(A)
2182.1444	2183.5334		391	408	2	(K)AYLLSTKKWPHNFAYSQK(T)
2196.1547	2197.5085		368	387	1	(K)ISSTSNGELEKFSFVPVQVK(G)
2205.1146	2206.4320		122	140	1	(K)AIQEQSDQELQEALYGVKR(E)
2209.9971	2211.5668	$1C_a$	351	367	2	(K)TYGRDPRTCYWTDDYVK(I)
2238.0853	2239.4992		260	277	1	(K)TDDEETFLFLSYLYPKEK(N)
2264.1095	2265.5056		424	443	1	(K)LGYGNQGFFTISTYSNPKNR(H)
2288.2186	2289.6587		399	418	1	(K)WPHNFAYSQKTANGLLSILK(D)
2311.1929	2312.6007		241	259	2	(K)EVTSLGREEYQRVFEDLLK(T)
2338.1707	2339.6458		40	62	2	(K)FSDDDERAKVMGAIGSLGTAIGK(F)
2354.1657	2355.6452	$1M_o$	40	62	2	(K)FSDDDERAKVMGAIGSLGTAIGK(F)
2416.3136	2417.8339		398	418	2	(K)KWPHNFAYSQKTANGLLSILK(D)
2439.1020	2440.6610		358	378	1	(R)TCYWTDDYVKISSTSNGELEK(F)
2459.3585	2460.9374		88	110	1	(K)DFAKFSPVFSILSLVVGLFSGTK(A)
2472.2055	2473.7269		424	445	2	(K)LGYGNQGFFTISTYSNPKNRHA(-)
2477.2956	2478.8851		63	87	0	(K)FQSGDPASIASGCLDILVGISSVLK(D)
2509.3297	2510.8675		368	390	2	(K)ISSTSNGELEKFSFVPVQVKGQK(A)
2510.1392	2511.8748	$1C_a$	358	378	1	(R)TCYWTDDYVKISSTSNGELEK(F)
2520.2405	2521.8063		419	441	1	(K)DVPSKLGYGNQGFFTISTYSNPK(N)
2548.3327	2550.0989	$1C_a$	63	87	0	(K)FQSGDPASIASGCLDILVGISSVLK(D)
2629.3398	2631.1675		194	215	2	(K)TESEIKRMLTMLELFTDLCSIR(D)
2636.4770	2638.1016		216	240	0	(R)DLILLDLHQLIATPGHSPNIASGIK(E)
2645.3347	2647.1668	$1M_o$	194	215	2	(K)TESEIKRMLTMLELFTDLCSIR(D)
2661.3296	2663.1662	$2M_o$	194	215	2	(K)TESEIKRMLTMLELFTDLCSIR(D)
2700.3769	2702.3813	$1C_a$	194	215	2	(K)TESEIKRMLTMLELFTDLCSIR(D)
2716.3718	2718.3806	$1M_o,1C_a$	194	215	2	(K)TESEIKRMLTMLELFTDLCSIR(D)
2728.3900	**2730.0253**		**154**	**178**	**0**	**(R)NEESDLRPTEVSALAANIPVYQGVR(F)**
2732.3667	2734.3800	$2M_o,1C_a$	194	215	2	(K)TESEIKRMLTMLELFTDLCSIR(D)
2790.3846	2792.0995		419	443	2	(K)DVPSKLGYGNQGFFTISTYSNPKNR(H)
2807.2829	2809.0562		355	378	2	(R)DPRTCYWTDDYVKISSTSNGELEK(F)
2814.4937	2816.2526		399	423	2	(K)WPHNFAYSQKTANGLLSILKDVPSK(L)
2825.4172	2827.2205		253	275	1	(R)VFEDLLKTDDEETFLFLSYLYPK(E)
2852.3625	2854.1187		260	282	2	(K)TDDEETFLFLSYLYPKEKNEQSR(K)
2866.4581	2868.1926		122	146	2	(K)AIQEQSDQELQEALYGVKREFAVSK(A)
2878.3200	2880.2700	$1C_a$	355	378	2	(R)DPRTCYWTDDYVKISSTSNGELEK(F)
2938.5230	2940.4066		63	91	1	(K)FQSGDPASIASGCLDILVGISSVLKDFAK(F)

MS-Digest in ProteinProspector 4.27.2basic
© Copyright (1995-2008) The Regents of the University of California

•

APPENDIX 8

Nucleotide and Deduced Amino Acid Sequences of *C. fleckeri* Venom Protein Expression Constructs

8.1 CfTX-1 in pProEXc

```
   1 - ATGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGAAAACCTGTATTTTCAGGGC - 72
   1 -  M  S  Y  Y  H  H  H  H  H  H  D  Y  D  I  P  T  T  E  N  L  Y  F  Q  G  - 24

  73 - GCCATGGGGATCCGGAATTCAAAGGCGGCTGAAAGCACGATTTCCTCCGGATTGAATTCATTAAAAACTAAG - 144
  25 -  A  M  G  I  R  N  S  K  A  A  E  S  T  I  S  S  G  L  N  S  L  K  T  K  - 48

 145 - ATAGATGCCAAAATGCCTTCAGGTAAACAACTTTTCGATAAAGTAGTGGAAATGCAAAAACAAATAGATGCC - 216
  49 -  I  D  A  K  M  P  S  G  K  Q  L  F  D  K  V  V  E  M  Q  K  Q  I  D  A  - 72

 217 - AAGTTTTCTAATGACGACGAAAGAGCCAAAGTAATGGGGGCTATTGGATCATTAAGCACTGCTGTAGGCAAG - 288
  73 -  K  F  S  N  D  D  E  R  A  K  V  M  G  A  I  G  S  L  S  T  A  V  G  K  - 96

 289 - TTCCAGTCAGGCGACCCGGCAAAGATTGCAAGTGGATGCTTGGATATTCTTGTTGGTATCTCCTCAGTTTTA - 360
  97 -  F  Q  S  G  D  P  A  K  I  A  S  G  C  L  D  I  L  V  G  I  S  S  V  L  - 120

 361 - AAAGACTTTGCAAAGTTTTCTCCAATCTTCTCGATCCTGTCACTGGTTGTCGGGCTGTTTTCTGGAACAAAA - 432
 121 -  K  D  F  A  K  F  S  P  I  F  S  I  L  S  L  V  V  G  L  F  S  G  T  K  - 144

 433 - GCCGAGGAGAGCGTCGGCTCTGTTGTTAAAAAGGCAGTCCAAGAACAATCTGATCAAGAGCTACAGGAAGCA - 504
 145 -  A  E  E  S  V  G  S  V  V  K  K  A  V  Q  E  Q  S  D  Q  E  L  Q  E  A  - 168

 505 - CTGTATGGTGTTAAAAGAGAATATGCAGTTTCCAAAGCATTCTTAGACGGAGTCAGAAACGAAACATCCGAT - 576
 169 -  L  Y  G  V  K  R  E  Y  A  V  S  K  A  F  L  D  G  V  R  N  E  T  S  D  - 192

 577 - CTTAGTCCTACTGAAGTCTCTGCGTTGGCTGCAAACGTTCCTATATACCAAGGCGTTCGTTTCATTGCAATG - 648
 193 -  L  S  P  T  E  V  S  A  L  A  A  N  V  P  I  Y  Q  G  V  R  F  I  A  M  - 216

 649 - GTAGTTCAAAGAATCAAATACATAAAACCAAAAACTGAAAGCGAGATCAAAAGAATGTTAACAATGCTGGAA - 720
 217 -  V  V  Q  R  I  K  Y  I  K  P  K  T  E  S  E  I  K  R  M  L  T  M  L  E  - 240

 721 - TTATTTACAGATCTCTGCAGTTTACGTGACTTGATACTTCTTGACTTATATCAACTAGTAGCTACCCCAGGC - 792
 241 -  L  F  T  D  L  C  S  L  R  D  L  I  L  L  D  L  Y  Q  L  V  A  T  P  G  - 264

 793 - CATAGTCCTAATATTGCATCTGGCATTAAAGAAGTCTCGAACCTTGGTCGAGAGGAATACAAGAAAGTTTTT - 864
 265 -  H  S  P  N  I  A  S  G  I  K  E  V  S  N  L  G  R  E  E  Y  K  K  V  F  - 288

 865 - GAAGATCTTTTGAAAAATGACGACAAGGAAACTTATTTGTTCTTATCTTACTTATATCCAAGAGAAAAAAAT - 936
 289 -  E  D  L  L  K  N  D  D  K  E  T  Y  L  F  L  S  Y  L  Y  P  R  E  K  N  - 312

 937 - GAACAAAGCCGAAAGATTTTCAATTTCTTTGATCTGATGAAAGTTAAATATGATGATCGTCTGAAACAAGAT - 1008
 313 -  E  Q  S  R  K  I  F  N  F  F  D  L  M  K  V  K  Y  D  D  R  L  K  Q  D  - 336

1009 - TTAACAGGAGTAAAAATCTTCTCCAATGTACATTGGCCAAATTACTTCATGTGCTCAAGTAATGACTACCTT - 1080
 337 -  L  T  G  V  K  I  F  S  N  V  H  W  P  N  Y  F  M  C  S  S  N  D  Y  L  - 360

1081 - GCTCTCATTTGCACCAAACCTTATGGGTCACTCAAGCTGGATAAATTAAATGATGGATATTATTCTATAAAA - 1152
 361 -  A  L  I  C  T  K  P  Y  G  S  L  K  L  D  K  L  N  D  G  Y  Y  S  I  K  - 384

1153 - ACAACCCAGCATGATCCAAAAATTTGCCATAGGTATGGTAATTATATTTTATTTACACATAAAAGGAATGAC - 1224
 385 -  T  T  Q  H  D  P  K  I  C  H  R  Y  G  N  Y  I  L  F  T  H  K  R  N  D  - 408

1225 - GATCTTGAGAAGTTTAATTTTGTTCCAGTGAAACTGGAAAAAAGAGAGATCTACCTACTGAGCTCAAAAGAA - 1296
 409 -  D  L  E  K  F  N  F  V  P  V  K  L  E  K  R  E  I  Y  L  L  S  S  K  E  - 432

1297 - AGTCCAAATAAGTTTGCATACGTACCGCAAAATGCAGATGGAGCCCTGTTTTTCGTGGATGGTATCCCAAGT - 1368
 433 -  S  P  N  K  F  A  Y  V  P  Q  N  A  D  G  A  L  F  F  V  D  G  I  P  S  - 456

1369 - AAAGTTGGCTACGGAAACCAAGGATATTTTACTCTTGTCGAGTAAGAGAATGAGCAAGTATTGTTGTTTTCC - 1440
 457 -  K  V  G  Y  G  N  Q  G  Y  F  T  L  V  E  *                             - 470
1441 - AAAAGAACACAGGCGTGTCAAATTATTTCATGCGAACAATTCAAAAGGACACTATATGGATAACTTCAAACG - 1512
1513 - CG                                                                       - 1514
```

8.2 CfTX-2 in pProEXc

```
   1 - ATGTCGTACTACCATCACCATCACCATCACGATTACGATATCCCAACGACCGAAAACCTGTATTTTCAGGGC - 72
   1 -  M  S  Y  Y  H  H  H  H  H  H  D  Y  D  I  P  T  T  E  N  L  Y  F  Q  G  - 24

  73 - GCCATGGGGATCCGGAATTCAAAGGCGAGTGAAAGCACAATTTCCTCTGGACTAGCTTCACTGAAAGCAAAG - 144
  25 -  A  M  G  I  R  N  S  K  A  S  E  S  T  I  S  S  G  L  A  S  L  K  A  K  - 48

 145 - ATAGATATAAAGAAACCAACTGGAAAGCAACTTTTTGATAAAGTAAAAAGTATGGAACAAGCTTTAGAAAAC - 216
  49 -  I  D  I  K  K  P  T  G  K  Q  L  F  D  K  V  K  S  M  E  Q  A  L  E  N  - 72

 217 - AAGTTTTCTGATGATGACGAACGAGCCAAAGTAATGGGGGCTATTGGATCATTAGGAACTGCTATAGGCAAG - 288
  73 -  K  F  S  D  D  D  E  R  A  K  V  M  G  A  I  G  S  L  G  T  A  I  G  K  - 96

 289 - TTCCAGTCAGGCGACCCGGCATCGATTGCAAGTGGATGCTTGGATATTCTTGTTGGTATCTCTTCCGTTTTG - 360
  97 -  F  Q  S  G  D  P  A  S  I  A  S  G  C  L  D  I  L  V  G  I  S  S  V  L  - 120

 361 - AAAGATTTTGCAAAGTTCTCTCCCGTCTTCTCAATCCTGTCACTGGTTGTTGGTCTATTTTCTGGAACCAAG - 432
 121 -  K  D  F  A  K  F  S  P  V  F  S  I  L  S  L  V  V  G  L  F  S  G  T  K  - 144

 433 - GCCGAGGAGAGTGTCAGCTCTGTAGTTACGAAGGCAATCCAAGAACAGTCCGACCAAGAGCTACAAGAAGCA - 504
 145 -  A  E  E  S  V  S  S  V  V  T  K  A  I  Q  E  Q  S  D  Q  E  L  Q  E  A  - 168

 505 - CTGTATGGTGTCAAAAGGGAATTTGCAGTTTCTAAAGCATTCTTAGACGGAGTAAGAAATGAAGAGTCCGAT - 576
 169 -  L  Y  G  V  K  R  E  F  A  V  S  K  A  F  L  D  G  V  R  N  E  E  S  D  - 192

 577 - CTTAGACCAACAGAAGTCTCTGCTTTAGCAGCAAACATTCCAGTATACCAAGGCGTTCGTTTCATTGCAATG - 648
 193 -  L  R  P  T  E  V  S  A  L  A  A  N  I  P  V  Y  Q  G  V  R  F  I  A  M  - 216

 649 - GTAGTTCAAAGAATCAAATATATAAAGCCCAAAACTGAAAGTGAGATCAAAAGAATGTTAACAATGCTGGAA - 720
 217 -  V  V  Q  R  I  K  Y  I  K  P  K  T  E  S  E  I  K  R  M  L  T  M  L  E  - 240

 721 - TTATTTACAGATCTCTGCAGTATACGTGACTTGATACTTCTTGACCTTCATCAACTAATAGCTACTCCAGGA - 792
 241 -  L  F  T  D  L  C  S  I  R  D  L  I  L  L  D  L  H  Q  L  I  A  T  P  G  - 264

 793 - CATAGTCCTAATATCGCATCTGGCATTAAAGAGGTTACGAGTCTCGGTAGAGAAGAATACCAAAGGGTTTTT - 864
 265 -  H  S  P  N  I  A  S  G  I  K  E  V  T  S  L  G  R  E  E  Y  Q  R  V  F  - 288

 865 - GAAGATCTTTTGAAAACTGACGACGAAGAGACTTTTTTGTTCTTATCCTACTTATATCCCAAAGAAAAAAAT - 936
 289 -  E  D  L  L  K  T  D  D  E  E  T  F  L  F  L  S  Y  L  Y  P  K  E  K  N  - 312

 937 - GAGCAAAGCCGAAAAATTTTCAAGTTCTTTGATCTGATTGAAGTTAAATATGATGACCGTTTCAAACTGGAT - 1008
 313 -  E  Q  S  R  K  I  F  K  F  F  D  L  I  E  V  K  Y  D  D  R  F  K  L  D  - 336

1009 - TTATCTGGTGGACAAGCTTTGTCCACTTTACAGTGGCCTAATTACTATTTGTGTCCTCACAATGACTACCTT - 1080
 337 -  L  S  G  G  Q  A  L  S  T  L  Q  W  P  N  Y  Y  L  C  P  H  N  D  Y  L  - 360

1081 - GCTAACAATTGCCACGATCTTCGTGTAGGTCTTAAGCTAGAAAAATTAAGCGATGGATTCTATACCATAAAG - 1152
 361 -  A  N  N  C  H  D  L  R  V  G  L  K  L  E  K  L  S  D  G  F  Y  T  I  K  - 384

1153 - ACCTATGGAAGAGATCCAAGAACATGCTATTGGACTGATGATTACGTCAAGATATCAAGCACAAGCAATGGG - 1224
 385 -  T  Y  G  R  D  P  R  T  C  Y  W  T  D  D  Y  V  K  I  S  S  T  S  N  G  - 408

1225 - GAGCTAGAAAAGTTTAGCTTTGTTCCCGTACAAGTGAAGGGTCAAAAAGCTTACTTGCTGTCGACTAAAAAA - 1296
 409 -  E  L  E  K  F  S  F  V  P  V  Q  V  K  G  Q  K  A  Y  L  L  S  T  K  K  - 432

1297 - TGGCCACATAATTTTGCATACTCACAAAAAACAGCTAATGGGCTTCTGTCAATACTAAAAGATGTCCCAAGC - 1368
 433 -  W  P  H  N  F  A  Y  S  Q  K  T  A  N  G  L  L  S  I  L  K  D  V  P  S  - 456

1369 - AAACTTGGTTACGGAAACCAAGGATTTTTCACAATTTCTACATACTCAAACCCAAAGAATAGACATGCTTAG - 1440
 457 -  K  L  G  Y  G  N  Q  G  F  F  T  I  S  T  Y  S  N  P  K  N  R  H  A  *  - 479

1441 - CTGTTGTTGAAAATAACACAGATGTATCTAGGTATTTTTTTAGACGGTTGAAAAGGGCACAGTTTTGATGG - 1512
1513 - TAAACTTCAAGTGTGTATTTATATCAATAACTTTCAGAAGAGGAAACGAGTTTGGCAGATGAGATTGTACTG - 1584
1585 - AACGTGTTTTGAAAAGCAGAATCGACATAAACGTAAAAAAAAAAAAAAA                        - 1633
```